UNIVERSITY
of
GLASGOW

Characterisation of the *Mycobacterium tuberculosis* DrrABC ATP-binding cassette transporter

By

Angus A. Nash B.Sc. (Hons)

A thesis presented for the degree of Doctor of Philosophy

in

The Faculty of Biomedical and Life Sciences

University of Glasgow

Faculty of Biomedical and Life Sciences

Division of Infection and Immunity

University of Glasgow

G12 8QQ

September 2003

The author was a recipient of an Engineering and Physical Sciences Research Council studentship. Except where stated otherwise all results reported in this thesis were obtained by the author's own efforts

Angus Alisdair Nash

20th September 2003

Acknowledgements

I would like to give my greatest thanks to Prof. Adrian R. Walmsley for giving me a chance and for his constant support and encouragement throughout this project.

None of the work presented here would have been achieved without the additional assistance of a huge number of people. My particular thanks must go to members of our research group past and present. Dr. Chris Gray, Dr. M. Ines Borges-Walmsley, Dr. Gareth Evans and Dr. Kenny McKeegan have all been there with words of wisdom and tremendous advice at times when they were most needed. Further thanks go to all the research and technical staff from the universities of Durham and Glasgow.

Finally I would like to thank all of my family and friends who have had to put up with me over the last few years. To my Mom, Dad and Brother, your love (and financial support!) is what has kept me going through even the hardest times. Without you none of this would have been possible.

"I went from animals to cells, from cells to bacteria,
from bacteria to molecules, from molecules to electrons.
This story has its irony since molecules and electrons
have no life at all.
Somewhere along the journey I let life run out between
my fingers."

Albert Szent-Györgi, 1972

This thesis is dedicated to my mother, Dr. Deirdre MacEachern, and to the memory of Alexander Wade, a great friend, dearly loved and sorely missed.

Summary

Mycobacterium tuberculosis, the aetiological agent of pulmonary tuberculosis (TB), is a re-emerging threat to global public health. TB continues to account for in excess of 2 million deaths annually with mortality and morbidity set to rise over the next decade as a consequence of the global HIV/AIDS pandemic. Many of the problems associated with treating this ancient disease are related to the intrinsic drug resistance of the causal organism, a factor that is compounded by the relatively small number of effective chemotherapeutic agents available. The current situation is further exacerbated by the emergence of multi-drug resistant strains of *M. tuberculosis* (MDR-TB) in certain parts of south-east Asia and eastern Europe. The global spread of such strains, fuelled by the increased availability of air travel, represents a serious threat to even the rich, industrialised nations of the West since successful treatment of MDR-TB requires the use of expensive and highly toxic drugs.

These various factors underline the urgent need for new and effective TB drug treatments. The development of such agents will undoubtedly require a much-improved understanding of *M. tuberculosis* at the most basic biochemical and genetic levels. To this end the completion of the TB genome sequencing project in 1998 represents a major breakthrough. The genome provides not only a detailed blueprint of every potential drug target that might be used in treatment of the disease, but also every conceivable drug resistance mechanism employed by the organism to evade our current treatments.

Active efflux is one of the key strategies used by microorganisms to avoid the deleterious effects of xenotoxic compounds, such transport processes being catalysed by an array of membrane-associated proteins. The ATP-binding cassette (ABC) transporters have been identified as one of the largest and most widely distributed families of such transmembrane transport systems. Virtually ubiquitous in nature, ABC transporters have been found in the genomes of every organism from the simplest archaea through to man. Whilst the vast majority of ABC transporters mediate standard cellular processes such as nutrient acquisition, a significant number have been shown to exhibit a broad substrate range and the capability to transport structurally diverse

molecules. It is the expression of these broad-substrate transporters that is most commonly linked with drug resistance phenotypes.

Little is currently known about the contribution of ABC transporter systems to the intrinsic drug resistance of *M. tuberculosis*. In fact, despite the presence of a large number of ABC transporters in the genome, very little is known about the biological role of *any M. tuberculosis* ABC transporter. Against this background it was the remit of this thesis to investigate the structure and function of just one such system. The *M. tuberculosis drr*ABC operon was identified during the genome sequencing project as encoding a potential antibiotic resistance mechanism. The basis of this assumption was a significant level of sequence homology between the *drr*ABC genes and those encoding a known antibiotic efflux ABC transporter from the evolutionarily related organism *Streptomyces peucetius*. The overall experimental strategy adopted was to clone and over-express the three genes of the *drr*ABC operon in a non-pathogenic heterologous host (*E. coli*). It was hoped that this approach might generate large enough quantities of the individual proteins to investigate not only the biological function of the DrrABC transporter but also the structure of isolated components.

Expression of the mycobacterial genes in *E. coli* proved to be a challenging undertaking, differences in the genetics of the host and donor organisms making co-expression of all three genes impossible. Expression of the individual proteins in isolation was somewhat more successful, with the ATP-binding sub-unit of the transporter, the DrrA protein, over-expressed and purified as a catalytically active fusion protein. The membrane-associated DrrB protein was successfully expressed in *E. coli* yet proved highly unstable and resistant to extraction from the membrane. The same problems applied to the second membrane-bound component of the transporter, DrrC. Issues of low expression and protein instability are not uncommon when working with membrane proteins and it appears that the *M. tuberculosis* DrrB and C proteins are no different in this respect. The over-expression and purification of DrrA allowed a partial characterisation of the catalytic activity of the protein using traditional biochemical methods. As a cation-dependent ATPase DrrA was shown to exhibit biological activity broadly consistent with a role as the catalytic sub-unit of an ABC transporter. As such DrrA is only the second *M. tuberculosis* protein for which such activity has been demonstrated.

The exact biological function of the DrrABC transporter remains a matter of some doubt. Research by one group of scientists suggest that the transporter is indeed a functional antibiotic efflux mechanism, whilst another group argue that its primary role is the transmembrane transport of an important mycobacterial cell wall component. Whilst these two roles are not mutually exclusive, they do suggest that DrrABC may yet prove to be an attractive target for drug intervention. Unambiguous demonstration of the roles of this transporter, and the many other *M. tuberculosis* transmembrane transport systems, will require a great deal of further research and the development of improved expression systems.

Table of Contents

Chapter 1 – General Introduction – Tuberculosis and the mycobacteria

1. 1. 1 The global burden of tuberculosis

In April 1993 the World Health Organisation (WHO) declared tuberculosis (TB) to be a "global health emergency". The basis for this unprecedented action was a world-wide resurgence in tuberculosis which, by 1998, was accounting for more than 2 million deaths annually (WHO Fact Sheet 104, 2002). Whilst the global HIV/AIDS pandemic is now believed to have overtaken TB as the leading cause of death due to an infectious disease, tuberculosis still kills more adults each year than diarrhoea, malaria and all of the tropical diseases combined. The vast majority of deaths due to TB still occur in the developing world (Raviglione *et al.*, 1995). However, the period 1985 to 1992 also saw an increase in the number of new TB cases presented in the industrialised nations of the West. For example, cases in the USA alone rose by 20% during this period (with New York City at the centre of several serious outbreaks), reversing a general decline in new cases of TB that had been taking place since 1953 (Cantwell *et al.*, 1994). A combination of factors, some of which are discussed later, are thought to be responsible for the observed global resurgence of TB. However, there is no doubt that the re-emergence of TB at the very heart of prosperous Western societies has prompted governments and scientists alike to once again focus significant efforts into the control and treatment of this disease.

1. 1. 2 Tuberculosis and the Mycobacteria

Famously described as "Captain of all the men of death" by evangelist Robert Bunyan, tuberculosis had been known as a deadly disease since antiquity. Despite this, it was not until 1882 that medical science was successfully able to identify and culture the etiological agent of the disease. This breakthrough came with the description of the '*Tubercule bacillus*' by Robert Koch, the man now widely regarded as the father of modern microbiology. We now know the organism as a member of the genus *Mycobacterium*, the vast majority of which are non-pathogenic, free-living, environmental saprophytes. However, a number of species of mycobacteria have evolved as either obligate or opportunistic pathogens capable of causing disease in both humans and animals. In addition to *Mycobacterium tuberculosis*, the causal agent of human TB, a second major pathogen of the genus is *Mycobacterium leprae*, the organism responsible for human leprosy (Sasaki *et al.*, 2001). Whilst the mycobacteria

form a relatively small and distinct genus, they are most closely related, in evolutionary terms, to the genus *Streptomyces*. Streptomycetes are a group of spore-forming, filamentous bacteria, many of which produce a variety of biologically active secondary metabolites e.g. antibiotics (Baltz, R.H., 1998).

1. 1. 3 The pathology of tuberculosis

Symptoms of 'active' TB include fits of coughing, sometimes bringing up blood, chest pain, exhaustion, fever, profuse sweating at night and weight loss (Tattevin *et al.*, 1999). Left untreated the disease causes progressive destruction of the lung tissues and eventually death. The disease is transmitted when infected aerosols are expelled from the lungs of patients and are subsequently inhaled by those around them (Bermudez and Goodman, 1996). This mode of transmission means that the lungs are the most frequently affected organs, although infection can then spread to other areas of the body including the brain, blood and heart (von Reyn *et al.*, 2001). These extra-pulmonary forms of TB are not widely regarded as infectious.

Whilst it is estimated that up to one third of the world's population may be infected with *Mycobacterium tuberculosis*, only 10% of those infected ever develop the active disease (Sudre *et al.*, 1992). This is because the immune system of a healthy individual is usually able to combat the infection. It does so by isolating infected cells, most frequently lung alveolar macrophages, into thick walled, fibrous structures known as *granulomas* (Bermudez and Goodman, 1996). Once isolated in this fashion the bacteria are unable to spread to other areas of the lung, nor are they expelled as in the case of active TB (Saunders *et al.*, 1999). The bacteria are not necessarily destroyed within the granuloma, and appear to have evolved mechanisms for avoiding clearance by the immune system. A thick, waxy cell wall and a number of scavenging molecules may play a role in the ability of mycobacteria to survive in a dormant state whilst encapsulated in this hostile environment (Couture *et al.*, 1999). The exact mechanisms of host-pathogen interaction within the granuloma may one day represent a target for drug intervention and treatment of TB.

1. 1. 4 Risk factors and the spread of TB

The WHO has identified a number of factors that have contributed to the global increase in new TB cases. Social factors such as homelessness, poverty, malnutrition, drug

addiction and incarceration all have long been recognised as risk factors in the development of TB. This may be because the immune system of these individuals is significantly weakened, compromising their ability to fight off initial infection, or facilitating the reactivation of a latent infection. Such social factors may also force many susceptible people to live together in large groups, thereby allowing rapid spread of infection via the aerosol route.

Severe immuno-compromisation due to HIV/AIDS is thought to be another of the major factors fuelling the global increase in deaths due to TB, particularly in sub-Saharan Africa and South East Asia where TB now accounts for 40% of AIDS related deaths (WHO fact sheet 104, 2002). The two infections seem to have a deadly synergistic effect, with AIDS leaving sufferers both more susceptible to initial infection and less able to fight it. HIV co-infection may also allow reactivation of a dormant mycobacterial infection, increase the chances of the disease spreading to other areas of the body, and can leave patients open to opportunistic infection by other mycobacterial species not usually recognised as major human pathogens (Zurawski *et al.*, 1997). A 31 month study amongst intravenous drug users in New York City has suggested that the likelihood of developing active TB increases from a 10% lifetime chance in immuno-competent patients, to a 7-10% *yearly* chance in HIV/TB infected patients (Selwyn *et al.*, 1989). These figures now make HIV infection the number one risk factor in the development of active TB disease.

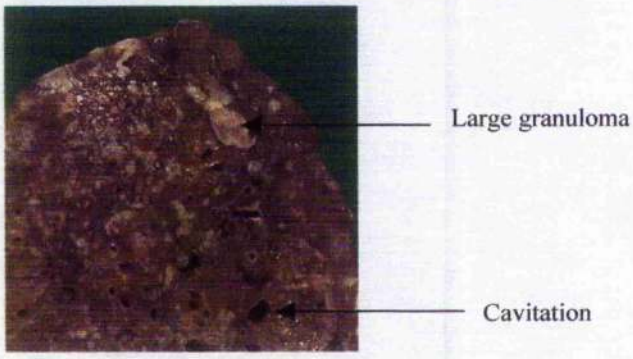
The increasing availability of air travel and large volumes of displaced peoples, such as refugees, have facilitated the spread of TB across national borders and throughout the world. These cases represent a significant reservoir of infection as they move from one part of the world to another. In the industrialised nations, where TB is relatively rare, a large proportion of TB cases are amongst foreign-born nationals. The WHO cites this as another reason why the affluent nations of the West cannot afford to ignore the growing burden of tuberculosis (WHO Fact Sheet 104, 2002).

1. 1. 5 Clinical diagnosis of *M. tuberculosis* infection

The identification of patients infected with *M. tuberculosis*, whether they exhibit active disease or not, is of vital importance in controlling the spread of TB on both local and global scales. The *Tuberculin Skin Tests* (TST), and variants the *Mantoux* and *Heaf*

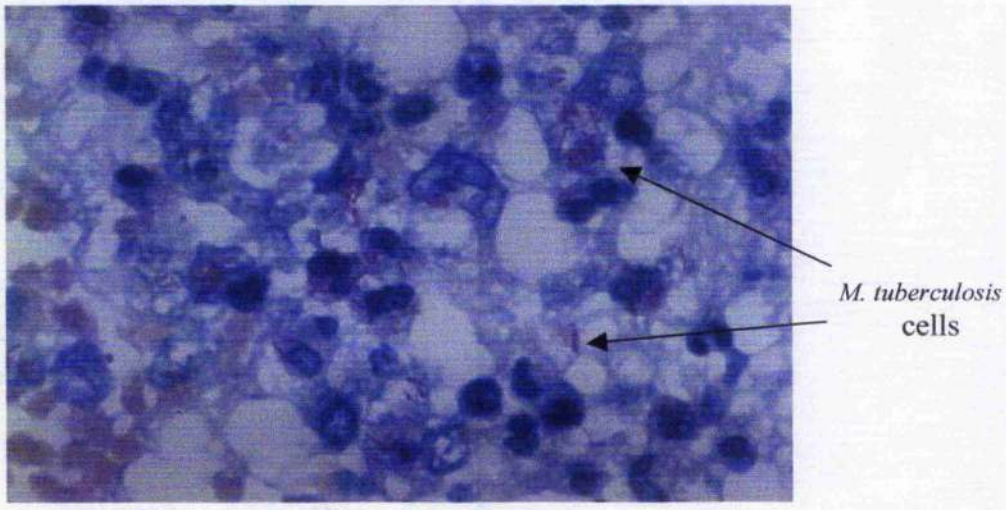
tests, are the most commonly used tools for screening large populations for potential exposure to *M. tuberculosis*. These tests rely upon a hypersensitivity reaction to an intradermal injection of a protein mixture derived from autoclaved cultures of *M. tuberculosis*. The extent of an individual's immune reaction to the injected protein is measured 48-72 hours after the injection. If a patient has latent *M. tuberculosis* infection, there will be an extensive reaction due to the presence of antibodies in the patient's blood. A significant limitation of hypersensitivity testing of this type is that false results, both positive and negative, are quite common. False positives may occur as a result of the cross-reactivity of antibodies that have been produced as a result of exposure to non-pathogenic, environmental mycobacteria. False negative results, a particularly dangerous form of failure, can occur in elderly people, HIV infected patients or very young children. This is because these groups of people display a generally weakened immune response. The TST may also fail in patients who have become infected with *M. tuberculosis* less than 10 weeks before testing. This is because the bacteria grow slowly within the body and it can take 8-10 weeks for the bacterial load to reach a threshold level capable of eliciting an immune response (Farr, B.M., 2001). Regardless of the result of a TST, those patients exhibiting the symptoms of active pulmonary TB should be given a chest X-ray. Usually, when a person has TB disease in the lungs, the chest X-ray will appear abnormal. It may show infiltrates (collections of fluid and cells in the tissues of the lung) or cavities (hollow spaces within the lung that may contain bacilli). It is important to note that an X-ray is not a definitive tool for diagnosis since other diseases can give rise to abnormalities, nor will an X-ray detect the early stages of *M. tuberculosis* infection. Another problem is that X-ray facilities are by no means widespread in certain areas of the developing world where TB is now endemic.

Figure 1.1 – Lung damage caused by TB infection



The 'gold standard' for clinical diagnosis of *M. tuberculosis* infection is a complete bacteriological examination. This entails collection of a clinical specimen (e.g. lung sputum for suspected pulmonary TB), staining and microscopic examination of the specimen, and culture of the infective organism. A useful diagnostic feature of the mycobacteria is that they are *acid fast* to particular stains, i.e. remain strongly stained even after a smear slide is washed in acid. The detection of many acid fast bacilli in a sputum smear means that the patient is highly likely to be contagious since it is these bacilli that may be aerosolised during fits of coughing. A variety of mycobacterial species are acid fast, and thus the test is not necessarily a direct indication of *M. tuberculosis* infection. However, since the results of a smear test are available to the clinician within a day, this test may allow potentially infectious patients to be isolated before they can pass the infection on to others. The final and definitive step of diagnosis is the culture of organisms from the clinical specimen under conditions specific for the growth of *M. tuberculosis*. Because of the slow growth of the organism a definite result may take 4-6 weeks to obtain, a period in which the affected individual should ideally be isolated in order that they do not infect anybody else. In certain instances drug susceptibility testing may also be necessary in order to identify a drug regime that will be effective in clearing the infection. Such investigations are performed by growing clinical isolates on solid media in the presence and absence of antibiotics and can take an additional 3 to 6 weeks to complete.

Figure 1.2 – Lung sputum smear showing acid fast bacilli



The slow growth of *M. tuberculosis* and the absence of reliable, specific immmunodiagnostic tests represent major problems in the diagnosis of pulmonary TB. The same problem also contributes to the high rates of morbidity and mortality associated with disseminated and meningeal TB, where rapid diagnosis and treatment are key factors in the prognosis of the patient. As a result of these problems, researchers are developing novel molecular assays for identifying and typing mycobacterial infections. It is hoped that these techniques, some of which are discussed below, will allow diagnosis to become far more rapid and specific.

1. 1. 6 Molecular diagnostics

The increased risk of TB infection in the developed world, and the concomitant increase in funding that basic research has received, has lead to some improved methods for the early detection and diagnosis of the disease. The advent of PCR in 1987 offered a potential means for the rapid amplification of specific nucleic acid sequences from clinical samples. Soon applied to TB, genomic techniques of this kind ultimately aim to identify and amplify species-specific sequences that are not present in other closely related bacteria. The first sequence successfully employed was that encoding the 65 kDa heat-shock protein (Hance *et al.*, 1989). Subsequently a variety of amplification targets have been described, including mycobacterial antigens, repetitive sequences and ribosomal RNA genes (Woods, G.L., 1999). One of the most common targets is a mobile genetic element unique to 5 species of mycobacteria that constitute the so-called *M. tuberculosis complex*. Insertion sequence *6110* has the major benefit of being present in multiple copies in the genome of these organisms. IS*6110* has also proven to be useful in strain typing, with the number of repeats and restriction fragment length polymorphisms associated with the sequence indicative of individual strains. The US FDA currently approves three nucleic acid amplification tests for the identification of *M. tuberculosis* in clinical samples. These tests have been found to be more sensitive than staining and microscopy techniques, yet tend to perform worse than the gold standard culture procedure (Bonington *et al.*, 1998; Brown *et al.*, 1999; Scarparo *et al.*, 2000). Whilst problems exist with the use of these systems in poorer countries, in an industrialised setting they offer the major advantage that the results are usually available within 24 hours and can detect the presence of mycobacteria in smear negative samples.

Progress has also been made in respect of culture techniques for the slow growing mycobacteria. A radiometric system based upon the release of $^{14}CO_2$ from labelled liquid growth media by the action of mycobacterial metabolism was one of the first systems to be licensed. More recently systems based upon oxygen consumption and redox indicators have been employed to detect the growth of mycobacteria in liquid culture. These latter methods enjoy the same advantages as the radiometric system without the need for the use of radiochemicals (Heifets *et al.*, 2000, Liu *et al.*, 1999). Such systems, combined with nucleic acid amplification tests, now allow the detection of *M. tuberculosis* in 12-20 days, comparing favourably with conventional culture and identification techniques which may take up to 6 weeks (Katila *et al.*, 2000). Liquid culture systems are also now being used for drug susceptibility profiling, replacing the traditional solid media technique (Rusch-Gerdes *et al.*, 1999).

Basic research has also spawned a genuinely novel system, currently under evaluation, for the monitoring of mycobacterial growth and drug susceptibility profiling. Engineered mycobacteriophages producing fire-fly luciferase are used to infect early stage liquid cultures growing in the presence and absence of antibiotics. Viable bacilli proliferate and produce light when luciferin is added to the culture media. The amount of light produced by each culture is measured and provides an accurate guide to the number of viable bacilli present. Susceptible cultures can exhibit as much as a 90% decrease in the light signal produced (Riska *et al.*, 1999).

As with all novel technologies, continual evaluation and innovation will be required if molecular diagnostics are ever to find widespread use. This is particularly relevant with regard to TB diagnosis since the bulk of the global TB burden is located in third world or developing countries where facilities, technical expertise and financial resources are limited.

1. 1. 7 Antimycobacterial chemotherapy

Less than 60 years ago there were no effective antimycobacterial drugs. However, the beginning of the antibiotic era was a time of great optimism since it heralded the introduction of the first effective treatments for TB. Streptomycin, discovered in 1944 by Selman Waksman, was initially found to be a highly successful treatment for those suffering from tuberculosis. Unfortunately, it soon became apparent that treating the

disease with this drug alone was insufficient. It was found that complete elimination of the bacillus was extremely difficult, requiring both high doses and long periods of treatment. This in turn lead to problems of streptomycin toxicity, and unpleasant side effects such as hearing loss. In addition it was found that resistant strains of the bacterium would develop rapidly, being present in approximately 80% of patients after just 3 months of treatment with streptomycin. The problem of streptomycin resistance was eventually overcome during the 1950's. It was noted that treatment of patients with p-aminosalycilic acid (PAS), a derivative of asprin, as well as streptomycin, was highly effective in suppressing the emergence of streptomycin resistant bacteria. This discovery was a key factor in establishing 'multidrug' or 'combination' therapy as the most effective means of treating tuberculosis.

The next breakthrough in TB chemotherapy came in 1952 with the introduction of the synthetic drug *Isoniazid*. This drug, now known to be a highly effective inhibitor of mycobacterial cell wall synthesis, is still the most commonly prescribed antimycobacterial agent. In fact, isoniazid forms the cornerstone of most modern multidrug treatment regimens used to treat tuberculosis and is still the most effective drug for targeting actively dividing bacilli. At this point in time it was discovered that a regimen of all three drugs, isoniazid, streptomycin and PAS, could be used to successfully treat pulmonary tuberculosis in just 18 months, without the emergence of resistant bacteria.

Rifampicin, a bacterial RNA polymerase inhibitor, was the next drug to be pressed into service against tuberculosis infection. Rifampicin had the major advantage of being active against dormant *M. tuberculosis* as well as actively dividing cells. It is the incomplete elimination of 'persisting' bacilli of this type that is most commonly responsible for patient relapses after the completion of a therapeutic course. A highly lipophilic molecule, rifampicin is able to cross the thick cell wall of *M. tuberculosis* to its site of action in the cytosol. Despite this, certain species of non-tuberculous mycobacteria are intrinsically resistant to rifampicin. This feature is often attributed to species-specific differences in cell wall structure and means that rifampicin and its derivatives are not suitable for the treatment of all mycobacterial infections. As with most antimycobacterial drugs, resistance arises rapidly if rifampicin is used alone, but,

introduced as part of a multidrug therapy regime allows the period of treatment to be shortened from 18 months to just 9.

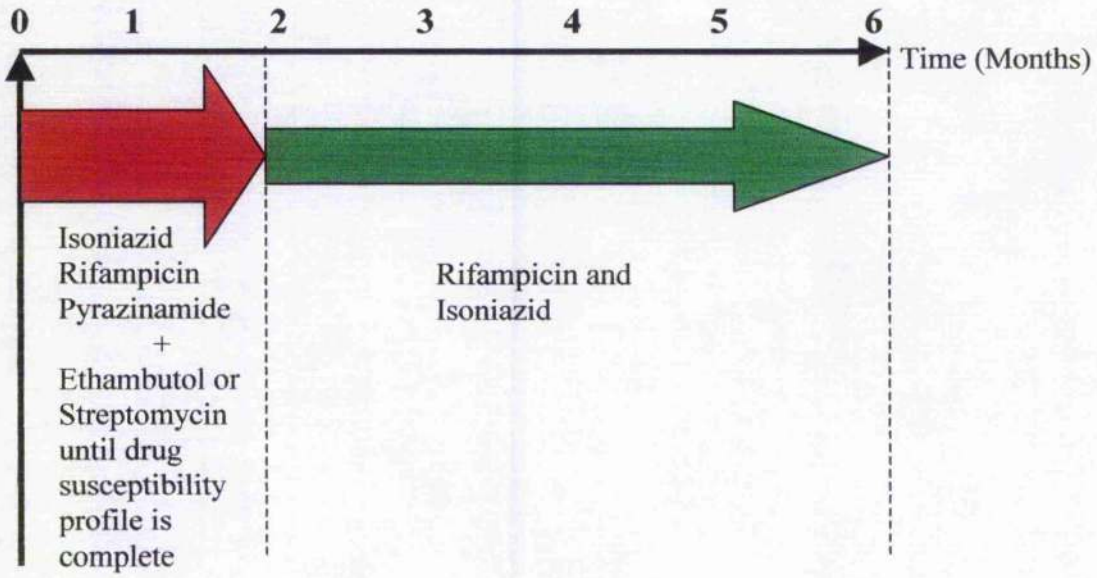
From the 1960's onwards, a number of other drugs were tested against TB infection and were found to have some efficacy. For example, pyrazinamide, like rifampicin, was found to be active against dormant bacilli and those present inside macrophages. The inclusion of pyrazinamide in the initial treatment regimen for pulmonary tuberculosis has further shortened the treatment cycle to 6 months. Ethambutol, like isoniazid, is an inhibitor of cell wall biosynthesis but is less effective at actually killing the bacterium. It is however useful for suppressing the emergence of resistance. Other drugs such as ethionamide and cycloserine have both poor bactericidal properties and troublesome side effects but are useful back-up treatments in cases where resistance to the front-line agents is a problem.

Cure rates of up to 95% can now be achieved through the use of the five front line anti-tubercular drugs, but only when they form part of a well-defined and organised treatment strategy. Failure to complete the entire course of chemotherapy can lead to patient relapses and the emergence of resistance. To help prevent this from happening, the WHO now recommend *Directly Observed Therapy* (DOT) where patients are administered the drug combinations in the presence of a health practitioner. Implemented throughout the world, DOT is now producing much improved results in the treatment of pulmonary tuberculosis, even in the poorest countries. DOT ensures that a far higher proportion of people complete the treatment schedule, and are thus cured of *M. tuberculosis* infection, thereby preventing the spread of the disease (WHO report on Global Tuberculosis Control, 2000). Table 1.3 and figure 1.4 show the front-line and secondary drugs used in modern TB treatment regimens and how they are used in a typical 6 month DOT setting.

Table 1.3 – Drugs used in the treatment of tuberculosis

Front-Line Drugs		**Secondary Drugs**	
Essential	Supplemental	Modern Era	Older
Isoniazid	Pyrazinamide	Quinolones	PAS
Rifampicin	Streptomycin	Ofloxacin	Ethionamide
	Ethambutol	Sparfloxacin	Cycloserine
		Ciprofloxacin	Capreomycin
		Macrolide	Kanamycin
		Clarithromycin	Thiocetazone

Figure 1.4 – A typical DOT drug regimen for the treatment of tuberculosis



1. 1. 8 Drug resistance and MDR-TB

Since the beginning of the widespread use of antibiotics in the 1940's the development of resistance amongst pathogenic organisms has been a significant problem in the effective treatment of clinical disease. As our arsenal of chemotheraputics has grown to include novel and semi-synthetic compounds, as well as naturally produced antibiotics, resistant organisms have emerged at every stage. In fact, the massive increase in the use of antibiotics over the last 60 years has almost certainly hastened the emergence of

resistance by placing susceptible organisms under increased evolutionary pressure. As molecular biology, biochemistry and genomics have provided insights into the mechanisms of antibiotic action, so they have also led to an increased understanding of how organisms evade antibiotic clearance.

The treatment of tuberculosis has not escaped the problems of drug resistance, a situation exacerbated by the relatively small number of effective agents. The use of combination therapy has however proven to be highly effective when adhered to correctly. Significant problems do however arise when treatment regimens break down. With the observed global increase in new TB cases, there is a major risk that the number of drug resistant cases may also be increasing. The widespread emergence of *Multiple Drug Resistant TB* (MDR-TB) is a real and serious threat. Minimally defined as resistance to both isoniazid and rifampicin, the two most effective drugs against TB infection, new cases of MDR-TB are currently relatively small in number compared with overall TB infection rates. In the year 2000 an estimated 273,000 cases of MDR-TB were identified out of a total of 8.7 million new cases, just 3.2% (Dye *et al.*, 2002). The global distribution of MDR-TB cases is also currently rather restricted, with the highest incidences in parts of China and Eastern Europe. Any serious spread of the disease is however potentially disastrous. This is because the treatment of MDR-TB is very difficult and expensive and utilises drugs which are poorly tolerated. The cost of treating a single case of MDR-TB in the USA will normally run to tens of thousands of dollars. Additionally, running at just 60-70%, cure rates for MDR-TB are significantly lower then those observed for the fully sensitive disease.

In order to prevent the possible spread of MDR-TB, research is being conducted in order to further our understanding of the underlying mechanisms of mycobacterial drug resistance. These efforts will hopefully yield new and effective drugs that circumvent the problems of resistance. Molecular biology and biochemistry have identified 4 basic mechanisms that contribute to drug resistance across all classes of pathogen (see table 1.5).

Table 1.5 – Mechanisms of drug resistance

(I) Enzymatic destruction or inactivation of the drug molecule

(II) Mutational alteration of the molecular target of the particular drug

(III) Decreased permeability of the cell towards drug molecules

(IV) Active efflux of drugs reducing cellular concentrations to sub-cytotoxic levels

All of these mechanisms have been shown to operate in mycobacteria and contribute, to a greater or lesser extent, to the high levels of drug resistance exhibited by *M. tuberculosis*. Of the above mechanisms mutational changes in drug targets are the most clinically significant. The aminoglycoside class of antibiotics, of which streptomycin is a member, operate by binding to the 30s ribosomal subunit and interfering with the transition between the initiation and elongation phases of protein synthesis. *M. tuberculosis* possesses only a single copy of the 16s rRNA gene and a number of mutations in this gene seem to prevent streptomycin from binding to the ribosome and exerting its action. Additionally, mutations in the ribosomal S12 protein have a similar effect and give rise to streptomycin resistance also (Finken *et al.*, 1993). Interestingly, a class of aminoglycoside detoxifying enzmes called *N-acetyltransferases* have been identified in the genomes of a number of mycobacterial species, including *M. tuberculosis*. As yet there is no direct evidence supporting a role for these enzymes in clinical drug resistance, and it seems likely that they perform some other physiological function (Ainsa *et al.*, 1997).

Target site mutation is also the major mechanism by which *M. tuberculosis* is able to escape the effects of the front-line agents rifampicin and isoniazid. Rifampicin is an inhibitor of bacterial RNA polymerase, a complex multi-subunit enzyme that mediates DNA transcription. Mutations in the *rpo*B gene that encodes the β-subunit of this enzyme are responsible for most rifampicin resistance phenotypes observed for *M. tuberculosis*. Most of the relevant mutations in fact occur in a very small locus of just 27 codons near the centre of the *rpo*B gene, and consist primarily of point mutations (Cole, S.T., 1994). In contrast, a great deal of literature regarding the exact mode of action of isoniazid has been accumulated without ever providing a complete, definitive explanation for its anti-tubercular activity. It is known with some certainty that isoniazid is a potent inhibitor of cell wall biosynthesis and that it inhibits one or more stages in

fatty-acid metabolism. Fatty-acid compounds are a known to be a vital component of the unique and highly complex *M. tuberculosis* cell wall structure (see next section). It is also known that isoniazid, in its administered form, is a pro-drug. i.e. it is converted into its active inhibitory form *in vivo* by the action of a bacterial enzyme. One proposed model of isoniazid action has the drug converted to its active form through the action of the catalase-peroxidase enzyme encoded by the *kat*G gene, whereupon it binds to and inhibits the InhA protein, an enoyl-ACP reductase (Mdluli *et al.*, 1998). This model is supported by the fact that mutations in both the structural gene *inh*A and the activator *kat*G have been correlated with isoniazid resistance (Rouse *et al.*, 1995). The importance of the mycobacterial cell wall is further underlined by the fact that another of the front-line anti-tubercular agents, ethambutol, is also a cell wall synthesis inhibitor. Ethambutol targets the production of the *arabinogalactan* component of the cell wall by inhibiting the activity of arabinosyl transferase enzymes (Telenti *et al.*, 1997). The product of the *emb*B gene is the most likely target of the drug in *M. tuberculosis* with over 60% of resistant organisms having a mutation in the codon for Met306 (Sreevatsan *et al.*, 1997). The fact that 40% of ethambutol resistant mutants have no mutation in the *emb*B gene raises the possibility that ethambutol, like isoniazid, acts at multiple points in the biosynthetic pathway.

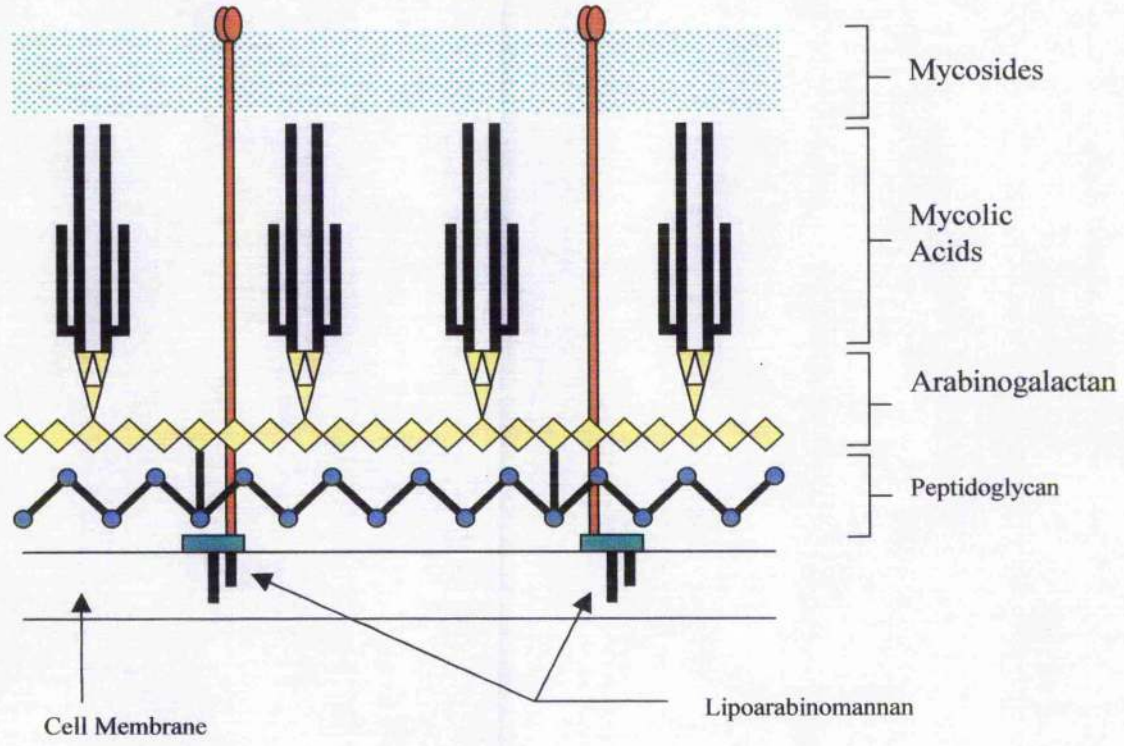
All of the above examples of clinically relevant resistance mechanisms are dependent upon target site mutations. Despite this, it is almost certain that multiple mechanisms contribute to *M. tuberculosis* drug resistance. For example, the unique structure of the lipid-rich cell wall represents a major permeability barrier to hydrophilic drug molecules, a feature which is discussed in the next section. Active efflux of drugs is another mechanism commonly employed by pathogenic organisms to avoid cytotoxic effects. Whilst active efflux has yet to be demonstrated to be a clinically relevant resistance mechanism in *M. tuberculosis*, the sequencing of the genome has shown that a number of potential drug pumps are encoded. The various types of active efflux pumps are discussed in greater detail in chapter 1.2, and it is the aim of this thesis to investigate the role of one such potential pump, the putative ABC transporter encoded by the *M. tuberculosis drr*ABC genes.

1. 1. 9 The mycobacterial cell wall

The mycobacteria fall into the class of Gram-positive prokaryotes, i.e. the cytoplasmic membrane of the cell is surrounded by a thick cell wall that lends the cell its shape and rigidity. However, significant differences exist between the mycobacterial cell wall and those of other Gram-positive organisms, such as the nocardia and corynebacteria, to which the mycobacteria are evolutionarily related. The differences are in fact so great that the classic Gram stain, used to differentiate between Gram-positive and Gram-negative organisms, works poorly on mycobacterial cells. The mycobacterial cell wall is the most complex of all organisms and is a structure essentially unique in nature. Its major distinguishing feature is the unusually high lipid content. Lipids may account for as much as 60% of the cell wall weight and include a number of molecules unique to the mycobacteria. The *acid-fastness* of mycobacterial cells, that is their resistance to destaining by weak organic acids, is another feature attributable to the unusual properties of the cell wall and is the basis of the Ziehl-Neelsen stain used to diagnose mycobacterial infection in lung sputum samples.

Significant amounts of research have been conducted into the physical structure and biochemical components of the cell wall. Much of this interest is stimulated by the fact the cell wall appears to play a large part in host-pathogen interactions and has a possible role in pathogenicity. The currently accepted model was first proposed by Minnikin in 1982 and later modified by McNeil and Brennan (Minnikin, 1982; McNeil and Brennan, 1991). This model, illustrated in figure 1.6, has the cell wall scaffold composed of three covalently linked structures: *peptidoglycans*, *arabinogalactans* and *mycolic acids*. In addition to these components are the *lipoarabinomannans* and the *mycosides*. The mycosides are a heterogeneous group of species-specific, surface exposed, lipid-containing molecules. Being the outermost layer of the mycobacterial cell wall they are believed to have a biological role analogous to the O-antigens of Gram-negative enteric bacteria such as *E. coli* i.e. account for such properties as agglutination serotype and colony morphology. It should be noted that the mycosides are not covalently attatched to the underlying scaffold structure, and for this reason they may sometimes be termed 'extractable lipids'. The various components of the cell wall are now considered in isolation and with reference to figure 1.6.

Figure 1.6 – Schematic representation of the mycobacterial cell wall



Peptidoglycan

The peptidoglycan layer of the cell wall is perhaps the simplest of the various components and is the only part of the cell wall to bear any significant similarity to the cell wall of other Gram-positive organisms. It is the innermost component of the wall, i.e. lies closest to the cytoplasmic membrane of the cell. Biochemically, the peptidoglycan consists of long chains of polysaccharide, cross-linked at various points by tetra-peptide chains. The overall effect is to create a mesh-work or lattice around the cell membrane that is responsible for cell morphology and rigidity. The polysaccharide chains consist of alternating residues of N-acetyl glucosamine and N-glycolyl muramic acid. This is very similar to the arrangement found in the peptidogylcan of other genera except for the incorporation of N-glycolyl muranic acid rather than the more common N-acetyl derivative.

Arabinogalactan

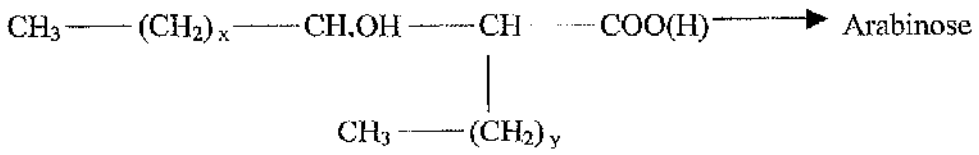
The arabinogalactan layer is the second innermost component of the cell wall and is linked to the underlying peptidoglycan by phosphodiester bridges. Arabinogalactan

constitutes the major polysaccharide component of the cell wall, but its structure is unusual in that it seems to lack repeating units (as is the case for the peptidoglycan). The structure is branched at many points, with the galactan portion of the molecule lying parallel to the underlying peptidoglycan, and the arabinan portion projecting perpendicularly from the branch points. The arabinose subunits branching from the galactan are arranged into groups of six and are further modified by the addition of mycolic acids to their reducing termini. The anti-tubercular ethambutol is believed to inhibit the arabinosyl transferase enzyme(s) that construct the hexa-arabinose units. The arabinogalactan portion of the cell wall essentially forms a 'bridge' between the peptidoglycan below and the major lipid component of the cell wall, the mycolic acids.

The Mycolic acids

The mycolic acids are a complex group of long-chain, branched fatty-acids. They are the largest (by weight) component of the mycobacterial cell wall, and can contain as many as 60-90 carbon units in *M. tuberculosis*. Mycolic acid synthesis has been shown to be the biochemical target of the anti-tubercular drug isoniazid (Mdluli *et al.*, 1998). The mycolic acids are attatched in groups of 4 to the reducing termini of hexa-arabinose units that form part of the underlying arabinogalactan layer. Only about two-thirds of these arabinose termini are modified by mycolation. The generalised structure of mycolic acids is illustrated below in figure 1.7. The two branches of the mycolic acid molecule are usually uneven in length with one arm containing significantly more hydrocarbon units than the other. In addition, the longer (or *meromycolate*) branch may contain functional group modifications such as oxygen containing groups, cyclopropane groups or unsaturated bonds. The mycolic acids are proposed to lie perpendicular to the cell membrane, and despite their modifications possess a low degree of unsaturation. Modifications occur at only two points in the meromycolate arm with the shorter arm always fully saturated. This allows the mycolic acids to pack tightly, forming a significant barrier to the passage of hydrophilic molecules, including drugs.

Figure 1.7 – mycolic acid



X = 40 – 60 (meromycolate arm)

Y = 20 – 25

Lipoarabinomannan (LAM)

Unlike the mycolic acids and the arabinogalactan, the liporabinomannans of the mycobacterial cell wall are not covalently linked with the basal peptidogycan layer. Rather, these very long, branched sugar molecules are anchored directly in the cell membrane by attachment to phosphatidyl-inositol phospholipids. Two sub-classes of LAM's are known to exist, those that terminate in arabinose residues, and those terminating with mannose. This distinction is relevant because AraLAM, but not ManLAM, is able to induce cytokine production by macrophage cells. In fact, LAM is one of the immunodominant antigens of *M. tuberculosis* and its structure may be an important modulator of host-pathogen interactions and virulence (Roach *et al.*, 1993). LAM's are believed to stretch all the way across the cell wall from the cell membrane to the outermost regions. In so doing, the LAM is believed to be an important 'anchor' molecule that helps to bind together components of the cell wall, such as the mycosides, that are not covalently linked with the underlying structures.

The Mycosides

Amongst the molecules exposed on the outer surface of the mycobacterial cell wall are a group of complex lipids collectively termed *mycosides*. These molecules can be sub-divided into two major classes, the *peptidoglycolipids*, and the *phenolic glycolipids*. The peptidoglycolipids are built up from amino acids, sugars and lipid moieties, whilst the phenolic glycolipids are lipid esters of the diol molecules phthiocerol or phenolphthiocerol lacking any further amino acid modifications. The lipid moiety of these molecules consists of lipids known as the mycoserosic acids. These are long chain multiple methyl-branched fatty acids synthesised by a single multifunctional enzyme encoded by the *mas* gene (Rainwater and Kolattukudy, 1985). The phenolic glycolipids

are of particular interest currently since they have been identified as cell wall components in only eight mycobacterial species, seven of which are strict or opportunistic pathogens and include *M. tuberculosis* (Azad *et al.*, 1997). Furthermore, it has been shown that mutations in the genes encoding the biosynthesis or transport of phenolic glycolipids result in attenuated virulence of *M. tuberculosis* in a mouse model (Cox *et al.*, 1999; Camacho *et al.*, 1999). Certain experimental data also suggest that phenolic glycolipids may interact with the host immune system and cause activation of human neutrophils. Once again, it appears that the lipid components of the mycobacterial cell wall are key determinants of virulence, and that successful disruption of cell wall biosynthesis may be an ideal target for new antitubercular agents.

In addition to the major structural components of the mycobacterial cell wall given above, significant quantities of other molecules have been identified. These are known to include triacyl-glycerols, acetylated trehalose molecules and sulpholipids. The exact roles of the many components of the cell wall are yet to be fully elucidated and much is still to be understood about how the various components contribute to physiological properties of the cell such as virulence and colony morphology. Moreover, researchers are still trying to isolate the various genes and biochemical pathways that are used in construction of the cell wall. The genome sequences of the various mycobacterial species completed so far will be of vital importance in future efforts to understand the uniquely complex nature of the cell wall structure. The above description of the mycobacterial cell wall is by no means exhaustive, and many excellent reviews on the subject have been written (e.g. see Liu *et al.* and Baulard *et al.* in Mycobacteria – Molecular Biology and Virulence, Eds. Ratledge and Dale, Blackwell, 2000).

1. 2. 1 Introduction to membrane transport

The cytoplasmic membrane forms a relatively impermeable barrier between a cell and its surrounding environment. The ability to selectively transport solutes across this barrier is a fundamental requirement of all cells in order to perform such basic functions as the acquisition of nutrients and the export of metabolic wastes. These transport processes are universally mediated by an array of integral membrane proteins collectively known as *membrane transport proteins*. Membrane transport proteins are currently the focus of intense research interest because a significant number have been implicated as mediators of drug and multidrug resistance in clinically important microbial pathogens.

We now know that membrane transport is dominated by two large protein "superfamilies", each one transporting solutes from one side of a membrane to the other by a fundamentally different mechanism. The first group to consider are known as *primary active transporters*, membrane transport systems that use a direct source of energy such as ATP hydrolysis or photon absorption to drive substrate translocation (Saier, 2000). Amongst the primary active transporters by far the largest sub-type are those energised by ATP hydrolysis. These transporters all share a common structural motif, the so-called *ATP-Binding Cassette*, a feature that has led this entire class of transporter to become known as the *ABC transporter superfamily* (Higgins, 1992). Members of this family have been identified in archaea, eukaryota and prokaryota alike, their universal distribution hinting at the essential role of these transporters in all forms of life. Indeed, ABC transporters are now known to contribute to such diverse phenomena as microbial antibiotic resistance (van Veen and Konings, 1998) and human inherited genetic disease (Riordan *et al.*, 1989). Structural and functional aspects of the ABC transporter superfamily, particularly their role as mediators of microbial drug resistance, are discussed in far greater detail in subsequent sections of this thesis.

The second major grouping of transmembrane transport proteins is known as the *Major Facilitator Superfamily* (MFS). These transporters operate by a variety of mechanisms including substrate uniport, substrate/cation symport and substrate/cation

antiport (Pao *et al.*, 1998). MFS transporters have been identified that catalyse both uptake and efflux reactions, the vast majority of which involve the transport of small molecules such as sugars, amino acids, drugs and metabolic intermediates. Whilst ABC transporters are known to be capable of transporting a similar range of small molecules, they are also known to catalyse the transport of much larger macromolecules such as protein toxins (Koronakis and Hughes, 1993). Another major difference between the ABC transporters and the MFS concerns the structure and genetic organisation of the systems. All MFS transporters are encoded by a single gene producing a polypeptide chain of approximately 400 amino acids. The chain is believed to traverse the cytoplasmic membrane multiple times (either 12 or 14) with the membrane-spanning regions adopting an α-helical conformation (Huang *et al.*, 2003). This is in stark contrast to the ABC transporters which are frequently encoded by multiple genes, the complete transporters often being the product of multiple, discrete protein subunits (see subsequent sections).

The MFS is now known to constitute at least 29 distinct sub-families of transporters, each family specific for a particular substrate or range of closely related substrates (Saier *et al.*, 1999). Two particular sub-families of MFS transporter are particularly relevant in the context of this work, these proteins all being involved in the active efflux of drugs via a drug/proton antiport mechanism. The two families differ in their transmembrane topology, one family constituting transporters which have 12 membrane-spanning helices, the other, smaller family including those transporters with 14 membrane-spanning regions. Well characterised members of each family include the TetA(B) tetracycline transporter of *E. coli* (12-helix), and the 14-helix TetA(K) tetracycline transporter of *Staphylococcus aureus* (see Borges-Walmsley *et al.*, 2003)

Whilst the ABC transporters and MFS account for more than 50% of all known active efflux systems, they are by no means the only membrane transport proteins that contribute to the damaging phenomenon of antimicrobial drug resistance. Of the approximately 250 identified families of membrane transport proteins at least three further families appear to play a significant role in the extrusion of toxic compounds from the cytoplasm. These families include the *Small Multidrug Resistance* (SMR) family, the *Resistance-Nodulation-Division* (RND) family and the newly

characterised *Multidrug And Toxic compound Extrusion* family (MATE transporters). The SMR and RND transporters seem to operate by a drug/proton antiport mechanism similar to that exhibited by MFS drug efflux systems (Borges-Walmsley *et al.*, 2003). However, the less well characterised MATE transporters, exemplified by the NorM multidrug transporter of *Vibrio parahaemolyticus*, function as drug/Na^+ antiporters (Morita *et al.*, 2000).

It is well beyond the scope of this introduction to embark upon a detailed analysis of the structural and functional aspects of *all* the different groups of transporters involved in antimicrobial drug resistance, excellent reviews of these matters being available elsewhere in the literature (for example see Borges-Walmsley *et al.*, 2003; McKeegan *et al.*, 2003). However, it is worth noting that significant advances in the field of membrane protein structural biology are beginning to shed light on the mechanisms by which certain drug efflux systems may operate. Foremost among these advances is the publication of the 3D crystal structure of the *E. coli* AcrB RND-family multidrug efflux pump (Murakami *et al.*, 2002). This structure not only provides researchers with their first glimpse of the detailed structure of a multidrug efflux pump, but also suggests a mechanism by which a single transporter may be able to recognise and extrude multiple, structurally unrelated compounds. The first MFS 3D crystal structures have also recently been published, those of the glycerol-3-phosphate and lactose import transporters of *E. coli* (Huang *et al.*, 2003; Abramson *et al.*, 2003) Whist neither of these transporters is involved in antimicrobial resistance, both structures confirm the α-helical nature of the membrane-spanning regions and also suggest possible transport mechanisms.

Much remains to be learned regarding the structure and mechanism of the many different transport systems involved in antimicrobial resistance, and many technical challenges remain to be resolved. The following sections will go on to describe in detail our current knowledge of the ABC transporter superfamily and the role that these systems play in drug resistance phenomena.

1. 2. 2 Identification of the ABC transporter superfamily

The identification of the first ABC transporters can be traced back to the investigations of Leon Heppel and co-workers into the nature of nutrient uptake in *Escherichia coli*. They noted that osmotically shocked cells exhibited dramatically reduced transport rates for certain amino acids, whereas the import of others was largely unaffected by this treatment. Subsequent studies revealed that the shock-sensitive systems were inactivated due to the loss of periplasmic *substrate binding proteins* (SBP's) that perform a vital role in the function of these multi-component transporters (Heppel, 1969). The shock-insensitive systems on the other hand appeared to function independently of SBP's, their activity being maintained in cell free systems such as isolated cytoplasmic membrane vesicles. This led to the idea that two distinct classes of transporter were operating in *E. coli*.

Further evidence for the existence of two distinct types of transporter was gathered when the mechanisms of energy coupling in the two groups of transporters were investigated (Berger, 1973; Berger and Heppel, 1974). The activity of the shock-insensitive transporters was found to be severely inhibited by chemical uncouplers of oxidative phosphorylation, whilst the SBP-dependent transporters were able to function in the presence of these agents provided that adequate supplies of glycolytic substrates were supplied. Together these data suggested to Heppel and others that the two groups of transporters operated by fundamentally different principles. The SBP-dependent systems required phosphate bond energy in order to drive the transport process, whilst the shock-insensitive systems were reliant on the presence of an "energised membrane state" which we now know to be the transmembrane ion gradients generated via oxidative phosphorylation.

1982 saw a major advance in the study of ATP driven membrane transport when the first complete nucleotide sequence of the genes encoding a SBP-dependent transporter was determined (Higgins *et al.*, 1982). Once again the breakthrough was made in studies of amino acid transport in a model prokaryote, in this case the histidine transporter of *Salmonella typhimurium*. The multi-component nature of the SBP-dependent transporters was confirmed by the presence of four structural genes encoded in a contiguous operon. The first gene in the operon had previously been shown to encode the periplasmic SBP, HisJ, whilst the remaining three structural

genes encoding the HisQ, M and P proteins were thought to form a complex present in the cytoplasmic membrane via which the substrate might be translocated. However, the sequences of the genes and their presumed protein products did not suggest any mechanism by which the transport process might be coupled to ATP hydrolysis.

Another SBP-dependent transport system under investigation at this time was the maltose permease of *E. coli*. This transporter was known to comprise of a minimum of five proteins encoded in two clustered but divergently transcribed operons (Raibaud, 1979; Silhavy, 1979). Biochemical data suggested that the *E. coli* maltose transport complex and the histidine transporter of *S. typhimurium* shared certain structural commonalties. Firstly there was the presence of a high affinity SBP encoded by the *mal*E gene. The products of two further genes in this operon, *mal*F and *mal*G, were located to the cytoplasmic membrane along with the product of the *mal*K gene present in the second operon (Bavoil *et al.*, 1980). The first suggestion that the histidine and maltose transport systems were related at a genetic level came with the sequencing of the *E. coli mal*K gene (Gilson *et al.*, 1982a). The *mal*K gene product was shown to share extensive homology with the HisP protein of *S. typhimurium* by alignment of their respective amino acid sequences. This process demonstrated that 32% of the residues occupying equivalent positions in the two proteins were identical, whilst an additional 35% represented conservative substitutions of similar amino acids (Gilson *et al.*, 1982b). Interestingly, neither MalK nor HisP exhibited a strongly hydrophobic character typical of integral membrane proteins, raising the prospect that they were peripherally associated with the transport complex. These observations of sequence homology led to the formation of a number of key hypotheses regarding the SBP-dependent transporters of prokaryotes. Firstly the data suggested that transporters of structurally diverse solutes may have evolved from a common evolutionary ancestor, and secondly that HisP and MalK might perform a conserved function in the two transporters, possibly coupling transport to ATP hydrolysis.

In 1985 a third protein sharing significant sequence homology to the HisP and MalK proteins was identified. The product of the *opp*D gene of *S. typhimurium* was known to form one component of the *S. typhimurium* oligopeptide permease. Once again this transporter was encoded by multiple genes and dependent upon a periplasmic SBP.

Close inspection of the sequences of the three proteins revealed two regions that were particularly well conserved. These sequences, now known as the "Walker A" and "Walker B" motifs, had previously been identified in a range of non-transport related ATP-binding proteins and were believed to contribute to an adenine nucleotide binding fold (Walker *et al.*, 1982). The identification of a consensus nucleotide binding motif on these three proteins, each involved in the transport of a different substrate, provided strong evidence that ATP was indeed the energy source driving the transport process (Higgins *et al.*, 1985). With all of this indirect evidence in place, it was not long before it was unambiguously demonstrated that HisP and MalK were capable of binding nucleotide analogues (Hobson *et al.*, 1984; Higgins *et al.*, 1985), and that ATP hydrolysis occurred during the transport cycle (Bishop *et al.*, 1989).

As more and more gene sequences were completed, it became apparent that the ATP-binding components of the bacterial SBP-dependent transporters were just part of a much larger family of proteins, all of which shared extensive amino acid homology along their entire length (Higgins *et al.*, 1986). Not all of these homologous proteins were involved in transmembrane transport. For example, the UvrA protein of *E. coli*, known to be involved in DNA repair, showed a similar level of homology to HisP as HisP had shown to MalK (Doolittle *et al.*, 1986). Despite their various different physiological functions, all of these proteins were thought to bind and/or hydrolyse ATP, and in doing so were providing a means of energy coupling to diverse cellular requirements. The ubiquitous nature of these ATPases was further underlined when the first eukaryotic homologue was identified, the mammalian multidrug resistance protein, P-glycoprotein (Gerlach *et al.*, 1986).

The term ABC (*ATP-Binding Cassette*) transporter was first used in 1990 in recognition of the similar structure adopted by the ATP-binding domains of the transport related proteins (Hyde *et al.*, 1990). Almost thirty years on from their identification in model prokaryotes, members of the ABC transporter superfamily are continually being discovered in archea, prokaryotes and eukaryotes alike. It is now clear that these proteins catalyse both export and import of solutes across cellular membranes of all kinds, and that their substrate specificity can vary from a single ion or small molecule to a broad range of large, structurally unrelated, hydrophobic molecules. Whilst much attention is currently being paid to those ABC transporters

implicated in human disease, studies in model bacteria continue to feature heavily in the literature. It is of note that the first complete atomic resolution structures of ABC transporter core domains have all been from prokaryotic systems, the first of them being the HisP ATP-binding protein of *S. typhimurium* (Hung *et al.*, 1998).

1. 2. 3 Universal distribution of the ABC transporters

Technical advances over the last decade in the fields of DNA sequencing and bioinformatics have resulted in an extraordinary increase in the rate at which gene sequences have been deposited in public databases. The complete sequencing of entire bacterial genomes is becoming almost "routine", with numbers of completed sequences growing monthly. Whilst this avalanche of raw data has presented certain challenges in terms of data storage, searching and retrieval, it has also spawned the new discipline of *comparative genomics*. Researchers hope that by comparing the genomes of two or more different organisms they will be able to extract important evolutionary and functional relationships between similar genes. Differences between genomes are likely to prove just as instructive, revealing sequences that are important in conferring such features as pathogenicity, host range or drug resistance.

The availability of complete genomes from all kingdoms of life has revealed the true extent of the ABC transporter superfamily, and the relative contribution that it makes to membrane transport as a whole. The ability to identify genes as components of ABC transport systems is made relatively simple by the highly conserved nature of the ATP-binding cassette (see next section). However, prediction of the physiological role of any given transporter is hampered by the lack of sequence conservation amongst the membrane-associated components of the transporter complex. The rate at which new transporters are being identified through bioinformatic methods is far outstripping the ability of researchers to verify computer-generated predictions of function by experimental means. Furthermore, in a large number of cases, functional assignment is impossible since no experimentally verified substrate or physiological role has been defined for a particular group of transporters. Nonetheless, a number of comparative genomics projects have sought to classify the ABC transporters into various families as well as identifying homologous systems across species (Tomii and Kaneshia, 1998).

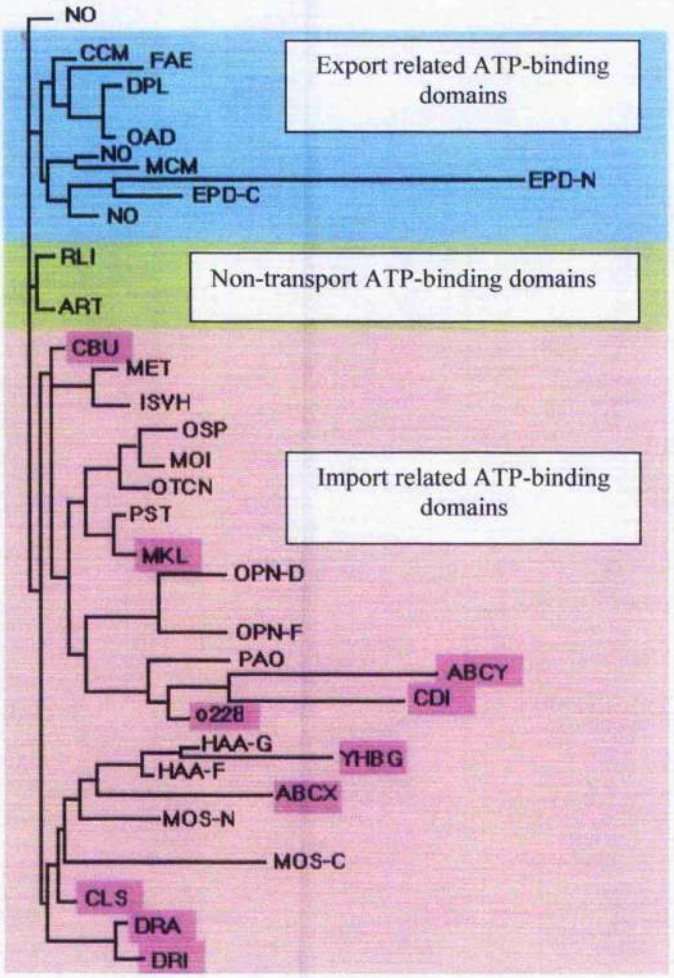
The complete genome of *Escherichia coli* K-12 (Blattner, 1997) revealed that genes encoding membrane transport systems accounted for almost 10% of the genes present. Analysis of the various types of transport systems present indicated that a full 4.9% of the genome encoded components of ABC transporters. As such, the ABC transporters were found to constitute the largest paralagous family of proteins present in the *E. coli* genome (Linton and Higgins, 1998). Whilst this figure undoubtedly establishes the importance of ABC transporters in *E. coli*, what about other microorganisms? An analysis of eighteen complete prokaryotic genomes revealed that whilst the ABC transporters were indeed the most numerous of the transport systems as a whole, both the total number of transporters in an organism and the relative proportion of secondary and primary transporters exhibited large variations (Paulsen *et al.*, 2000). For example, *Mycoplasma genitalium* was found to encode just twenty membrane transport systems in total, whilst *E. coli* and *Bacillus subtilis* encoded 304 and 265 respectively. This large variation is not merely a function of genome size since *E. coli* and *B. subtilis* also exhibited the highest numbers of transport systems per megabase of genomic DNA. Paulsen has suggested that the increased transport capabilities of *E. coli* and *B. subtilis* may allow these organisms to adapt readily to changes in their external environment. The ability of these heterotrophs to make full use of all available nutrients might contribute to the rapid growth of these organisms in complex media. In this respect these commonly studied "model" organisms may not necessarily be representative of all prokaryotes.

Paulsen also noted that the relative proportions of ABC transporters and MFS transporters within an organism vary in relation to the organisms primary means of energy generation. Strict aerobes such as *Mycobacterium tuberculosis* generate ATP via an electron transport chain and subsequent oxidative phosphorylation. This form of metabolism generates a strong transmembrane proton gradient that can be used to drive secondary transporters directly. Conversely, those organisms with a strictly fermentative metabolism (e.g. *Mycoplasma genitalium* and *Treponema pallidum*) generate ATP via substrate level phosphorylation only. In the absence of strong ion gradients these organisms are far more reliant on ATP driven ABC transporters. Thus, the number of ABC transporters maintained in the genomes of different organisms is likely to vary considerably, variations being a reflection of an organism's mode of metabolism and the ecological niche that it occupies.

The same bioinformatic methods that are used to identify members of the ABC transporter superfamily can also be used to examine the evolutionary relationships between transporters from different organisms. Such phylogenetic analysis can be used to subdivide the large number of transporters into groups based on features such as sequence similarity and substrate specificity. By comparing the sequences of ATP-binding components of ABC transporters from many different species, both prokaryotes and eukaryotes, a number of general features have emerged (Saurin *et al.*, 1999; Dassa and Bouige, 2001). Firstly, the transporters divide into two broad groups based upon their general function i.e. import vs. export (see figure on the next page). The family of importers is dominated by prokaryotic SBP-dependent systems exemplified by the *E. coli* maltose importer. The absence of such nutrient uptake systems from eukaryotic genomes probably accounts for the far larger numbers of ABC transporters identified in prokaryotes (e.g. 67 in *E. coli* vs. 29 in *S. cerevisiae*). On the other hand, the export related ABC transporters, have been found in all genomes so far analysed. The high level of sequence conservation amongst ABC modules involved in both import and export suggests that these proteins are under significant evolutionary pressure. The clustering together of homologous systems from archea, eubacteria and eukaryotes would suggest that these proteins are evolutionarily ancient and that the functional division of the ABC transporters into importers and exporters took place before the three kingdoms of life diverged. The only other possible hypothesis consistent with these observations is massive horizontal gene transfer between organisms. The fact that ATP-binding proteins associated with transport of similar substrates cluster together in these phylogenetic analyses is a good indication that sequence similarity is of predictive value in assigning function to newly identified ABC transporters. This assertion is strengthened when it is considered that, despite a much reduced level of sequence homology, a similar clustering of related proteins is observed when the hydrophobic components of ABC transporters are considered in isolation (Saurin and Dassa, 1994). Saurin has proposed that the many different specificities and architectures of ABC transporters have arisen through multiple gene duplication and fusion events.

Phylogenetic analysis reveals that the importers and exporters can each be subdivided many times over into groups of closely related transporters sharing similar function.

Phylogenetic analysis of ABC transporter ATP-binding domains



The phylogenetic tree shown above is taken from Saurin *et al.*, 1999. The tree clearly shows early segregation of ATP-binding domains into two large families, those associated with an export function (blue background), and those associated with an import function (pink background). Each large family is then further sub-divided to give a total of 34 distinct ATP-binding domain families, each one associated with the transport of a particular substrate or range of substrates. A third class of ATP-binding domain is shown in the tree (green background). These proteins are associated with non-transport cellular processes such as DNA repair and regulation of gene expression.

The next section will go on to describe just a few of these groups in order to illustrate the many roles of the ABC transporters.

1. 2. 4 Physiological roles of the ABC transporters

The enormous extent of the ABC transporter superfamily precludes an extensive discussion of all of the various roles of these transporters. However table 1.8, giving just a few examples of well characterised ABC transporters, illustrates the huge substrate diversity observed amongst members of this superfamily. Substrates include ions, amino acids, oligopeptides, proteins, mono- and disaccharides, lipids, vitamins, steroids, pigments and complex carbohydrates. Most of the transporters listed are specific for a single substrate or a range of closely related compounds. However a few ABC transporters, particularly those mediating export reactions, will transport a variety of structurally unrelated molecules e.g. mammalian P-glycoprotein or the LmrA transporter of *Lactococcus lactis*. It is broad-specificity transporters of this kind that are most frequently associated with multidrug resistance phenotypes in microorganisms and mammalian cancer cells (see below). Not only do the various transport substrates possess markedly different physico-chemical properties, but they are also very different in size. Understanding how the ABC transporters recognise and transport such a widely differing array of molecules is one of the key challenges facing researchers in this area.

In accordance with their wide range of substrates, the ABC transporters contribute to a number of physiological processes within a cell. A few of these processes are described below using examples taken from table 1.8.

Nutrient uptake

The eubacterial and archeal SBP-dependent ABC transporters can be thought of as low capacity, high affinity transporters that perform a vital scavenging role for organisms growing and competing in a changing environment (Higgins, 1992). Under conditions where nutrients are scarce the potential growth advantage conferred by such systems may be significant. For example the energetic cost of *de novo* histidine synthesis is 41 ATP equivalents, whereas the requirements of ABC transporters are generally believed to be 1-2 ATP molecules consumed per transport cycle (Ames, 1986). The high affinity of the SBP-dependent importers also makes them ideal for

Table 1.8

ABC*	Organism	Substrate	Function(s)	Reference
MalK	*E. coli*	Maltose/maltodextrins	Nutrient uptake	Higgins *et al*. (1985)
HisP	*S. typhimurium*	Histidine	Nutrient uptake	Higgins *et al*. (1982)
OppD	*S. typhimurium*	Oligopeptides	Nutrient uptake	Hiles *et al*. (1987)
BtuD	*E. coli*	Vitamin B12	Nutrient uptake	Friedrich *et al*. (1986)
ModC	*E. coli*	Molybdenum	Nutrient uptake	Maupin-Furlow *et al*. (1995)
MsbA	*E. coli*	Lipid A	Cell membrane synthesis	Zhou *et al*. (1998)
KpsT	*E. coli (K5)*	Capsular polysaccharide	Capsule synthesis	Pavelka *et al*. (1994)
TagH	*B. subtilis*	Teichoic acids	Cell wall synthesis	Lazarevic (1995)
HlyB	*E. coli*	Haemolysin	Virulence	Goebel *et al*. (1982)
YbdA	*B. subtilis*	Unknown	Sporulation	Isezaki *et al*. (2001)
NodI	*R. leguminosarum*	Lipo-oligosaccharides	Intercellular signalling	Spaink *et al*. (1995)
DrrA	*S. peucetius*	Anthracycline antibiotics	Antibiotic self protection	Kaur (1997)
LmrA	*L. lactis*	Multidrugs	Multidrug resistance	van Veen *et al*. (1996)
Pgp	*H. sapiens*	Multidrugs	Multidrug resistance	Riordan *et al*. (1985)
Pdr5p	*S. cerevisiae*	Multidrugs	Multidrug resistance	Kolaczkowski *et al*. (1996)
ALDP	*H. sapiens*	Long-chain fatty acids	Intracellular transport	Aubourg *et al*. (1993)
Atm1p	*S. cerevisiae*	Iron-Sulphur clusters	Intracellular transport	Kispal *et al*. (1997)
CFTR	*H. sapiens*	/	Chloride ion regulation	Riordan *et al*. (1989)

* In the case of multicomponent systens only the ABC domain is listed

the uptake of essential ions or co-factors that may be very scarce in the external environment. Examples of such dedicated transporters include the molybdenum (ModBC) and vitamin B_{12} (BtuCD) importers *of E. coli* (see table 1.8 for references). The large numbers of SBP-dependent systems identified in bacterial genomes account for a significant proportion of all known ABC transporters.

Protein export

A number of proteins perform their biological functions extracellularly. The majority of these secreted proteins are synthesized with a cleavable N-terminal leader peptide which directs their export via the Sec pathway (Mori, 2001). However, some proteins lack the characteristic signal peptide and are thought to be transported via a non-classical pathway involving ABC transporters. Proteins exported in this Sec-independent manner are often associated with 'secondary' functions such as virulence. The best characterised of these is the haemolysin export ABC transporter of *E. coli*, HlyB (Holland *et al*, 1990). The *hly* operon encodes a number of other proteins required for production and activation of the toxin, as well as an additional protein facilitating passage of the toxin molecule, HlyA, across the periplasmic space. Such accessory proteins are commonly required in Gram-negative organisms since the transport substrate has to be exported across both membranes that form the cell envelope of these bacteria. A phylogenetically related system is utilised by the phytopathogen *Erwinia chrysanthemi* to secrete several extracellular metalloproteases involved in the virulence of this organism. It is believed that the secretion signal for both HlyA and the *E. chrysanthemi* proteases lies in a structural motif at the C-termini of these proteins (Koronakis *et al*., 1989; Binet *et al*., 1997).

Export of structural and cell surface components

Both Gram-positive and negative eubacteria produce extracellular structures that they present to the external environment. Some of these components are known to be of importance in bacterial pathogenicity and interaction with hosts e.g. cells of the mammalian immune system. The Gram-positive bacteria possess a thick peptidoglycan cell wall interwoven with teichoic acids, whereas the Gram-negative bacteria have a much thinner cell wall surrounded by a second membrane, the outer leaflet of which consists of lipopolysaccharides (LPS). Additionally, some species of bacteria secrete an extracellular matrix, or capsule, primarily composed of

carbohydrates. Many of the molecules that make up these extracellular structures are synthesised in the cytoplasm and then transported across the plasma membrane by dedicated ABC transporters. *E. coli* is known to produce a number of antigenically distinct polysaccharide capsule types, one of which, the K5 serotype, has been well characterised. Rigg and co-workers were able to identify as many as twelve genes in three closely located operons that are responsible for biosynthesis and export of the capsular carbohydrate. One of the operons, consisting of the *kps*M and *kps*T genes, encodes an ABC transporter mediating export of the newly synthesised molecule from the cytoplasm. Immunolocalisation experiments suggested that the biosynthetic proteins and the ABC transporter formed a large macromolecular complex on the inner surface of the cytoplasmic membrane (Rigg *et al.*, 1998). This system is another example of genes encoding a biosynthetic function being closely linked to those for export of the end product, similar to the *E. coli hly* operon. Export ABC transporters have also been characterised for Lipid A, a major component of Gram-negative LPS, and teichoic acids of the Gram-positive cell wall (table 1.8, MsbA and TagGH).

Cellular signalling and differentiation

The Gram-positive organism *Bacillus subtilis* is one of a number of bacterial species that produce hardy spores as a stress response to harsh environmental conditions. Such a response requires an array of systems that are constantly able to monitor the presence of nutrients and the changing state of the external environment. *B. subtilis* encodes an ABC transporter, YbdAB, that is believed to be responsive to external conditions and plays a role in the initiation of sporulation. Whilst the exact nature of the substrate is unclear, YbdA and B mutations result in decreased spore formation and down regulation of a number of genes important in sporulation. These data suggest that YbdAB may be a component of, or interact with, cellular signalling pathways (Isezaki *et al.*, 2001).

Members of the genus Rhizobium are nitrogen fixing symbiotic bacteria which form nodules on the roots of leguminous plants. The formation of nodules requires intercellular signalling between the bacteria and their host. In the case of the rhizobia this signalling is mediated by modified sugars (lipo-chitooligosaccharides) known as Nod factors. Secretion of these signalling molecules from the cell requires the action of an ABC transporter encoded by the *nod*IJ genes (Spaink *et al.*, 1995). Similar to the

capsular polysaccharide exporter of *E. coli*, KpsMT, the *nod*IJ genes are closely linked in the genome to those required for Nod factor synthesis and modification. Phylogenetic analysis suggests that these two systems are closely related, possibly a reflection of their similar substrate specificities (Reizer *et al.*, 1992).

Self protection in antibiotic producing organisms

Bacteria of the family Actinomycetales produce a large number of bioactive secondary metabolites, many of which are medically useful antibiotics. In many instances circumventing the toxicity of these compounds to the organism producing them is achieved by the active transport of the antibiotic substance out of the cell. ABC transporters are frequently employed as a means of extruding these metabolites, maintaining their concentrations at a level below which they become autotoxic (Mendez and Salas, 2001). DNA sequencing of antibiotic production loci frequently identifies the presence of an ABC transporter associated with the biosynthetic genes, many of which have yet to be biochemically characterised. One of the better characterised systems is the DrrAB transporter of *Streptomyces peucetius*. This organism produces the anthracycline antibiotics doxorubicin and daunorubicin, both of which are used in the treatment of human cancers. The DrrAB pump has been experimentally confirmed as a resistance determinant by expressing the genes in a heterologous host normally sensitive to the action of these compounds (Kaur, 1997).

Multidrug resistance

The phenomenon of multidrug resistance (MDR) is commonly defined as the ability of a cell to withstand doses of more than one drug that are normally lethal to non-resistant cells (Brennan, 2001). MDR is of significant clinical importance as a cause of chemotherapeutic failure in the treatment of bacterial infection and cancer. Research into the molecular determinants of MDR in mammalian cells frequently pointed to the activity of a 170 kDa plasma membrane protein, known as *P-glycoprotein* (Pgp), the expression of which was found to be amplified in multidrug resistant cell lines (Riordan *et al.*, 1985). When the gene for Pgp (MDR1) was cloned it was found to bear significant homology to bacterial transport genes that we now know encode members of the ABC transporter superfamily (Gerlach *et al.*, 1986). Since its discovery Pgp has been the focus of extensive research. Whilst most ABC transporters are relatively specific, Pgp has been shown to pump an array of diverse

molecules including anthracyline antibiotics, vinca alkaloids and paclixatel, all of which are used in the treatment of cancer (Dean, 2001). Interestingly, Pgp has been shown to pump drugs directly from the inner leaflet of the plasma membrane, rather than from the cytoplasm suggesting that this transporter may interact with its substrates on the basis of their hydrophobicity (Shapiro and Ling, 1997). Normally expressed in the liver and at the blood-brain barrier, a full characterisation of the natural substrates of the MDR1 gene product has yet to be completed. However, we now know that MDR1 is just one of a multigene family in humans, several of which are involved in the cellular transport of hydrophobic compounds such as phospholipids (MDR3) and bile salts (van Helvoort *et al.*, 1996; Borst *et al.*, 2000). It has been suggested that the ability of this family to confer multidrug resistance upon cells may be a result of cross-specificity towards many different hydrophobic molecules.

MDR phenotypes are not limited to human cancer cells. Analogous systems operate in lower eukaryotes and prokaryotes. The best characterised of the prokaryotic ABC transporters conferring multidrug resistance is the LmrA pump of *Lactococcus lactis*. Identified by van Veen in 1996, LmrA shows remarkable conservation of both sequence and function with mammalian Pgp (van Veen *et al.*, 1996). Similar to Pgp, LmrA is known to extrude multiple, structurally unrelated, hydrophobic drugs directly from the cytoplasmic membrane and has proven to be of great use as a bacterial model of Pgp function (van Veen and Konings, 1997). Homologues of LmrA have been identified in a number of pathogenic and non-pathogenic bacteria raising the prospect that broad-substrate multidrug resistance determinants may be widely distributed amongst prokaryotes (van Veen and Konings, 1998). This fact has significant implications for the future effectiveness of antibiotic treatment of human infections.

Yeasts have frequently been used as model organisms in the study of eukaryotic cellular processes. The phenomenon of *pleiotropic drug resistance* (PDR) in yeast is essentially analogous to mammalian MDR. PDR is dependent upon a complex network of interacting proteins and genes which combine to control the expression of a variety of ABC transporters. At the heart of this network are the *S. cerevisiae* Pdr5p and Snq2p ABC transporters, homologues of Pgp, which have been shown to confer

resistance to literally hundreds of diverse drugs including the important azole class of antifungals. Importantly, homologues of Pdr5p have been identified in opportunistic pathogenic yeasts such as *Candida albicans*, *Cryptococcus neoformans* and *Aspergillus nidulans*. As such, these pumps have the potential to give rise to clinical drug resistance (Bauer *et al.*, 1999). Additionally, work in our laboratory has recently identified two ABC transporters from the thermally dimorphic fungal pathogen *Paracoccidioides brasiliensis*. Expression of the genes encoding these transporters is upregulated in response to exposure to azole antifungals, strongly suggesting a potential function as drug resistance determinants (Gray *et al.*, 2003).

Intracellular transport

A number of eukaryotic ABC transporters are associated with intracellular membranes. This is perhaps unsurprising given the proposed evolution of modern-day organelles from putative endosymbiotic bacteria. The first mitochondrial ABC transporter to be identified was the Atm1p transporter of *S. cerevisiae*. This protein is believed to function as mitochondrial exporter, probably transporting chelated Fe/S clusters synthesised in the mitochondrial matrix and subsequently utilised by cytosolic enzymes (Lill and Kispal, 2001). A human homologue of Atm1p, ABC7, has been linked with inherited X-linked sideroblastic anemia, a disease characterised by the abnormal accumulation of iron within mitochondria (Allikmets *et al.*, 1999). Thus, ABC transporters appear to play an important role in iron homeostasis in both yeast and humans.

Mutations in one of three peroxysomal ABC transporter genes gives rise to the disease X-linked adrenoleukodystrophy. This is a neurodegenerative disorder characterised by defective oxidation of very long-chain fatty acids, a process that is initiated in the peroxysome. This suggests that the ALDP ABC transporter may play a role in transporting fatty acids into the peroxysome (Liu *et al.*, 1999). Interestingly the ATP-binding domains of the ALDP transporter are exposed to the cytosol rather than the interior of the peroxysome. This finding suggests that transport into the peroxysome is functionally equivalent to an export reaction taking place at the cell surface (Contreras *et al.*, 1996). The best characterised of the human intracellular ABC transporters is the *transporter associated with antigen processing* (TAP). This ER-resident protein is responsible for transporting viral or bacterial peptides from the

cytosol into the lumen of the ER where they are complexed with MHC class I molecules. The antigen-MHC complex is then trafficked from the ER to the cell surface where the foreign antigen is presented to the cellular immune system (Lankat-Buttgereit and Tampe, 2001).

CFTR - An unusual transporter

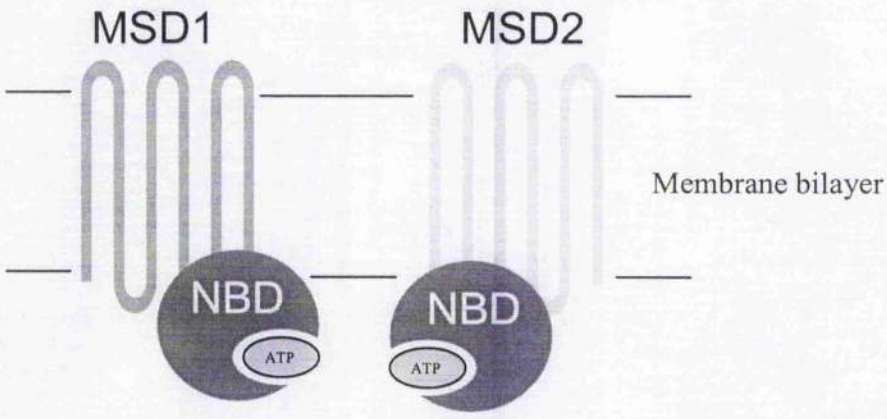
In addition to TAP and Pgp another human ABC transporter has become the focus of significant research attention, the *Cystic fibrosis transmembrane conductance regulator* (CFTR). Mutations in this gene are a major cause of the most common inherited disease amongst Caucasians, cystic fibrosis (CF). Cloning of the CF gene showed it to be a member of the ABC transporter family, and yet a significant body of evidence had built up suggesting that CF was caused by a defect in epithelial cell chloride ion regulation (Riordan *et al.*, 1989). We now know that CFTR is an unusual member of the ABC transporter superfamily in that it operates as an ATP-regulated channel, rather than a transporter *per se*. The passage of chloride ion across the apical membrane of epithelial cells via CFTR is not dependent upon stoichiometric ATP hydrolysis. However, the open/closed state of the channel is thought to be influenced by ATP binding as well as the action of intracellular kinases and phosphatases that interact with the CFTR protein (Gadsby and Nairn, 1999). It remains to be seen whether other 'apparent' ABC transporters are in fact nucleotide gated ion-channels.

1. 2. 5 Structural organisation of the ABC transporters

After thirty years of research we now understand the ABC transporter complex to be highly modular in nature, with the generally accepted model consisting of four "core" domains (Higgins *et al.*, 1992). Firstly, two hydrophobic domains span the membrane multiple times and contribute to a transmembrane pathway via which solutes are transported from one side of the membrane to the other. These integral membrane components of the complex are thought to possess one or more substrate binding sites which affect the specificity of the transporter. The transmembrane topology of these domains gives rise to their description as 'membrane-spanning domains' or MSD's. The remaining components of the complex are two relatively hydrophilic domains that bind and hydrolyse ATP, providing energy for the transport process. These two 'nucleotide-binding domains', or NBD's, are closely apposed to the MSD's at the cytoplasmic face of the membrane barrier. The NBD's contain the highly conserved

ATP-binding cassette that lends its name to this class of transporter. This minimal model of the ABC transporter complex is illustrated in figure 1.9. No ABC transporters have yet been identified that are able to function without this minimal complement of four domains.

Figure 1.9 – The four core domains of an ABC transporter

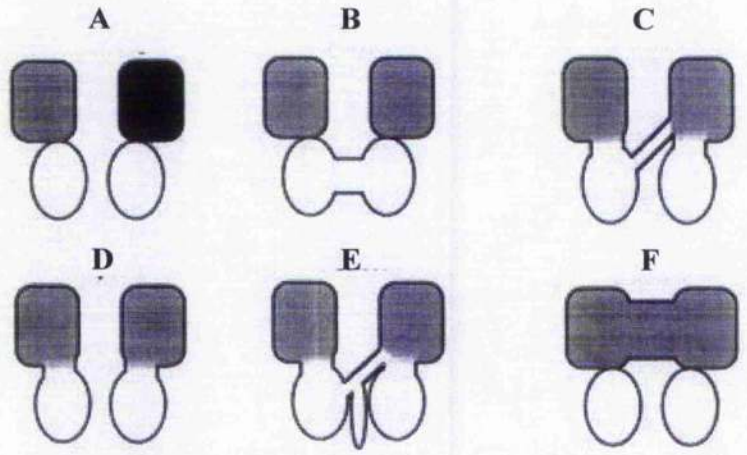


The arrangement illustrated in figure 1.9 indicates the situation most commonly encountered in prokaryotic systems, each of the four domains being encoded separately and the overall complex formed by four distinct polypeptide chains. However, in many cases, particularly in eukaryotic systems, two or more of the domains may be fused together and encoded by a single gene. This situation is exemplified by the mammalian MDR1 gene, the product of which, P-glycoprotein, is a single multidomain polypeptide. Another common arrangement of domains, once again prominent in eukaryotic systems, is the so-called *half-transporter*. In these cases a single gene encodes a two-domain protein in which one MSD is fused to a single NBD. In fact, examples of virtually all conceivable arrangements of MSD and NBD have been identified amongst members of the ABC transporter superfamily, a few examples of which are illustrated in figure 1.10.

Genetic loci encoding ABC transporters may initially seem to contain fewer genes than necessary to account for all four core domains of the functional transport complex. For example, the core of the well-characterised *S. typhimurium* histidine importer is encoded by just 3 genes, a NBD encoded by *his*P, and two MSD's encoded by *his*Q and *his*M. In such cases experimental evidence suggests that one of

the domains, in this case the NBD, operates as a homodimer, thereby generating the complete four domain structure (Kerppola *et al.*, 1991). A similar situation exists regarding the maltose transporter of *E. coli* where only one NBD gene is present in the *mal* operon (Kennedy and Traxler, 1999). It should be noted that whilst both the histidine and maltose transporters are SBP-dependent systems, the SBP itself is reversibly associated with transport complex and does not constitute one of the four core domains. Half-transporters, such as the LmrA multidrug transporter of *L. lactis*, only possess two of the four domains required for a functional transport complex. Consequently, these proteins must dimerise in order to complete the structure of the transporter. Such systems may be homodimeric, as in the case of LmrA, or even heterodimeric through combination with another half-transporter encoded elsewhere in the genome e.g. the Tap1/Tap2 transporter (van Veen *et al.*, 2000; Antoniou *et al.*, 2002).

Figure 1.10 – Molecular architectures of the ABC transporters



A	=	$HisQMP_2$	All modules encoded individually
B	=	RbsAC	Fused NBD's
C	=	P-glycoprotein	All 4 domains fused, single polypeptide
D	=	HylB	Homodimeric half-transporter
E	=	CFTR	4 fused domains plus regulatory domain
F	=	FhuBC	Fused MSD's

After its initial proposal, and in the absence of high resolution structural data, various biochemical and genetic means were used to test and refine the four domain

hypothesis. For example, the use of computer algorithms to predict the topology of MSD's, backed-up by experimental techniques such as epitope mapping, suggested that each MSD contributes six membrane-spanning regions to the translocation pathway, a total of twelve forming the overall complex (Pearce *et al.*, 1992; Gentschev and Goebel, 1992). By analogy with the archetypal membrane protein bacteriorhodopsin, the membrane-spanning segments were predicted to be α-helical in nature. The basic 4 domain model has also come to be known as the 'two-times-six' model.

Unsurprisingly amongst such a large class of proteins, a number of ABC transporters exhibit deviations from the four domain, two-times-six model. For example, the *malF* encoded MSD of the *E. coli* maltose importer seems to possess eight membrane-spanning segments. However, deletion of two of the α-helices at the N-terminal of this protein can be achieved without affecting maltose accumulation. It has been suggested that these additional helices may form a separate domain, not directly involved in transport, perhaps facilitating membrane insertion, packaging of the helices or correct orientation of the protein within the membrane (Froshauer *et al.*, 1988). CFTR, the mammalian chloride channel, possesses an additional fifth domain present between the first NBD and second MSD of this multidomain protein (figure 1.10). Reversible phosphorylation of this so-called 'R-domain' is thought to play an important role in regulating the function of the channel (Gadsby and Nairn, 1999). In fact, an entire sub-family of human ABC transporters, the MRP (multidrug resistance protein) family, possess an additional N-terminal membrane-bound domain known as TMD0. Bakos and colleagues were able to demonstrate that TMD0 contains as many as four or five transmembrane α-helices that do not directly participate in substrate transport (Bakos *et al.*, 1998). The presence of additional domains, even relatively large ones such as TMD0, do not seem to affect the core function of ABC transporters. These domains have probably evolved much later in evolutionary time, performing some specific role related to individual transporters.

As the various structural and functional models of ABC transporters were tested, a great deal of biochemical and genetic data were accumulated in order to understand how the various domains of the transporter contribute to its function. Site-directed

mutagenesis and deletion experiments were particularly important in defining regions and sub-domains of the transporter involved in key processes such as ATP binding, substrate binding and protein-protein interactions. The following sections will go on to describe what is known about the individual components of the transporter complex, and how recently available structural data have shed light on how the various components interact to transport such a diverse array of molecules.

1. 2. 6 Conserved Sequences within the NBD

The central role of the NBD within the ABC transporter complex is to bind and hydrolyse ATP, thereby providing the energy required for the transport of solutes against potentially large concentration gradients. When two ABC transporters are compared, the NBD's are invariably the most highly conserved modules. Levels of identity ranging from 30% to 50% are commonly observed between NBD's from evolutionarily diverse sources, much of this homology lying within a 215-230 amino acid region that comprises the ATP-binding cassette. Some of the most highly conserved and intensely studied regions of the NBD are described here.

The Walker A sequence is a glycine-rich motif with the consensus sequence GxxGxGKS(T), where G represents glycine, x any amino acid and the final position being occupied by either serine or threonine (Walker *et al.*, 1982). Present in a variety of nucleotide-binding proteins, residues within the Walker A motif are key to stabilising interactions between the protein and the bound nucleotide. The Walker A residues form a loop around the tri-phosphate moiety of the nucleotide and extensive hydrogen bonding interactions occur between main-chain nitrogen atoms of the protein and oxygen atoms in the β- and γ-phosphate groups of the NTP. Alternatively known as the *phosphate* or *P-loop*, this structure has been identified in a variety of non-transport related proteins including proto-oncogene rasP21, RecA and the F1-ATPase (Saraste *et al.*, 1990). The fine crystal structure of HisP, the NBD of the *S. typhimurium* histidine ABC transporter, indicates an identical role for the Walker A motif in the ABC transporters (Hung *et al.*, 1998).

The Walker B motif is always found further towards the C-terminal of the protein than Walker A. The consensus sequence of this motif is far less rigid than that of

Walker A, but generally conforms to 'ffffD' where f represents a hydrophobic amino acid and D is a highly conserved aspartic acid residue. The role of the Walker B motif, and the aspartic acid residue in particular, is to co-ordinate a Mg^{2+} ion in the active site of the NBD. This Mg^{2+} ion plays a vital role in catalysing ATP hydrolysis. Together the Walker A and B motifs constitute a nucleotide-binding fold present in many nucleotide-binding proteins. Mutations of the residues within these highly conserved sequences have differing effects, but are in general poorly tolerated and often lead to the loss of ATP binding and/or hydrolysis. This is particularly true of the lysine residue in Walker A and the aspartate of Walker B (Schneider and Hunke, 1998).

Whilst the Walker motifs are present in many different classes of proteins, the NBD's of ABC transporters possess a third highly conserved motif that is unique to the ABC transporter superfamily. The so-called 'ABC signature sequence' or 'linker peptide' is found between the Walker A and B motifs, immediately upstream of Walker B. The general consensus of this ABC transporter specific sequence is 'LSGGQ' and has proven to be of use in identifying new members of the ABC transporter superfamily (Higgins *et al.*, 1988). The residues between Walker A and B in the primary structure of NBD's are thought to form a separate α-helical sub-domain that folds independently of the nucleotide-binding core. Recently published NBD crystal structures, such as that of MJ0796 ATP-binding cassette, support the presence of this helical domain (Yuan *et al.*, 2001). Specific mutations within the signature sequence of CFTR have been linked to cystic fibrosis, clearly suggesting that this motif plays an important role in the function of the protein (Welsh and Smith, 1993). The exact role of the helical domain and the signature sequence within it have been the source of some controversy and will be discussed in the context of recently derived crystal structures.

A number of shorter regions of homology, sometimes consisting of only one or two highly conserved residues, have been described. One such motif is known as the 'switch region'. An invariant histidine residue is present at the end of this motif, mutation of which has been shown to destroy the transport capabilities of a variety of ABC transporters, for example the *S. typhimurium* histidine permease. Mutations in

the switch region do not affect ATP hydrolysis by the NBD, suggesting that residues in this region play some alternative yet vital role in the transport cycle. It has been proposed that the switch region may operate as a signal transduction domain, propagating conformational changes induced by ATP hydrolysis to different components of the complex (Shyamala *et al.*, 1991).

1. 2. 7 Biochemical properties of the NBD

Demonstration that NBD's of ABC transporters are able to bind and hydrolyse ATP, concomitant with transport, was a key stage in understanding how substrate translocation is powered. Conformational changes occurring in the nucleotide-binding domains of ABC transporters during their catalytic cycle are of central importance in driving substrate translocation. These changes are induced by ATP-binding and hydrolysis and propagated to components of the translocation pathway through inter-domain interactions. The biochemical parameters of ATP hydrolysis and the nature of conformational change have been studied extensively using both isolated NBD's and intact transport complexes.

Studying the biochemistry of ABC transporters is dependent upon the ability of researchers to develop systems for their expression and purification. This has been by no means routine since ABC transporters are naturally present in low abundance within the membrane, and recombinant systems for the overexpression of these proteins often result in the production of insoluble, inactive protein (Wang *et al.*, 1999). The production of stable, soluble protein that maintains its biological activity has been the limiting factor in the study of many ABC transporter systems. This is particularly true with respect to structural determination using biophysical techniques such as X-ray crystallography and NMR which require relatively large amounts of highly pure protein (Duffieux *et al.*, 2000).

Biochemical data regarding the activity of both isolated NBD's and intact transport complexes has been accumulated for a number of systems including the prokaryotic histidine and maltose importers, and the medically important eukaryotic systems of P-glycoprotein and CFTR (Schneider and Hunke, 1998). The affinity of ABC transporters for ATP, as measured by their Michaelis constants, vary considerably, as do the maximal rates of ATP hydrolysis. Expressed in the absence of their cognate

MSD's, purified NBD's frequently exhibit a spontaneous, constitutive ATPase activity (Nikaido *et al.*, 1997; Morbach *et al.*, 1993). For reasons described below, this constitutive activity is unlikely to be a true reflection of the behaviour of NBD's in the intact transporter complex and care should be taken in interpreting results derived from experiments using isolated components.

Using the well-characterised histidine permease of *S. typhimurium* as an example, the ATPase activity the of NBD's is under somewhat tighter control. Purified HisP has an intrinsic rate of ATP hydrolysis that is not exhibited by the intact transporter complex. This observation suggests that the membrane-spanning components of the transporter complex impose a negative regulation upon the ATPase activity of the NBD. Negative regulation is only removed and ATP hydrolysis initiated when the transporter is able to interact with histidine, liganded by the SBP HisJ (Liu and Ames, 1997). The necessity for such down regulation of constitutive activity becomes apparent in a physiological context. Unless controlled, the NBD's of ABC transporters would continually hydrolyse intracellular supplies of ATP in an unproductive fashion. Such regulation in the intact transporter complex is indicative of a series of inter-domain signalling processes. Firstly, histidine-liganded HisJ is present in the periplasmic space, yet ATP hydrolysis is initiated by the NBD's which are located at the *interior* face of the membrane. Interaction between the transporter complex and the loaded SBP must take place at the exterior face of the complex, implying the presence of a substrate binding site in the MSD's of the transporter. Secondly this interaction must somehow be transmitted to the NBD's which then hydrolyse ATP and set in motion the conformational changes required to transport the released histidine molecule from one side of the membrane to the other. An essentially analogous series of events has been proposed by Davidson and colleagues regarding the activity of the *E. coli* maltose transporter (Davidson *et al.*, 1992).

Induction of ATP hydrolysis through interaction with a substrate is not restricted to the prokaryotic nutrient uptake ABC transporters. A similar phenomenon is also observed in the mammalian multidrug transporter P-glycoprotein. The rate of ATP hydrolysis by the NBD's of P-glycoprotein is stimulated several times over when the transporter is exposed to a number of the drugs that it expels (Scarborough, 1995). Once again these interactions are only observed when all four domains of the

multidomain transporter are reconstituted together. When only two of the four domains are expressed a basal level of ATPase activity can be measured, but this activity is not stimulated by drugs (Loo and Clarke, 1994). It seems that the ATPase activity of P-glycoprotein is regulated in a similar manner to that of bacterial nutrient importers, and that interaction between transporter and substrate occurs via the MSD's.

Conflicting evidence exists as to whether the NBD's are able to interact directly with the substrate of the transporter. For example, the ATPase activity of purified HisP is largely unaffected by free histidine, the SBP HisJ, or the histidine-HisJ complex. Nor are the activity of P-glycoprotein NBD's affected by the presence of drugs when they are expressed in isolation. However, Buche and colleagues have presented data showing a direct interaction between a purified NBD and its transport substrate. OleB constitutes the NBD of a multi-component ABC transporter from the antibiotic producing organism *Streptomyces antibioticus*. OleB binds directly to the transport substrate, oleandomycin, the antibiotic produced and secreted by *S. antibioticus* (Buche *et al.*, 1997). Another recent example is the PotA NBD of the *E. coli* spermidine-preferential uptake system. Kashiwagi and co-workers have shown that the ATPase activity of PotA decreases in the presence of the substrate, spermidine, and that this interaction takes place specifically with a C-terminal region of the protein. This mechanism has been proposed as a form of feedback inhibition imposed upon the activity of the transporter, mediated via the NBD's (Kashiwagi *et al.*, 2002). Whilst observations of direct interactions of this kind are uncommon, they do perhaps suggest that different types of ABC transporter may interact with their substrates in differing ways. This is maybe unsurprising when it is considered that some ABC transporters are importers, some export molecules from the lipidic environment of the membrane interior, and others still export molecules directly from the cytoplasm.

Obviously the NBD must be able to bind and hydrolyse ATP. However, these domains also commonly exhibit an affinity for other nucleotide triphosphates including GTP, CTP and UTP. In general the affinity for these other molecules is lower than that shown towards ATP, and the accompanying rate of hydrolysis is usually significantly lower. Nonetheless, the ability of NBD's to utilise alternative NTP's is indicative of at least some degree of conformational flexibility within their

highly conserved structure. A number of inhibitors of ATP hydrolysis have been found to interact directly with NBD's. Many ABC transporters are strongly inhibited by the presence of ortho-vanadate at micromolar concentrations. The mechanism of inhibition by vanadate involves the trapping of an ADP molecule in the active site of the NBD. Unable to exchange ADP out of the active site, all subsequent steps in the binding and hydrolysis cycle of the NBD are blocked (Urbatsch *et al.*, 1995). Another inhibitor known to act directly on the NBD is N-ethylmaleimide (NEM). This molecule covalently modifies the free sulphydril group present on cysteine residues. A number of ABC transporters, including P-glycoprotein, possess a cysteine residue in, or close to, the Walker A motif of the NBD. By modifying this Walker A cysteine, NEM is able to block the access of ATP to the catalytic active site of the protein (Al-Shawi *et al.*, 1994). Replacing the reactive cysteine by mutagenesis or preincubating the protein with high concentrations of ATP have both been shown to protect P-glycoprotein from the effects of NEM. Perhaps the most unusual inhibitor identified to date is the macrolide antibiotic Bafilliomycin A1. Shown to inhibit the activity of intact transport complexes such as P-glycoprotein and the *S. typhimurium* maltose importer, the mechanism of inhibition by this molecule is yet to be elucidated (Sharom *et al.*, 1995; Hunke *et al.*, 1995). In some instances both dinucleotides, such as ADP, and non-hydrolysable nucleotide analogues, such as ATPγS, will act as competitive inhibitors of ATP hydrolysis (Urbatsch *et al.*, 1994).

The binding and hydrolysis of nucleotides by the NBD are thought to induce a series of conformational changes within the transporter complex that are central to the transport mechanism. A number of techniques have been used to probe these changes, but perhaps the most commonly used is fluorescence spectroscopy. The fluorescence of individual tryptophan residues within the structure of a protein are uniquely sensitive to the immediate environment surrounding them. Thus, as the global conformation of the NBD is altered through interaction with nucleotides, so the local environment of individual tryptophan residues changes. Typically, tryptophan residues exposed to the aqueous environment exhibit a lower fluorescence yield than those buried in the interior of the protein. This arises because buried residues are less susceptible to the effects of collisional quenching by surrounding water molecules. Additionally, the maximum wavelength of tryptophan fluorescence may also be blue-

shifted as residues move towards a less polar environment. This technique has been used to demonstrate conformational changes in the structure of the -helical domain of MalK, a region of the molecule thought to be of importance in interactions with the MSD's of the complex (Schneider *et al.*, 1994).

One of the major questions regarding the NBD's of ABC transporters is not how they interact with inhibitors, transport substrates or ATP, but whether or not the two NBD's present in the intact transporter complex interact with one another. Conflicting data have emerged for different transporters, dependent upon whether isolated NBD's are studied or complete transporter systems. However, the emerging consensus, supported by recent structural data, is that in the complete membrane-associated structure of the ABC transporter, the two nucleotide binding domains interact with one another and that this interaction is intimately linked with the transport mechanism. In their studies of purified HisP, Nikaido and colleagues proposed that the active form of this NBD was most likely to be a dimer, consistent with two copies of HisP being present in the functional transporter complex $HisQMP_2$ (Nikaido *et al.*, 1997). Positive cooperativity for ATP hydrolysis between the two NBD's was also observed in the maltose transport complex $MalFGK_2$ (Davidson *et al.*, 1996). Further evidence supporting interaction between the two NBD was presented for P-glycoprotein on the basis of inhibitor studies. Urbatsch and colleagues showed that vanadate-induced nucleotide trapping in the active site of one NBD completely inhibits transport, suggesting that neither NBD can drive transport independent of the other, and also that the two NBD's are mechanistically linked (Urbatsch *et al.*, 1995). Interestingly, purified MalK is insensitive to the action of vanadate whereas the mature transport complex is strongly inhibited. It is possible that the action of vanadate is dependent upon the formation of the complete transporter complex. The necessity for two functional NBD's has also been demonstrated using genetic means. Mutations in either one of the two Walker A motifs present in the NBD's of P-glycoprotein completely abrogated the biological activity of the protein (Azzaria *et al.*, 1989). A similar requirement for two completely functional NBD's has been demonstrated for the $MalFGK_2$ transporter and the Ste6p a-pheromone transporter of yeast (Davidson and Sharma, 1997; Berkower and Michaelis, 1991). In addition to biochemical and genetic suggestion that the two ABC transporter NBD's interact with one another mechanistically, physical interaction of

the two domains has been demonstrated. Chemical cross-linking agents can be used to identify residues in close proximity to one another. Such agents demonstrated that both MalK subunits of the maltose importer are in close proximity with one another and with the two MSD's, MalF and MalG. Moreover, upon the addition of nucleotides, the relative positions of the various components were found to shift, as reported by a change in the pattern of chemical cross-linking. These results strongly support the idea that the two NBD's of ABC transporters are closely linked, and that nucleotide binding/hydrolysis is accompanied by global conformational changes in all components of the transporter (Hunke *et al.*, 2000).

The stoichiometric coupling of ATP hydrolysis and substrate translocation has been investigated. Some estimates have suggested that as many as fifty ATP molecules might be necessary to drive the export of a single molecule of the dye Hoechst 33342 by P-glycoprotein (Shapiro and Ling, 1997). However, such a high ATP requirement seems rather unlikely. As progress has been made in the development of artificial systems for reconstituting complete ABC transporters, the estimates of ATP requirement have steadily been revised downwards. Current best estimates place the number of ATP molecules hydrolysed per transport cycle at between one and two (e.g. see Patzlaff *et al.*, 2003). Such estimates are far more consistent with the presence of two NBD's per transporter.

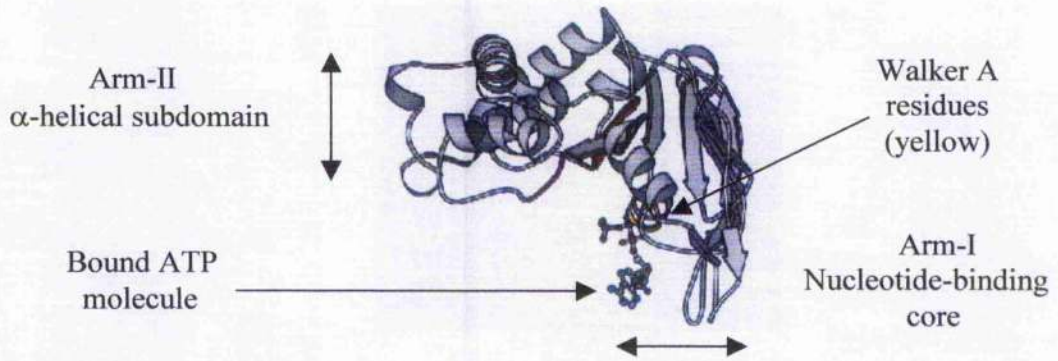
All of these various biochemical and genetic data have been used to propose a number of models which link the events of nucleotide binding and hydrolysis to the process of substrate translocation. The mechanistic details of these various models are beyond the scope of this introduction and are reviewed elsewhere (van Veen *et al.*, 2001; Senior *et al.*, 1995). The details of the schemes differ for the various proteins for which models have been presented. Whilst the conserved four domain structure of the ABC transporters would imply a conserved mechanism of transport, this may not necessarily be the case. CFTR is a prime example of a protein that to all intents and purposes resembles an ATP driven transporter, yet functions as a nucleotide-gated ion channel. Consequently it is possible that the alternative models proposed for different transporters represent genuine differences in the mechanisms of these transporters, rather than an incomplete understanding of the processes taking place.

1. 2. 8 3D Structure of the NBD

Until relatively recently, high resolution structural data for nucleotide binding domains has been scarce. This is perhaps a reflection of the difficulties encountered in producing the large quantities of highly purified, soluble protein required for the initiation of crystallisation trials. However, exhaustive efforts by a number of groups and modern techniques for high-throughput expression/purification/crystallisation have yielded crystals of several isolated NBD's. These have in turn allowed the structure of the NBD's to be solved by X-ray crystallography.

The first fully described structure of an ABC transporter NBD was that of the HisP protein from *Salmonella typhimurium*, published in 1998 by Kim and colleagues (Hung *et al.*, 1998). The molecule was shown to adopt an overall 'L-shape' with two distinct arms. Arm-I, as it is described in the original publication, is a domain composed of a mixture of α-helices and β-sheets and contains the invariant residues of the Walker A and Walker B motifs. The overall fold in this region of the molecule strongly resembles the nucleotide-binding core or the F1-ATPase. Interposed between the Walker A and B sequences is a separate, predominantly α-helical, subdomain. This region forms Arm-II of the L-shaped structure and is unique to the NBD's of ABC transporters. Furthermore, it is this region of the molecule that contains the residues of the ABC signature sequence. Crystallised in the presence of 10 mM ATP, the bound nucleotide is clearly visible within the structure of HisP, and as predicted interacts strongly with the residues of the Walker A sequence. The structure of the HisP monomer, showing Arm-I, Arm-II and the position of the bound nucleotide is illustrated in figure 1.11 below.

Figure 1.11 - The HisP monomer

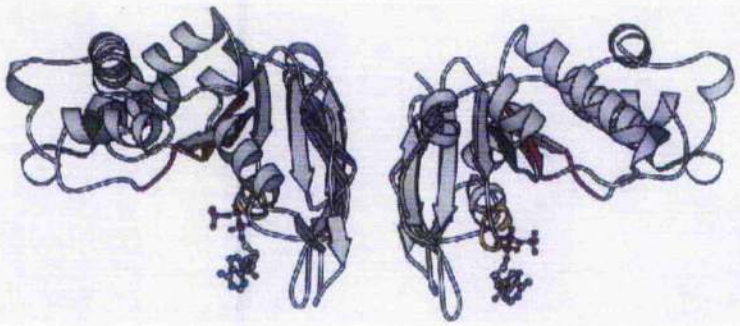


The completion of crystal structures for a number of other ABC transporter NBD's, including MalK from the archeon *Thermococcus litoralis* and two from *Methanococcus jannaschii*, establishes the validity of the two domain structure observed in the HisP monomer (Diederichs *et al.*, 2000; Karpowich *et al.*, 2001; Yuan *et al.*, 2001). Comparison of nucleotide-free and ADP bound forms of the *M. jannaschii* NBD's MJ0796 and MJ1267 with the ATP bound structure of HisP led the authors to propose that a series of intramolecular conformational changes occur during nucleotide hydrolysis. It is proposed that these changes result in movement of the two domains with respect to one another and that they may facilitate nucleotide exchange in and out of the catalytic site.

Whilst intramolecular conformational changes within the NBD may be important during the ATP binding/hydrolysis cycle, it is unlikely that these changes are, on their own, large enough to drive substrate translocation via the MSD's of the intact transporter complex. Rather, it is interactions between the modules of the transporter that are believed to drive transport. Biochemical data has long suggested that the two NBD's of ABC transporters interact not only with the MSD's, but also with one another (Jones and George, 1999; Fetsch and Davidson, 2002). The publication of the various NBD structures has raised a number of questions regarding details of any interaction between the two NBD's, and also raised the possibility that NBD:NBD interactions may differ between transporters.

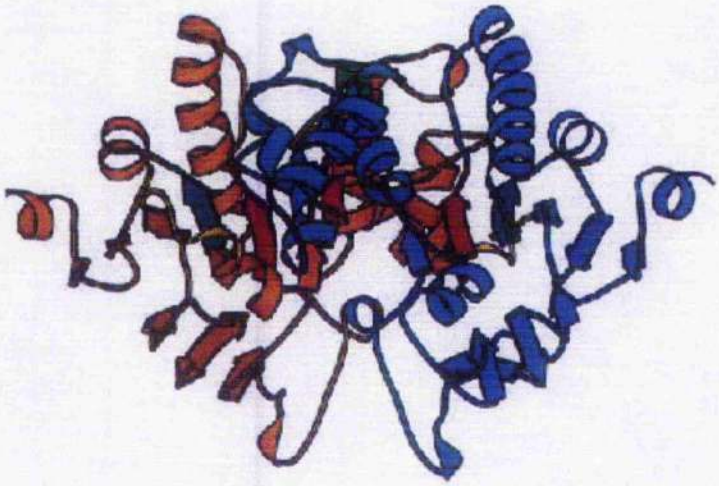
When expressed and purified independently of the other components of an ABC transporter complex, NBD's are monomeric. This is true for HisP, MalK and the MJ NBD's. Even in the presence of ATP, stable NBD dimers have yet to be purified in the absence of MSD's. It is against this background of either weak or transient interaction between isolated NBD's that the crystal structures of the NBD's must be considered. For example, the crystallographic unit cell of HisP contained a dimer, with the two L-shaped molecules oriented in a 'back-to-back' manner and the nucleotide-binding folds facing away from one another (sees figure 1.12 below). In the view of the dimer shown below, it is the upper surface of the dimer that is proposed to interact with the MSD's.

Figure 1.12 – The HisP dimer



The dimer observed in the MalK crystal structure is however very different from that in HisP. In the MalK dimer the two nucleotide- binding folds are facing one another, and the monomers contribute to a much more compact, interlocking structure. Unlike the HisP dimer, the two subunits of the MalK dimer are not related by a perfect 180° two-fold symmetry. The MalK dimer is shown below in figure 1.13, and as with HisP, the dimer surface proposed to interact with the MSD's is shown uppermost. Each monomer is shown in a different colour.

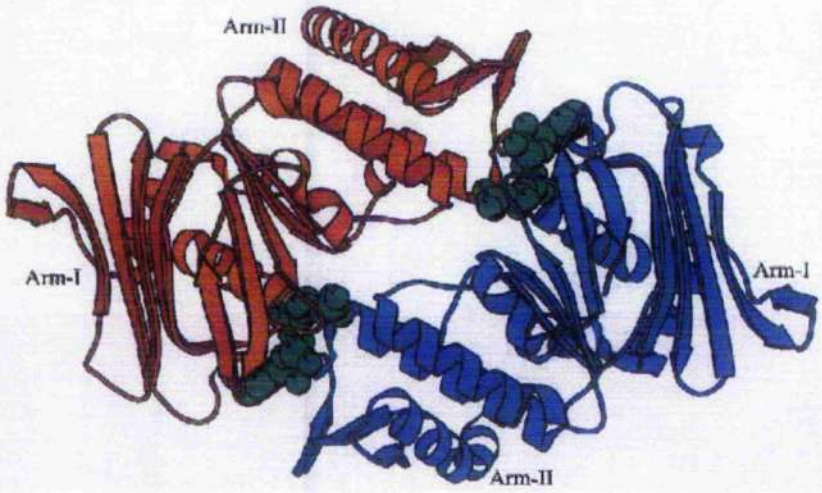
Figure 1.13 – The MalK dimer



The enormous amounts of biochemical data obtained with respect to both the maltose transporter and the histidine transporter can be pressed into service in defence of both of these possible dimerisation models. However, perhaps the best dimer model for the relevant physiological NBD:NBD interaction is based upon ABC domains not from a

transporter, but from a DNA repair protein called Rad50. The structure of the Rad50 dimer of *Pyrococcus furiosus* is illustrated below (Hopfner *et al.*, 2000). Each monomer is coloured differently, and 2 bound ATP molecules are shown in green.

Figure 1.14 – The Rad50cd dimer



Note that in the Rad50 dimer structure, the bound ATP is present at the dimer interface. In fact ATP is not only bound by residues of the Walker A and B motifs from one NBD monomer, but also by residues present in the signature sequence motif of the opposite NBD. This observation is consistent with the fact that unlike other NBD's the Rad50 dimer forms and remains stable in the presence of ATP and its analogues (Hopfner, *et al.*, 2000). The interlocking domain structure is more similar to that observed for the MalK dimer than for HisP. It has been suggested that the MalK and HisP dimers observed in their respective crystal structures are artifactual, that is they do not represent physiologically relevant dimers but form in response to intermolecular forces present within the crystal. As discussed in the final section of this chapter, the Rad50 type NBD:NBD interaction has received significant experimental support with the publication of the complete structure of the BtuCD transporter.

1. 2. 9 The membrane-spanning domains

The membrane-spanning domains of ABC transporters have been far harder to study than their cognate NBD's. Naturally present at low levels, significant difficulties

surround purification of membrane proteins from the far more abundant and numerous cytoplasmic proteins. Added to this problem is the fact that genetic systems designed to overexpress cytoplasmic proteins are far less efficient at overexpressing membrane proteins. Purification of membrane proteins necessitates their removal from the membrane environment, a process which requires the use of mild detergents. Such detergents may not sufficiently stabilise the structure of the protein which can lead to aggregation or disruption of the protein's native structure. Despite these technical challenges regarding the study of membrane proteins in general, significant progress has been made in understanding the contribution of the MSD's to the function of the ABC transporter complex.

Primary amino acid sequence data is of little use in predicting the detailed function of ABC transporter MSD's. This likely reflects the ability of many different amino acid combinations to support the α-helical architecture of the MSD within the membrane. Levels of homology between MSD's are often higher between the two MSD's of an individual transport complex than between those of different transporters. This supports the hypothesis that the two MSD's of a transporter have often evolved through gene duplication events. Despite the limited level of homology between ABC transporter MSD's, it is believed that some structural constraints are imposed upon these domains by the nature of the substrate which they transport. This is reflected in the fact that the MSD's of ABC transporters with similar substrates often cluster together in phylogenetic analyses.

The prokaryotic permeases, those ABC transporters involved in nutrient uptake, possess one moderately conserved region known as the 'EAA-Loop'. This region extends over approximately twenty amino acids and conforms to the consensus 'EAA---G---------I-LP'. Mutations in this motif were shown to have variable effects when introduced individually into the MalF and MalG MSD's of the *E. coli* maltose permease. However, simultaneous mutation of identical residues in both EAA motifs resulted in decreased maltose accumulation (Mourez *et al.*, 1997). Located in a cytoplasmic loop of the MSD, this sequence is thought to be involved in the interaction of the MSD with the α-helical sub-domain of the NBD. This notion is supported by the observation that specific mutations in the helical domain of MalK

can suppress the phenotypic effects of mutations in the EAA motif.

In order to transport a molecule from one side of the membrane to the other, the MSD's of an ABC transporter must physically interact with their substrate. Interactions between substrate and MSD's are still poorly understood for most ABC transporters. These interactions are probably dependent on the overall structure of the translocation pathway, rather than the presence of locally conserved sequence motifs. In this way residues distant from one another in the primary sequence of the MSD's may become juxtaposed with one another in the fully folded, membrane-bound complex and contribute to a composite substrate binding site. A detailed understanding how the two MSD's of the ABC transporter interact with one another and their substrate is of particular importance with regard to the multidrug transporters such as LmrA and P-glycoprotein. Any pharmaceutical agent aimed at blocking or modifying the activity of these transporters must presumably interact specifically with their MSD's.

Exactly how MSD's interact with their substrates is likely to vary significantly between different families of ABC transporters. For example P-glycoprotein is known to export many amphipathic substrates directly from the interior leaflet of the cytoplasmic membrane. This implies that the substrate binding site(s) of P-glycoprotein must be exposed to the interior of the lipid bilayer. Qu and Sharom have recently demonstrated the presence of such a lipid-exposed substrate binding site (Qu and Sharom, 2002). On the other hand, the substrate binding sites of the bacterial permeases must be exposed to the exterior of the cytoplasmic membrane in order to interact with their liganded SBP's.

Lower resolution techniques have been employed to study the structure of the MSD's of certain ABC transporters. 2D cryo-electron microscopy and single particle analysis have been used to construct projection maps for mammalian P-glycoprotein, mammalian MRP1, and the YvcC ABC transporter of *Bacillus subtilis* (Rosenberg *et al.*, 1997; Rosenberg *et al.*, 2001; Chami *et al.*, 2002). The map of P-glycoprotein indicates the presence of two arc shaped electron densities, related to one another by pseudo-two-fold symmetry. At the centre of the structure lies an asymmetric 'pore' that is exposed to the extracellular face of the membrane. Each of the two arc shaped

domains consists of three major electron densities that are proposed by the authors to represent paired clusters of transmembrane helices. This observation is consistent with a model in which each homologous half of P-glycoprotein contributes six MSD's to a translocation pathway, and the overall structure being generated by a total of twelve helices. It should be noted that in this structure the two are shaped domains are not arranged totally symmetrically and that gaps between the domains may allow access to the central cavity from the interior of the lipid bilayer. Once again, these observations are consistent with the ability of P-glycoprotiens to extrude drugs directly from the inner leaflet of the membrane.

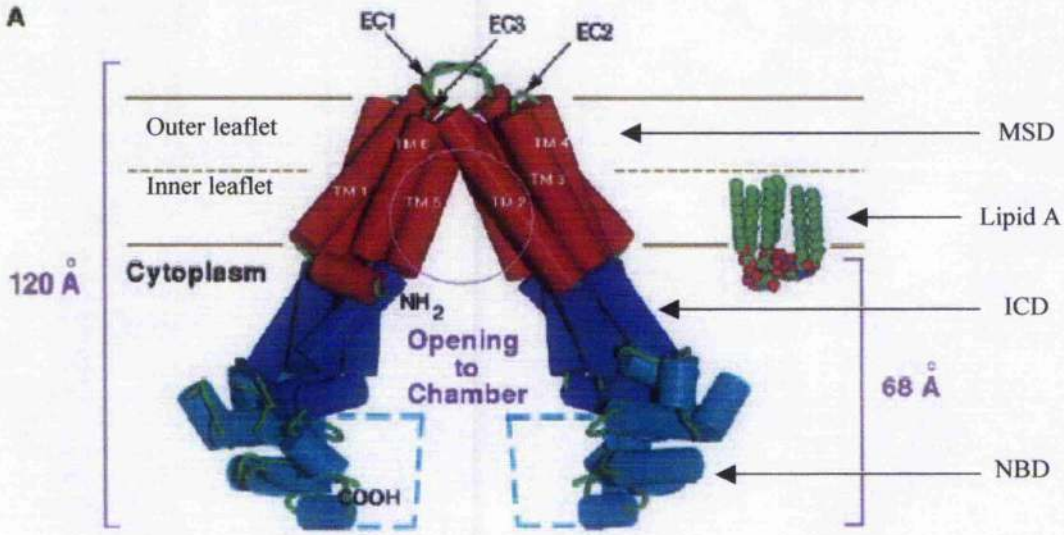
An alternative model to the 'two-times-six' paradigm for the arrangement of ABC transporter MSD's was proposed by Jones and George (Jones and George, 1998). In this radically different model the authors proposed that the MSD's interact to generate two membrane-embedded 16-stranded beta-barrels which form the transmembrane translocation pathway. However, the recent publications of the structures of the BtuCD ABC transporter and the MsbA transporter seem to discount this hypothesis and support the two-times-six-model. Indeed, these two structures, discussed in the final section of this chapter, have provided the greatest ever insight into how the various components of the ABC transporter complex interact with one another and even suggested possible mechanisms by which the transporters may operate.

1. 2. 10 Structure of the ABC transporter complex

The recent publication of two complete ABC transporter structures represent major breakthroughs in our understanding of these proteins. The first publication was the structure of the *E. coli* MsbA transporter, a lipid 'flippase' that catalyses the transport of lipid A from the cytoplasmic side of the bacterial cell membrane to the outer leaflet (Chang and Roth, 2001). The extreme difficulties encountered when trying to crystallise membrane transport proteins are well illustrated by the work leading up to the publication of the MsbA structure. A total of 20 different proteins from 14 organisms were tested for expression and purification before the authors conducted an astonishing 96,000 crystallisation trials! Initial screens produced MsbA crystals diffracting to just 6.5 Å, insufficient to trace the complete backbone of the protein. After a further round of refinement produced crystals diffracting to 4.5 Å resolution which were then used to solve the structure.

The MsbA structure confirmed such fundamental hypotheses as the α-helical nature of the membrane-spanning domain, and that the MSD forms the translocation pathway across the membrane. Although less well resolved than the MSD's, electron densities corresponding to the NBD's could be seen at positions near to the cytoplasmic face of the membrane. Furthermore, two MsbA subunits were required to generate the complete 4 domain transporter, consistent with predictions, and the fact that the MsbA protein is a half-transporter. The crystal structure of the MsbA dimer is shown in figure 1.15. It illustrates that the two α-helical MSD's are tilted towards one another at an angle of 30-40° with respect to the plane of the membrane generating an overall cone shape. The dimer interface occurs primarily between transmembrane helices 2 and 5 of each monomer at a position corresponding to the outer leaflet of the membrane. No crystal contacts between monomers are observed at the base of the cone shaped structure, the NBD's apparently separated by a distance of approximately 50 Å. As a consequence of their tilted orientations, the MSD's of the two monomers form a conical chamber within the inner leaflet of the membrane bilayer. A concentration of positively charged residues in this region suggests that the interior of the chamber may be a relatively polar environment.

Figure 1.15 – The MsbA homodimer



An unexpected feature of the crystal structure is the presence of a bulky 'intracellular domain (ICD)' folded between the MSD and NBD of each monomer. The ICD is

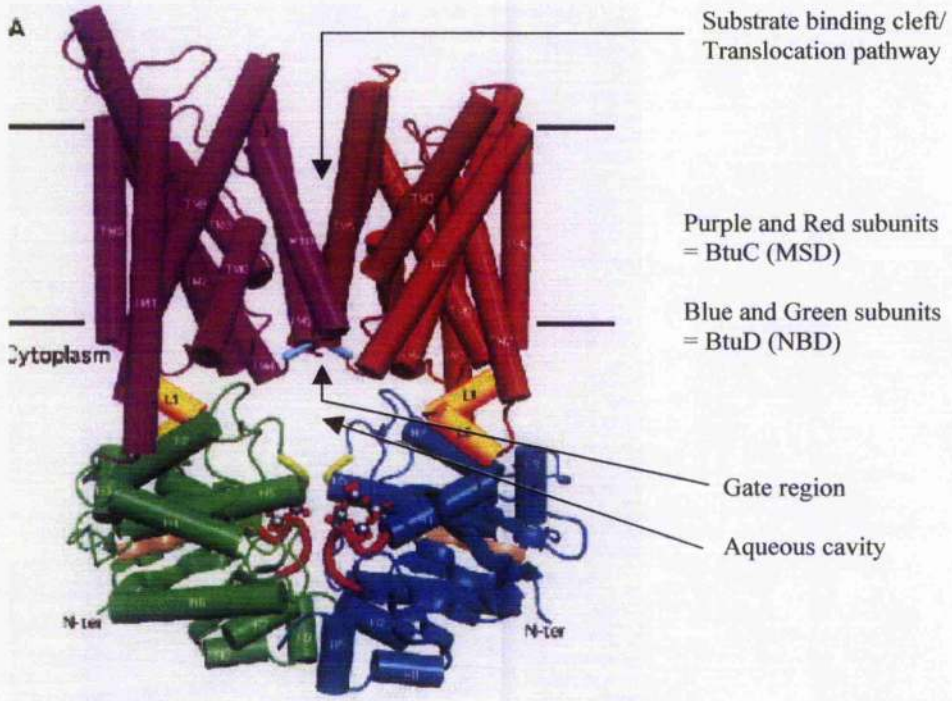
formed from extended loops that connect the α -helical segments of the MSD at the intracellular side of the membrane. The ICD forms extensive crystal contacts with both the NBD below and the MSD above, and packs directly above the ABC signature-sequence motif of the NBD. This raises the possibility that the ICD plays a part in signal transduction within the dimer, connecting the molecular events of ATP binding and hydrolysis to structural rearrangements within the MSD. The presence of the ICD's causes the NBD's to project away from the cytoplasmic surface of the membrane, rather than packing directly against it. The ICD's effectively extend the cone shaped chamber formed by the MSD's well into the cytoplasam. At its base the conical cavity is now wide enough to allow access of lipid A to the interior chamber of the transporter.

Chang and Roth have suggested a possible transport mechanism based upon the structure of MsbA. They propose that lipid A enters the base of the conical cavity from the inner leaflet of the bilayer, triggering ATP binding/hydrolysis by the NBD's. This process drives structural rearrangements within the dimer, possibly involving association of the NBD's, eventually allowing lipid A access to the outer leaflet of the membrane. The final proposed stage involves nucleotide exchange at the NBD's and a 'resetting' of the dimer complex to its initial state. The authors even propose that a similar mechanism may drive multidrug transporters. Such a mechanism would be consistent with the ability of MDR pumps such as P-glycoprotein and LmrA to extrude drugs directly from the inner leaflet of the cytoplasmic membrane. Additionally, the presence of a relatively large chamber able to accommodate molecules of varying sizes and shapes may account for the observed broad substrate range of the MDR pumps.

The second complete ABC transporter structure to be published was that of the *E. coli* vitamin B_{12} transporter, BtuCD (Locher *et al.*, 2002). Unlike MsbA, which is a homodimeric exporter/flippase, BtuCD is a SBP-dependent importer. The transport complex consists of two molecules of BtuC (forming the MSD's) in association with two molecules of BtuD, the ABC protein. Once again the crystallisation process was a tour-de-force effort involving 28 different ABC transporters, the BtuCD complex eventually yielding the best results. The published structure of BtuCD is resolved to 3.2 Å with all important regions of the molecule clearly visible within the structure.

Shown in figure 1.16 below, the BtuCD structure is somewhat different from that of MsbA.

Figure 1.16 – The BtuCD transporter



The first immediate difference of note is that the BtuC subunits each traverse the membrane 10 times, rather than the 6 times more commonly predicted for ABC transporter MSD's. In common with MsbA, the membrane-spanning regions are α-helices. The 10 helices of each BtuC protein are not all oriented in the same direction, as is the case in MsbA, and pack together in a rather intricate manner. Each BtuC monomer forms contacts with the opposite monomer, thereby contributing to the presumed translocation pathway, and also with the BtuD subunit below. In the structure of BtuCD it should be noted that there is no region analagous to the ICD observed in the structure of MsbA. Each BtuD subunit packs directly against the MSD above it, close to the plane of the membrane.

Perhaps the most significant difference between the two structures regards the nature of the association between the NBD's in the BtuCD structure. In MsbA, the two NBD's are separated by as much as 50 Å whereas in BtuCD the two BtuD monomers are packed in direct contact with one another. Moreover, the nature of the association between the two NBD's is very similar to that observed in the Rad50cd

crystallographic dimer. In both structures, the ABC signature motif of one monomer is juxtaposed with the Walker A motif of the second monomer. This results in the formation of a 'composite' nucleotide-binding site which likely stabilises the dimer structure. Crystal contacts between the two monomers include a number of moderately conserved motifs, including residues of the so-called D-loop and switch regions. Such contacts involving conserved residues suggest that all ABC transporter NBD's may interact in this manner. The dimer interface between the two BtuD subunits is relatively small when compared with other dimeric complexes, just 740 $Å^2$. However, the association of both NBD's with the MSD above is likely to further stabilise the interaction. The authors also suggest that the small dimer interface may be the reason why isolated, purified NBD's do not purify as dimers. In the absence of their MSD's, the weak forces holding the dimers together may be insufficient to maintain the dimer structure and explain the different crystal forms of the MalK and HisP dimers.

The interface between each BtuD monomer and the BtuC subunit above is an interaction worthy of note. A relatively long cytoplasmic loop between transmembrane helices 6 and 7 of BtuC folds into two short α-helices called L1 and L2 (yellow in figure 1.16). L1 and L2 contribute the bulk of the interfacial area between the two proteins. Interestingly, the L1 and L2 regions align with the conserved 'EAA' sequence motif present in the MSD's of the bacterial importers, as well as residues that are known to be important in the function of P-glycoprotein and CFTR. Locher and co-workers hypothesise that regions of other MSD's corresponding to the position of L1 and L2 helices may be important mediators of inter-subunit interactions.

The most likely region of the complex to form the translocation pathway is a V-shaped cleft between the two BtuC subunits (see figure 1.16). The cleft spans approximately two–thirds of the width of the membrane, and is formed by an anti-parallel association of helices 5 and 10 of each monomer (4 helices in total). It should be noted that this is only a putative potential binding site since the crystals were generated in the absence of bound substrate. The cleft is however exposed to the periplasmic side of the membrane, facilitating interaction with the substrate-loaded

SBP. The cleft is closed at its base by two serine residues and two threonine residues, a pair of each contributed by each monomer. This 'gate region' is believed by the authors to control access of the substrate to a large hydrophilic cavity at the interface of all four subunits.

Once again, the authors propose a perfectly plausible transport mechanism based upon a series of binding and structural rearrangements coupled to ATP hydrolysis by BtuD. The differences in the structures of BtuCD and MsbA highlight one of the limitations of X-ray crystallographic techniques as mechanistic tools. Each crystal structure is essentially a snap-shot in time of one of the structures that the complex adopts during the catalytic cycle. So, whilst the two structures presented may look very dissimilar at certain points during transport, they may also adopt similar structures at other times. As such, crystallography may provide clues to the molecular events occurring during the transport cycle, but it is the more traditional biochemical and biophysical techniques that will eventually complete the mechanistic picture.

Chapter 2

Materials and Methods

2. 1 Laboratory reagents and equipment

2. 1. 1 Sources of reagents

General laboratory reagents such as salts and buffers were purchased in the highest available grade from Sigma (UK) Ltd. Microbiological growth media were variously sourced from Oxoid, Difco and Melford Laboratories. Except where otherwise stated enzymes used in molecular biology protocols were purchased from Promga (UK) Ltd.

The sources of additional reagents used in specialised applications are detailed in association with specific protocols.

2. 1. 2 Preparation of solutions

All solutions, including microbiological growth media and molecular biology reagents, were prepared using high quality distilled, deionized water. Gross quantities of reagents were weighed out on a Sartorius digital balance whilst smaller quantities were measured using a Mettler Toledo digital analytical balance.

2. 1. 3 Sterilisation techniques

Where possible materials were sterilised by autoclaving. A typical protocol, such as that used for the sterilisation of growth media, involved heating to 121 °C for 15 minutes at 4 atmospheres pressure in a Priorclave electrically heated autoclave set on liquid cycle. Solutions of heat labile reagents, such as antibiotics, were sterilised by filtration through 0.22µm syringe filters. Certain plasticware unsuitable for autoclaving was sterilised by washing thoroughly in 100% ethanol before drying in a laminar flow hood.

2. 1. 4 Centrifugation

Standard room temperature centrifugation of 1.5 ml microcentrifuge tubes was performed in a Sigma benchtop microfuge fitted with a 24 place fixed angle rotor. Where there was a requirement for cooling a Jouan CR3 refrigerated benchtop centrifuge was used. Larger

volumes of liquids were processed in a Beckman-Coulter Avanti J-E refrigerated centrifuge which accommodated a several different fixed angle rotors. A Beckman JA-10 rotor was used to process volumes up to 3 litres in 500 ml plastic containers, whilst a Beckman JA-21 rotor was used to process up to 400 ml in 50 ml tubes. High-speed ultracentrifugation was performed in a Beckman L8M ultracentrifuge using a Beckman Ti50 fixed angle rotor.

2. 2 Growth and storage of bacteria

All bacterial strains used and their associated genotype are listed in table 2.1 on the next page.

2. 2. 1 Growth media

Liquid cultures of the *Escherichia coli* strains listed were propagated using two main types of growth media: -

Luria-Bertani (LB) medium		2x YT medium	
NaCl	10g/litre	NaCl	5g/litre
Tryptone	10g/litre	Tryptone	16g/litre
Yeast extract	5g/litre	Yeast extract	10g/litre

Solid growth medium was prepared by adding 1.5% w/v agar to liquid LB medium prior to autoclaving. LB medium was used for the growth of most bacterial strains, whilst the richer 2x YT medium was used in expression cultures because of its ability to support higher cell densities.

2. 2. 2 Growth conditions

Liquid cultures were initiated by inoculation of a suitable volume of sterile medium with a single bacterial colony transferred from solid medium. Except where otherwise indicated, liquid *E. coli* cultures were grown at 37 °C with rotary shaking between 170 and 225 rpm. Incubation times varied from several hours to overnight. Solid cultures of

Table 2.1 – Bacterial strains used in these studies

Strain	Genotype	Origin	Application
E. coli NovaBlue	*end*A1 *hsd*R17(r_{K12}^- m_{K12}^+) *sup*E44 *thi* –1 *rec*A$^-$ *gyr*A96 *rel*A1 *lac* [F', *pro*AB *lac*I^{q}ZΔM15 Tn10 (Tet^R)]	Novabiochem	General cloning
E. coli JM109	*end*A1 *hsd*R17(r_{K12}^- m_{K12}^+) *sup*E44 *thi rec*A1 *gyr*A96 *rel*A1 λ- Δ(*lac-pro*AB) [F', *tra*D36, *pro*AB, *lac*I^{q}ZΔM15]	Promega	General cloning Site directed mutagenesis
E. coli TOP10	F' *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) L80 *lac*ZΔM15 Δ*lac*X74 *deo*R *end*A1 *rec*A1 *nup*G *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L	Invitrogen	General cloning Protein expression
E. coli BL21 (DE3)	*E. coli* B F$^-$ *omp*T *had*S (r_B^- m_B^-) *dcm*$^+$ Tet^R *gal end*A	Stratagene	Protein expression
E. coli LMG194	F$^-$ Δ*lac*X74 *gal*E *thi rps*L Δ*pho*A (*pvu* II) Δ*ara*714 *leu*::Tn10	Invitrogen	Protein expression
E. coli BMH 71-18 *mut*S	*thi sup*E Δ(*lac-pro*AB) [*mut*S::Tn10] [F', *pro*AB, *lac*I^{q}ZΔM15]	Promega	Site directed mutagenesis
E. coli WD2	*msb*A mutant	Gift of W.T Doerrler	Complementation studies
E. coli N43	Δ*acr*A mutant	*E. coli* Genetic Stock Centre	Complementation studies

E. coli were prepared by spreading or streaking organisms onto the surface of the solidified medium using disposable plastic microbiological loops or spreaders. Plates were allowed to dry in a laminar flow hood before being covered and incubated at 37 °C in an inverted position overnight.

2. 2. 3 Storage of bacterial strains

Stock cultures of each bacterial strain used were prepared in order to maintain their long term integrity and viability. Stocks were typically prepared by mixing 800µl of an overnight liquid culture with 200µl of sterile 80% glycerol. The mixture was then transferred to a sterile cryovial for storage at -80 °C. Strains stored in this manner were retrieved by scraping a small amount of the frozen stock from the cryovial using a sterile loop, and then streaking this material onto solid growth medium for overnight incubation at 37 °C.

2. 2. 4 Use and preparation of antibiotics

Many bacterial strains, particularly those bearing recombinant vectors, required the presence of antibiotics in their growth media in order to retain their genetic integrity and prevent contamination by extraneous microorganisms. Antibiotic stocks were routinely prepared at 1000 times concentration in water or ethanol as required and filter sterilised before use. Such stocks were stored in aliquots at -20 °C. Stock concentrations: ampicillin 100mg/ml, carbenicillin 100mg/ml, kanamycin 100mg/ml, tetracycline 25mg/ml (50% ethanol), doxorubicin 2mg/ml (DMSO).

2. 3 Nucleic acid purification and analysis

2. 3. 1 Agarose gel electrophoresis

The results of nucleic acid manipulations were routinely analysed by agarose gel electrophoresis. High purity agarose (Roche Molecular Biochemicals) was dissolved in TAE buffer* at an appropriate concentration (typically in the range 0.8-2% w/v) by heating the mixture in a microwave. The molten agarose solution was allowed to cool slightly before the addition of the DNA staining reagent ethidium bromide at a final concentration of 0.1µg/ml. The solution was then poured into a horizontal gel-tray and

allowed to set solid by cooling. Wells for sample loading were set in to the agarose slabs by fitting a comb into one end of the molten agarose solution. Once set, the agarose slabs were placed into electrophoresis tanks and submerged in TAE buffer prior to sample loading. Each DNA sample to be analysed was mixed with 6x loading buffer* and loaded into a separate well in the agarose gel. A sample of a commercially available DNA marker was usually loaded into one well of the gel. Comparison of experimental samples with such markers facilitates estimation of size and quantity of DNA present. Application of an approximately 100V potential difference across the gel promotes migration of the DNA in the samples towards the anode pole of the electrophoresis tank. Progress of DNA through the gel was monitored by the migration of the xylene cyanol and bromophenol blue dyes present in the loading buffer. Once the samples had migrated a sufficient distance, DNA was visualised by exposure to a UV light trans-illuminator. In those instances where a permanent record of the experiment was required the entire gel was photographed using a Polaroid MP4 camera fitted with UV filters and Polaroid 677 black and white film.

* Buffer compositions

1x TAE	6x DNA loading buffer
4.84g Tris base	100mM EDTA (pH 8.0)
1.142ml glacial acetic acid	1% w/v SDS
2ml 0.5M EDTA (pH 8.0)	0.1% w/v bromophenol blue
Made up to 1L with distilled water	0.1% w/v xylene cyanol
	50% glycerol

2. 3. 2 Purification of plasmid DNA

Small-scale purification of plasmid DNA from *E. coli* host strains was performed using the QIAprep Spin Miniprep Kit (Qiagen) according to the protocols provided. A modification of the well known alkaline lysis procedure, the technique is summarised below.

Cells from a 5ml overnight liquid culture are harvested by centrifugation and the culture medium decanted away from the cell pellet. The cell pellet is resuspended in 250µl of buffer P1 supplied with the kit and transferred to a 1.5ml microcentrifuge tube. The cells are then lysed by the addition of 250µl of buffer P2 containing SDS and NaOH. The suspension is thoroughly mixed by inverting the tube several times, thereby ensuring complete lysis. The mixture is then neutralised by the addition of 350µl of buffer N3. The addition of this solution causes cellular proteins and genomic DNA to precipitate out of solution, whilst plasmid DNA remains soluble. Insoluble and soluble components are separated by centrifugation of the mixture at 12,000 rpm for 10 minutes. The supernatant fraction, containing the soluble plasmid DNA, is next loaded onto a spin miniprep column mounted in a 2ml collection tube. The mini-columns contain a positively charged silica-based matrix upon which plasmid DNA becomes immobilised. The supernatant is spun through the resin bed and into the collection tube by centrifugation at 12,000 rpm for 1 minute. The immobilised plasmid DNA is then washed with 750µl of buffer PE (buffered 80% ethanol) by repeating the loading and centrifugation steps. Excess remaining ethanol buffer is removed from the immobilised plasmid by decanting the contents of the collection tube and performing a further 1 minute centrifugation step. Finally, the immobilised plasmid is eluted by transferring the spin column to a clean 1.5ml microcentrifuge tube and adding 50µl of molecular biology grade water. The column is allowed to sit at room temperature for 5 minutes before a final 1 minute centrifugation. The eluted plasmid collects in the microfuge tube whilst the spin column discarded. Purified plasmid DNA is then analysed by agarose gel electrophoresis or used in subsequent procedures. Plasmid DNA purified in this manner is stable for long periods of time if stored frozen at –20 °C.

2. 3. 3 Extraction of DNA from agarose gels

A number of common molecular biology procedures, e.g. cloning reactions, require the extraction of DNA fragments from agarose gels after their analysis. Extraction was commonly performed using the QIAquick Gel Extraction Kit (Qiagen). Once again the manufacturer's protocols were adopted as summarised below.

DNA bands to be extracted were visualised by exposing the gel on a UV transilluminator. Slices of the gel containing a desired DNA fragment were excised using a clean scalpel blade, weighed, and transferred to a microfuge tube for extraction. The QIAquick Kit allows for processing of gel slices of up to 400mg in weight. The gel slice is mixed with a solubilisation buffer (buffer QG), with 30µl of buffer added per 10mg of gel. The slice is then dissolved in the QG buffer by incubation in a 50 °C water bath for 10 minutes with occasional mixing. Next, the DNA/buffer mixture is applied to a spin column containing a silica matrix upon which the dissolved DNA becomes immobilised. The column, in a 2ml collection tube, is subjected to 1 minute of centrifugation at 12,000 rpm. The DNA becomes immobilised whilst the buffer passes through to the collection tube. The DNA is treated with a further 500µl of QG buffer to remove any remaining traces of agarose which may interfere with subsequent steps. The immobilised DNA is washed with 500µl of an ethanol wash buffer (buffer PE) by applying this to the column and repeating the centrifugation process. The collection tube is emptied and the column centrifuged for 2 minutes more in order to remove traces of ethanol from the immobilised DNA. Finally, the DNA is eluted by the addition of 50µl of high purity water to the spin column. The column is incubated at room temperature for 5 minutes before being transferred to a clean 1.5ml microfuge tube. The column is centrifuged for 1 minute to collect the purified DNA solution. Once again, for the majority of subsequent applications, DNA purified in this manner is stable if stored frozen at –20 °C.

2. 4 DNA amplification, cloning and sequencing

2. 4. 1 The Polymerase Chain Reaction (PCR)

PCR was used extensively throughout this work as a means of amplifying and modifying specific genes from *M. tuberculosis* genome. The technique requires the use of oligo-deoxynucleotides as a means of priming PCR amplification. All oligonucleotides used were obtained from a commercial source (Invitrogen Custom Primers) and supplied on a 50nmol scale. Stock solutions of oligonucleotides, supplied in a desalted, lyophilised form, were prepared by resuspending the DNA in high purity water at a concentration of 60pmol/µl. Working solutions of primers were prepared from stocks by dilution to a concentration of 12pmol/µl. Both stocks and working solutions were stored frozen at –20

°C. All possible care was taken to avoid cross-contamination of primer solutions. Primers were generally designed to be between 18 and 27 nucleotides in length and have melting temperatures in the range 55-70 °C.

The names and sequences of all oligonucleotides used in this work are presented in table 2.2 on the next page.

Table 2. 2 – Oligonucleotides

Primer Name	Sequence (5' to 3')	Use (reference)
drrAF	CAT ATG CGC AAC GAC GAC ATG GCG GTG	PCR (chapter 3)
drrAR	AAG CTT TCG CGC GGA CCC CGA CAC CAG	PCR (chapter 3)
drrBF	CAT ATG AGC GGC CCG GCC ATA GAT GCG	PCR (chapter 3)
drrBR	AAG CTT TGG CCG CCT AGC CAA AAC AAT	PCR (chapter 3)
drrCF	CAT ATG ATC ACG ACG ACA AGT CAG GAA	PCR (chapter 3)
drrCR	AAG CTT ATG CGT GCT GGC CCG TCG GTA	PCR (chapter 3)
K48E	GGG CCC CAA CGG GGC CGG CGA GAC GAC CAT G	Mutagenesis (chapter 6)
G47S	GGG CCC CAA CGG GGC CAG CAA GAC GAC CAT G	Mutagenesis (chapter 6)
Y17W	TTA ACG GGG TTC GCA AAG CCT GGG GCA AGG A	Mutagenesis (chapter 6)
W110G	CCG CGG ACT TGC TCA GTC CCC CCA GAC GAC C	Mutagenesis (chapter 6)
W180G	CGG CAA GCT ATT GGG GAT CTG GTG	Mutagenesis (chapter 6)
Y141W	GTG GGC ACC TGG TCC GGC GGA ATG CGC	Mutagenesis (chapter 6)
Trx2W	CTG GTT GAT TTC GGG GCA CAC GGG TGC GGT CCG	Mutagenesis (chapter 6)

Oligonucleotide names ending with 'F' and 'R' represent pairs of PCR primers, *forward* and *reverse*, used for the amplification of specific *M. tuberculosis drr* genes. The underlined sequences 'CAT ATG' and 'AAG CTT' correspond to *Nde*I and *Hin*dIII restriction enzyme sites that were incorporated into amplicons in order to facilitate downstream manipulation of the amplified sequences.

2. 4. 2 Preparation of PCR experiments

PCR reactions were performed using HotStar Taq™ DNA polymerase, a recombinant, thermostable, derivative of *Thermus aquaticus* DNA polymerase (Qiagen). The enzyme is supplied with buffers and additional reagents optimised for use with HotStar Taq. The enzyme requires a 15 minute incubation performed at 95 °C before it becomes active. This feature of HotStar Taq minimises the chances of non-specific priming and extension during the preparation of PCR reactions. Cycling parameters used in PCR reactions were variable depending upon the primer pairs used and the nature of the template DNA. However, the examples of reaction mixture preparation and cycling parameters detailed below are typical of those used throughout this work.

All PCR reactions were performed in sterile, thin walled, 0.5ml capacity PCR tubes free of DNase and RNase contamination (Axygen). Where stocks and dilutions were prepared water of the highest available quality was used.

PCR reaction component	Volume	Stock concentration
10x PCR Buffer (Qiagen)	10µl	/
dNTP's*	2µl	10mM each
Primer 1	5µl	12pmol/µl
Primer 2	5µl	12pmol/µl
HotStar Taq	0.5µl	5 units/µl
Template DNA†	Variable	Variable
Water	Up to 100 µl	/

*deoxyribonucleotides dATP, dTTP, dCTP and dGTP were obtained from Promega (UK) Ltd. as 100mM stock solutions. A mixture of the 4 dNTP's was prepared with each one at a concentration of 10mM.

†A typical PCR reaction used less than 1µg of template DNA.

In certain instances the proprietary reagent *Q-Solution* (supplied with HotStar Taq DNA polymerase) was added to reaction mixtures at a concentration of 20% v/v. This reagent alters the melting properties of duplex DNA and was found to be useful when PCR amplification using standard reaction conditions was poor or failed altogether.

Typical Cycling Parameters

Step 1	1 cycle	95 °C	15 minutes (activation)
Step 2	30 cycles	94 °C	45 seconds
		50-65 °C*	45 seconds
		72 °C	1-3 minutes
Step 3	1 cycle	94 °C	45 seconds
		50-65 °C	45 seconds
		72 °C	10 minutes

*The temperature at which the annealing phase of the PCR was performed was the most highly variable parameter since it was dependent upon the particular primer pair in use. Wherever possible primer pairs with similar melting temperatures (T_m) were designed and used. In initial experiments with any given primer pair an annealing temperature was chosen at 5 °C below that of the primer T_m's.

Thermal cycling was performed in either an Eppendorf Gradient Thermocycler or a Techne Cyclogene Thermocycler. Both of these PCR machines were fitted with heated lids to prevent evaporation of the reaction mixture during the cycling process. The lids are pre-heated to 105 °C before cycling commences. Exact details of reaction parameters are provided in the text associated with specific experiments.

2. 4. 3 Cloning of PCR products

The success of PCR reactions was analysed directly by agarose gel electrophoresis. In the event that a PCR product of the desired size was obtained, the product was purified by gel extraction in order to produce a solution of DNA free of interfering substances (such as protein and free nucleotides). The purified DNA was then cloned into a multi-copy vector, allowing the amplified gene sequence to be maintained in a stable form in a heterologous host (*E. coli*).

The vector of choice for the initial cloning of PCR products was the pGEM-T Easy system (Promega). Taq DNA polymerase exhibits a non-template directed activity that results in all PCR products having a single base overhang at the 3' end. This additional base is virtually always an adenine residue. As such, PCR products readily ligate into linearised vectors possessing complimentary 3' single base thymidine overhangs. The pGEM-T Easy vector used here is just one of many such commercially available vectors. Reactions were performed in 0.5ml microfuge tubes and contained the following components:-

Linear pGEM-T Easy vector DNA	1μl (=50ng)
2x Rapid Ligation Buffer	5μl
Purified PCR amplicon	3μl
T4 DNA Ligase	1μl (5 Units)

Ligation is highly efficient in this particular type of cloning reaction and a large number of recombinant vector molecules are often produced after incubation of the reaction mixture at room temperature for just 1 hour. Alternatively, ligation reactions were allowed to proceed overnight at 16 °C.

2. 4. 4 DNA sequencing

DNA sequencing of recombinant vectors was performed by BaseClear (Leiden, The Netherlands). Sequencing reactions were performed on ABI377 or LiCor automated

DNA sequencers. Sequence data were returned as both text files and chromatogram files. Data was analysed using Vector NTI 6.0 bioinformatics software (InforMax Inc.).

2. 5 Enzymatic DNA manipulation

2. 5. 1 Use of restriction endonucleases

The use of restriction endonucleases was central to the production and analysis of recombinant DNA molecules. Except where otherwise indicated, DNA digestion was performed in sterile 0.5ml microfuge tubes at a temperature of 37 °C. Enzymes and their appropriate reaction buffers were obtained from Promega (UK) Ltd.

Typically, single enzyme restriction digests were performed in a total volume of 20µl and contained the components listed below: -

10x restriction enzyme buffer	2µl
Experimental DNA sample	2-5µl (up to 1.5µg)
Restriction enzyme	1µl (2-10 Units)
MilliQ water	Up to 20µl

Digestion reactions were allowed to proceed for variable lengths of time depending upon downstream processes. For example, mapping of plasmids by restriction digestion (to verify the presence of a recombinant fragment) could typically be achieved with digestion times of 1 hour. However, where DNA was to be sub-cloned from one vector to another, digestion times of up to 4 hours were used in order to ensure that the reaction had proceeded to completion. Additionally, directional sub-cloning of DNA fragments often required the use of two restriction enzymes simultaneously. Such reactions were also allowed to proceed for longer periods of time, once again up to 4 hours. Reactions were terminated by the addition of DNA loading buffer and subsequently analysed by agarose gel electrophoresis.

2. 5. 2 5'-Dephosphorylation of vector DNA ends

5'-Dephosphorylation of vector DNA ends has been shown to reduce the background of false positives in cohesive end cloning reactions. To this end, pre-digested vector DNA to be used in cohesive end cloning was routinely treated with *shrimp alkaline phosphatase* (SAP). Appropriate quantities of enzyme and associated buffer were added to gel purified vector DNA solutions and incubated for 10 minutes at 37 °C. The enzyme was then inactivated by incubating the solution for a further 15 minutes at the elevated temperature of 65 °C. This second incubation prevents carry-over of phosphatase activity into the subsequent ligation reaction.

2. 5. 3 Cohesive end DNA cloning

The production of recombinant vectors for heterologous protein expression required that PCR amplifed *M. tuberculosis* DNA be cut out of the pGEM-T Easy vector and ligated into several different *E. coli* expression vectors. DNA excised from pGEM-T Easy by restriction enzyme digestion was ligated into linearised, dephosphorylated expression vectors that had been digested with the same enzyme(s) such that the two DNA fragments possessed complimentary ends. The ligation of the two fragments was performed using T4 DNA ligase (Promega, 10 Units/µl). Reactions in a total volume of 10µl contained the following components: -

10x Ligase buffer	1µl
T4 DNA ligase	1µl
Vector DNA	1µl
Insert DNA	3µl
MilliQ water	4µl

The relative proportion of vector and insert DNA's used in ligation reactions is known to affect efficiency of the reaction. Whilst the exact amounts of DNA used in ligation reactions were rarely formally quantified, a 3:1 molar excess of insert over vector was approximated and generally found to be sufficient. Cohesive end cloning reactions were

incubated at 16 °C overnight to maximise the number of recombinant DNA molecules generated.

2. 5. 4 5'-Phosphorylation of oligonucleotides

5'-Phosphorylated oligonucleotides were generated for use in site-directed mutagenesis protocols since they are known to improve the efficiency of mutagenesis (see chapter 6). Unphosphorylated mutagenic oligonucleotides were purchased from Invitrogen Custom Primers and then modified using T4 polynucleotide kinase (Promega). Reactions were performed in 0.5ml thin-walled PCR tubes at 37 °C for 30 minutes. The reaction mixtures contained the following components: -

Unphosphorylated oligonucleotide	100pmol
10x T4 kinase buffer	2.5µl
T4 polynucleotide kinase	2.5µl (5 Units)
10mM ATP	2.5µl
MilliQ water	Up to 25 µl

After incubation, the T4 polynucleotide kinase was inactivated by heating the reaction to mixture to 70 °C for 10 minutes. Phosphorylated oligonucleotides were used immediately for the mutagenesis reaction.

2. 6 Use of chemically competent *E. coli*

2. 6. 1 Sources and production of chemically competent *E. coli*

Recombinant plasmids used for both maintenance of heterologous DNA and the expression of heterologous protein were transformed into chemically competent *E. coli* strains. In many instances these strains were obtained from commercial sources in single use aliquots (see table 2.1). However, in other instances it was necessary to produce chemically competent cells in the laboratory. The chemical treatment procedure used to induce competence is outlined below: -

Cells of the desired strain were grown in 50ml liquid LB medium at 37 °C in a baffled conical flask with rotary shaking at approximately 225rpm. When the cell suspension reached an OD_{600} of 0.5 they were immediately transferred to an ice bath and left to cool for two hours. During the cooling period a fresh 50ml preparation of *competence solution*† was made, filter sterilised and allowed to cool in the same ice bath. After two hours the cooled cells were pelleted in a refrigerated centrifuge and the culture medium discarded. The cell pellet was then resuspended in 45ml of the ice cold competence solution and returned to ice for a further 45 minutes. After this final incubation period the cells pelleted once more in a refrigerated centrifuge and the excess solution poured off. The final stage of the procedure entailed resuspending the cells in 3ml of ice cold competence solution plus 2ml of sterile 80% glycerol. This mixture was then split into 50µl aliquots (in thin-walled 500µl PCR tubes) and individual aliquots flash-frozen in a dry ice/ethanol bath. Competent cells produced in this manner were stored at -80 °C.

†Competence solution – 100mM calcium chloride, 70mM manganese chloride, 40mM sodium acetate, pH 5.5.

2. 6. 2 Transformation of chemically competent *E. coli*

Recombinant plasmids were generally transformed into competent cells according to the manufacturer's protocols. Whilst each protocol differs slightly, an outline of the general procedure used is given below. Competent cells produced in the laboratory by the previously described procedure were also transformed according to this protocol.

Aliquots of cells were removed from storage at –80 °C and allowed to thaw on ice. 50-500ng of plasmid DNA (or ligation mixture) was then added to the cells which were gently stirred with the tip of a pipette. The mixture was incubated on ice for 30 minutes. After the incubation period the cells were subjected to heat-shock in a 42 °C water bath for a period of exactly 45 seconds. After heat-shock the cells were returned to ice for 2 minutes. 300µl of liquid LB growth medium was then added to the cell suspension and the newly transformed cells incubated at 37 °C in a rotary incubator shaking at approximately 225rpm for a period of 1 hour. Subsequently various volumes of cells

were spread onto solid growth medium containing appropriate selection antibiotics/reagents and incubated in an inverted position overnight at 37 °C.

2. 6. 3 Selective screening for recombinant bacterial clones

All plasmid vectors used in these studies carried antibiotic resistance determinants as a means of screening for transformed clones. Only successfully transformed bacterial clones carry these determinants, conferring upon them the ability to grow in the presence of the selective antibiotics, whereas non-transformed cells are unable to do so. After this initial screening stage plasmids were isolated from the cells and analysed by restriction digestion and/or DNA sequencing to confirm the presence of the recombinant insert. The most commonly used vectors carried the gene for ampicillin resistance and transformed cells were selected on LB-agar medium containing 50-100μg/ml of the antibiotic.

Blue/white screening for recombinant pGEM-T Easy bearing clones

One of the advantages of using the pGEM-T Easy system for the cloning of PCR products was the ability to screen for recombinant plasmids on the basis of insertional inactivation of the β-galactosidase gene. Cells bearing recombinant plasmids are unable to hydrolyse the artificial lactose analogue X-Gal (5-bromo-4-chloro-3-indolyl- β-D-galactoside), whereas those cells with non-recombinant plasmids (where the β-galactosidase gene has not been disrupted) are able to do so. Cells metabolising X-Gal contain a blue pigment, a by-product of the metabolic process, whilst cells with a disrupted β-galactosidase gene remain white in colour. *E. coli* transformed with pGEM-T Easy constructs were plated onto solid media, the surface of which had been spread with 40μl of 2% w/v X-Gal. Ampicillin (100μg/ml) was added to the medium to select for successful transformants, whilst IPTG was included at a concentration of 1mM to induce expression of β-galactosidase. After 16 hours of growth cells bearing recombinant plasmids remained white in colour and were selected for further analysis.

2. 7 Topoisomerase mediated T-A cloning

The pTrx-DrrA expression construct (see chapter 5) was generated using the 'pBAD/TOPO ThioFusion Expression Kit' (Invitrogen). This kit allows rapid generation of expression constructs through topoisomerase mediated T-A cloning of *Taq*-polymerase-amplified PCR products. Briefly, a fresh PCR reaction was used to amplify the *M. tuberculosis drr*A gene from a plasmid template. The approximately 1 Kb PCR amplicon isolated from an agarose gel slice by standard techniques following electrophoretic separation. The amplified gene was then cloned into the pBAD/TOPO ThioFusion vector in a 5 minute, room temperature reaction performed according to the manufacturer's protocol:-

TOPO cloning reaction components

Fresh PCR product (*drr*A gene)	3µl
Salt solution*	1µl
TOPO vector	1µl
dH_2O	1µl

*Salt solution = 1.2M NaCl, 60mM $MgCl_2$

After completion of the TOPO cloning reaction 2µl of the reaction mixture was used to transform chemically competent *E. coli* TOP10 cells according to standard protocols. Recombinant colonies were isolated by plating the transformation reaction on to solid LB-agar medium supplemented with 100µg/ml carbenicillin. Transformation plates were incubated overnight at 37 °C.

2. 8 Site-directed mutagenesis

The 'GeneEditor *in vitro* Site-Directed Mutagenesis System' (Invitrogen) was used to perform all sited directed mutagenesis experiments described in this work. There follows a brief summary of the mutagenesis protocol:-

A – Preparation of template DNA

Single stranded plasmid DNA was the template for all mutagenesis reactions. This was produced by alkaline denaturation of double stranded plasmid. The denaturation reaction contained the following components:-

Plasmid DNA	5µl (approx. 2µg)
2M NaOH, 2mM EDTA	2µl
dH_2O	13µl

This mixture was incubated at room temperature for 5 minutes before single stranded DNA was precipitated by the addition of 2µl of 2M ammonium acetate (pH 4.6) and 75µl of 100% ethanol. The mixture was then incubated at –80 °C for 30 minutes. Precipitated DNA was collected by centrifugation at 13,000rpm in a bench-top centrifuge for 15 minutes. The ssDNA pellet was washed in 200µl of 70% ethanol, re-collected by centrifugation and then air-dried. The pellet was dissolved in 100µl of dH_2O and a 10µl aliquot was analysed by agarose gel electrophoresis in order to estimate yields.

B – Annealing of mutagenic oligonucleotides

Up to 4 mutagenic oligonucleotides could be annealed to the single stranded plasmid DNA simultaneously. All mutagenic oligonucleotides were 5'-phosphorylated as described in section 2.5.4. Each annealing reaction contained a 'Selection oligonucleotide' provided in the kit which introduces specific mutations into the vector encoded β-lactamase gene. These mutations alter the substrate range of the β-lactamase enzyme and allow selection of clones carrying mutated plasmids on the basis of

resistance to a proprietary antibiotic mixture. The annealing reactions were prepared as follows:-

Template DNA	10µl
Selection oligonucleotide	1µl
Mutagenic oligonucleotide	xµl (1.25pmol each)
10x annealing buffer	2µl
dH_2O	up to 20µl

The annealing reaction was heated to 75 °C and then allowed to cool to room temperature slowly. During this period the oligonucleotides anneal to the template DNA.

C – Mutant strand synthesis and ligation

Various components were added to the annealing mixture in the order shown below. These reagents completed synthesis and ligation of the mutated plasmid DNA.

dH_2O	5µl
10x synthesis buffer	3µl
T4 DNA polymerase	1µl (10 Units)
T4 DNA ligase	1µl (3 Units)

The reaction mixture was incubated at 37 °C for 90 minutes to allow mutant strand synthesis.

D – Transformation of BMH 71-18 *mut*S *E. coli*

This *E. coli* strain, provided with the kit, has a mutation in the *mut*S DNA repair gene which allows the cells to tolerate plasmids with small double strand mismatches. The cells were transformed with 3µl of the mutant plasmid mixture prepared in second C

above. Transformation was performed by a standard heat-shock protocol. Rather than plating the transformed cells onto solid medium, the entire mixture was incubated in 4ml of LB medium containing the proprietary selection antibiotic overnight (37 °C). This allows time for plasmid replication and segregation in such a way as to enhance the number of cells carrying the mutated plasmid.

E – Plasmid Miniprep and selection of mutants

The final stage of the selection procedure was to prepare total plasmids from the overnight culture. This was done by standard plasmid mini-prep (see section 2.3.2). Mutant plasmids were transformed into chemically competent JM109 *E. coli* and cells carrying mutated plasmids were selected on solid LB-agar medium containing the selection antibiotic (overnight incubation, 37 °C).

JM109 clones resistant to the antibiotic selection mixture were used to prepare plasmids by mini-prep. The successful introduction of specific mutations was determined by DNA sequencing.

2. 9 Recombinant Protein Expression

2. 9. 1 Expression of recombinant proteins from pET series plasmids

A number of experiments were performed to express recombinant *M. tuberculosis* proteins in the heterologous host *E. coli* using the pET series expression plasmids (Novagen). pET expression constructs were typically transformed into the host strain BL21(DE3), the DE3 lysogen encoding an IPTG-inducible bacteriophage T7 RNA polymerase required for transcription from the pET-series plasmids. A single recombinant clone, from either a fresh transformation or streaked from a glycerol stock, was selected after overnight incubation at 37 °C on solid LB-agar medium supplemented with the appropriate selection antibiotic. The clone was used to inoculate a liquid starter culture (typically 5ml in LB medium) which was grown overnight at 37 °C with vigorous shaking. Starter cultures were used to inoculate larger expression cultures grown at a variety of temperatures depending upon the experiment in question (16-37 °C). All expression cultures were grown in baffled flasks with rotary shaking between 170-

220rpm to promote aeration. Cultures were supplemented with the appropriate antibiotics required for maintenance of pET expression plasmids. Growth of all expression cultures was monitored spectroscopically by changes in optical density at 600nm.

When each culture reached an OD_{600} of approximately 0.6 (corresponding to mid-log phase), expression of recombinant protein was initiated by the addition of IPTG to a final concentration ranging between 100µM and 1mM. The addition of IPTG to *E. coli* BL21(DE3) initially results in expression of the chromosomally encoded bacteriophage T7 RNA polymerase. This enzyme then drives expression of heterologous genes cloned in front of the T7 promoter sequence present in the recombinant pET plasmid. Where necessary, for example during time-course experiments, 1ml samples were removed aseptically from expression cultures and heterologous protein expression analysed by SDS-PAGE.

2. 9. 2 Expression of recombinant protein from pBAD plasmids

Recombinant expression constructs based upon the pBAD-TOPO ThioFusion vector (Invitrogen) were transformed into one of the two host strains recommended by the manufacturer, *E. coli* TOP10 or *E. coli* LMG194. Recombinant clones were isolated by plating onto solid LB-agar medium supplemented with carbenicillin at 100µg/ml. Single clones were used to inoculate overnight starter cultures in 5ml of LB medium. Starter cultures were again grown at 37 °C with vigorous shaking. These cultures were then used to inoculate large-scale expression cultures in either LB or 2x YT medium supplemented with 100µg/ml carbenicillin. Cultures were grown at a variety of temperatures between 17 and 37 °C in baffled flasks with rotary shaking at 170-220rpm. As with the pET expression cultures, growth was monitored by an increase in OD_{600}.

Expression of recombinant protein from pBAD plasmids was initiated by the addition of a filter-sterilised arabinose solution. Various arabinose concentrations between 0.0002% and 2% (w/v) were tested in order to identify the optimal concentration required for expression of soluble protein. Protein expression was induced when the cells reached mid-log phase (OD_{600} approx. 0.6). Once again expression of recombinant protein was

monitored by withdrawing 1ml samples from the cultures at various time-points before and after induction. Samples were then analysed by SDS-PAGE.

2. 10 Analysis of recombinant expression

2. 10. 1 Preparation of total cell protein for analysis by SDS-PAGE

In order to monitor expression of recombinant proteins 1ml samples of culture were removed both prior to induction and at various time points thereafter. Cells were harvested by centrifugation at 13,000rpm in a bench-top microcentrifuge and the supernatant discarded. Pelleted cells were solubilised directly in an appropriate volume of SDS-PAGE sample buffer (typically 100-300µl). Each sample was boiled for 5 minutes in a water bath to ensure complete solubilisation. Samples were loaded into separate wells of an SDS-PAGE gel and analysed as described below.

2. 10. 2 SDS-PAGE

SDS-PAGE was used extensively as a means of monitoring both recombinant protein expression and the efficacy of protein purification procedures. All SDS-PAGE experiments were performed using pre-cast Bis-Tris polyacrylamide gels obtained from a commercial source (NuPage gels, Invitrogen). Two different gel types were used dependent upon the protein samples under analysis; 4-20% polyacrylamide gradient gels for samples containing proteins of widely variant molecular weights, and uniform 12% polyacrylamide gels for less complex samples.

Protein samples were mixed with an appropriate volume of 10x sample buffer (see below) and loaded into the wells of a vertically mounted gel. Pre-stained molecular weight standards were loaded into one well of each gel to allow estimation of protein sizes and to allow visualisation of protein separation (SeeBlue Markers, Invitrogen). The entire gel, once loaded, was clamped vertically into an electrophoresis apparatus and both the inner and outer buffer chambers filled with 1x NuPage MOPS running buffer (see below). Samples were separated by the application of a 200V current for approximately 45 minutes.

After completion of the electrophoresis procedure each gel was removed from the apparatus and rinsed gently for 10 minutes in de-ionised water to remove the running buffer. Separated proteins were visualised by staining with 'Gelcode Blue' protein stain (Pierce). After rinsing each gel was immersed in 20ml of Gelcode Blue contained in a 10cm square petri dish. The gel was placed on a rocker-table for approximately 1hr to allow staining to reach completion. After 1hr the stain was discarded and the gel re-immersed in 20ml of de-ionised water for destaining. Following the destaining procedure protein bands could be clearly visualised and photographed.

10 x Sample Buffer	20 x NuPAGE MOPS Running Buffer	
4 ml dH_2O	1M MOPS	(104.6g)
1ml 0.5M Tris-HCl (pH 8.6)	1M Tris base	(60.6g)
1.6ml 10% SDS (w/v)	69.3mM SDS	(10.0g)
0.4ml β-mercaptoethanol	20.5mM EDTA	(3.0g)
0.05% bromophenol blue (w/v)	dH_2O up to 500ml	

2. 10. 3 Western blotting

A number of Western blots were performed during this work in order to detect the expression of recombinant proteins. Firstly, the protein sample under analysis was separated by SDS-PAGE as described above. Following electrophoresis the gel was washed extensively in several aliquots of dH_2O to remove traces of running buffer. The gel was then soaked in 1 x NuPAGE transfer buffer prepared according to the manufacturer's protocol:-

1 x NuPAGE transfer buffer

Bicine	4.08g	Methanol	100ml
Bis-Tris	5.23g	dH_2O	900ml
EDTA	0.29g		

Whilst the gel was soaking several pieces of Whatman 3M filter paper were cut to the same size as the separating gel (8cm x 10cm). A similar sized piece of PVDF blotting

membrane (Amersham) was also prepared. Both the membrane and filter papers were soaked in transfer buffer.

The blotting apparatus was assembled as follows. Firstly two abrasive nylon fibre pads were soaked in transfer buffer and placed into the cathode chamber of the blotting tank. Two pieces of pre-soaked filter paper were layered on top of the blotting pads. Next the SDS-PAGE gel was placed on top of the filter papers and all air bubbles removed. The gel was then covered with the PVDF blotting membrane followed by two further pieces of filter paper. Finally two more nylon fibre blotting pads were placed on the gel/membrane sandwich. The blotting cassette was sealed and placed into an empty electrophoresis tank. Both the anode and cathode chambers of the blotting apparatus were filled with 1x transfer buffer and transfer of proteins from the gel to the membrane was achieved by the application of a 30V current for a period of 1 hour. After transfer the apparatus was disassembled and transfer of proteins to the membrane verified by the effective transfer of the pre-stained SeeBlue molecular weight standards.

Following transfer of separated proteins to PVDF membrane all blots were processed using the 'Immun-Star Chemiluminescent Protein Detection System' (Bio-Rad) and associated protocols. The membrane was first washed at room temperature with 25ml of Tris-buffered saline (TBS; 20mM Tris base, 500mM NaCl, pH 7.5). After 10 minutes the wash solution was decanted and the washing procedure repeated with a second 25ml aliquot of TBS. After washing the membrane was 'blocked' with 25ml of 0.2% w/v non-fat dried milk in TBS. Blocking was performed over a period of 1 hour at room temperature with gentle agitation provided by a rocker table. After the blocking stage of the procedure the membrane was washed twice with 25ml of TBS for ten minutes at room temperature to remove unbound protein.

All blots performed in this work were designed to detect the presence of the His-Tag sequence. This was achieved through the use of a commercially available mouse monoclonal antibody raised against the His-Tag (Sigma). The primary antibody was diluted 1 in 3000 into 25 ml of blocking solution and incubated with the transfer membrane for 2 hours. This period was generally long enough to allow efficient binding

of the mouse monoclonal antibody to His-Tagged proteins. After binding the primary antibody solution was decanted and the membrane washed three times with separate 25ml aliquots of TBS. The second stage of the detection procedure used a goat anti-mouse IgG secondary antibody conjugate to amplify the signal from the primary antibody. The secondary antibody conjugate was diluted 3 parts in 10 in blocking buffer and incubated with the membrane for up to 2 hours. Finally the membrane was washed with three more 25ml aliquots of TBS.

Blot development was performed as follows. The membrane was removed from the TBS wash buffer and placed on a flat surface. The surface of the blot was covered with 3ml of a chemiluminescent substrate solution and incubated for 5 minutes. The membrane was then drained of excess liquid and sealed in a heat-sealable plastic bag. The bag was fixed to the interior of an X-ray film cassette and a piece of film placed over the blot under darkroom conditions. Finally, the film was exposed for time periods of between 1 and 10 minutes until a suitably strong signal was observed.

2. 11 Protein purification and miscellaneous methods

2. 11 .1 Cell lysis and sub-cellular fractionation by differential centrifugation

Sub-cellular fractionation was frequently performed as an initial crude purification procedure and also as a means of identifying the cellular location of particular proteins. Bacterial cells from expression cultures were initially separated from the growth medium by centrifugation at 4000-6000rpm (Beckman JA-10 rotor). The bacterial pellet was then re-suspended in an isotonic buffer to prevent osmotic lysis of the cells. Such re-suspension buffers typically contained 100-400mM NaCl, 10% glycerol and a buffering salt (e.g. Tris-HCl) at 50mM. The pH of such solutions was generally maintained at pH 8.0 so as to inhibit the action of cellular acid proteases. Protease activity was further prevented by the addition of EDTA-free protease inhibitor tablets to the re-suspension buffer (Roche Molecular Biochemicals, 1 tablet/50ml of buffer). Cell pellets were typically resuspended with 5ml of ice-cold buffer per gram (wet weight) of cells.

Once re-suspended, cells were lysed by mechanical disruption in a Constant Systems Cell Disruption apparatus. This device uses a hydraulic ram to force bacterial cells through a very small aperture at high pressure (20-30 KPa). Upon passing through the aperture the

cells impact upon a fixed target which is believed to cause lysis. The cell lysate then drains by gravity flow into a separate chamber where it is collected. At all stages after disruption cell lysates was kept on ice to reduce proteolysis.

The initial crude lysate thus generated consists of three basic components that can be separated by differential centrifugation; dense insoluble material such as fragments of the cell wall, soluble cytosolic components of the cell, and an emulsion of small hydrophobic lipid/protein/carbohydrate vesicles derived from cellular membranes. Cell wall material was not required in any of the studies presented here and was removed from lysates by centrifugation at 4000-6000rpm in a Beckman JA-10 rotor. The supernatant liquid, containing soluble proteins and membrane vesicles, was carefully decanted so as not to disturb the pelleted material. The supernatant was further fractionated where necessary into membrane and soluble components by high-speed ultracentrifugation at 100,000x g for 1.5 hrs (43,000rpm, Beckman Ti-50 rotor). This procedure resulted in the generation of a supernatant containing cytosolic proteins and small molecules and a pellet of membraneous material. The supernatant was decanted off and analysed by SDS-PAGE or used as the starting material for protein purification. The membrane pellet could be emulsified in a small volume of buffer by repeated passage through a syringe needle. At this point integral membrane proteins could be extracted from membrane lipids through the addition of buffer containing a biological detergent (e.g. DDM). The detergent solubilised membrane proteins were then used as the starting material for purification procedures. Alternatively, total membrane protein was analysed by SDS-PAGE after solubilisation in SDS-PAGE sample buffer.

2. 11. 2 Purification of His-Tagged proteins by IMAC chromatography

Throughout this work IMAC chromatography was used as the primary means of protein purification. All heterologous expression constructs were designed in such a way as to encode a string of six consecutive histidine residues at either the N- or C-terminus of the recombinant protein. This so-called 'His-Tag' sequence exhibits a high affinity for immobilised divalent cations and forms the basis of the IMAC purification procedure (Porath *et al.*, 1975).

The basic IMAC purification protocol relies upon the specific binding of recombinant His-Tagged proteins to a metal affinity resin loaded with Ni^{2+} ions. In this work two different metal affinity resins were used, Ni-NTA agarose resin (Qiagen), and Ni-IDA sepharose resin (Amersham-Pharmacia). The vast majority of non-tagged proteins do not interact with these materials and thus pass straight through the resin bed. Non-specifically bound proteins can be removed by washing the resin with buffers containing low concentrations of imidazole, a structural analogue of the histidine residues that constitute the affinity tag. Finally, His-Tagged proteins are eluted from the resin material by the addition of buffers containing a much higher imidazole concentration.

An enormous number of variations on the basic IMAC procedure are possible and include the use of different affinity resins, binding methods, washing protocols and buffer compositions. In fact, all of these factors were found to be important variables in the optimisation of the purification procedures used here. The exact experimental details of the various purification protocols used in this work are described in association with individual experiments (see chapters 4 and 5). The results of the various protein purification experiments were analysed by SDS-PAGE.

2. 11. 3 Protein dialysis and concentration

In a number of instances, particularly following IMAC chromatography, dialysis was used to remove small molecules from protein preparations. This was necessary because even low levels of contaminants were found to have a significant bearing on the outcome of downstream experiments. For example the presence of even tiny amounts of inorganic phosphate ions was found to interfere with ATPase assays, and imidazole molecules gave rise to high background readings in measurements of intrinsic protein fluorescence. Purified proteins were dialysed in 'Slidalyzer' dialysis cassettes (Pierce, 10 kDa MWCO). The loaded cassettes were generally dialysed against approximately 3 litres of dialysis buffer cooled to 4 °C. Dialysis was continued for 3 hours in a cold room with continuous stirring of the buffer to promote rapid equilibration. Following dialysis, protein solutions were either concentrated (see below), used directly for biochemical experiments, or snap-frozen in dry-ice/ethanol and stored at –80 °C.

Concentrated protein stocks were prepared using 'Centricon' centrifugal filter devices (Millipore). Protein concentration in these devices is achieved by ultrafiltration of the protein sample through a membrane that allows the passage of only low molecular weight solutes. In this way solvents and small molecules pass through the membrane, whilst high molecular weight molecules such as proteins are retained in a much reduced volume. The ultrafiltration process is driven by centrifugal force. Generally up to 2ml of dilute protein solution was transferred into the upper chamber of a device and subjected to centrifugation at 5000 x g in a Beckman JA-21 fixed angle rotor. Centrifugation times were varied depending upon the desired protein concentration and the ability of the protein to remain in solution. At 4 °C 2ml of the initial solution could be reduced to approximately 200µl in a period of 1 hour. After the concentration procedure protein concentration was determined using the BCA protein assay (see next section) or frozen at –80 °C.

2. 11. 4 Protein concentration measurements – The BCA protein assay

All protein concentration determinations were performed using the 'BCA Protein Assay Kit' (Pierce). This is a colorimetric assay in which proteins react quantitatively with a reagent solution to produce an increase in absorbance at 562nm. The assay maintains linearity over a wide range of protein concentrations from 20-2000µg/ml. Briefly, 100µl of the protein sample under investigation is transferred to a 2.5ml plastic cuvette and mixed with 2ml of the BCA working reagent. The reaction is then incubated at 37 °C for 30 minutes to allow colour development. After the incubation period reactions were cooled to room temperature and Abs_{562nm} measured against a dH_2O blank. Protein concentrations were determined by comparison with a series of standard reactions containing known concentrations of BSA (0-2mg/ml). This allowed the generation of a calibration curve of protein concentration vs. Abs_{562nm}. All standard and experimental reactions were performed in duplicate and corrected for background absorbance.

2. 12 Biochemical Assays

2. 12. 1 ATPase assay – EnzChek Phosphate Assay System

The EnzChek Phosphate Assay System (Molecular Probes) was used to conduct a number of ATPase activity assays (see chapter 6). The system relies on the

spectrophotometric detection of inorganic phosphate (P_i) released in to the assay medium by the activity of ATPase enzymes. P_i release is enzymatically coupled to phosphorolysis of an artificial substrate (MESG). The phosphorolysis reaction generates a molecule with a strong absorption at 360nm, the production of which can be monitored in real-time. The components of the reaction mixture (see next page) were prepared in 1.5ml plastic cuvettes. Generally, ATPase activity was initiated by the addition of $MgCl_2$ to the reaction mixture. ATPase activity was monitored as an increase in absorbance of the reaction mixture at 360nm over time (using a dH_2O blank as the reference sample).

ATPase reaction components

MESG substrate solution	200µl
Purine nucleotide phosphorylase	10µl (1 Unit)
ATP	x µl
$MgCl_2$	y µl
Experimental enzyme	z µl
MOPS buffer (50mM)	790 – x – y – z µl

It is the purine nucleotide phosphorylase component of the reaction mixture which couples release of P_i by the ATPase enzyme to phosphorolysis of the MESG substrate.

2. 12. 2 The 'Malachite green' P_i release assay

This assay was used extensively throughout this work. The methodology described below is essentially a modification of the ATPase assay developed by Harder and colleagues (Harder *et al.*, 1994). The technique once again relies on the spectrophotometric detection of P_i released by ATPase enzyme activity. Inorganic phosphate is found to react quantitatively with an acidified mixture of ammonium molybdate and the dye malachite green to produce a phosphomolybdate-malachite green complex which absorbs light strongly in the range 600-630nm.

ATPase reaction mixtures were prepared in 800µl volumes containing the following components at appropriate concentrations (see chapter 6): ATP, buffer, experimental ATPase enzyme, $MgCl_2$. Typically each reaction would be prepared in the absence of $MgCl_2$ and pre-incubated for 10 minutes at 37 °C in a thermostatically controlled heating block. ATPase activity would then be initiated by the addition of Mg^{2+} to the reaction mixture. Immediately after induction of ATPase activity, and at various time points thereafter, 45µl samples of the reaction mixture were withdrawn and mixed with 5µl of 500mM EDTA in a separate well of a microtitre plate. The purpose of the high EDTA concentration is to rapidly chelate Mg^{2+} ions and immediately halt further ATPase activity.

After collection of a complete set of samples (typically 12 samples collected over a 10 minute period) the amount of inorganic phosphate in each sample was analysed by the addition of 100µl of malachite green detection reagent (components shown below).

Malachite green P_i detection reagent

3 parts 0.045% w/v malachite green (in dH_2O)

1 part 4.2% w/v ammonium molybdate (in 6N HCl)

0.01% v/v Tween 20

The addition of detection reagent to P_i-containing samples resulted in the rapid formation of a green coloured malachite green-phosphomolybdate complex. The absorbance of each sample was measured at 610nm in a plate reader. Measurement of absorbance was performed as rapidly as possible after the addition of the detection reagent since the low pH of this reagent might cause acid hydrolysis of un-hydrolysed ATP. The absorbance of the zero time point sample was subtracted from all subsequent absorbance readings to correct for non-enzymatic P_i release and P_i contamination in reagents.

P_i in each sample was accurately quantified by the use of standard curves prepared in parallel with each set of ATPase reactions. Various known quantities of P_i were mixed with 5µl of 500mM EDTA in a separate well of a microtiter plate and the total volume

brought up to 50µl with MOPS reaction buffer (50mM MOPS, 100mM NaCl, pH 7.4). The amount of P_i in the standards was typically 0-1500 picomoles. 100µl of detection reagent was added to each standard sample and the absorbance at 610nm measured at the same time as the experimental samples derived from the ATPase assay reaction mixture. A typical standard curve produced by this technique is shown in figure 6.1.

The amount of phosphate present in the various experimental samples was used to calculate the enzymatic ATPase activity in terms of nanomoles of P_i released/minute/mg of protein.

2. 12. 3 Steady-state protein fluorescence measurements

Measurements of intrinsic protein fluorescence were performed at room temperature in a Jasco FP-750 fluorescence spectrophotometer. Excitation light at 285nm was selected via a series of chromating mirrors and fluorescence of the sample measured at right-angles to the direction of incident light using a photomultiplier tube to amplify the signal.

For the purpose of the experiments described here reaction mixtures were prepared in a total volume of 200µl and transferred to a Hellma Suprasil quartz fluorescence cuvette. Protein fluorescence spectra were collected at wavelengths between 295 and 400nm. Both the excitation and emission band-widths were set to 5nm. The detection sensitivity of the photomultiplier tube was set to medium.

The effects of ligand binding to proteins were measured as changes in protein fluorescence spectra resulting from the addition of small volumes of concentrated ligand to the protein solution. The ligands used during the studies were ATP (100mM stock solution in dH_2O) and $MgCl_2$ (100mM stock solution in dH_2O). ATP is known to absorb both incident light and emitted light and can cause a significant decrease in the intensity of fluorescence emission spectra. This phenomenon, known as an inner filter effect, was a complicating factor in measuring the effect of ATP binding on intrinsic protein fluorescence. The magnitude of the inner filter effect caused by various concentrations of ATP was determined by placing a second fluorescence cuvette in the excitation beam

between the light source and the protein sample. This cuvette was filled with reaction buffer and an emission spectrum collected in the absence of ligand. ATP was then added at various concentrations to the control cuvette and the effect on emission spectra determined as a percentage decrease in fluorescence. Such data was used to correct the fluorescence spectra observed when ATP was mixed with protein directly.

2. 12. 4 Stopped-flow fluorescence spectroscopy

Stopped-flow fluorescence data were collected in an Applied Photophysiscs stopped-flow device. Experiments were designed to measure transient changes in protein fluorescence caused by the binding of ligands such as ATP and magnesium ions. Interactions between protein and ligand occurred in a mixing chamber maintained at 22 °C. Delivery of protein and ligand to the mixing chamber was achieved by forcing equal volumes of each solution (approx. 100µl) in to the chamber from external reservoirs using nitrogen-driven hydraulic plungers. Excitation light at 285nm was provided to the mixing chamber by means of an argon lamp and two serially connected monochromators. A photomultiplier tube was used to collect and measure the intensity of emission light at right-angles to the incident excitation beam. Cut-off filters placed over the window of the detection photomultiplier tube were used to select the wavelengths above which fluorescence data were collected. When measuring changes in tryptophan fluorescence a 320nm cut-off filter was used. When measuring changes in the fluorescence of the external probe MANT-ATP a 420nm filter was chosen. Changes in fluorescence were observed as an alteration in the output signal from the photomultiplier tube. Fluorescence data following mixing of protein with ligand were collected over logarithmic time-bases ranging from 10 seconds to 1000 seconds.

Chapter 3

Initial cloning and sequence analysis of the *Mycobacterium tuberculosis drr*ABC operon

The *drr*ABC operon was initially identified by Cole and colleagues during the *M. tuberculosis* genome sequencing project (Cole *et al.*, 1998). During the annotation phase of the project all gene sequences, and the proteins encoded therein, were scanned for homology to sequences of known function. The genes of the *drr* operon were identified as being likely components of a transmembrane efflux system, with *drr*A bearing sequence motifs typical of ABC transporters, and the *drr*B and C genes exhibiting characteristics typical of transmembrane proteins. Moreover, the genes seemed to encode proteins homologous to those of a known efflux transporter, the DrrAB system of *Streptomyces peucetius*. In this organism two genes, *drr*A and B, encode the ATP-binding and transmembrane domains of an ABC transporter mediating producer self-protection against the antibiotic molecules *daunorubicin* and *doxorubicin* (Kaur, 1997).

These observations raise a number of interesting questions with regard to the biology of *M. tuberculosis*. For example, is it possible that *M. tuberculosis* is producing and secreting one or more as yet unidentified bioactive secondary metabolites? Is it possible that ABC transporter efflux systems contribute to the high level of intrinsic drug resistance exhibited by *M. tuberculosis*? Could it be that the *drr*ABC system has some alternate but uncharacterised physiological function?

This chapter reports the initial stages of an investigation into the function of the *M. tuberculosis drr*ABC genes. The genes were first amplified from *M. tuberculosis* genomic DNA, cloned, re-sequenced and then compared with the published sequence data from the *M. tuberculosis* genome sequencing project. An *in silico* analysis of the protein sequences encoded by the *drr*ABC genes was used to identify homologous systems in related mycobacterial pathogens and suggest possible physiological roles for the transporter. Finally, the amplified *drr* genes were used to generate recombinant expression vectors suitable for the production of Drr proteins in the heterologous host *Escherichia coli*.

3. 1 PCR amplification of the *drr*A, B, and C genes from *M. tuberculosis* genomic DNA

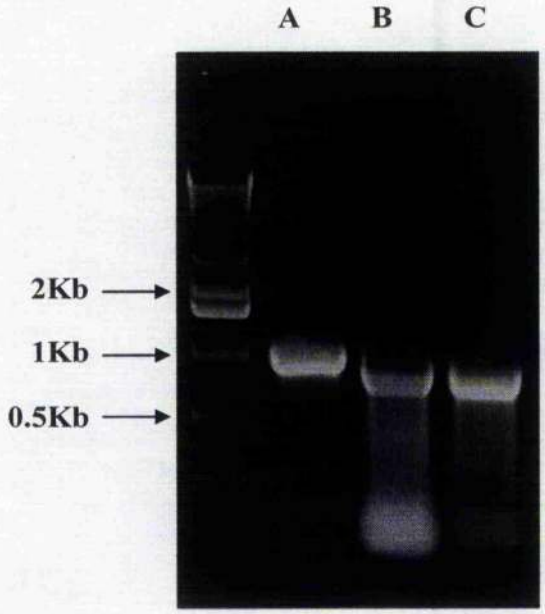
The published sequences of the *drr*A, B and C genes were used to design oligonucleotide PCR primers for the specific amplification of the genes from genomic DNA. The three genes constituting the *drr*ABC operon are located on segment 129 of the *M. tuberculosis* genome, deposited in the GenBank database under accession number Z83857. The operon occupies a total of 2689bp of the genome (*drr*A = 996bp, *drr*B = 870bp, *drr*C = 831bp). *M. tuberculosis* (strain H37Rv) genomic DNA was a kind gift from Dr. S. Ali (Dept. of Biochemistry, University of Glasgow).

The PCR primers used for amplification, detailed in table 2.2, were also used to introduce additional nucleotides at the 5' and 3' ends of each gene. These additional sequences encoded restriction enzyme recognition sites to allow excision of the sequences from the initial cloning vector. This was done in order to simplify the process of sub-cloning the amplified sequences from the initial cloning vector into expression vectors. Each of the 3' primers were designed to eliminate the native stop codon of the gene and introduce a 3' *Hin*dIII recognition site. The 5' primers maintained the native ATG start codon of each gene as part of an *Nde*I site.

Figure 3.1a shows the successful amplification of three genomic sequences with sizes consistent with those published for *drr*A, B and C. Figure 3.1b shows the amplification of the complete region of the *M. tuberculosis* genome predicted to contain all three *drr* genes. Once again, the approximately 2.7Kb amplicon is consistent with the predicted size of the *drr*ABC operon.

In order to maintain and propagate the amplified sequences, each one was extracted from the agarose gel and ligated into the commercially available T-A cloning vector pGEM-T Easy. The ligation reaction was then used to transform competent *E. coli* NovaBlue cells. Transformed cells were selected on the basis of ampicillin resistance, and recombinant clones identified by blue/white screening. For each of the four sequences cloned, up to six recombinant colonies were selected for further analysis.

Figurere 3. 1a **PCR amplification of the *drr*A, B and C genes**



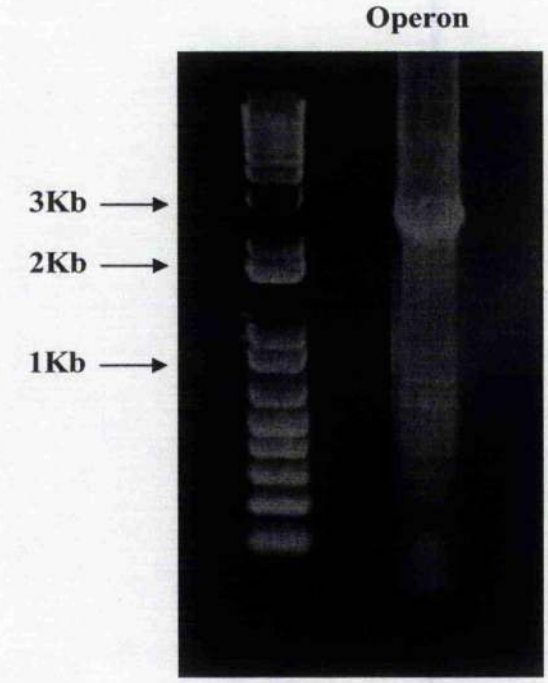
Lane A = *drr*A amplified using primers AF/AR

Lane B = *drr*B amplified using primers BF/BR

Lane C = *drr*C amplified using primers CF/CR

94°C 1 min – denaturation
54°C 1 min – annealing
72°C 1 min - extension
30 Cycles

Figure 3.1b **PCR Amplification of the *drr*ABC operon**



Operon amplified using primers AF/CR

94°C 1 min – denaturation
54°C 1 min – annealing
72°C 2.5 min – extension
30 Cycles

Plasmids were recovered from recombinant clones by mini-prep and further analysed by restriction digest to check for the presence of the inserted sequence. Digestion of isolated pGEM-T Easy plasmids with *Eco*R1 allowed for simple identification of recombinant clones due to the presence of *Eco*R1 sites either side of the T-A cloning site. Figure 3.2a illustrates the analysis of 6 pGEM-T Easy plasmids digested with *Eco*R1, all of which contain the inserted *drr*A amplicon. The slightly increased size of the excised fragment compared with the amplified sequence is due to the presence of small amounts of vector DNA between the T-A cloning site and the *Eco*R1 sites. Figure 3.2b shows isolated recombinant plasmids bearing the *drr*B and C amplicons and the complete *drr*ABC genomic region. Plasmids were submitted for automated DNA sequencing using standard vector directed sequencing primers (M13 Forward and Reverse). The data generated was further analysed to confirm the identity of the cloned sequences.

3. 2 Sequence analysis of the cloned *drr* genes

The nucleotide sequences of the four cloned PCR amplicons were entirely consistent with those of the published *drr* genes (with the exception of the 5' and 3' mutations introduced by oligonucleotide directed mutagenesis). These data confirmed the success of the initial genomic PCR reaction and allowed the amplified sequences to be maintained as *E. coli* glycerol stock cultures. A detailed *in silico* analysis of each cloned sequence was performed and used to test the hypothesis of Cole and colleagues that the *M. tuberculosis* genome encodes a daunorubicin/doxorubicin efflux ABC transporter.

Sequence analysis of *drr*A

The *drr*A gene, in its native form, is 996bp long encoding a protein of 331 amino acids. There is a marked G+C bias in the coding sequence of *drr*A of 63%. This is close to the overall GC-bias in the *M. tuberculosis* genome of 65% (Cole *et al.*, 1998). The translated sequence of *drr*A contains a number of highly conserved protein motifs strongly suggesting that the protein is the ATP-binding subunit of an ABC transporter (see figure 3.3a). The first major indicator that DrrA is a nucleotide-binding protein is the presence of the Walker A motif (PROSITE entry: PS00017). This glycine-rich motif has been identified in a large number of ATP/GTP binding proteins from such evolutionarily diverse species as bacteria and man.

Figure 3. 2a **Restriction analysis of 6 pGEM /*drr*A plasmids**

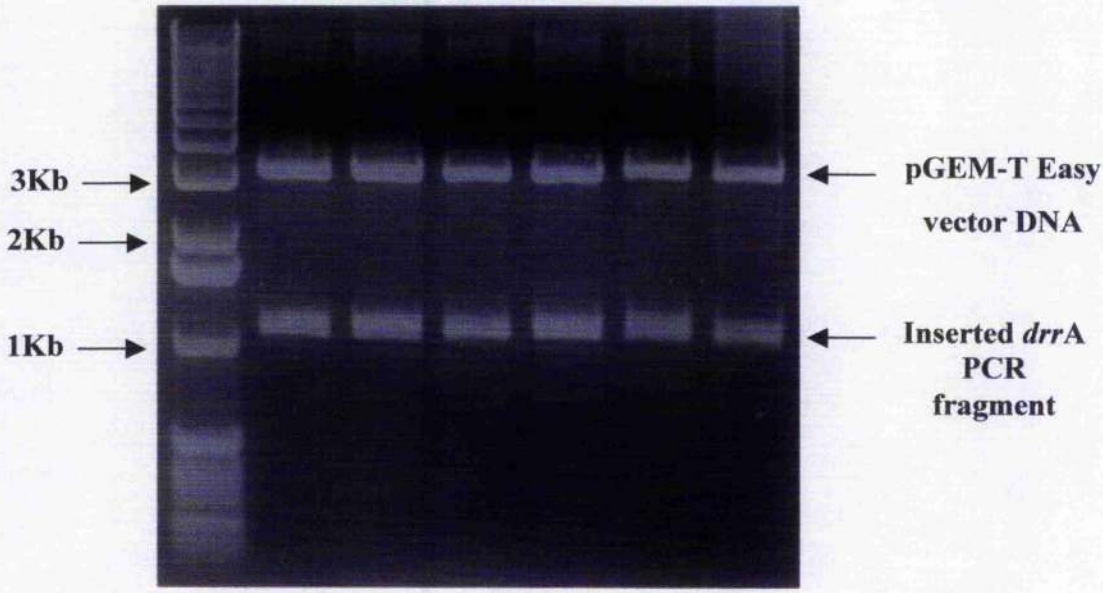
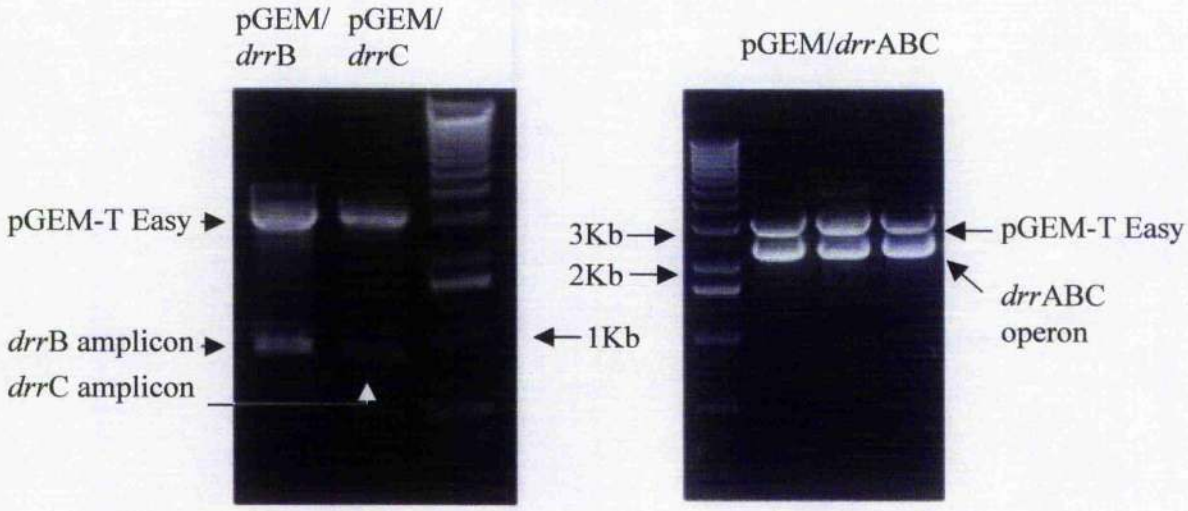


Figure 3. 2b **Restriction analysis of *drr*B, *drr*C and *drr*ABC clones**



The EcoR1 digests above were performed in a total volume of 20µl using 10U of enzyme. Digests were incubated at 37 °C for 1hour.

Crystallographic studies of proteins containing the Walker A sequence show that the residues of the motif form a flexible loop that interacts extensively with bound NTP's through a network of hydrogen bonds (e.g. Hung *et al.*, 1998). The Walker A motif, highlighted in red in figure 3.3a, is found towards the N-terminus of the *M. tuberculosis* DrrA protein. The most highly conserved residues of the motif include the invariant glycines at positions 42 and 47 and the lysine at position 48. The second part of the ATP-binding site is formed by the Walker B motif, always found down-sequence from Walker A. Walker B includes an invariant aspartic acid residue (D172 in DrrA) that is believed to co-ordinate a Mg^{2-} ion in the active site of the protein. The magnesium ion bound at this site is a vital component catalytic complex and must be present for ATP hydrolysis to occur. The Walker A and B motifs have been identified in a variety of non-transport related proteins, e.g. the *E. coli* DNA repair enzyme UvrA, and are not necessarily definitive markers of ABC transporters (Thiagalingam and Grossman, 1991). However, located between the Walker A and B sites of *M. tuberculosis* DrrA, approximately 100 residues downstream of Walker A, is the so-called 'ABC transporter signature sequence' (PROSITE entry: PS00211). The presence of this motif in DrrA, combined with the presence of Walker A and B sites, is a very strong indication that DrrA is the ATP-hydrolysing subunit of an ABC transporter complex. The ABC transporter signature sequence is often described in the literature as the 'LSGGQ' motif since these five amino acids are most commonly found at the start of the sequence. It should be noted however that a degree of degeneracy is observed at all positions of the motif, except for a conserved glycine residue at position four. This glycine residue is conserved in DrrA as part of the sequence 'YSGGM' (see figure 3.3a). As described in chapter 1.2, the exact function of the signature sequence is a matter of some debate. Some studies suggest that it may form part of a composite ATP-binding site in a dimeric model of the ABC transporter NBD's, whilst others suggest a role in inter-domain communication. ABC transporter NBD's often exhibit 30-40% sequence homology to one another, irrespective of their substrates, due to the highly conserved nature of the Walker A, B and signature sequence motifs. However, much higher degrees of sequence homology that extend outwith the core ATP-binding region can be a good indication that NBD's share a similar function. A BLAST search was used to screen protein sequence databases for homologues of the *M. tuberculosis* DrrA protein in order to identify proteins that may be evolutionarily or functionally related (Altschul *et al.*, 1997).

3. 3a Translation of the *M. tuberculosis drrA* gene showing conserved sequence motifs.

```
(1)    MRNDDMAVVV NGVRKTYGKG KIVALDDVSF KVRRGEVIGL  (40)
(41)   LGPNGAGKTT MVDILSTLTR PDAGSAIIAG YDVVSEPAGV  (80)
(81)   RRSIMVTGQQ VAVDDALSGE QNLVLFGRLW GLSKSAARKR  (120)
(121)  AAELLEQFSL VHAGKRRVGT YSGGMRRRID IACGLVVQPQ  (160)
(161)  VAFLDEPTTG LDPRSRQAIW DLVASFKKLG IATLLTTQYL  (200)
(201)  EEADALSDRI TLIDHGIIIA EGTANELKHR AGDTFCEIVP  (240)
(241)  RDLKDLDAIV AALGSLLPEH HRAMLTPDSD RITMPAPDGI  (280)
(281)  RMLVEAARRI DEARIELADI ALRRPSLDHV FLAMTTDPTE  (320)
(321)  SLTHLVSGSA R                                 (331)

Walker A     Walker B    ABC transporter signature sequence
```

3. 3b BlastP search results using *M. tuberculosis* DrrA as the query sequence.

Accession #	Description	Organism	P-Value
gi|2145829	Daunorubicin resistance protein drrA*	*M. leprae*	1e-146
gi|4416482	Daunorubicin resistance protein A*	*M. avium*	3e-74
gi|4808357	ABC transporter ATP-binding component*	*S. coelicolor*	7e-66
gi|29831196	ABC transporter ATP-binding protein*	*S. avermitilis*	9e-60
gi|11346307	Probable ABC-type transport protein*	*S. coelicolor*	1e-59
gi|10129760	ABC transport system ATP-binding protein*	*S. coelicolor*	1e-59
gi|20520964	ABC transporter ATP-binding protein*	*S. coelicolor*	2e-59
gi|11141834	Putative ATP binding protein*	*S. viridochromo*	2e-59
gi|29832484	ABC transporter ATP-binding component*	*S. avermitilis*	4e-59
gi|399405	Daunorubicin resistance ATP-binding protein	*S. peucetius*	2e-57

* Putative functional assignment

A search of the NCBI non-redundant sequence database using the BlastP algorithm produced hits to a large number of putative ABC transporter NBD's. A modified version of the BlastP output showing the top ten database hits is shown in figure 3.3b. In common with *M. tuberculosis* DrrA, the vast majority of the hits were uncharacterised proteins identified during genome sequencing projects. As such, these data cannot be used to infer an unambiguous function for the *M. tuberculosis* DrrA protein. The top hit was to a *Mycobacterium leprae* protein which exhibited an astonishing 85% identical residues, the overall homology rising to 90% when conservative amino acid substitutions were taken into account. The second most homologous protein was also from a pathogenic mycobacterial species, *M. avium*. This protein shared 51% identity to *M. tuberculosis* DrrA and 63% similarity. These levels of homology would suggest that the proteins almost certainly perform a similar function in all three species of mycobacteria. The remainder of the top ten database hits were all proteins from various *Streptomyces* species, including a number of antibiotic producing organisms. One of these hits, the *S. peucetius* DrrA protein, is the only homologue which has been experimentally characterised. This protein, known to energise the transmembrane efflux of the daunorubicin and doxorubicin antibiotics, is 42% identical to the *M. tuberculosis* DrrA protein. A gapped sequence alignment of the *M. tuberculosis* and *S. peucetius* DrrA proteins is shown in figure 3.4. From these data it is difficult to draw firm conclusions as to whether the sequence homology between the mycobacterial and streptomycete proteins is a genuine reflection of conserved function, i.e. antibiotic transport. The homology may equally well reflect the common evolutionary roots of the streptomycetes and the mycobacteria. It is entirely possible that the ABC transporters with which these NBD's are associated may have evolved different functions *after* divergence of the two genera.

Sequence analysis of the *drr*B and *drr*C genes

The *drr*B and *drr*C genes are 870 and 831bp in length respectively. Each of the open reading frames display a positive GC-bias (*drr*B – 59%, *drr*C - 57%) that is just slightly lower than that of the complete *M. tuberculosis* genome (65%). The proteins encoded by the *drr*B and *drr*C genes are predicted to be 289 and 276 amino acids long. Both proteins contain more than 50% hydrophobic amino acids (A, I, L, F, W, V) strongly suggesting that they are associated with the cell membrane. This observation is consistent with a potential role for the DrrB and DrrC proteins as the MSD's of an ABC

Figure 3. 4 **Sequence alignment of the *S. peucetius* and *M. tuberculosis* DrrA proteins**

```
S.peu DrrA     (1) MNTQPTRAIETSGLVKVYNG--TRAVDGLDLNVPAGLVYGILGPNGAGKS
 M.tb DrrA     (1) -MRNDDMAVVVNGVRKTYGKGKIVALDDVSFKVRRGEVIGLLGPNGAGKT
 Conserved     (1)        A    G  K Y      A D     V  G V G LGPNGAGK

S.peu DrrA    (49) TTIRMLATLLRPDGGTARVFGHDVTSEPDTVRRRISVTGQYASVDEGLTG
 M.tb DrrA    (50) TMVDILSTLTRPDAGSAIIAGYDVVSEPAGVRRSIMVTGQQVAVDDALSG
 Consensus    (51) T    L TL RPD G A   G DV SEP  VRR I VTGQ   VD  L G

S.peu DrrA    (99) TENLVMMGRLQGYSWARARERAAELIDGFGLGDARDRLLKTYSGGMRRRL
 M.tb DrrA   (100) EQNLVLFGRLWGLSKSAARKRAAELLEQFSLVHAGKRRVGTYSGGMRRRI
 Consensus   (101)   NLV  GRL G S   AR RAAEL   F L  A  R   TYSGGMRRR

S.peu DrrA   (149) DIAASIVVTPDLLFLDEPTTGLDPRSRNQVWDIVRALVDAGTTVLLTTQY
 M.tb DrrA   (150) DIACGLVVQPQVAFLDEPTTGLDPRSRQAIWDLVASFKKLGIATLLTTQY
 Consensus   (151) DIA   VV P   FLDEPTTGLDPRSR   WD V      G   LLTTQY

S.peu DrrA   (199) LDEADQLADRIAVIDHGRVIAEGTTGELKSSLGSNVLRLRLHDAQSRAEA
 M.tb DrrA   (200) LEEADALSDRIILIDHGIIIAEGTANELKHRAGDTFCEIVPRDLKDLDAI
 Consensus   (201) L EAD L DRI  IDHG  IAEGT  ELK   G         D

S.peu DrrA   (249) ERLLSAELGVTIHRDSDPTALSARIDDP---RQGMRALAELSRTHLEVRS
 M.tb DrrA   (250) VAALGSLLPEHHRAMLTPDSDRITMPAPDGIRMLVEAARRIDEARIELAD
 Consensus   (251)    L   L          P         P   R    A         E

S.peu DrrA   (296) FSLGQSSLDEVFLALTGHPADDRSTEEAAEEEKVA
 M.tb DrrA   (300) IALRRPSLDHVFLAMTTDPTESLTHLVSGSAR---
 Consensus   (301)   L   SLD VFLA T  P
```

The sequence alignment above shows the degree of amino acid homology between the DrrA proteins of *M. tuberculosis* and *Streptomyces peucetius*. Conserved residues are shown in red. The regions of highest homology include the Walker A nucleotide-binding motif, the Walker B region and the ABC transporter signature-sequence. The overall homology between the two proteins is 42% identical residues rising to 57% when conservative substitutions are taken into account

transporter complex. The original 'two-times-six' model of ABC transporter MSD's predicted that the proteins possess multiple, hydrophobic, membrane-spanning α-helices. This model was recently confirmed by the high resolution structures of the MsbA and BtuCD ABC transporters. In order to assess whether the DrrB and C proteins might possess similar structural features, their sequences were submitted for analysis by the TopPred sequence analysis program. This web-based sequence analysis server makes predictions of membrane protein topology based upon primary amino acid sequence data (http://bioweb.pasteur.fr/cgi-bin/seqanal/toppred.pl, von Heijne, G., 1992). The output data for the DrrB and C proteins is summarised in figures 3.5 and 3.6. The topology predictions for the proteins suggest that they both have six membrane spanning helical segments. These data further support a model of the *M. tuberculosis* DrrABC transporter in which the DrrB and C proteins contribute to a heterodimeric transmembrane pathway.

As described in chapter 1.2, the sequence homology between ABC transporter MSD's is often very limited, even when two transporters share a similar substrate. Despite this, a search of the NCBI non-redundant protein sequence database was performed in order to identify any potential homologues of the DrrB and DrrC proteins. The searches produced a relatively small number of database hits, with two or three results being of particular interest. The genomes of the pathogenic mycobacteria *M. leprae* and *M. avium* appear to encode homologues of both DrrB and DrrC. The *M. leprae* proteins exhibit an extraordinary level of sequence conservation for ABC transporter MSD's (DrrB – 77%, DrrC – 75%). The *M. avium* homologues have a lower, but still significant level of homology, 42% for DrrB and 49% for DrrC. These results reflect those obtained when the database was searched using the DrrA protein sequence. It seems very likely that both *M. avium* and *M. leprae* encode a DrrABC transport system similar to *M. tuberculosis*, the very high level of sequence homology suggesting that all three systems perform a similar physiological function. Exactly what the function of the mycobacterial DrrABC systems is remains unclear. Hits to proteins with a characterised function were few, but, in the case of DrrB, included the ORF2 protein of *S. rochei* and the DrrB protein of *S. peucetius*. Both of these proteins are known MSD components of ABC transporters associated with antibiotic resistance (Fernandez-Moreno *et al.*, 1998; Kaur, 1997). These data would seem to support the hypothesis that the *M. tuberculosis* DrrB and C genes constitute the MSD's of a novel antibiotic efflux ABC transporter.

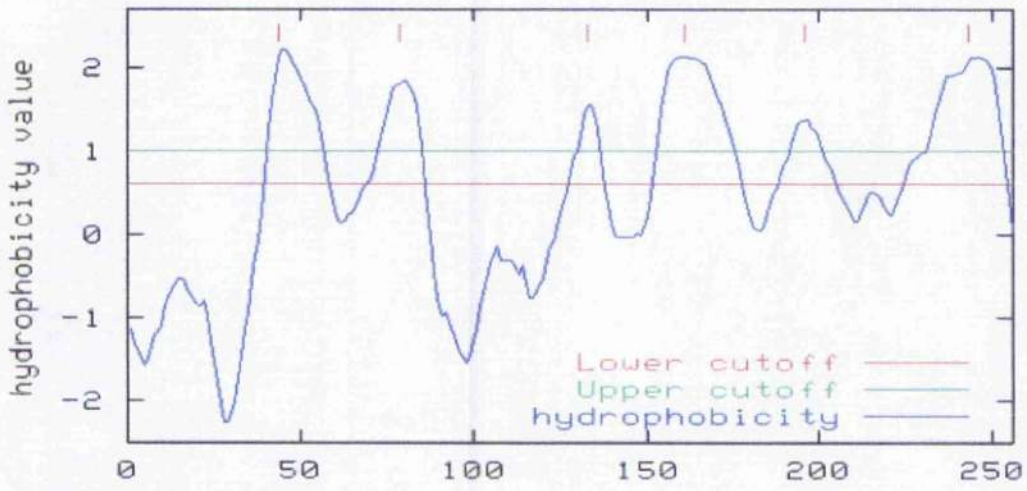
Due to the low level of sequence homology between the MSD's of ABC transporters conserved sequence motifs are rare. The major exception to this rule is the so-called 'EAA' motif identified in a subset of substrate-binding protein dependent nutrient uptake ABC transporters (Mourez *et al.*, 1997). However, a second motif known as the 'ABC-2 integral membrane protein signature' (PROSITE entry: PS00890), has also been identified in a small number of bacterial ABC transporters associated with the export of antibiotics and cell wall components (Reizer *et al.*, 1992). Interestingly, this motif is present in both the DrrB and C proteins of *M. tuberculosis* and in the *S. peucetius* DrrB protein. The motif itself is somewhat extended and has a high degree of degeneracy throughout, the only absolutely conserved residue being a proline ten residues from the C-terminal end. In the *M. tuberculosis* DrrB protein the ABC-2 motif spans residues 203 to 239 inclusive, with the conserved proline at position 230. In the case of DrrC, the motif extends from residues 181 to 220 and the conserved proline occurs at position 211. The presence of this motif in DrrB and C is further evidence the DrrABC system is an efflux-type ABC transporter.

Figure 3. 5 **TopPred sequence analysis of *M. tuberculosis* DrrB**

Candidate membrane-spanning segments: 6 segments found

Helix	Begin	End	Certainity
1	49	69	Certain
2	88	108	Certain
3	143	163	Certain
4	167	187	Certain
5	199	219	Certain
6	266	286	Certain

Hydropathy Analysis



Primary sequence of DrrB – putative transmembrane helices shown in red

(1) MSGPAIDASP ALTFNQSSAS IQQRRLSTGR QMWVLYRRFA (40)

(41) APSLLNGE**VL TTVGAPIIFM VGFYIPFAI**P WNQFVGGASS (80)

(81) GVASNLG**QYI TPLVTLQAVS FAAIGSGF**RA ATDSLLGVNR (120)

(121) RFQSMPMAPL TPLLARVWVA VD**RCFTGLVI SLVCGYVIGF** (160)

(161) **RFH**RGA**LYIV GFCLLVIAIG AVLSFAA**DLV GTVTRNPD**AM** (200)

(201) **LPLLSLPILI FGLLSIGLM**P LKLFPHWIHP FVRNQPISQF (240)

(241) VAALRALAGD TTKTASQVSW PVMAP**TLTWL FAFVVILALS** (280)

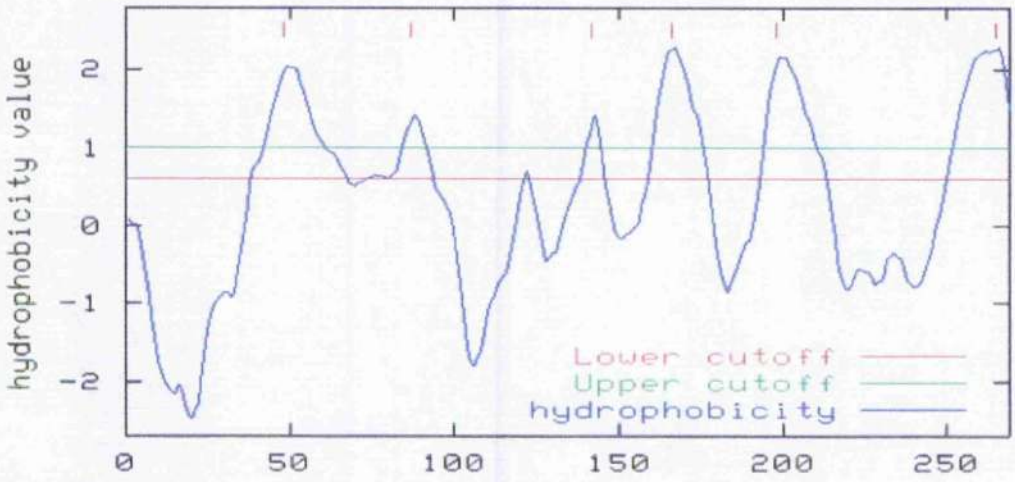
(281) **STIVLA**RRP (289)

Figure 3. 6 **TopPred sequence analysis of *M. tuberculosis* DrrC**

Candidate membrane-spanning segments: 6 segments found

Helix	Begin	End	Certainity
1	45	65	Certain
2	80	100	Certain
3	134	154	Certain
4	162	182	Certain
5	197	217	Certain
6	244	264	Certain

Hydropathy Analysiss



Primary sequence of DrrC – putative transmembrane helices shown in red

(1) MITTTSQEIE LAPTRLPGSQ NAARLFVAQT LLQTNRLLTR (40)

(41) WARD**YITVIG AIVLPILFMV VLNIV**LGNLA YVVTHDSGL**Y** (80)

(81) **SIVPLIALGA AITGSTFVAI** DLMRERSFGL LARLWVLPVH (120)

(121) RASGLISRIL ANA**IRTLVTT LVMLGTGVVL GFRF**RQGLIP (160)

(161) S**LMWISVPVI LGIAIAAMVT TV**ALYTAQTV VVEGVE**LVQA** (200)

(201) **IAIFFSTGLV PLNSYPG**WIQ PFVAHQPVSY AIAAMRGFAM (240)

(241) GGP**VLSPMIG MLVWTAGICV VCAV**PLAIGY RRASTH (276)

3. 3 Sub-cloning of the *M. tuberculosis drr* genes

With a significant amount of sequence data supporting the hypothesis that the cloned *drr* genes do indeed encode a novel efflux ABC transporter, the next stage of the investigation involves analysing the biochemical properties of the proteins themselves. *M. tuberculosis* is a particularly slow growing organism and, due to its pathogenicity, requires special facilities for its growth and handling. These features make working with live *M. tuberculosis* particularly undesirable. As such, it was decided that a biochemical study of the DrrA, B, and C proteins would require their expression in a heterologous host. *E. coli* was selected as the heterologous host since it has been used to successfully overexpress an enormous number of proteins from biologically diverse species. In addition, genetic systems for the expression of heterologous protein in *E. coli* are well developed and characterized.

Myriad different plasmid based expression systems are available for use with *E.coli*. The commercially available pET-series plasmids (Novagen) were chosen for the expression of the *drr*A, B and C genes since these vectors possess a number of highly desirable features. Firstly, the transcription of the heterologous gene is under the control of the very strong bacteriophage T7 promotor. Since *E. coli* RNA polymerase does not interact with this promoter, background or 'leaky' expression of the target protein is minimized in the uninduced state. The transcription of the heterologous gene is achieved by transforming the expression construct into a host cell that possesses a chromosomal copy of the T7 RNA polymerase, itself under control of the *lacUV5* promotor. The expression of the T7 RNA polymerase is induced by the addition of IPTG to a log phase culture. This in turn results in massive transcription of the target gene sequence. The manufacturers suggest that virtually all of the cell's resources become involved in expression of the heterologous protein and that it may constitute up to 50% of the total cell protein after just a few hours (pET System Manual, ninth edition, Novagen). In addition to the tightly controlled and high level of expression achieved through use of the pET plasmids, the vectors also encode one or more 'His-Tag' sequences. This is a sequence of six consecutive histidine residues placed at either the N- or C-terminus of the recombinant protein. The His_6 sequence acts as both an epitope-tag, enabling the expression of the recombinant protein to be monitored by immunogenic means, and an affinity purification tag. The six consecutive histidine residues bind strongly to certain divalent cations, particularly Ni^{2+}, enabling efficient purification of the recombinant protein using 'immobilised metal

affinity chromatography' (IMAC). Detection and purification of the recombinant Drr proteins will be discussed in greater detail in following chapters.

The basic sub-cloning strategy used for all of the cloned *drr* sequences is illustrated in figure 3.7, and the techniques involved are described in chapter 2. Unique restriction enzyme sites incorporated into the 5' and 3' ends of the *drr* genes during the initial PCR reaction enabled the genes to be precisely and selectively excised from the pGEM-T Easy constructs using *Nde*I and *Hin*dIII. Isolated *drr* gene fragments were then purified from vector DNA by agarose gel electrophoresis and isolated by gel extraction (figure 3.8a). The non-recombinant pET vector, with unique *Nde*I and *Hin*dIII sites forming part of the multiple cloning site (MCS), was treated in an identical fashion. After digestion, the ends of the linearised plasmid have non-complimentary, single-stranded overhangs which prevent the vector ends from ligating to themselves. However, incomplete digestion of the vector is difficult to detect. In order to reduce the chances of vector self-ligation, the isolated vector DNA was treated with shrimp alkaline phosphatase. The dephosphorylated vector DNA was then purified by electrophoresis and gel extraction (see figure 3.8a).

The two isolated DNA fragments, possessing complimentary 5' and 3' single stranded overhangs, were ligated together using T4 DNA ligase. Recombinant expression vectors generated in this way were recovered by transforming the ligation mixture into high efficiency NovaBlue cells. Transformants were selected on the basis of vector encoded antibiotic resistance. Recombinant expression plasmids were isolated by mini-prep and checked for the presence of the inserted *drr* gene by restriction digest (see figure 3.8b). Four recombinant expression vectors were generated by these reactions, one for each of the amplified *drr* sequences. They were termed pDrrA, pDrrB, pDrrC and pDrrABC. The details of each of these plasmid constructs are described in subsequent chapters.

Figure 3. 7 Generalised sub-cloning strategy

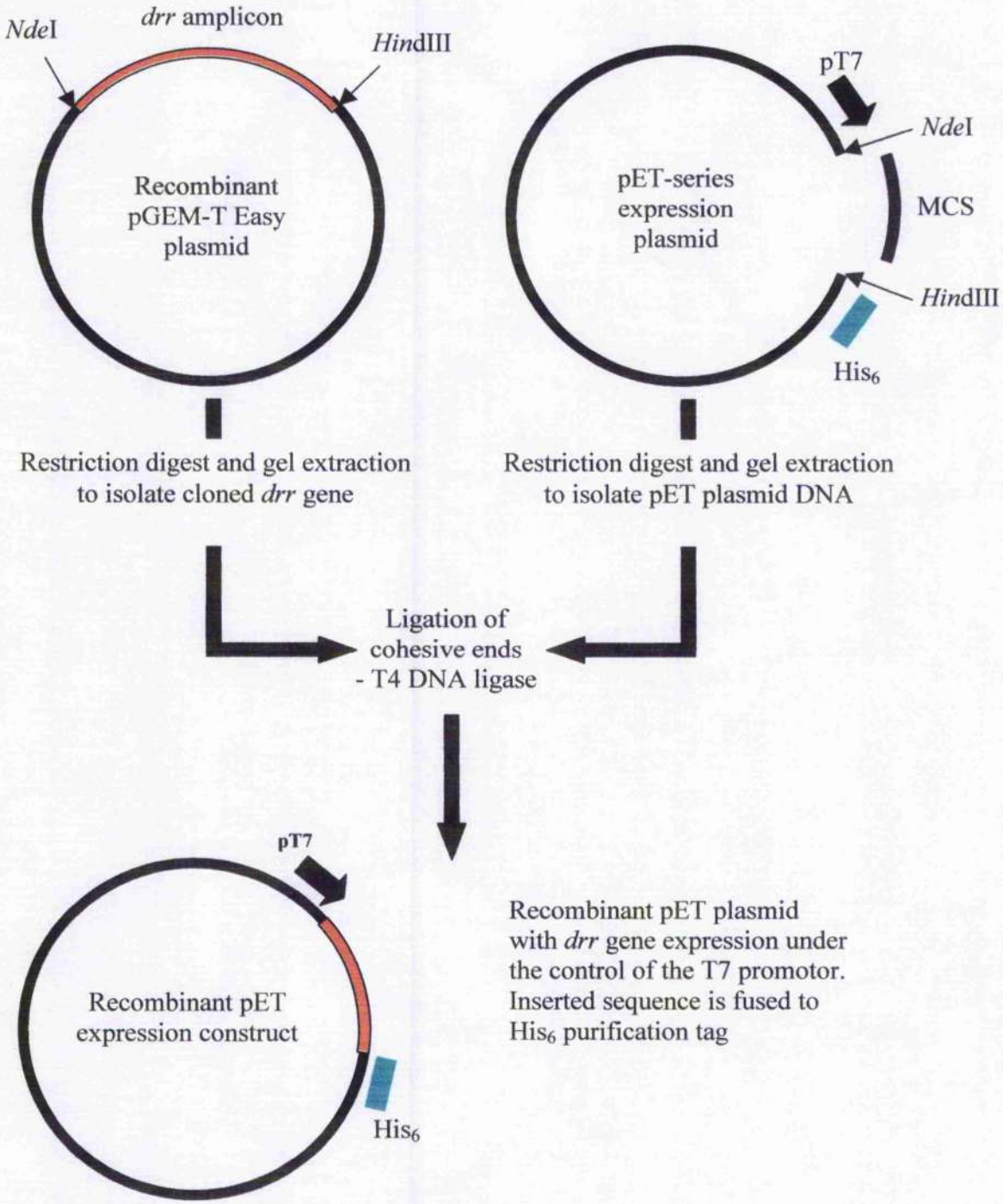
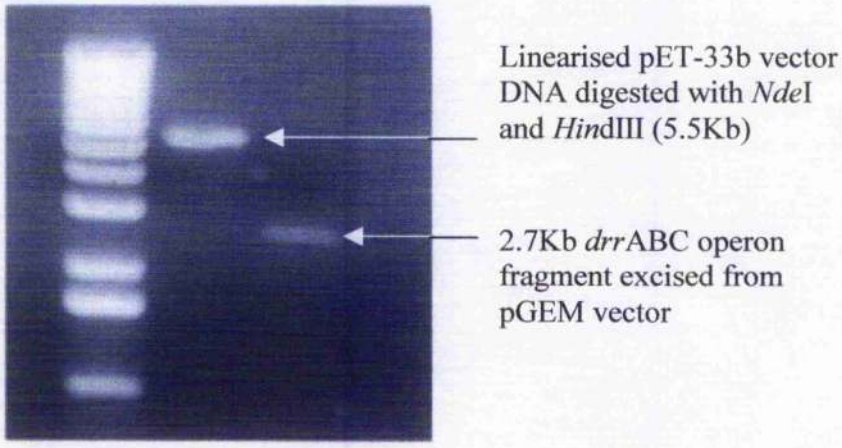
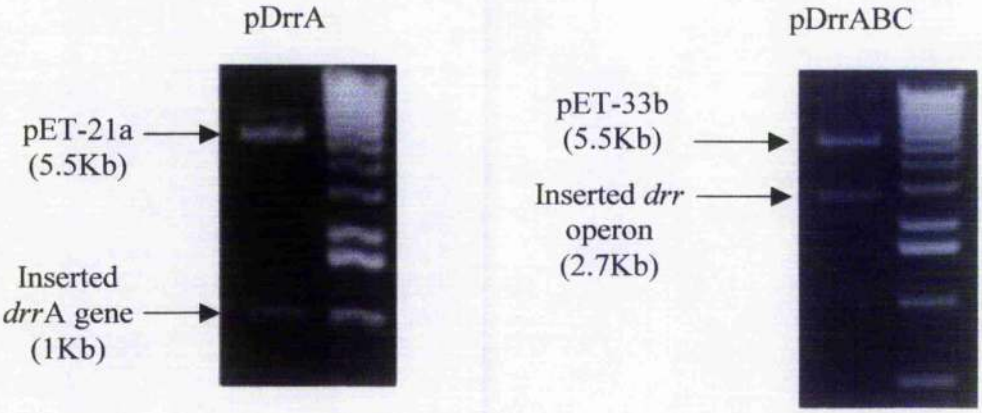


Figure 3. 8a **Gel purification of linearised pET-33b and *drr*ABC insert after restriction digestion with *Nde*I and *Hin*dIII**



The fragments shown above were ligated together using T4 DNA ligase and the mixture was transformed into competent *E. coli* NovaBlue. Recombinant clones were isolated and checked for the presence of the inserted fragment by restriction enzyme analysis (see below).

Figure 3.8b **Restriction analysis of recombinant pET expression plasmids**



*Nde*I/*Hin*dIII double digests above were performed in a total volume of 20µl using 10U of each enzyme. Digests were incubated at 37 °C for 1hour.

3.4 Discussion

With tuberculosis still a major threat to global public health, the completion of the *M. tuberculosis* genome sequence represents a major step forward in our understanding of the biology of this organism. Unsurprisingly, given the almost ubiquitous distribution of these systems in nature, the *M. tuberculosis* genome appears to encode a significant number of ATP-binding cassette transporters. The active transmembrane efflux of chemotherapeutic agents is a known contributor to the growing levels of drug resistance observed amongst pathogenic microorganisms and human cancer cells. Whilst active efflux of drugs by ABC transporters has yet to be demonstrated as a clinically relevant in the treatment of mycobacterial disease, the presence of such systems in the *M. tuberculosis* genome is an indication that this may yet prove to be the case.

Analysis of the *M. tuberculosis* genome reveals the presence of three genes, *drr*A, B, and C, that bear many of the characteristics of known efflux ABC transporters. The three genes are arranged in an operon structure in which the stop codon of *drr*A overlaps the start codon of *drr*B, and the stop codon of *drr*B overlaps the start codon of *drr*C. This feature of the operon, known as *transcriptional-linkage*, is common amongst bacterial systems in which the protein products of the linked genes are functionally related. The current models of bacterial ABC transporters comprise a minimum of four protein domains, two nucleotide-binding domains which energise the transport process, and two membrane spanning regions that comprise the translocation pathway. We propose a model of the *M. tuberculosis* DrrABC transporter in which two copies of the DrrA protein function as NBD's, whilst the *drr*B and C genes encode the membrane-spanning translocation pathway. A similar organisation of genes and proteins, in which just three genes encode all four functional domains of an ABC transporter complex, has been demonstrated for a number of systems including the $MalFGK_2$ maltose permease of *E. coli* and the $HisQMP_2$ histidine permease of *Salmonella typhimurium*.

In support of this model of DrrABC, and indeed its role as an ABC transporter, are a number of protein sequence motifs encoded by the genes. Firstly, the DrrA protein features the Walker A and B sequences common in a wide variety of nucleotide-binding proteins. Perhaps more importantly, DrrA also exhibits the characteristic ABC transporter signature sequence, a protein motif unique to the nucleotide-binding domains of ABC transporters. The *drr*B and C genes on the other hand appear to encode highly

hydrophobic proteins, each of which possesses a possible six membrane-spanning α-helical segments. These features are consistent with the ability of DrrB and C to intercalate into the cell membrane and form a trans-bilayer protein complex. Additionally both the DrrB and C proteins contain one of the few conserved protein motifs yet identified amongst the MSD's of ABC transporters, the so-called ABC-2 signature sequence.

In an effort to identify functional homologues of *M. tuberculosis* DrrABC, database searches revealed the presence of highly similar systems in the genomes of both *Mycobacterium leprae* and *Mycobacterium avium*. The presence of *drr*ABC genes in these pathogenic mycobacterial species suggests that this transporter plays a vital and conserved role in the biology of all three organisms. Sequence homologues of DrrA, B and C are rare outside of the mycobacteria, the most similar proteins being found amongst various species of *Streptomyces*. Whilst the exact function of the streptomycete homologues has yet to be unambiguously demonstrated in most cases, many are implicated in the role of antibiotic and secondary metabolite export. This further raises the possibility that the *M. tuberculosis* genome has retained antibiotic transport capabilities that were present in ancestral actinomycetes. Amongst the sequence homologues with known function are the *Streptomyces peucetius* DrrA and B proteins. These are the NBD and MSD components respectively of an ABC transporter responsible for the efflux of the antibiotics daunorubicin and doxorubicin from the organism which produces them. Whilst the *S. peucetius* DrrAB system is encoded by just two genes, and is believed to function as a $DrrA_2B_2$ tetramer (Kaur and Russell, 1998), the homology between the two systems is enough to at least suggest a conserved function. In particular, the *S. peucetius* DrrB protein, in common with the *M. tuberculosis* DrrB and C proteins contains the ABC-2 signature motif.

Perhaps the greatest insight into the function of the *M. tuberculosis* DrrABC system comes from examination of its location within the genome, and additionally from gene knockout experiments. The 5' end of the *drr*A gene is found just 11bp downstream of a very large biosynthetic system, the *pps*ABCDE operon. This cluster of genes occupies almost 50Kb, making it one of the largest operons in the entire *M. tuberculosis* genome. The *pps* genes encode the components of a modular (type I) polyketide synthase that is involved in the biosynthesis of the molecules phenolphthiocerol and phthiocerol. These

two diol molecules are found at the exterior face of the mycobacterial cell wall esterified to mycoserosic acid (Azad *et al.*, 1997). The very short distance between the *pps*ABCDE operon and the *drr*ABC operon, and the absence of any obvious intervening promoter sequence to drive the transcription of the *drr*ABC genes, suggests that the two gene clusters may be transcribed in a single polycistronic mRNA. The clustering together of biosynthetic and transport functions has been observed in other bacterial species, particularly in the biosynthesis and export of cell wall components. One well characterised example of this type genetic linkage is the *kps* gene cluster of *E. coli* that encodes both the synthesis of polysialic acid, and an ABC transporter for its export to the cell surface (Rigg *et al.*, 1998). A role for the DrrB and C proteins in the export of cell wall components is supported by the presence of the ABC-2 motif. This sequence is primarily found in the integral membrane components of ABC transporters associated with the export of capsular polysaccharides and cell surface O-antigens (see PROSIE entry: PS00890).

Random transposon mutagenesis experiments aimed at identifying important virulence gene clusters in the *M. tuberculosis* genome implicated the region containing the *drr*ABC and *pps*ABCDE operons as an important mediator of *in vivo* pathogenicity. Knockouts in the *pps* promoter region and also in the *drr*C gene itself were shown to markedly reduce the growth and viability of the organism in a mouse model of pulmonary tuberculosis (Camacho *et al.*, 1999, Cox *et al.*, 1999). The knockouts in the *pps* promoter region were shown to reduce the synthesis of the phthiocerol esters that are found in the cell wall, suggesting that these complex lipids may be important virulence factors. It should be noted that knockouts in a number of other genes found in this region of the genome also resulted in attenuated virulence. This data suggests that the synthesis and export of phthiocerol esters is a complex process involving a large number of proteins, any of which may, in the future, represent useful targets for drug intervention.

The fact that the DrrABC transporter may play a physiological role in the export of complex lipid species to the mycobacterial cell wall does not preclude the possibility that the transporter has a dual function. For example, much evidence suggests that the primary function of the human MDR1 gene product, P-glycoprotein, is as a phospholipid flipase controlling the asymmetric distribution of phospholipids in the cell plasma membrane (Romsicki and Sharom, 2001). As well as its primary function, this ABC transporter is an

important mediator of multidrug resistance because of its ability to excrete a wide variety of structurally diverse, amphipathic molecules. Included amongst the substrates of P-glycoprotein are the important anti-cancer agents doxorubicin and daunorubicin. As such, the DrrABC transporter of *M. tuberculosis* may prove be of use as a prokaryotic model of the dual function ABC transporters that cause failure of cancer chemotherapy in humans.

This chapter reports the successful cloning of the *M. tuberculosis drr*ABC genes, an analysis of their potential function, and finally their transfer into vectors suitable for overexpresion of the proteins in a heterologous host. Expression and purification of large quantities of these proteins are necessary in order to further our understanding of the biochemical, structural and physiological properties of ABC transporters in general and the function of the *M. tuberculosis* DrrABC system in particular.

Chapter 4

Heterologous expression of the DrrA, B and C proteins using pT7 based plasmids

Bioinformatic methods have proven to be of great use in identifying potential members of the ABC transporter superfamily and in the assignment of *putative* functions. Despite this, a more complete understanding of ABC transporters at a biochemical and structural level requires the expression and purification of the proteins concerned. Such studies take on an added importance when particular ABC transporters are implicated in clinically and biologically significant processes such as drug resistance, microbial pathogenesis, and virulence. Since the completion of the *M. tuberculosis* genome sequence very little functional research has emerged regarding the role of ABC transporters in the biology of this pathogen. Two notable exceptions include partial characterisation of ABC transporters involved in phosphate and sulphur acquisition (Sarin *et al.*, 2001; Wooff *et al.*, 2002).

The potential importance of the DrrABC transporter as a mediator of virulence and/or antibiotic resistance marks it out as a system worthy of further study given the urgent medical requirement for new anti-tubercular agents and drug targets. This chapter reports a series of experiments in which recombinant expression was used as a tool to express and purify the individual proteins of the *M. tuberculosis* DrrABC transporter using the heterologous host *E. coli*. Results of these experiments were contrasting, and in many respects mirrored the results of other groups in their efforts to overexpress the components of ABC transporter complexes. The presumed ATP-binding protein, DrrA, was expressed at very high levels in *E. coli*, but unfortunately the vast majority of the protein produced was in the form of biologically inactive aggregates known as *inclusion bodies*. In contrast, successful over-expression of the membrane associated components of the transporter, DrrB and DrrC, proved difficult to establish and maintain. These latter results reflect the continuing difficulties associated with over-expression of membrane-bound proteins, a process which has long been a bottle-neck in research aimed at increasing our understanding of transmembrane transport. These results, and the problems encountered, are analysed in the wider context of ABC transporter research in general in the discussion section at the end of this chapter.

4. 1 Properties of the pDrrA expression plasmid

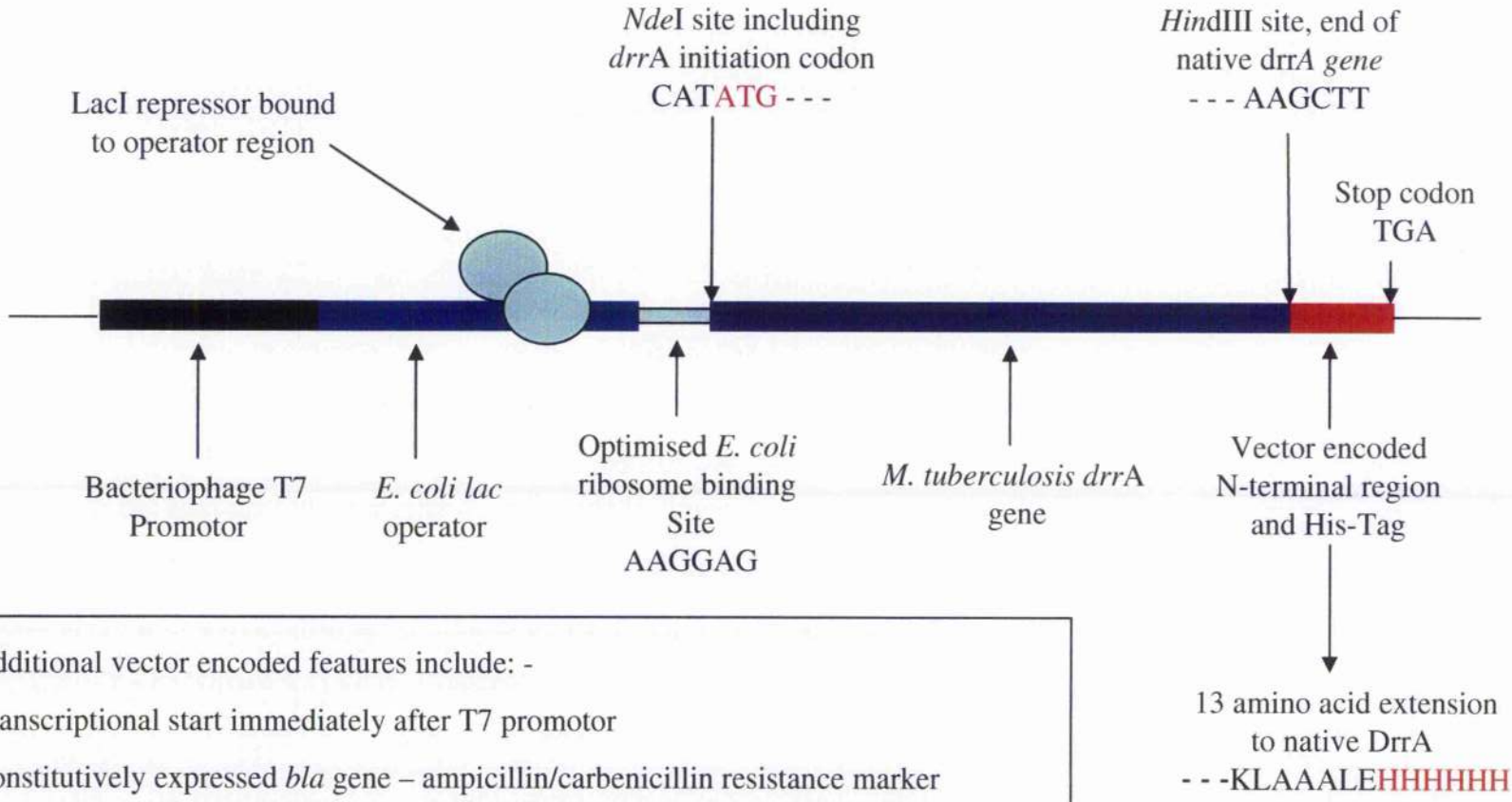
The pDrrA expression plasmid was generated by sub-cloning a fragment of pGEM/*drr*A, containing the *M. tuberculosis drr*A gene bounded by *Nde*I and *Hin*dIII restriction sites, into the MCS of a pET-21a expression plasmid. This resulted in the generation of a 1032bp open reading frame in which the native *drr*A gene (lacking the 5' termination codon) is fused to a short 5' sequence encoding a His-Tag. The 3' ATG start codon of the *drr*A gene is located downstream of a bacteriophage T7 promoter, a *lac* operator sequence and an optimised *E. coli* ribosome binding site. These elements together ensure maximal repression of protein expression in the uninduced state, whilst driving high levels of transcription when the plasmid is transformed into a suitable host strain and induced with IPTG. A schematic representation of the pDrrA expression plasmid is shown in figure 4.1.

The *drr*A open reading frame encodes a 344 amino acid fusion protein with a predicted molecular weight of 37.4 KDa. The vector encoded N-terminal region of the protein adds an additional 13 amino acids to the native DrrA sequence, the last six residues constituting the His_6-Tag. The protein encoded by this open reading frame is subsequently referred to as DrrA-His_6.

4. 2 Recombinant expression of DrrA-His_6

Small-scale expression trials were initiated in order to test for expression of the recombinant DrrA protein. The pDrrA expression construct was first transformed in to the manufacturer's recommended host strain, *E. coli* BL21(DE3). Certain features of the BL21(DE3) strain make it ideal for the expression of recombinant proteins from pET-based expression plasmids. Firstly, the strain lacks two major cellular proteases, Lon and OmpT, which may otherwise interfere with the expression and purification of intact, full-length, recombinant protein. Secondly, the DE3 lysogen, stably incorporated into the *E. coli* chromosome, encodes an IPTG inducible bacteriophage T7 RNA polymerase. It is induction of this T7 RNA polymerase that ultimately drives transcription of the recombinant gene sequence. Freshly transformed BL21(DE3) cells were selected on the basis of carbenicillin resistance (vector encoded) and used to prepare 5ml starter cultures for expression trials. A similar culture of BL21(DE3) cells, transformed with the non-

Figure 4. 1 Key features of the cloning and expression region of pDrrA



Additional vector encoded features include: -

Transcriptional start immediately after T7 promotor

Constitutively expressed *bla* gene – ampicillin/carbenicillin resistance marker

Bacteriophage T7 transcriptional terminator downstream of His-Tag sequence

Constitutively expressed *lac*I gene – maximal repression in uninduced state

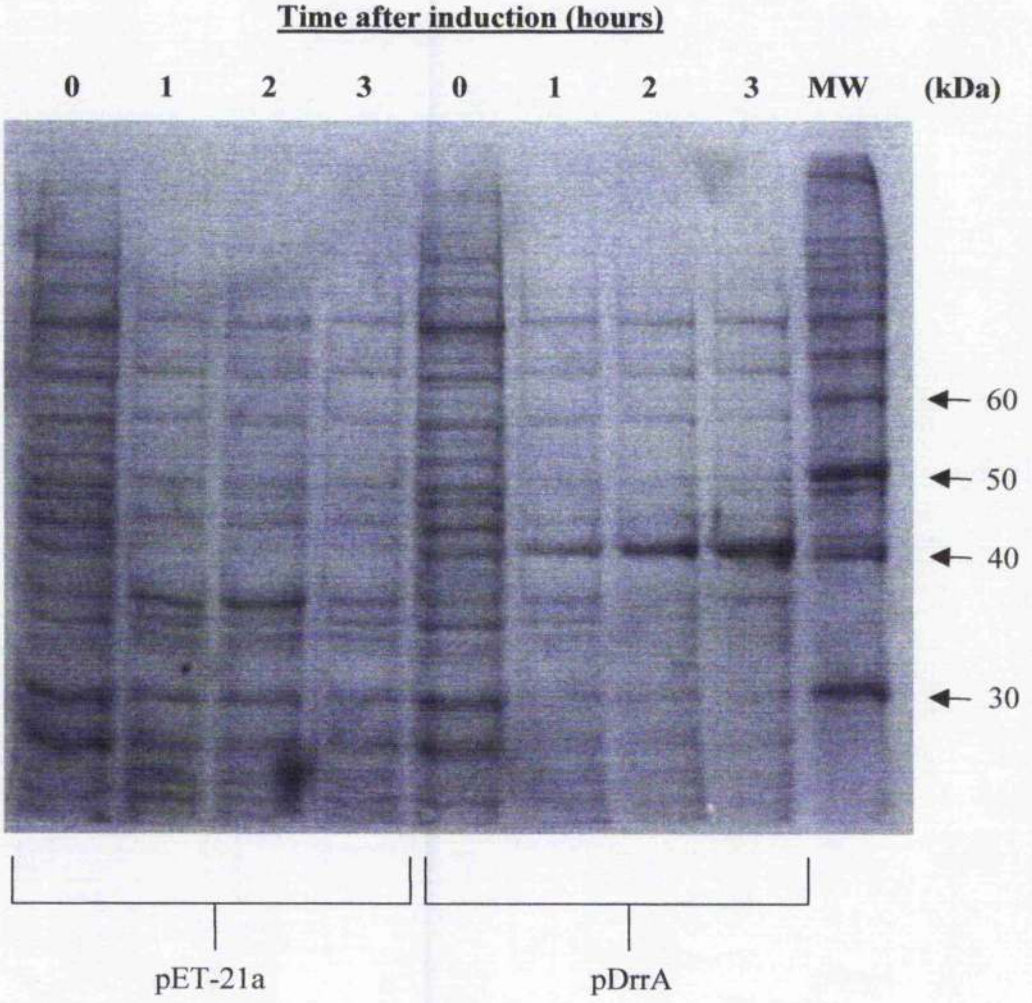
recombinant pET-21a plasmid, was grown in order to provide an expression negative control. Trials were started by diluting 1ml of each overnight starter culture into 50ml of LB medium containing fresh selection antibiotic. Cells were grown at 37°C with rotary shaking until the cultures reached an $O.D._{600}$ of 0.6. At this point, 1ml samples of each culture were withdrawn as zero time-point controls. Immediately afterward, expression of recombinant DrrA was induced by the addition of 1mM IPTG to each culture. 1ml samples were withdrawn from the two cultures at 1 hour intervals, up to three hours, and the cells harvested by centrifugation. Each sample was prepared for analysis by SDS-PAGE. Figure 4.2 illustrates the SDS-PAGE results for one of the first expression trials.

Cells carrying the pDrrA plasmid appear to accumulate a protein of approximately 40kDa over the course of the experiment that is not expressed in the control culture. This protein is found in neither of the cultures prior to induction and appears only in response to the addition of IPTG. Taking into account experimental limits, the accumulated protein is well within the size range of that expected for recombinant DrrA-His_6 (37.4 kDa). Considered together, these results make a strong case for the apparent 40kDa protein being DrrA-His_6.

4. 3 Western blot identification of DrrA-His_6

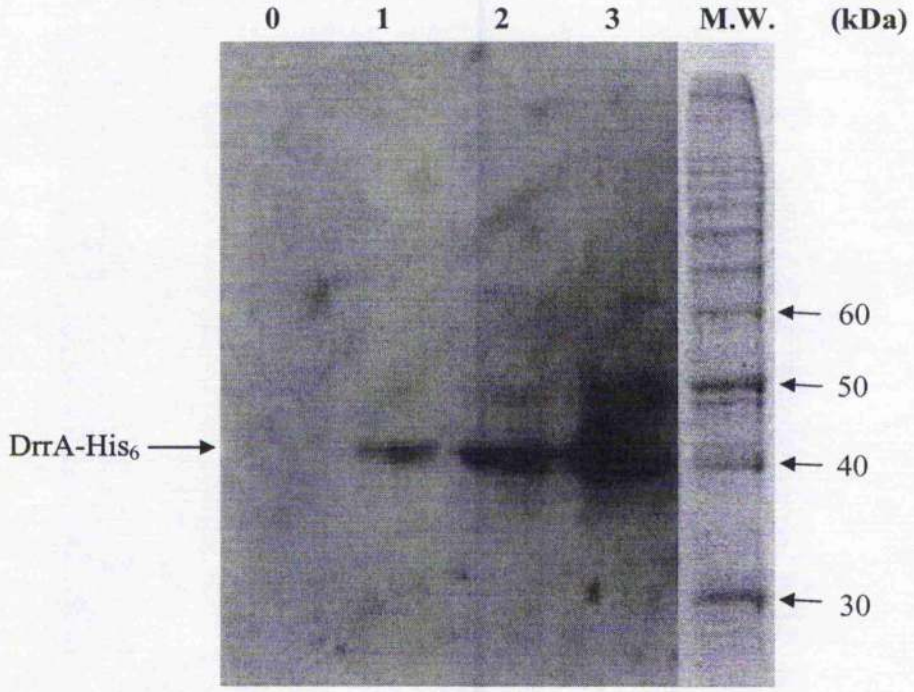
In order to further test the hypothesis that the candidate 40kDa was indeed DrrA-His_6, use was made of the immunogenic nature of the His-Tag epitope to perform a Western blot. The same samples from induced pDrrA bearing cultures shown in figure 4.2 were separated by SDS-PAGE, transferred to a PVDF membrane by electroblotting and then probed using a mouse monoclonal primary antibody raised against the His-Tag epitope. The blot was developed using an enzyme conjugated secondary antibody in conjunction with a chemiluminescent substrate following the manufacturer's recommended protocol (BioRad Immune-Star Chemiluminescent Detection Kit, see chapter 2). The results of this Western blot are shown in figure 4.3 and clearly indicate that the 40kDa candidate protein reacts strongly with the anti-His_6 monoclonal antibody. Moreover, the reaction seems to be highly specific and does not reveal the presence of any further candidate proteins. Taken in conjunction with the expression profile of this protein, and its absence from control cultures, these results strongly indicate the successful over-expression of a recombinant form of the *M. tuberculosis* DrrA protein.

Figure 4. 2 **Expression time course of DrrA-His$_6$ in *E. coli* BL21(DE3) analysed by SDS-PAGE**



Control (pET-21a) and experimental (pDrrA) cultures were induced with 1mM IPTG and samples withdrawn at the time points indicated. Amount of protein in each lane was normalized using O.D.$_{600}$ of the culture at the time each sample was taken. A candidate DrrA-His$_6$ protein can be seen accumulating over time in the experimental culture. The same protein is absent in the control cells transformed with non-recombinant pET-21a plasmid. Molecular weight markers (right hand lane, MW) show the candidate protein to be approximately 40 kDa in size, similar in size to that predicted for the recombinant DrrA-His$_6$ protein (37.4 kDa).

Figure 4. 3 **Anti-His_6 Western blot to confirm expression of the recombinant DrrA protein**



The Western blot above illustrates the strong reaction between an anti-His_6 monoclonal antibody and an approximately 40 kDa protein that accumulates after induction of expression cultures bearing the pDrrA plasmid. The same reaction is not observed prior to induction (lane 0), nor is it observed in negative control cultures (data not shown). These data serve as a positive identification of the accumulated protein as recombinant *M. tuberculosis* DrrA.

4. 4 Sub-cellular fractionation of cells expressing DrrA-His_6

In a number of instances the overexpression of ABC transporter NBD's, in the absence of their associated MSD's, has caused large amounts of the NBD to accumulate as a soluble cytosolic protein (e.g. HisP, Nikaido *et al.*, 1997; ArsA, Ilsu and Rosen, 1989). This effect is particularly desirable since it enables efficient purification of the protein for further biochemical and structural studies. In order to determine the sub-cellular localisation of recombinant DrrA-His_6, cells expressing the protein were lysed by mechanical disruption and fractionated by differential centrifugation (see Chapter 2). An initial low-speed centrifugation step (16,000 rpm; Beckman JA-21 rotor) was used to pellet dense materials, such as cell wall fragments and large insoluble proteins, whilst retaining cytosolic proteins and membrane vesicles in the supernatant. These two initial 'crude' fractions were then analysed for protein content by dissolving in SDS-PAGE sample buffer and subsequent electrophoresis. The results of the initial analysis are shown in figure 4.4. SDS-PAGE of the two fractions indicated that the vast majority of the overexpressed DrrA-His_6 protein sedimented during the low-speed centrifugation phase, whilst a much smaller amount may be present in either the cytosolic fraction or associated with membrane vesicles. This suggested that the recombinant Drr-His_6 protein was aggregating into high molecular weight complexes during expression. The formation of such complexes, often described as inclusion bodies, is not uncommon when attempting to express large quantities of a recombinant protein using strong promoters and multi-copy plasmid vectors. Inclusion bodies are thought to form when the host cell's capacity for correctly folding proteins is overwhelmed by the very large amounts of recombinant protein being synthesised. The incorrectly folded proteins rarely retain any biological function and are therefore of little use in biochemical studies. Inclusion body formation has been observed during attempts to express a number of ABC transporter NBD's, e.g. the *S. typhimurium* MalK protein (Walter *et al.*, 1992).

Despite the fact that the majority of the DrrA-His_6 was found to be present as inclusion bodies, the SDS-PAGE analysis showed that a small amount of the protein expressed may have survived in a soluble, or possibly membrane associated form (figure 4.4., left arrow). This phenomenon has been observed previously, and, through further manipulation of the expression conditions, has allowed useful quantities of soluble protein to be expressed and purified.

Figure 4. 4 **Preliminary fractionation of BL21(DE3) cells to determine sub-cellular localisation of DrrA-His$_6$**

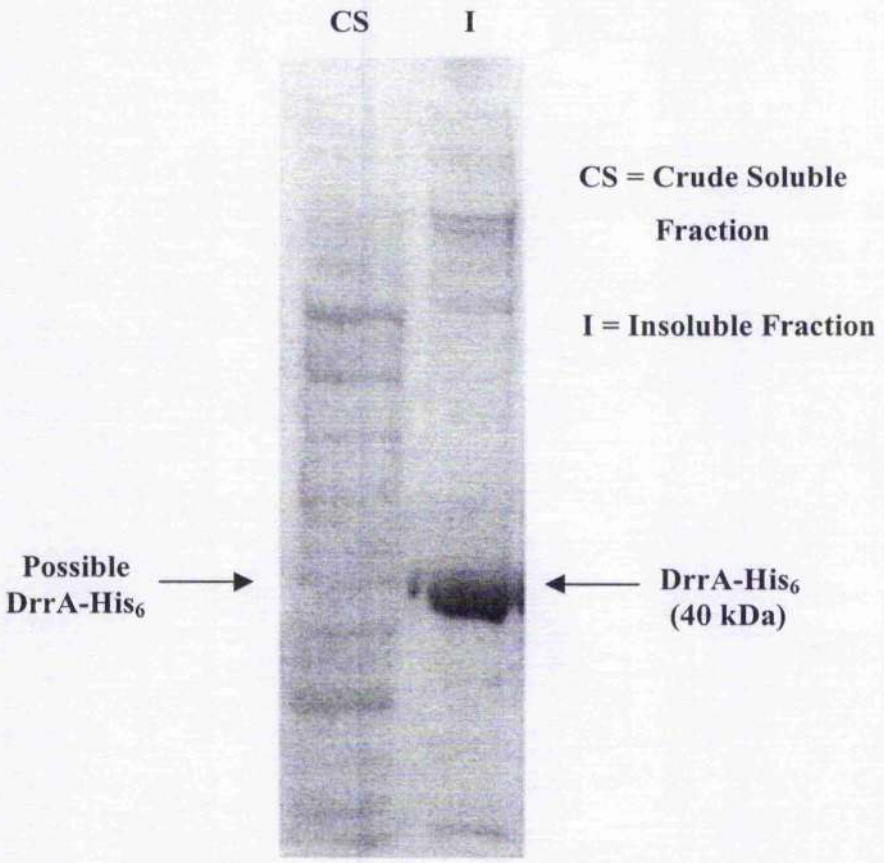
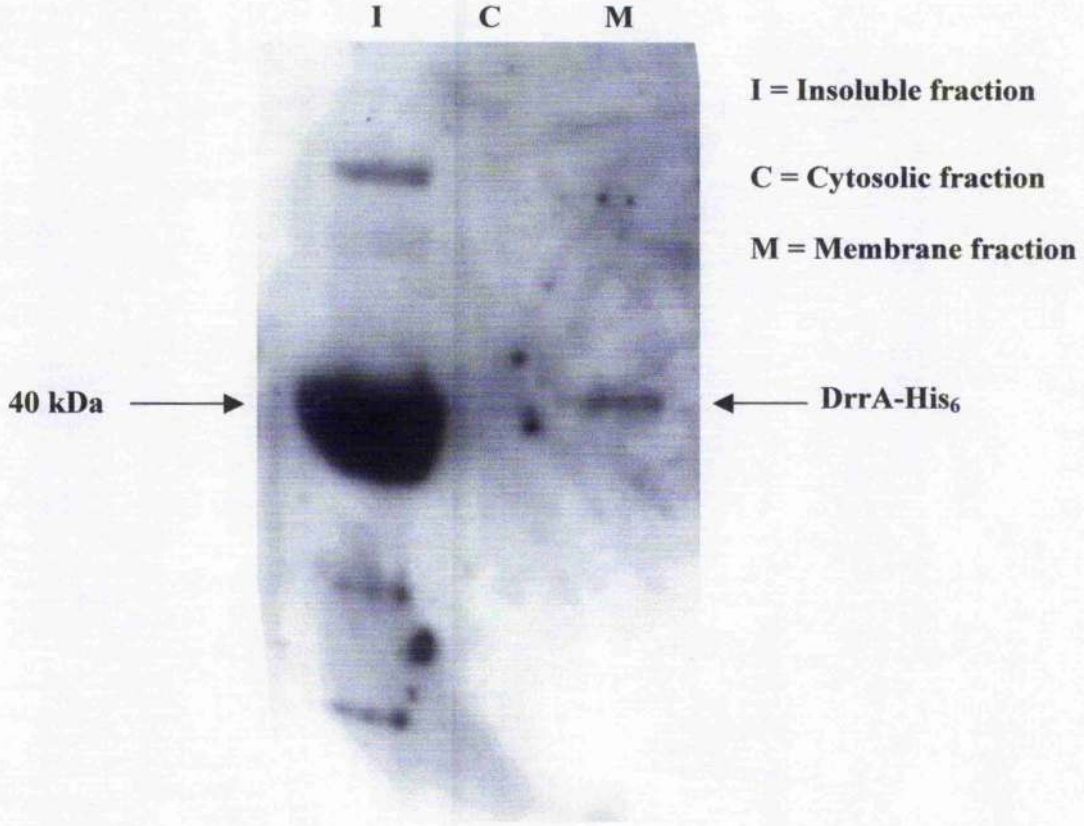


Figure 4.4 illustrates an SDS-PAGE analysis of protein fractions derived from BL21(DE3) cells harboring the pDrrA expression plasmid. Lane CS shows proteins present in a crude cytosolic extract (including membrane fragments), whilst Lane I shows the proteins present in the 'insoluble fraction'. The two fractions are derived from the supernatant and pellet fractions, respectively, following low-speed centrifugation of disrupted cells. DrrA-His$_6$ is seen to accumulate primarily in the insoluble fraction where it accounts for greater than 90% of the total protein present. This is a strong indication that the protein has formed inclusion bodies. The arrow to the left of the figure indicates the position of a small amount of DrrA-His$_6$ that has potentially remained soluble in the cytosol or is associated with membrane vesicles.

Since these initial results were the product of low-speed fractionation experiments, the possibility existed that any DrrA-His_6 observed in the crude soluble fraction might be an experimental artefact caused by cross-contamination of the fractions. In order to further assess this possibility, and more accurately determine the sub-cellular localisation of DrrA-His_6, a second round of fractionation experiments were performed.

Disrupted cells were first subjected to two rounds of low-speed fractionation in order to reduce the possibility of insoluble proteins mixing with the crude soluble extract. The soluble extract was then further fractionated into cytosolic and membranous components by high-speed ultracentrifugation (100,000 x g, 1 hour). This technique generated three protein fractions; insoluble, cytosolic and membrane associated. A sample of each was dissolved in SDS-PAGE sample buffer and analysed by electrophoresis. A Coomassie blue stain of the SDS-PAGE gel once again showed the presence of large amounts of the DrrA-His_6 protein in the insoluble fraction. Strong bands corresponding to DrrA-His_6 were observed in neither the cytosolic nor membrane associated protein fractions (data not shown). This suggested that expression of soluble DrrA-His_6 was either very low, and consequently masked by the protein background, or that a soluble form of the protein was not expressed at all. As a further stage in the analysis, a Western blot of each fraction was performed using an anti-His_6 monoclonal antibody to detect the presence of DrrA-His_6. Since immuno-affinity detection techniques are significantly more sensitive and selective than simple protein staining, it was hoped that a Western blot might succeed in detecting even very low levels of soluble DrrA-His_6 expression. The results of the Western blot are shown in figure 4.5. The results confirm that DrrA-His_6 expressed under these conditions accumulates almost exclusively in the insoluble fraction. Interestingly, absolutely no reactive protein was detected in the cytosolic fraction, whilst a small amount was found to be associated with the membranes. Since the experiments were performed in such a way as to minimise cross contamination of fractions, the presence of small amounts of DrrA-His_6 associated with the membrane was considered a genuine possibility despite the predicted hydrophilic nature of the protein. In fact, Kaur noticed a similar phenomenon when examining the sub-cellular localisation of the *S. peucetius* DrrA protein over-expressed in *E. coli*. Whilst those experiments yielded a significant amount of soluble, cytosolic DrrA, as much as 60% of the total DrrA protein was found associated with the membrane (Kaur, 1997). Together these results suggest that the DrrA proteins retain

Figure 4. 5 **Western blot of protein fractions from BL21(DE3)/pDrrA cells expressing DrrA-His_6**



Western blot of the three primary sub-cellular fractions generated by differential centrifugation (probed using an anti-His_6 monoclonal antibody). The overwhelming majority DrrA-His_6 was detected in the insoluble fraction (lane I), whilst a small amount could be detected in the membrane associated protein fraction (lane M). The experiment indicates that, under the conditions used, expression of soluble DrrA-His_6 cannot be detected, even by very sensitive techniques.

some degree of hydrophobicity, a feature which may contribute to the association of the NBD with membrane embedded, hydrophobic MSD's.

4. 5 Modulation of DrrA-His_6 expression conditions to improve yield

The above experiments showed that, under standard conditions for growth and induction of *E. coli* expression cultures, insignificant quantities of soluble DrrA-His_6 were produced. Since the ultimate aim of the expression experiments was the generation of stable, soluble, active DrrA-His_6 suitable for biochemical and structural studies, a series of expression trials were performed under different conditions aimed at increasing the yield of soluble protein.

The documentation supplied with the pET series vectors suggested a number of routes that might lead to increased yield of soluble protein (pET System Manual, Ninth Edition, Novagen). The methods described include lowering the temperature at which expression cultures are grown and induced, and also lowering the concentration of inducer (IPTG) added to the expression cultures. Both of these methods are thought to slow down the rate at which the heterologous protein is synthesised, thereby allowing the protein to fold correctly and assume a stable, globular conformation.

Initial trial cultures had been grown at 37 °C, induced with 1mM IPTG and grown over time periods of 3 hours. Optimisation experiments were performed at lowered growth temperatures of 30 °C (6 hours) and 20 °C (overnight). Reduced IPTG concentrations of 300μM and 50μM were also tested. Whilst some improvement of soluble DrrA-His_6 yield was observed, particularly when the expression cultures were grown at low temperature (20 °C) and induced overnight, yields did not improve to the point that even large-scale expression experiments might provide suitable quantities of the soluble protein (data not shown). Given the failure of these techniques to provide the basic material required for further studies, an alternative route of investigation was pursued.

4. 6 Purification of DrrA-His_6 under denaturing conditions

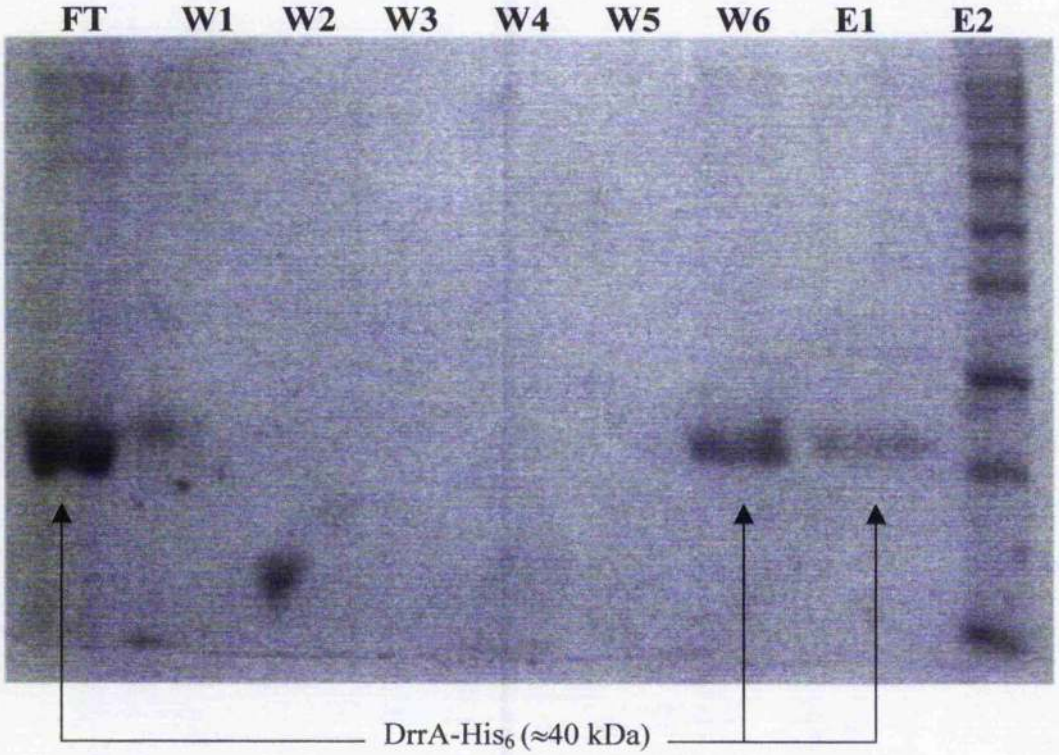
Despite the inactive nature of proteins isolated from inclusion bodies, their formation does give rise to certain advantages. Firstly, inclusion bodies represent a very concentrated and relatively pure form of the expressed protein (see figures 4.4 and 4.5). Secondly, proteins sequestered in inclusion bodies are largely protected from degradation

and proteolytic attack, and thirdly, inclusion bodies are easily isolated from cultures in large quantities by low-speed differential centrifugation. Techniques for solubilisation of inclusion bodies, and strategies for refolding the protein into an active conformation, have advanced significantly over the last decade and have been successfully employed in a number of instances. Perhaps the best example in the context of this work is the successful application of refolding techniques in the isolation of soluble, active MalK, the NBD component of the *S. typhimurium* maltose ABC-transporter (Walter *et al.*, 1992).

The inclusion bodies isolated from DrrA-His_6 expression cultures appeared to contain greater than 90% of the target protein, with relatively few contaminating proteins (see figure 4.4). However, since the conformation of the His-Tag sequence is not an important factor in determining its affinity for Ni^{2+} affinity resins, an initial purification step performed under denaturing conditions was tested as a means of further enhancing the purity of the initial inclusion body preparation. Inclusion bodies were prepared by growing a 1L culture of BL21(DE3)/pDrrA cells at 37 °C for 3 hours after induction with 1mM IPTG. Cells were harvested by centrifugation and resuspended in 50ml of disruption buffer (50mM Tris-HCl, 300mM NaCl, 10% glycerol; pH 8.0). The cells were then lysed by a single round of mechanical disruption. Inclusion bodies were recovered from the pellet following low-speed centrifugation of the crude lysate (16,000rpm; Beckman JA-21 rotor). Typical yields of DrrA-His_6 inclusion bodies were in the range of 100-300mg per litre of culture.

Inclusion body solubilisation requires the use of strong denaturants to completely separate and unfold the aggregated protein. Reagents employed for this purpose include buffered solutions of either 6M guanidine-HCl or 8M urea. In these experiments a solution of 8M urea in 50mM Tris-HCl, 300mM NaCl (pH 8.0) was used as the denaturant. 50mg of inclusion body pellet was solubilised in 10ml of this buffer by stirring at room temperature for 1 hour. A further round of centrifugation at 16,000rpm removed material that remained insoluble. The inclusion body solution was then mixed with 1ml of IMAC Ni^{2+} affinity resin (Ni-NTA, Qiagen) and agitated gently for 1 hour at room temperature in order for the His_6-tagged protein to bind to the resin. The solution was then loaded into a 0.8cm diameter gravity fed chromatography column and the resin allowed to settle. The flow-through solution was collected for further analysis.

Figure 4. 6 **Purification of denatured DrrA-His$_6$ by IMAC column chromatography**



FT	=	**Column flow-through**
W1 + W2	=	**20mM imidazole washes**
W3 + W4	=	**30mM imidazole washes**
W5 + W6	=	**40mM imidazole washes**
E1 + E2	=	**300mM imidazole elutions**

Base Buffer (pH 8.0)	
50mM	Tris-HCl
300mM	NaCl
8M	Urea

A Coomassie blue stained SDS-PAGE gel showing various stages in the purification of DrrA-His$_6$ by IMAC chromatography. The source material used was DrrA-His$_6$ inclusion bodies, and the entire purification procedure was performed under denaturing conditions (8M urea). Lane FT clearly shows that a large quantity of solubilised DrrA-His$_6$ did not bind to the IMAC resin, possibly as a result of saturation. Lanes W1-W6 show the removal of background contaminants under a stepped imidazole gradient. The final two lanes show the elution of DrrA-His$_6$ from the resin using a buffer containing 300mM imidazole. The two elutions contained highly purified, denatured DrrA-His$_6$.

The resin was washed with 10ml of the initial solubilisation buffer in order to remove any unbound protein. Extensive washing was performed at this and subsequent stages since the amount of protein loaded onto the resin (nominally 50mg) far exceeds the known binding capacity of 1ml of Ni-NTA resin (≈10mg/ml). The resin was washed a further 3 times using 10ml of buffers supplemented with 20mM, 30mM, and 40mM imidazole. These washes were aimed at removing non-specifically bound background proteins and loosely bound aggregates. At all stages of the washing procedure 5ml fractions were collected for analysis. Finally, specifically bound, denatured DrrA-His_6 was eluted from the column using 10ml of solubilisation buffer supplemented with 300mM imidazole. The eluate was collected in two fractions of 5ml each. At the end of the procedure all fractions collected were analysed for protein content by SDS-PAGE.

The results of the denaturing purification are illustrated in figure 4.6. The chromatography procedure was essentially successful, yielding a highly purified solution of denatured DrrA-His_6 ideal for use in subsequent refolding experiments. However, as can be judged from the SDS-PAGE analysis, significant losses in yield occurred, and only a small proportion of the starting material was recovered in the final elutions. The yield losses incurred were almost certainly due to the limited binding capacity of the small resin volume used.

4. 7 DrrA-His_6 refolding experiments – Part I - IMAC

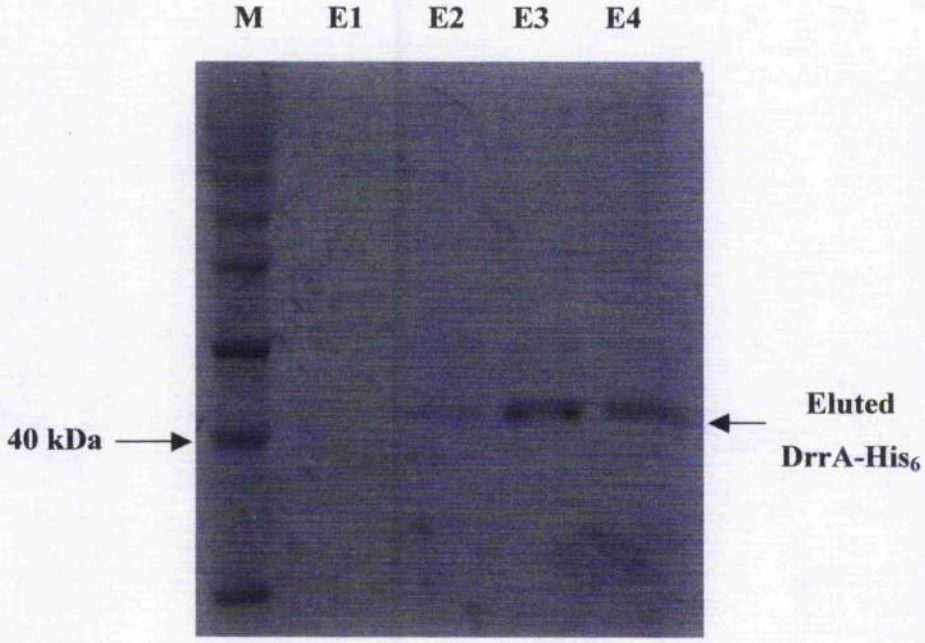
The demonstration that inclusion body-derived DrrA-His_6 was able to bind to the IMAC resin in a denatured form was the starting point for the first set of refolding experiments undertaken. Most forms of protein refolding require a reduction in the concentration of denaturant, allowing time for individual protein molecules to adopt a number of intermediate conformations before reaching their final, fully folded state. However, many of the intermediate protein conformations are likely to be unstable. Interactions between these unstable folding intermediates can lead to protein aggregation and/or eventual precipitation. Physical separation of the molecules, for example by immobilising them on an IMAC resin, is thought to reduce the possibility of unfavourable protein-protein interactions occurring during the refolding process. This technique was used successfully during the early MalK refolding experiments, albeit using a different method of immobilisation (Walter *et al.*, 1992), and has even been applied to the refolding of membrane proteins (Rogl *et al.*, 1998).

Urea solubilised DrrA-His_6 inclusion bodies were prepared for refolding as described in section 4.5 with the exception of two changes. Firstly, the solubilisation process and all subsequent steps were performed at 4 °C rather than at room temperature in order to reduce the possibility of protein hydrolysis or damage. Secondly, the NaCl concentration in all buffers was raised from 300mM to 0.5M. The increase in ionic strength of the buffer was designed to reduce non-specific protein-protein interactions during refolding, particularly as the urea concentration was reduced. A batch binding procedure was once again used to immobilise the solubilised DrrA-His_6, this time onto a somewhat larger volume of resin (4ml) in order to improve yields. After 1 hour of binding at 4 °C the IMAC resin was loaded into a chromatography column and the resin allowed to pack under gravity flow.

The refolding procedure began by washing unbound protein out of the resin using 20ml of solubilisation buffer. Subsequent stages involved washing the resin in 16 separate 5ml aliquots of buffer, the urea concentration in each buffer being reduced by 0.5M each time (from 8 to 0M). The aim was to generate a stepped gradient that gradually reduced the urea concentration to zero whilst allowing the immobilised protein time to refold into a native conformation. Finally, refolded, soluble DrrA-His_6 was released from the column by washing with a number of 1ml aliquots of elution buffer (0M urea, 300mM imidazole).

One of the major problems encountered during the various refolding trials performed was a dramatic reduction in the flow rate of the column that occurred as the urea concentration approached the 3-4M range. This was believed to be caused by protein aggregation within the matrix, sometimes manifesting itself as a white coloured precipitate. Despite a number of modifications to the procedure, including shallower gradients and allowing extended periods of time between buffer changes, yields of soluble protein were rarely significant. A typical result from one of these refolding experiments, illustrating the low yield of DrrA-His_6 recovered, is shown in figure 4.7. Furthermore, whilst SDS-PAGE analysis of the elution fractions shows monomeric, apparently soluble DrrA-His_6, no biological activity could be detected for the protein (for example in ATPase activity assays). This suggested that the elutions were merely releasing inactive protein aggregates from the column matrix which were then being re-solubilised by the detergent present in SDS-PAGE sample buffer.

Figure 4. 7 **IMAC-based refolding of immobilised DrrA-His_6**



Gradient Buffer: 50mM Tris-HCl
500mM NaCl
8-0M Urea (in 0.5M steps)
pH 8.0

Elution Buffer: 50mM Tris-HCl
500mM NaCl
300mM Imidazole
pH 8.0

Example of a Coomassie blue stained SDS-PAGE gel showing 4 x 1ml elution fractions (E1-E4) recovered from a typical IMAC refolding experiment. Whilst the eluted DrrA-His_6 runs as a monomeric protein of high purity, it displays no detectable ATPase activity. This data, combined with evidence of protein precipitation during the urea gradient, suggests that the eluted protein is unable to adopt a native conformation under the experimental conditions used.

The IMAC based method of protein refolding described above failed to recover any biological activity from urea solubilised DrrA-His_6 inclusion bodies. Nevertheless, only a single buffer system was tested during the procedure. Given that large quantities of a relatively pure, easily isolated starting material were available, an alternative approach was adopted in order to identify optimal refolding parameters.

4. 8 DrrA-His_6 refolding experiments – Part II – The FoldIt Screen

The efficiency of protein refolding is known to be influenced by a large number of external conditions, as well as by the individual physico-chemical properties of the protein under investigation. Since there is, as yet, no rational framework for predicting ideal refolding conditions, it is necessary to establish the parameters empirically. A number of kits are commercially available that help to identify optimal refolding conditions. One such kit, the FoldIt Screen (Hampton Research), was used in an effort to identify the various conditions that might affect the efficiency of DrrA-His_6 refolding. The screen is factorial in nature, using 16 different basic buffer conditions in combination with a number of additives known to promote successful refolding. The identification of the basic parameters can then be used to further optimise the refolding process. The various optimisation parameters tested in the refolding of DrrA-His_6 are shown in table 4.8 on the next page.

Inclusion bodies for use in the FoldIt screen were prepared as described in section 4.5 and solubilised in a buffer of 8M Urea, 50mM Tris-HCl, 50mM NaCl, pH 8.0. The reduced salt concentration in the solubilisation buffer was deemed necessary because ionic strength is a variable in the screen. DrrA-His_6 inclusion bodies were solubilised at a nominal concentration of 20mg/ml and insoluble material was removed from the preparation by centrifugation at 12,000rpm in a bench-top microfuge. In contrast to the IMAC based refolding protocol described in the previous section, where the concentration of denaturant was reduced gradually, the FoldIt screen utilises rapid dilution in order to reduce denaturant concentration. Reactions were performed in 1.5ml micro-centrifuge tubes and initiated by the addition of a small volume of the

Table 4. 8 - DrrA-His_6 refolding parameters tested

Protein concentration	0.1mg/ml vs. 1.0mg/ml
Presence of polar additive	+/- 550mM L-Arginine
Presence of detergent	+/- 30mM Lauryl Maltoside
pH	pH 6.5 vs. pH 8.2
RedOx potential	100mM DTT vs. 100mM GSH
Presence of chaotropic salt	+/- 550mM Guanidine Hydrochloride
Ionic strength	264mM NaCl +11mM KCl vs. 10.56mM NaCl + 0.44mM KCl
Presence of cations or chelator	2.2mM CaCl + 2.2mM MgCl vs. 1.1mM EDTA
Presence of osmolyte	+/- 0.055% (w/v) PEG 3350
Presence of non-polar additive	+/- 440mM Sucrose

concentrated DrrA-His_6 inclusion body preparation to a much larger volume of the various refolding buffers. Typically either 5 or 50µl of inclusion body solution was added to 950µl of refolding buffer in order to achieve final DrrA-His_6 concentrations of 0.1mg/ml and 1mg/ml. The tubes were capped and incubated at 4 °C with gentle agitation for a period of 4 hours.

After 4 hours each of the 16 refolding reactions were examined visually for signs of protein precipitation. A degree of precipitation was observed in all cases, particularly at higher protein concentration. However the extent the visible precipitate varied widely between different tubes, suggesting that certain refolding solutions represented a more beneficial environment for refolding than others. The polar additive L-arginine, present in 8 of the 16 reactions, appeared to exert a particularly positive influence in terms of minimising the amount of visible precipitate present.

At this stage, the ideal situation would have been to remove precipitates by centrifugation, and then assay each solution for the predicted ATPase activity of correctly folded DrrA-His_6. Unfortunately, a number of the reagents present in the FoldIt buffers, for example L-arginine, divalent cations and EDTA, are known to interfere with both

spectroscopic protein quantitation techniques (A_{280}, BCA assay) and ATPase assays. This made a meaningful comparison of enzymatic activity in each of the 16 samples impossible at this stage. In order to make such a comparison possible, all 16 samples were transferred into separate dialysis tubing and dialysed against 2L of a buffer consisting of 50mM Tris-HCl (pH 8.0), 150mM NaCl, 10% glycerol (v/v). This dialysis procedure appeared to result in the removal of reagents that were enabling DrrA-His_6 to remain soluble, leading to further extensive precipitation. The FoldIt screen documentation suggests that yields of 30-70% soluble protein after the dialysis procedure are typical where refolding has been successful. Unfortunately, such yields were not observed for DrrA-His_6 and no ATPase activity could be detected in any of the 16 samples tested (using the Molecular Probes spectroscopic Pi assay, see chapter 2). The failure of this screen to identify suitable refolding conditions for DrrA-His_6 led to the abandonment of inclusion bodies as a suitable source material for further biochemical studies of the recombinant *M. tuberculosis* DrrA protein. An alternative method for obtaining correctly folded, native DrrA was sought, as described in subsequent chapters.

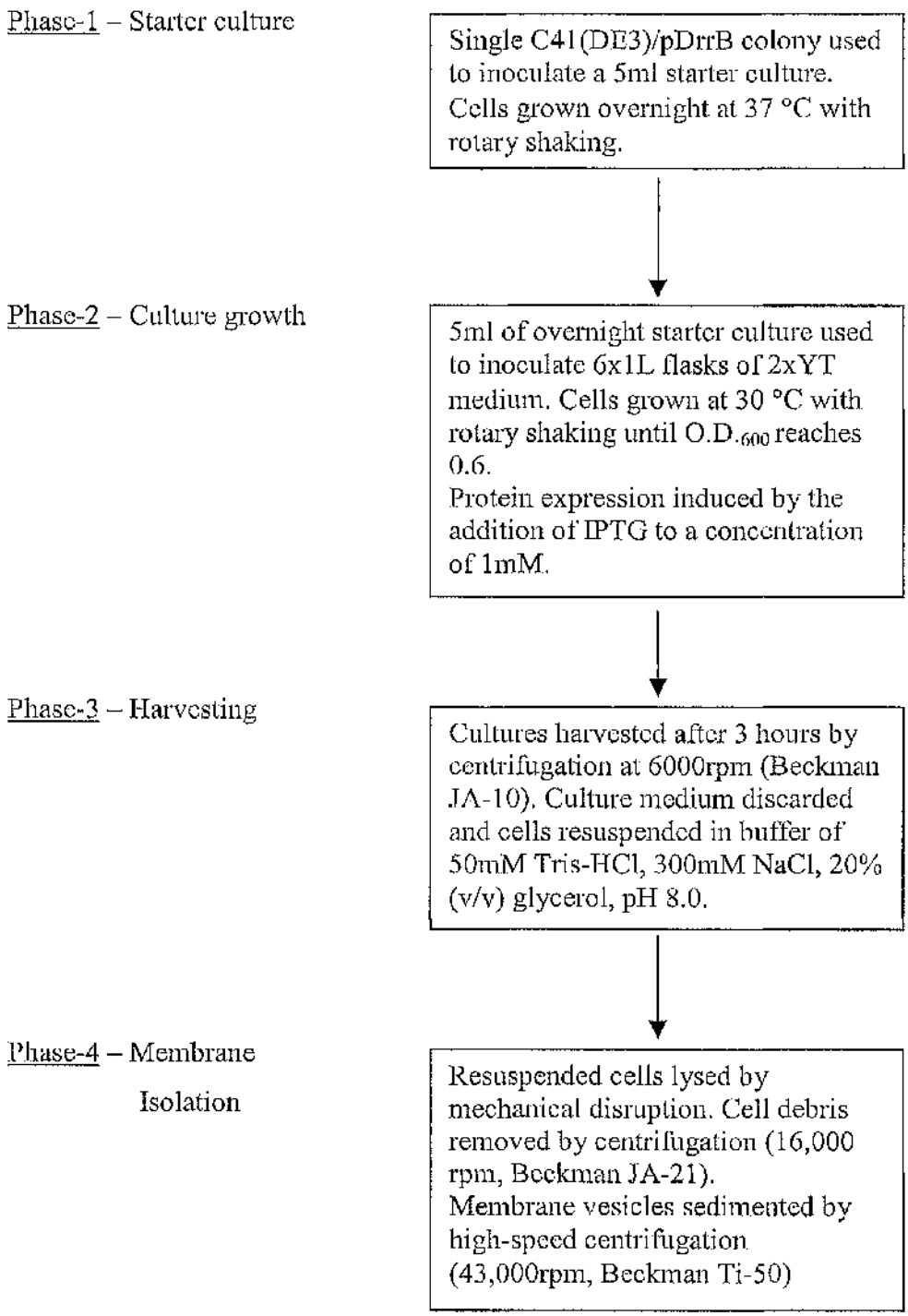
4. 9 Heterologous expression of the *M. tuberculosis* DrrB and DrrC proteins – the pDrrB and pDrrC expression constructs

Integral membrane proteins, such as those mediating the majority of transmembrane transport processes, are notoriously difficult to over-express and purify. Ordinarily present at only low concentrations within cellular membranes, their over-expression can result in toxic side effects for both heterologous and homologous hosts. Another problem frequently encountered is the failure of heterologously expressed membrane proteins to be correctly targeted to the appropriate compartment, a process that may require host specific signal-sequences or protein factors. Furthermore, the conformation of transmembrane proteins is intimately linked to the hydrophobic environment of the membrane interior. Removal of integral membrane proteins from their lipidic environment, for example by the use of detergents, can dramatically affect their structure and function. The difficulties associated with membrane protein over-expression and purification largely account for the paucity of high-resolution structural information available for this biologically important class of proteins. The problems are however not insurmountable, as demonstrated by the recent publication of 3D crystal structures for the *E. coli* MsbA and BtuCD ABC transporters (Chang and Roth, 2001; Locher *et al.*,

2002), and the AcrB multidrug resistance protein (Murakami *et al.*, 2002). Advances in detergent chemistry have created whole series of molecules specifically designed for the gentle extraction of proteins from biological membranes whilst maintaining their functional conformation. Methods for reconstituting transport proteins into artificial membrane environments, for example proteoliposomes, have enabled the development of novel membrane transport assays. Hand-in-hand with these advances have come a variety of new reagents used to probe the conformation and transport capabilities of membrane embedded proteins. With this knowledge in hand, attempts were made to over-express and purify the DrrB and DrrC components of the *M. tuberculosis* DrrABC transporter.

Expression constructs for both the DrrB and DrrC proteins were generated by subcloning *NdeI/HindIII* fragments from recombinant pGEM-T Easy plasmids into the pET-21a expression vector. The two plasmids thus generated, pDrrB and pDrrC, are essentially analogous to the pDrrA expression construct used for the expression of DrrA. The pDrrB plasmid encodes a 302 amino acid protein with a predicted molecular weight of 32.5 kDa. The first 289 amino acids are encoded by the cloned *drr*B gene, whilst the remaining 13 amino acids at the C-terminal end are vector encoded and include a His-Tag sequence. The pDrrC plasmid encodes a slightly smaller protein of 289 amino acids (M.W. 28.1 kDa) with the same 13 amino acid extension appended to the C-terminal of the native *drr*C coding sequence.

Both the pDrrB and pDrrC plasmids were transformed in to the expression strain C41(DE3). This strain is an undefined mutant derived from *E. coli* BL21(DE3) that possesses certain advantages for the expression of membrane proteins. Identified by Miroux and Walker, C41(DE3) appears to be more resistant to the toxic effects of IPTG than its parent strain (Miroux and Walker, 1996). As such, C41(DE3) cultures grow to higher cell densities in liquid media and thus generate a higher yield of recombinant protein per litre of culture. This is advantageous when the target protein under investigation expresses only at low levels, as is the case for most membrane proteins. Because of the lower expression levels expected for the recombinant DrrB and C proteins, expression experiments were performed using larger culture volumes, typically 3-6L, and cells were grown in the rich 2xYT medium to enhance growth. A flow chart showing the progress of an expression experiment from cell culture to isolation of membrane material is shown on the next page.



Following high-speed centrifugation the supernatant fraction (containing cytosolic proteins) was discarded. Membrane pellets were then resuspended in an ice-cold, detergent free buffer (50mM Tris-HCl, 300mM NaCl, 20% glycerol, pH 8.0) supplemented with EDTA-free protease inhibitor tablets. Resuspension of the hydrophobic membrane material in an aqueous buffer required the pellet to be broken up

and passed through a syringe needle many times in order to generate a smooth emulsion of vesicles. Membranes were resuspended using 5ml of buffer per litre of culture. At this point membrane suspensions were either stored at –80 °C or processed immediately.

4. 10 Detergent extraction of membrane proteins

Purification of membrane proteins requires their extraction from the membrane environment. This is generally performed using biological detergents at concentrations high enough to disrupt the membrane structure whilst not destroying protein conformation. Solubilisation is a complex, multi-phasic process, the end result being the generation of both mixed lipid/detergent micelles, and lipid/detergent/protein micelles. It is the hydrocarbon 'tails' of detergent molecules that are thought to associate with integral membrane proteins, thereby mimicking the hydrophobic environment of the membrane interior, and also preventing protein aggregation. The detergent *dodecyl-maltoside* (DDM) has been used successfully in the extraction of a number of membrane proteins, including human MDR1 (Howard and Roepe, 2003), and was used here in attempts to solubilise DrrB and DrrC.

Resuspended membrane vesicles were solubilised by the addition of an equal volume of ice-cold resuspension buffer supplemented with 4% w/v DDM and protease inhibitors. A final concentration of 2% DDM was found to be more than enough to completely solubilise a typical membrane preparation. Solubilisation was enhanced by gentle mixing of the solution for a period of 1 hour at 4 °C. Low temperatures, protease inhibitors and pH of 8.0 were used at all times during protein extraction. These factors are all known to inhibit the action of proteases (to which solubilised membrane proteins become particularly susceptible). After detergent extraction any remaining insoluble material was removed by a further round of ultracentrifugation (100,000 x g, 1 hour). The quantity of sedimented material at this stage was used to judge the efficiency of the solubilisation process. The supernatant fraction, containing detergent-solubilised proteins, was used as the starting material for the purification procedure.

4. 11 IMAC purification protocols for the DrrB and DrrC proteins

Both the pDrrB and pDrrC expression constructs were designed to append a C-terminal His-Tag sequence to the recombinantly expressed proteins. This provided a means of selectively purifying the proteins by IMAC using Ni-NTA affinity resin. The purification

procedure was started by adding 1ml of Ni-NTA-agarose to the solubilised membrane protein solution and allowing binding to proceed for 1 hour at 4 °C with gentle stirring. The mixture was then loaded into a 1.5cm diameter gravity flow chromatography column and the resin allowed to settle. The flow-through fraction containing unbound proteins was collected for analysis. The following washing and elution steps, containing different concentrations of imidazole, were then used as a starting point to identify fractions containing the recombinant proteins: -

Wash 1.	20ml buffer* containing 0mM imidazole to wash out unbound protein.
Wash 2.	20ml buffer containing 25mM imidazole, removal of loosely bound contaminants.
Wash 3.	20ml buffer containing 40mM imidazole, removal of more tightly associated background proteins.
Elutions	5x 1ml buffer containing 300mM imidazole to elute specifically bound His-Tagged proteins.

* A variety of Tris-based buffers were tested during purification trials. Glycerol concentration was maintained at 20% v/v in all buffers tested in order to promote membrane protein stability. NaCl concentrations tested varied between 50mM and 400mM. Detergent concentration and formulation were two further variables tested. After the solubilisation process, in which the majority of lipids are stripped away from membrane proteins, much lower concentrations of detergent (e.g. 0.1% w/v) were used to maintain proteins in solution. Different detergents with alternative head groups and chain lengths were also tested for their ability to promote protein stability and reduce background contamination e.g. Octyl-glucoside and nonyl-maltoside as well as DDM.

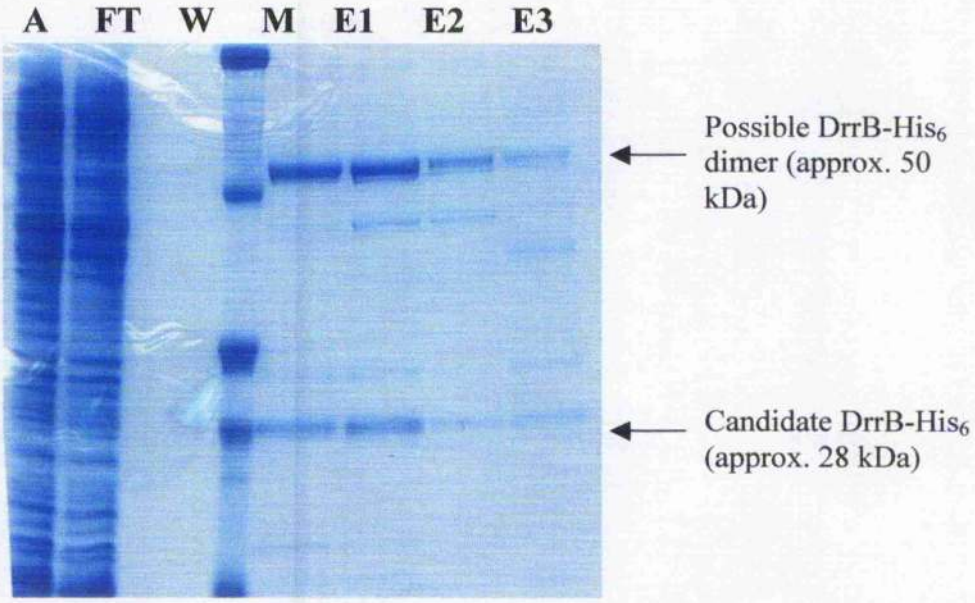
4. 12 Purification of recombinant DrrB-His_6

Unlike recombinant DrrA-His_6, induction of DrrB-His_6 expression with 1mM IPTG did not lead to the formation of inclusion bodies. In fact, simple SDS-PAGE analysis of the various sub-cellular fractions obtained from DrrB-His_6 expression cultures did not reveal

the obvious presence of a protein exhibiting particularly high expression levels (data not shown). Attempts to demonstrate the expression of DrrB-His_6 by Western blot also proved unsuccessful, primarily hindered by the large number of background proteins in the crude fractions. Together these observations led to the conclusion that DrrB-His_6 expression levels were far lower than those observed for the recombinant DrrA protein. Such a result might be predicted based upon the generally low levels of expression observed for integral membrane proteins. An alternative explanation was that no DrrB-His_6 was being expressed at all. Causes of such negative results might include toxicity of the DrrB-His_6 protein. Another potential reason for a low expression might be the large evolutionary separation between the Gram-negative expression host *E. coli* and the Gram-positive donor organism *M. tuberculosis*. Such evolutionary separation can lead to differences in codon usage between the host and donor organisms, which in turn are reflected in the relative abundance of tRNA species. This may be relevant when attempting to express GC-rich mycobacterial genes in *E. coli*, an organism that possesses a slightly AT-biased genome.

Nonetheless, working on the hypothesis that DrrB-His_6 would be localised in the cell membrane, attempts to purify the protein from this fraction were initiated using the protocols described in section 4.10. After a number of different trials, a candidate DrrB-His_6 protein was partially purified from DDM solubilised membranes. An SDS-PAGE analysis of one successful procedure is shown in figure 4.8. The figure illustrates the absence of any obviously over-expressed protein in the solubilised membranes prior to purification, but also the specific elution of two major protein species in the 300mM imidazole fractions. The molecular weight of the smaller protein was approximately 28 kDa (judged by comparison with molecular weight standards). This value is slightly lower than the calculated size for DrrB-His_6, 32.5 kDa. However, since many hydrophobic proteins display a greater than expected electrophoretic mobility, the possibility remained that this protein was indeed DrrB-His_6. The larger protein present in the eluate fractions purified in similar quantities to the lower band, suggesting that it too was specifically bound to the affinity resin, rather than being present as a

Figure 4. 8 **IMAC purification of recombinant DrrB-His_6**



Coomassie blue stained SDS-PAGE gel showing partial IMAC purification of candidate DrrB-His_6 protein(s) from DDM solubilised membranes. Lanes A and FT show the samples applied to the IMAC column and flow-through collected, respectively. Lane W shows the fraction obtained after extensive washing of the column to remove non-specifically bound protein (60ml of buffer containing 40mM imidazole). Lanes E1-E4 show elution fractions (4x 1ml) generated by washing the column with buffer containing 300mM imidazole. Arrows mark the positions of candidate proteins.

Membrane solubilisation buffer - 50mM Tris-HCl (pH 8.0)
150mM NaCl
20% (v/v) glycerol
2% (w/v) DDM

Column buffers: - 50mM Tris-HCl (pH 8.0)
150mM NaCl
20% (v/v) glycerol
0.1% (w/v) Nonyl-maltoside
Imidazole: 60 mM (wash), 300mM (elution)

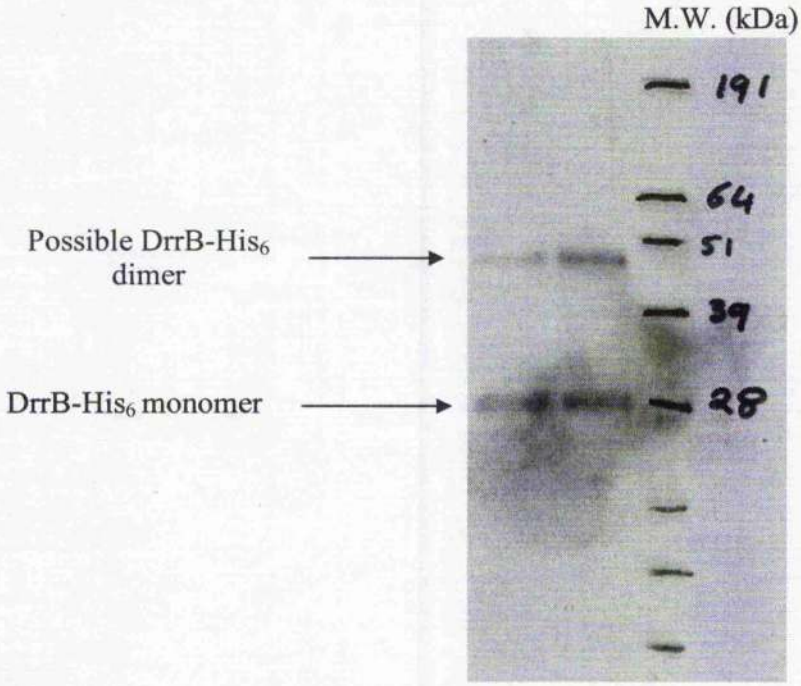
contaminant. The observed molecular weight of this protein, approximately 50 kDa, suggested that it might be a dimeric or aggregated form of the smaller protein. Since ABC transporter MSD's are known to exist as dimers in the functional complex, the possibility remained that DrrB-His_6 might form a stable homodimer in the absence of DrrC. This would account for co-purification of this higher molecular weight species.

In order to further characterise the candidate proteins, a Western blot was performed using an anti-His_6 monoclonal antibody. This experiment was aimed at identifying which, if either, of the two major protein species eluted from the IMAC column possessed a His-Tag. The results of this Western blot are shown in figure 4.9.

Both of the major protein species reacted strongly with the anti-His_6 antibody. Similar positive reactions were not observed for the low abundance contaminants released from the IMAC column at the same time. This result provided confirmation not only of the identities of both protein species as DrrB-His_6, but also of the membrane localisation of the recombinantly expressed DrrB protein. Such a result was consistent with the hypothesis that the *M. tuberculosis drr*B gene encodes a membrane-spanning component of a novel mycobacterial ABC transporter.

Despite the successful expression and purification of DrrB-His_6, further manipulation of the protein proved difficult. Such basic technical procedures as dialysis, necessary to remove imidazole from the preparation, or concentration of the protein led to precipitation. This suggested that the isolated DrrB-His_6 protein became unstable after its extraction from the cell membrane. A search for more suitable elution buffers that might stabilise the protein ultimately proved fruitless. Attempts were made to reconstitute the protein into artificial proteoliposomes in order to increase its stability. Unfortunately the reconstitution process required extensive dialysis of the protein preparation in order to remove detergent and promote incorporation of the protein into lipid vesicles. The dialysis procedure invariably resulted in aggregation and precipitation. The presence of imidazole in the preparations, necessary for elution of the protein from the IMAC column, presented a further experimental problem. Imidazole exhibits strong absorbance and fluorescence in the same range of wavelengths as native proteins. Present at 300mM in a typical DrrB-His_6 preparation, imidazole would mask any of the subtle spectroscopic effects that might be induced by interaction of DrrB-His_6

Figure 4. 9 Western blot of heterologously expressed DrrB-His_6



An anti-His_6 monoclonal antibody was used to determine the presence of His-Tagged proteins in elutions from an IMAC purification column aimed at purifying DrrB-His_6. The above Western blot illustrates that the two major protein species eluted from the IMAC column (see figure 4.8) react strongly with the antibody. Molecular weights of the reactive proteins are consistent with monomeric and dimeric forms of the heterologously expressed DrrB-His_6 protein.

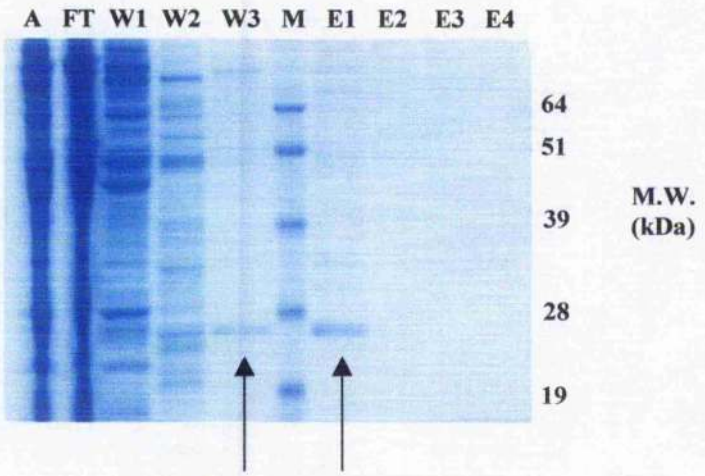
with potential substrates (such as doxorubicin). Given these experimental problems, a further characterisation of DrrB function proved impossible. It was hypothesised that DrrB stability might be dependent upon the co-expression and purification of its presumed partner, DrrC.

4. 13 Attempted expression and purification of DrrC-His_6

The sequence homology between the *M. tuberculosis* DrrB and DrrC proteins suggested that a similar set of experimental parameters might be of use in purifying both of the proteins. As such, initial DrrC-His_6 expression trials were performed using identical conditions to those used for the expression and purification of the DrrB protein. C41(DE3)/pDrrC cells were grown at 30 °C in 2x YT medium and induced by the addition of 1mM IPTG. As had been the case with recombinant DrrB, no obviously over-expressed protein could be identified by simple fractionation the cells. The decision was made to progress directly to a series of IMAC purification trials. *E. coli* membranes were prepared from C41(DE3)/pDrrC cells as described in section 4.8.

By using a slight modification of the DrrB-His_6 purification protocol a potential DrrC-His_6 protein was soon identified. An SDS-PAGE analysis of the IMAC purification procedure is shown in figure 4.10 along with the experimental conditions. The candidate DrrC-His_6 protein eluted as a single species from the IMAC column exhibiting an apparent molecular weight of slightly less than 28 kDa. In common with the previously identified DrrB protein, the recombinant form of DrrC exhibited a marginally higher electrophoretic mobility than might be expected for a 28.1 kDa protein. Once again this discrepancy was put down to the likely hydrophobic nature of the DrrC-His_6 protein. The candidate DrrC-His_6 protein was, as expected, slightly smaller than the previously purified DrrB-His_6 protein. Unlike the DrrB protein however, the candidate DrrC-His_6 showed no evidence of dimerisation or formation of higher order aggregates. Additionally, the candidate protein showed a lower affinity for the Ni-NTA resin than DrrB-His_6 had done. Figure 4.10 illustrates how some of the target protein was lost during washing of the IMAC resin with buffer containing 60mM imidazole. Under similar washing conditions DrrB-His_6 had remained fully bound to the column matrix. The decreased yield of target protein compared with DrrB-His_6 was deemed an acceptable trade-off given the increased purity achieved and the elution of only a single protein species.

Figure 4. 10 **Purification and SDS-PAGE analysis of a candidate DrrC-His_6 protein**



Solubilisation buffer: -

50 mM Tris-HCl (pH 8.0)

150mM NaCl

20% glycerol (v/v)

2% DDM (w/v)

IMAC purification buffers:-

50 mM Tris-HCl (pH 8.0)

150mM NaCl

20% glycerol (v/v)

0.1% Nonyl-maltoside (w/v)

Imidazole (20mM, 40mM, 60mM: washes)

Imidazole 300mM (elution buffer)

A Coomassie blue stained SDS-PAGE gel showing the purification of a candidate DrrC-His_6 protein. The purification shown used total cell membranes isolated from 3L of C41(DE3)/pDrrC cell culture. The candidate protein elutes as a single species of high purity with an apparent M.W. slightly lower than the predicted 28.1 kDa. Lanes A and FT represent applied and flow-through fractions respectively. Lanes W1-W3 are 20ml wash fractions containing increasing concentrations of imidazole (20mM, 40mM and 60mM). Lanes E1-E4 are 4x 1ml eluate fractions of a buffer containing 300mM imidazole. The candidate protein is present in only the first eluate fraction. However a significant quantity of the target protein appears to have been stripped from the IMAC column by the 60mM imidazole wash (lane W3).

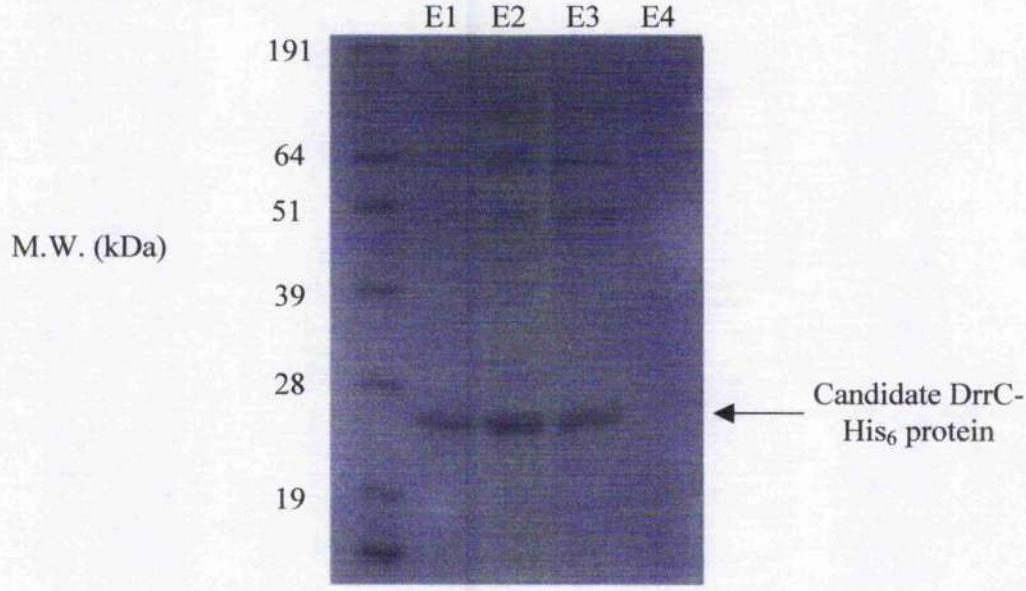
The candidate DrrC-His_6 protein proved to be considerably more stable than its counterpart, DrrB. The protein could be dialysed safely to remove imidazole without significant precipitation. This observation was somewhat surprising given the highly unstable nature of the homologous DrrB-His_6 protein. Minor modifications to the purification protocol produced much larger yields of the target protein. These modifications included processing larger quantities of isolated membranes in a single purification step (e.g. total membranes isolated from 6 litres of cell culture). Additionally, less stringent column washes were employed in an attempt to reduce target protein losses during the purification. The lower stringency washes only caused a minor increase in the levels of contaminants co-eluting with the target protein. Elution fractions from the enhanced purification protocol are shown in figure 4.11.

The high purity of the eluted protein and the increased quantities available enabled the protein to be identified by direct N-terminal amino acid sequencing. Samples of the target protein were sent to the University of Leicester for analysis by automated N-terminal Edman degradation. 10 cycles of analysis were performed generating the following results: -

Cycle	1	2	3	4	5	6	7	8	9	10
Residue	M	K	V	A	L	D	L	V	V	S

Comparison of this N-terminal sequence with that encoded by the native *M. tuberculosis drr*C gene reveals similarity at only 1 position, the initiation methionine at the extreme N-terminus. The predicted N-terminal sequence of DrrC-His_6 should have read MTTTTSQEIE. This data suggested that a major contaminant protein was purifying in preference to any DrrC-His_6 that might be present, and that DrrC-His_6 expression levels were very low. An examination of the NCBI protein database revealed the identity of the contaminating protein, and also provided an explanation for its high affinity for the Ni-NTA resin. The 10 N-terminal amino acids of the contaminant exhibited 100% homology with those of the native *E. coli* protein SlyD (GenPept accession no.P30856). This protein normally functions as a *peptidylproline cis-trans-isomerase*, one of a number of chaperone-like proteins that facilitate correct protein folding (Hottenrott *et al.*, 1997). The protein has been identified as a major contaminant in a number of Ni-NTA based IMAC

Figure 4. 11 **Modification of the purification procedure enhances the yield of the candidate DrrC-His_6 protein**



Buffer compositions : -
50mM Tris-HCl (pH 8.0)
150mM NaCl
20% (v/v) glycerol
0.1% (w/v) Nonyl-maltoside
Imidazole (40mM in wash buffer, 300mM in elution buffer)

SDS-PAGE analysis showing elution of increased quantities of the candidate DrrC-His_6 protein. Total membranes were isolated from 6L of C41(DE3)/pDrrC, solubilised in buffer containing 2% DDM, and the target protein purified by IMAC. The increased yield shown in this preparation (c.f. figure 4.10) was also enhanced by a reduction in the imidazole concentration present in the column washing buffer. The 1ml IMAC resin bed volume was washed just once with 20ml of a 40mM imidazole wash buffer prior to elution of the target protein. Elution was achieved by the addition of 4x 1ml aliquots of buffer containing 300mM imidazole (lanes E1-E4). The purified protein shown in this figure was sent for N-terminal protein sequencing.

1999 and Mukherjee *et al.*, 2003). SlyD exhibits a high affinity for Ni^{2+} ions based upon the presence of an extremely histidine rich C-terminal region. SlyD is just 196 amino acids in length, yet includes 15 histidine residues amongst the last 45 amino acids. It is the histidine-rich region of the protein that allows SlyD to compete with His-Tagged proteins for chelated Ni^{2+} ions present in the Ni-NTA IMAC resin. In the event that no His-Tagged proteins are present in a given sample, SlyD will bind to Ni^{2+}-charged IMAC resin preferentially.

The undesired purification of SlyD was taken as a strong indication that little or no DrrC-His_6 expression was occurring. Modification of basic cell growth and induction parameters, such as lowered temperature and inducer concentrations, failed to produce any improvement in the observed results. Potential reasons for the failure of this protein to express effectively include toxicity of the protein to host *E. coli* cells, or extreme instability of the protein which might prevent its extraction from the cell membrane. Instability within the membrane due to the absence of its heterodimeric partner DrrB is another potential reason for the difficulties associated with the purification of this protein.

4. 14 Discussion

Sequence analysis of the *M. tuberculosis drr*ABC operon suggests that these genes encode an efflux-type ABC transporter playing an essential role in mycobacterial cell wall biosynthesis and/or drug resistance. However, a deeper understanding of the structure and function of this pump requires the study of its protein components at a biochemical level. This chapter describes various attempts to over-express and purify the DrrA,B and C proteins using the heterologous host *E. coli*, the aim being to provide suitable quantities of the proteins for basic biochemical analysis. The commonly used pET vectors were chosen for cloning and expression of the *M. tuberculosis* proteins since they have been used to successfully over-express and purify proteins from a wide variety of heterologous organisms. Moreover, the pET vectors also contain features that provide for high levels of target protein expression combined with specific immuno-affinity and purification tags. Each of the three mycobacterial genes were cloned into the pET-21a vector downstream of the strong bacteriophage T7 promoter and upstream of a short C-terminal fusion sequence encoding a His-Tag. The His-Tag provides a simple and efficient means of detecting and purifying heterologously expressed proteins against the background expression of *E. coli* proteins.

Expression of DrrA-His_6, the presumed ATP-binding protein of the *M. tuberculosis* ABC transporter, was achieved at high levels in the heterologous host. The high expression levels occurred despite a large separation, in evolutionary terms, between the Gram-negative host organism *E. coli*, and source of the heterologous protein, the Gram-positive *M. tuberculosis*. This suggested that codon usage differences between the two organisms would not be a barrier to heterologous expression, and that *E. coli* would be a suitable host for expression and purification of all three DrrABC transporter components. In a number of previous studies, over-expression of both eukaryotic and prokaryotic ABC transporter NBD's in isolation, that is in the absence of their cognate MSD's, has lead to the stable cytosolic accumulation of the NBD protein (e.g. ArsA, the ATP-binding component of the *E. coli* arsenical transporter; Hsu and Rosen, 1989). A similar sub-cellular localisation might also have been expected for the heterologously expressed DrrA protein which, sequence analysis suggests, exhibits a predominantly hydrophilic character. However, upon investigation, the DrrA-His_6 fusion protein was found to accumulate almost exclusively in the form of insoluble protein aggregates, known in *E. coli* as inclusion bodies. The protein present in inclusion bodies, whilst relatively pure and easy to isolate, is believed to be incorrectly folded, and thus devoid of biological activity. In their 2001 paper reporting the crystal structure of the *Methanococcus janschii* MJ0796 ATP-binding cassette, Yuan and colleagues suggest that certain structural features of ABC transporter NBD's may make these protein particularly prone to aggregation. A key structural element present in the F_1-type ATP-binding core of the NBD, a conserved β-sheet element, is able to form non-physiological pairings with similar elements in other molecules, eventually leading to aggregation (Yuan *et al.*, 2001). Such events might be favoured in a situation where the NBD, normally a component of a stable multisubunit complex, is expressed in isolation. Under such circumstances the protein could expose to the environment certain cavities and structural elements that would normally be sequestered away from non-physiological interactions. It should be noted that sub-cellular localisation experiments identified a small quantity of DrrA-His_6 as being associated with the cell membrane, a phenomenon that was also observed for the *S. peucetius* DrrA protein. This suggests that the mycobacterial DrrA protein may possess certain surface-exposed hydrophobic regions. These regions may be involved in association of the NBD with the predominantly hydrophobic MSD's in the intact transporter complex. Attempts to solubilise and refold the DrrA-His_6 inclusion

body protein ultimately proved fruitless, as did attempts to modify the growth and expression conditions in such a way as to promote the synthesis of a soluble, active form of the protein. These experiments did however demonstrate that a C-terminal located His-Tag was a suitable affinity tag for the purification of the DrrA fusion protein by IMAC chromatography on Ni-NTA resin. Given the failure of the pET system to provide a biologically active form of heterologously expressed DrrA for biochemical analysis, an alternative route to soluble protein expression was sought. These experiments are described in subsequent chapters.

Analysis of the DrrB primary sequence prior to expression suggested that this hydrophobic protein would locate to the *E. coli* cell membrane, the normal location of ABC transporter MSD's. Low expression levels are typical of transmembrane proteins, even when their expression is driven by a strong promoter such as the bacteriophage T7 promoter of the pET system. A low expression level for the DrrB fusion protein may have been responsible for the difficulties associated with detecting its expression in crude sub-cellular fractions. Even when Western blotting was used to probe for the presence of the DrrB His-Tag in these fractions, definitive results were hampered by the large number of native *E. coli* proteins present in the crude fractions, leading to a high background in the blot results. Despite problems detecting DrrB-His_6 expression in crude fractions, the protein was eventually partially purified from detergent solubilised *E. coli* membranes by IMAC affinity chromatography. Positive identification of DrrB-His_6 was achieved by Western blotting, using an anti-His_6 monoclonal antibody to probe for the presence of the His-Tag sequence in IMAC column eluate fractions. The DrrB protein purified as two predominant species, a monomeric and dimeric form, both species being present in approximately equal quantities. It was unclear as to whether the dimeric species formed in a response to the absence of the DrrC protein. Bacterial ABC transporter MSD's are believed to function as either homo- or heterodimeric complexes, each protein contributing six transmembrane helices to an overall twelve helix substrate translocation pathway. This hypothesis was confirmed by the publication of the complete X-ray crystal structures of the BtuCD and MsbA ABC transporters (Locher *et al.*, 2002; Chang and Roth, 2001). Unfortunately, further attempts to manipulate either form of the purified DrrB-His_6 protein proved unsuccessful due to problems with protein stability. Any attempt to dialyse or concentrate the protein for further experiments caused the detergent-solubilised DrrB-His_6 to aggregate and precipitate. This problem persisted despite a

number of trials in which different detergents and buffer conditions were used. Membrane protein instability has been a major problem associated with ABC transporter studies, especially where heterologous expression has been the route to obtaining the starting material. It is of note that much of the detailed structural and biochemical data in the literature with regard to prokaryotic ABC transporters has been obtained using model ABC transporter systems homologously expressed in *E.coli* (e.g. MalFGK; Lippincott and Traxler, 1997). Even when such conditions are met, as was the case for the *E. coli* MsbA transporter, an enormous number of trials were necessary in order to identify suitable conditions for expression and purification of the protein for structural studies. The difficulty experienced in working with isolated DrrB-His_6 perhaps reflects the multisubunit nature of prokaryotic ABC transporters. Expressed in the absence of the other domains, individual subunits may become unstable and prone to aggregation, making further analysis of their properties impossible. This intrinsic instability, and the interdependence of ABC transporter subunits on one another for correct function, was shown for the *S. peucetius* DrrAB doxorubicin efflux system. Kaur demonstrated that maintenance of the DrrB protein within the host cell membrane was entirely dependent upon the presence of the DrrA ATP-binding subunit. Additionally, the *S. peucetius* DrrA protein appeared to be unable to bind ATP in the absence of DrrB, suggesting that the conformation of DrrA was linked to its association with the DrrB protein (Kaur, 1998). Whilst this does not appear to be the case for *M. tuberculosis* DrrB-His_6, as evidenced by the ability to purify it from the membrane, its stability once extracted seems to be severely impaired.

The interdependence of the DrrABC transporter subunits for stability may also account for the failure of experiments aimed at expressing a recombinant DrrC protein. No evidence was found for the expression of this protein anywhere within the *E. coli* host cell. In the absence of any His-Tagged proteins, IMAC chromatography led to the purification of a histidine rich *E. coli* protein called SlyD. This protein has been identified as a major contaminant in a number of IMAC procedures where *E. coli* has been used as the host organism for heterologous expression (e.g. see Finzi *et al.*, 2003; Mitterauer *et al.*, 1999).

The various experimental problems described here required that alternative approaches be investigated as a means of purifying sufficient, active and stable forms of the *M.*

tuberculosis DrrA, B and C proteins. These methodologies included attempts to co-express all three components of the transporter in order to improve expression levels and stability of the individual proteins. The high expression levels associated with DrrA-His_6 were viewed as a positive sign that an alternative expression system may be of use in producing significant quantities of this NBD for biochemical analysis.

Chapter 5

Soluble expression of the *M. tuberculosis* DrrA protein by construction of a N-terminal fusion to *E. coli* thioredoxin

The inefficiency of the pET-based expression system with regard to generating soluble, recombinant DrrA necessitated a search for an alternative expression system. An *in vitro* expression system was considered as a potential means of producing DrrA for further studies, but was eventually rejected because of the high costs and relatively low yields associated with the technique. A system based upon the use of a more closely related Gram-positive expression host, *Streptomyces lividans,* was also investigated. Unfortunately a number of technical difficulties associated with the growth, handling and genetic manipulation of these organisms, combined with a lack of expertise and experience within the laboratory, led to the curtailment of these experiments. Instead, a series of *E. coli* expression vectors possessing features believed to promote protein solubility were selected for further investigation.

Transcriptional fusions of heterologous proteins to native *E. coli* proteins have been used as a means of improving both the expression level and solubility of the heterologous protein. Commonly used fusion partners include glutathione S-transferase (GST), maltose binding protein (MBP) and thioredoxin. All of these proteins are non-toxic when expressed at elevated levels in *E. coli* and provide the additional benefit of being highly soluble. Moreover, certain fusion partners may act as an affinity tag to facilitate purification of the protein by chromatographic means (e.g. purification of MBP fusions by amylose affinity chromatography). Many proteins have been expressed and purified in this way but a particularly pertinent example is the OleB protein of *Streptomyces antibioticus*. This protein, the ATP-binding domain of the oleandomycin efflux ABC transporter, was successfully characterised when expressed as a MBP-fusion (Aparicio *et al.*, 1996). A similar MBP-fusion was used to produce a soluble form of the first NBD of the human MDR1 gene product (Wang *et al.*, 1999).

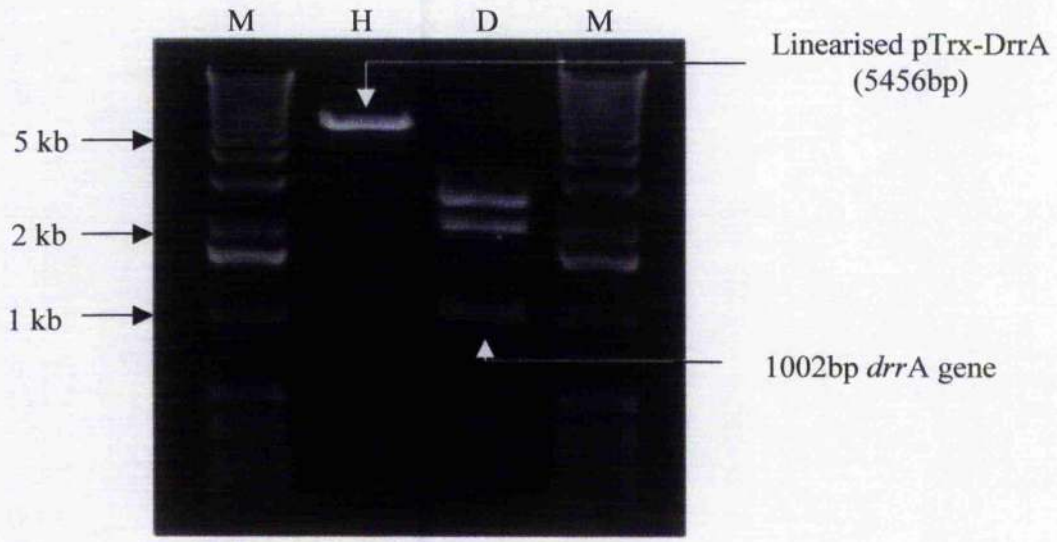
Each of the various fusion protein vectors available differ not only in the size and nature of the fusion partner, but also in the promoters used to drive expression and the presence of additional affinity tags. Of all of the systems investigated, the "pBAD-TOPO ThioFusion" vector (Invitrogen) was considered to offer the largest number of positive benefits. *E. coli*

thioredoxin is able to accumulate to approximately 40% of the total cell protein whilst remaining soluble, improves the efficiency of translation when present at the N-terminal, and is also relatively small in size compared with DrrA (11.7 kDa). An additional benefit of using thioredoxin is that the protein has no known nucleotide/drug/lipid binding functions that may have interfered with subsequent biochemical assays of the fusion protein. A relatively small fusion partner for DrrA was preferred because of the proximity of the Walker A nucleotide-binding residues to the N-terminal of the protein. A large fusion partner, such as MBP(42 kDa), may have had the potential to interfere with correct folding of the DrrA moiety, or perhaps have caused steric effects with regard to DrrA's presumed nucleotide-binding activity. The use of the pBAD-ThioFusion system also presented the opportunity to test a different promoter to drive DrrA expression (the *ara*BAD promoter) whilst maintaining a C-terminal His-Tag sequence for effective purification of the fusion protein.

5. 1 The pTrx-DrrA expression construct

The pBAD/TOPO ThioFusion vector was obtained from a commercial source (Invitrogen) and used to generate a recombinant expression construct for DrrA according to the manufacturer's instructions (see chapter 2). The key features of this construct, pTrx-DrrA, are shown in figure 5.1. The vector contains a 1455bp open reading frame encoding a protein of 485 amino acids with a predicted molecular weight of 52.5 kDa. The first 110 amino acids of the fusion protein constitute the *E. coli* thioredoxin moiety. There then follows a short 13 amino acid 'linker sequence' before the beginning of the *M. tuberculosis* DrrA moiety (334 amino acids). The remaining 28 amino acids of the protein are a C-terminal extension to DrrA containing the His-Tag sequence. One of the major differences between the new construct and those generated previously was that expression of the heterologous protein was under the control of a promoter native to *E. coli*, the *ara*BAD promoter. This promoter has the advantage that it is regulated by a non-toxic inducer, arabinose, and that levels of transcription from the promoter are tightly linked to inducer concentration (Guzman *et al.*, 1995). It was hoped that by using this promoter the levels of fusion protein production could be more finely tuned than had been possible with the pET-21a construct. Under such control it may be possible to identify those conditions in which soluble protein expression is enhanced and inclusion body formation reduced or prevented altogether.

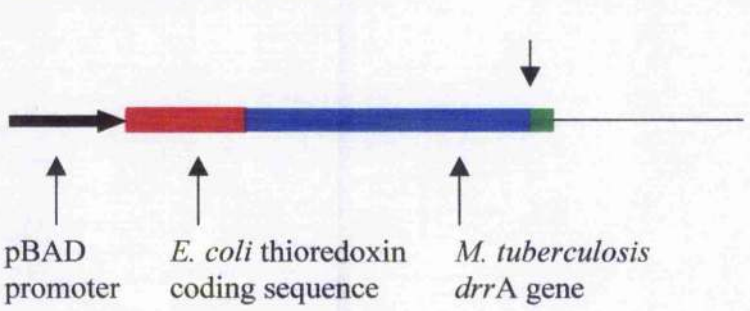
Figure 5. 1 **The pTrx-DrrA expression construct**



Restriction enzyme digestion of the pTrx-DrrA expression construct:

The pTrx-DrrA expression plasmid totals 5456bp in size (4454bp vector, 1002bp *drr*A insert). The vector contains only a single *Hin*dIII restriction site which was introduced by PCR to the 3' end of the *drr*A amplicon. Digestion of the plasmid with this enzyme (Lane H above) results in linearization of the plasmid. The plasmid contains two *Nde*I restriction enzyme sites, one at the 5' end of the *drr*A insert, the second within the vector itself. A double digestion of the plasmid with both *Nde*I and *Hin*dIII (Lane D above) results in the generation of three DNA fragments, two of which are derived from the vector, the third being the 1002bp *drr*A amplicon. The digests above were performed in a total volume of 20µl using 10U of each enzyme. Digests were incubated at 37 °C for 1hour.

Schematic representation of pTrx-DrrA

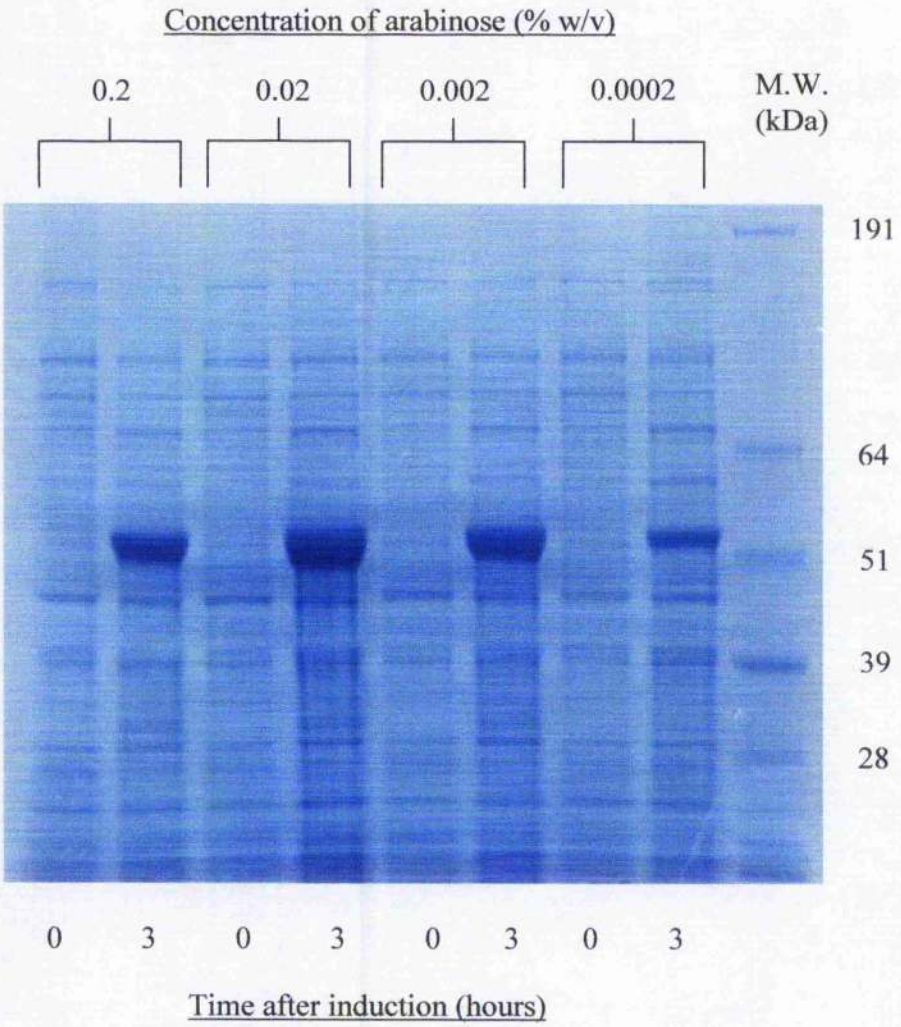


5. 2 Expression trials using pTrx-DrrA

The newly generated expression construct was transformed into the manufacturer's recommended host strain *E. coli* TOP10 (see table 2.2, chapter 2 for genotype). Transformed colonies were selected on the basis of vector-encoded ampicillin resistance, and isolated plasmids screened for the presence of the inserted DrrA sequence by restriction digestion (see figure 5.1). A single recombinant colony was used to initiate expression trials and to determine those conditions leading to maximal expression of the Trx-DrrA fusion protein. The recombinant colony selected was used to prepare a 5ml liquid starter culture which was grown overnight at 37 °C in LB medium with vigorous shaking. The next day 0.1ml samples of the starter culture were used to inoculate four separate 5ml cultures, each of which was grown to an OD_{600} of 0.6 before induction with arabinose. Each culture was induced with a different concentration of arabinose ranging between 0.2% w/v and 0.0002% w/v. The 1000-fold difference in inducer concentration was designed to identify optimal induction conditions. After 3 hours of growth samples were withdrawn from each culture and changes in protein expression were analysed by SDS-PAGE and by comparison with uninduced samples. The results of this analysis are presented in figure 5.2.

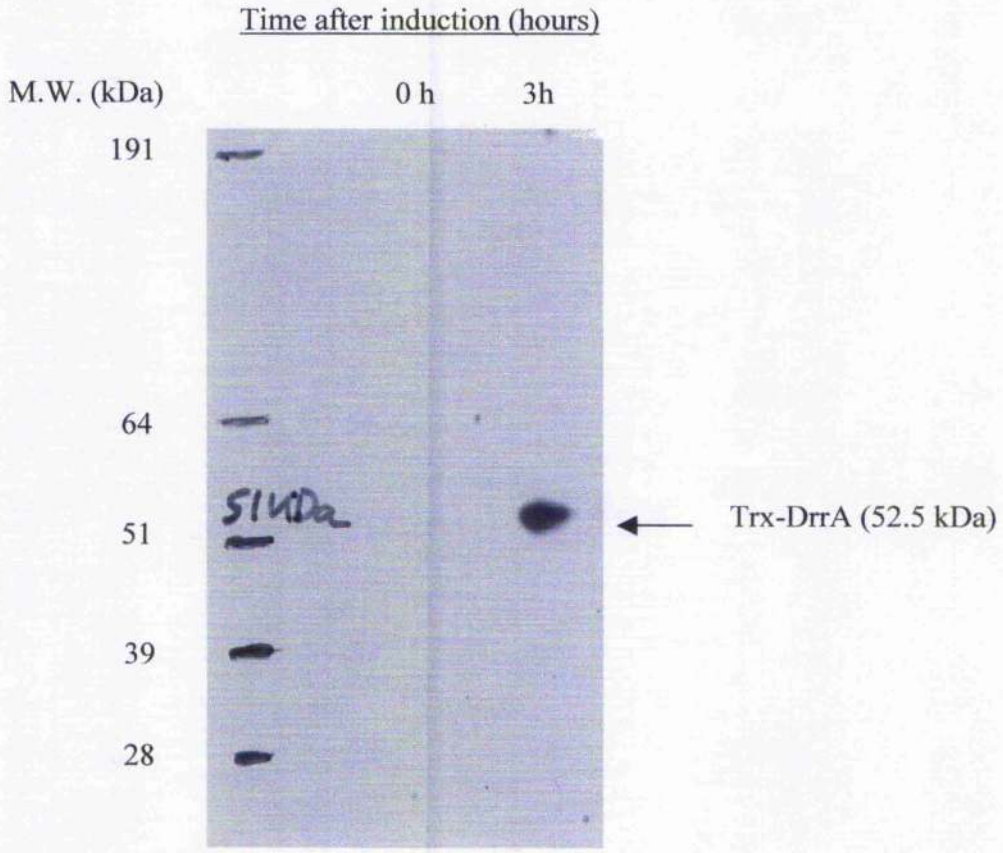
The expression trials demonstrated the accumulation of a protein with a size consistent with that predicted for Trx-DrrA. The same protein was expressed in all four trial cultures and was apparently absent prior to induction with arabinose. The concentration of arabinose added to each culture had a moderate effect, the greatest level of expression being observed in cultures induced with either 0.02% or 0.002% arabinose. In all future experiments using the pTrx-DrrA expression construct 0.02% w/v arabinose was used to induce expression. A Western blot against the C-terminal His-Tag was used to confirm the identity of the 52 kDa protein as Trx-DrrA. Samples of total cell protein taken before and after induction with 0.02% arabinose were separated by SDS-PAGE and transferred to a PVDF membrane before being probed with an anti-His_6 monoclonal antibody. The results of the blot, shown in figure 5.3, clearly show the positive reaction of the 52 kDa protein, and that it is only present in the culture *after* induction. This result was interpreted as a positive indication that the expressed 52 kDa protein was indeed recombinant Trx-DrrA.

Figure 5. 2 **Expression trials using the pTrx-DrrA fusion protein construct**



A Coomassie blue stained SDS-PAGE gel showing induction of Trx-DrrA expression. Each lane shows total cell protein isolated from *E. coli* TOP10/pTrx-DrrA prior to induction, or 3 hours post induction. Amounts of protein loaded in each lane were nomalised by OD_{600} at the time the samples were taken. The concentration of inducer (arabinose) added to each culture varied over a 1000-fold range. A protein of approximately the predicted size for Trx-DrrA (52 kDa) can be seen to accumulate in all four cultures 3 hours after induction. Maximal levels of Trx-DrrA synthesis were observed in cultures induced with 0.02% and 0.002% arabinose.

Figure 5. 3 **Western blot to show expression of Trx-DrrA**



Anti-His_6 Western blot to show arabinose induced expression of the 52.5 kDa Trx-DrrA fusion protein. The blot was performed on total cell protein isolated from *E. coli* TOP10/pTrx-DrrA cells immediately prior to induction with 0.02% w/v arabinose (Lane 0), and after growth at 37 °C for three hours (Lane 3). Cells were grown in LB medium supplemented with 100μg/ml carbenicillin and induced at OD_{600} = 0.6.

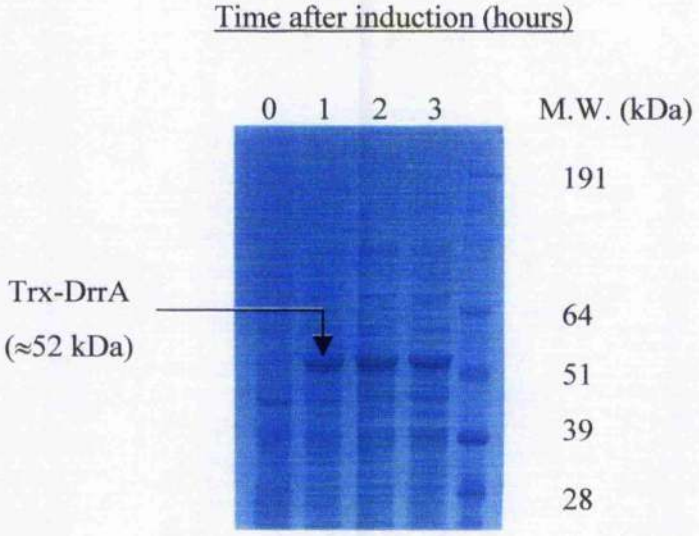
5. 3 Sub-cellular localisation of Trx-DrrA

Sub-cellular fractionation experiments, in line with those described in chapter 4, were used in order to determine the physical location of Trx-DrrA within the cell. A 1 litre culture of TOP10/pTrx-DrrA was grown in LB medium at 37 °C with shaking. Trx-DrrA expression was initiated with 0.02% arabinose when the culture reached an OD_{600} of 0.6. 1ml samples were removed from the culture prior to induction, and at hourly intervals thereafter, in order to monitor expression of Trx-DrrA over the course of the experiment. After 3 hours of growth the cells were harvested by centrifugation and resuspended in a buffer of 50mM Tris-HCl, 300mM NaCl, 10% glycerol w/v, and lysed by mechanical disruption. Cytosolic, membrane and insoluble fractions were prepared as described previously and solubilised in SDS-PAGE sample buffer for analysis. The results of the expression time-course and fractionation experiments are shown in figures 5.4a and 5.4b.

The time-course experiment demonstrated that maximal levels of Trx-DrrA were achieved after just one hour (figure 5.4a). However, as the culture continued to grow over the entire length of the experiment, monitored by an increase in OD_{600}, it was assumed that Trx-DrrA expression was non-toxic. As such, the overall yield of Trx-DrrA was believed to increase throughout the entire growth period. Fractionation experiments (figure 5.4b) revealed that a significant quantity of the Trx-DrrA synthesised was once again present in the insoluble fraction. This result was surprising given the known solubility of *E. coli* thioredoxin, but was perhaps a consequence of the high levels of expression seen within the first hour of growth after induction. A small quantity of Trx-DrrA was found to be present in the cytosolic fraction, an indication that the N-terminal fusion of thioredoxin to DrrA was having a positive effect on DrrA solubility, even when the cells were grown at 37 °C.

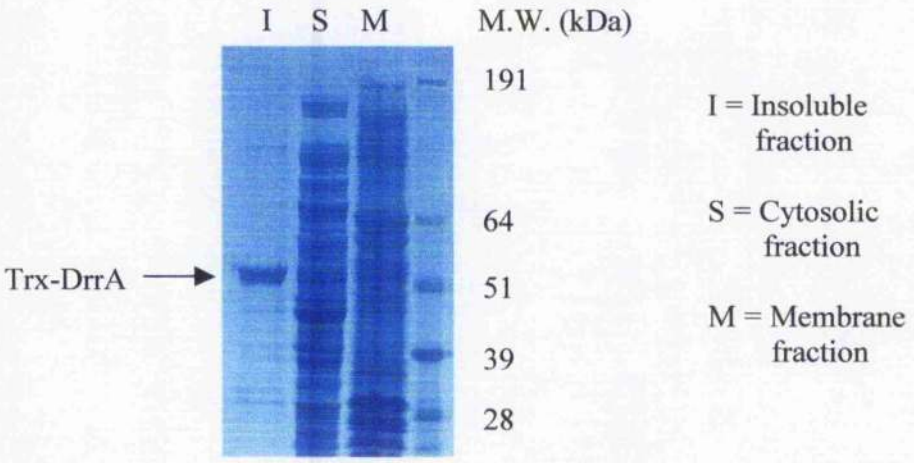
Following identification of soluble Trx-DrrA a number of optimisation experiments were performed in order to improve the yield of soluble protein. These experiments primarily involved induction of Trx-DrrA expression at much lower temperatures and using much longer periods of growth. These measures were designed to slow down the rate of Trx-DrrA synthesis after induction, and to allow the fusion protein to fold into a native conformation. At the same time the growth medium was changed from LB to the richer 2x YT medium, promoting cultures to grow to greater cell densities.

Figure 5. 4a **Time-course of Trx-DrrA expression**



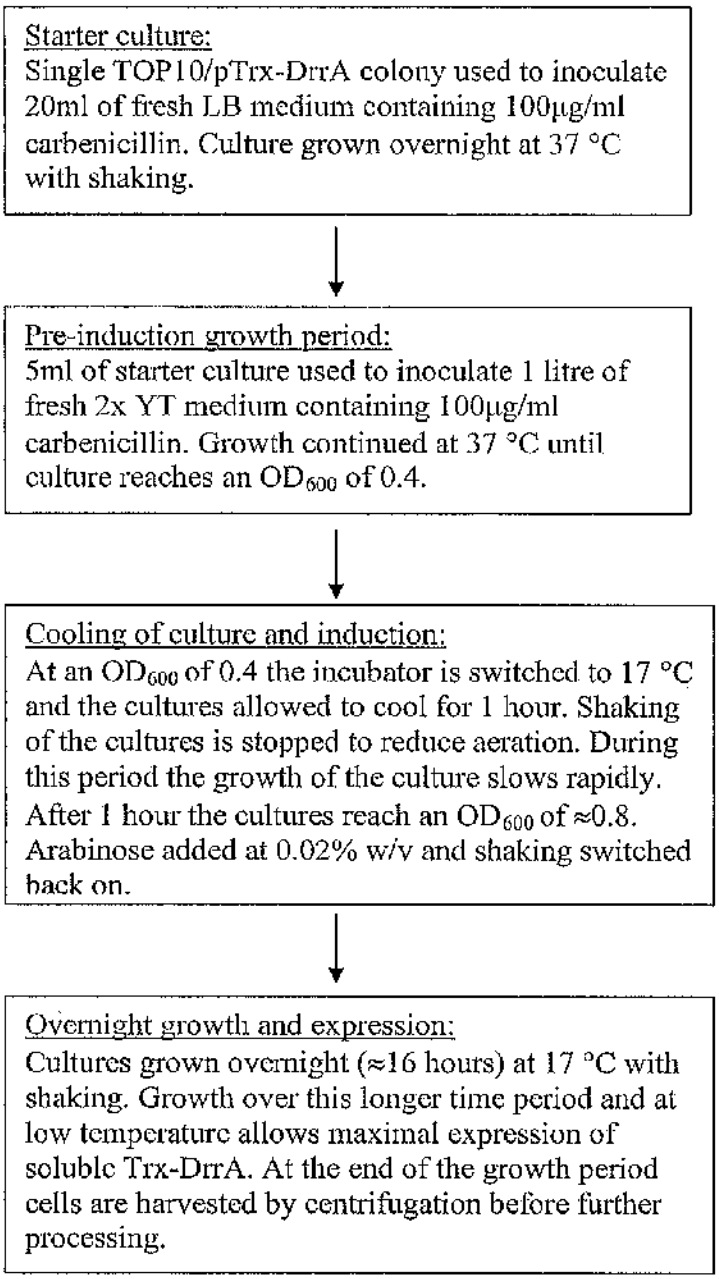
Total cell protein isolated from TOP10/pTrx-DrrA before and after induction with 0.02% w/v arabinose. Trx-DrrA expression is fully induced after 1 hour.

Figure 5. 4b **Sub-cellular fractionation of cells expressing Trx-DrrA**



Sub-cellular fractions isolated from cells expressing Trx-DrrA at 37 °C. The majority of Trx-DrrA is found in the insoluble fraction (lane I), whilst a small amount is visible in the soluble or cytosolic fraction (lane S).

Optimal expression of soluble Trx-DrrA was found to occur when using the following growth and induction scheme: -



Following harvesting the cells were resuspended in a buffer of 50mM Tris-HCl (pH 8.0), 300mM NaCl, 10 % glycerol (v/v). Cells were resuspended using 40ml of buffer per litre of cell culture. At this point cells were either stored at -80 °C for future use, or used immediately for the purification of soluble Trx-DrrA (see next section).

5. 4 Initial purification of soluble Trx-DrrA by IMAC chromatography

The presence of the C-terminal His-Tag on Trx-DrrA had previously been demonstrated by Western blot (figure 5.3). This feature of Trx-DrrA was then used in attempts to isolate and purify the fusion protein by IMAC chromatography. Resuspended cells were lysed by mechanical disruption and insoluble material removed from the lysate by centrifugation (16,000 rpm, Beckman JA-21 rotor). Cells and lysates were kept on ice at all times in order to reduce proteolysis. EDTA-free protease inhibitors were added to the cleared lysate in order to further reduce proteolytic damage to the fusion protein. 1ml of Ni-NTA IMAC resin was added to the lysate along with imidazole at a final concentration of 10mM. The addition of imidazole at this stage was designed to reduce non-specific binding of *E. coli* proteins to the IMAC resin. Trx-DrrA was allowed to bind to the resin for a period of 2 hours at 4 °C whilst the mixture was gently agitated on a blood wheel. At the end of the binding period the mixture was loaded into a 1.8cm diameter chromatography column and the resin allowed to settle under gravity flow.

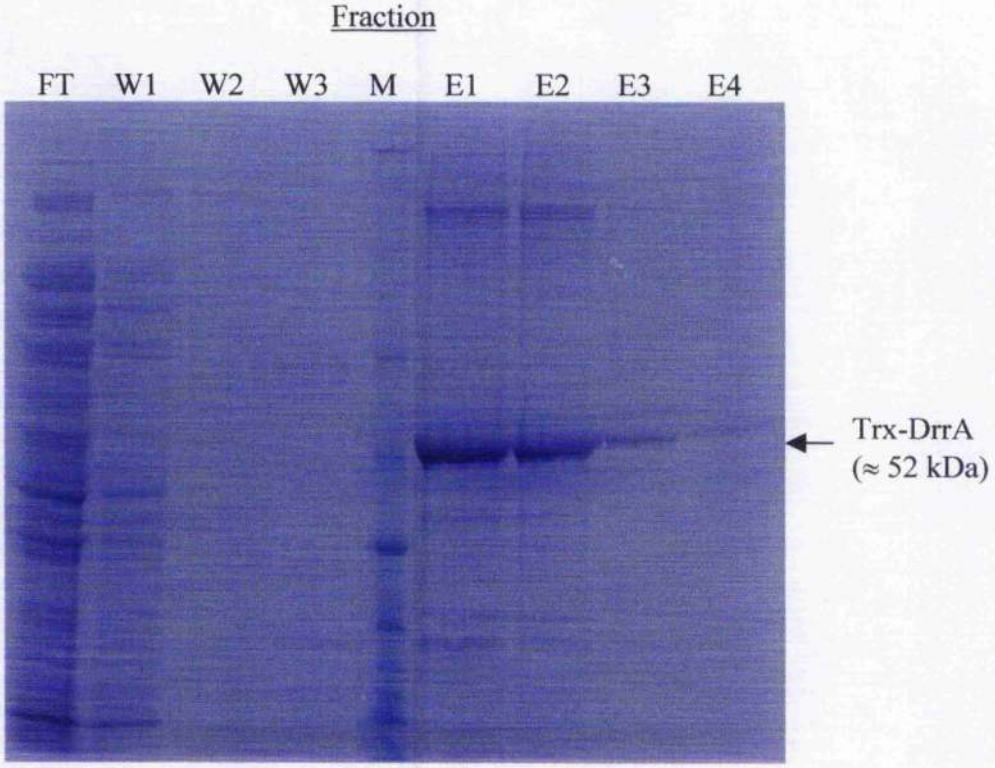
The IMAC resin was then washed with a series of buffers containing increasing concentrations of imidazole in order to remove non-specifically bound proteins, followed by elution of Trx-DrrA using a buffer containing a high concentration of imidazole. The washing and elution conditions used to prepare Trx-DrrA are shown below:-

Buffer A (same as resuspension buffer):-
50 mM Tris-HCl (pH 8.0)
300mM NaCl
10% glycerol (v/v)

Wash I	:	50ml Buffer A
Wash II	:	50ml Buffer A + 25mM imidazole
Wash III	:	25ml Buffer A + 50mM imidazole
Elution	:	5 x 1ml Buffer A + 300mM imidazole

Protein samples were collected throughout the purification procedure for analysis by SDS-PAGE. Results obtained using the purification protocol detailed above are shown in figure 5.5.

Figure 5. 5 **Initial purification of soluble Trx-DrrA**



SDS-PAGE analysis of protein fractions obtained during initial purification trials of soluble Trx-DrrA. Column chromatography was performed according to the protocol described in section 5.4. Fractions: FT = column flow-through, W1-W3 = washes to remove non-specifically bound protein, E1-E4 = Elutions of the Trx-DrrA fusion protein.

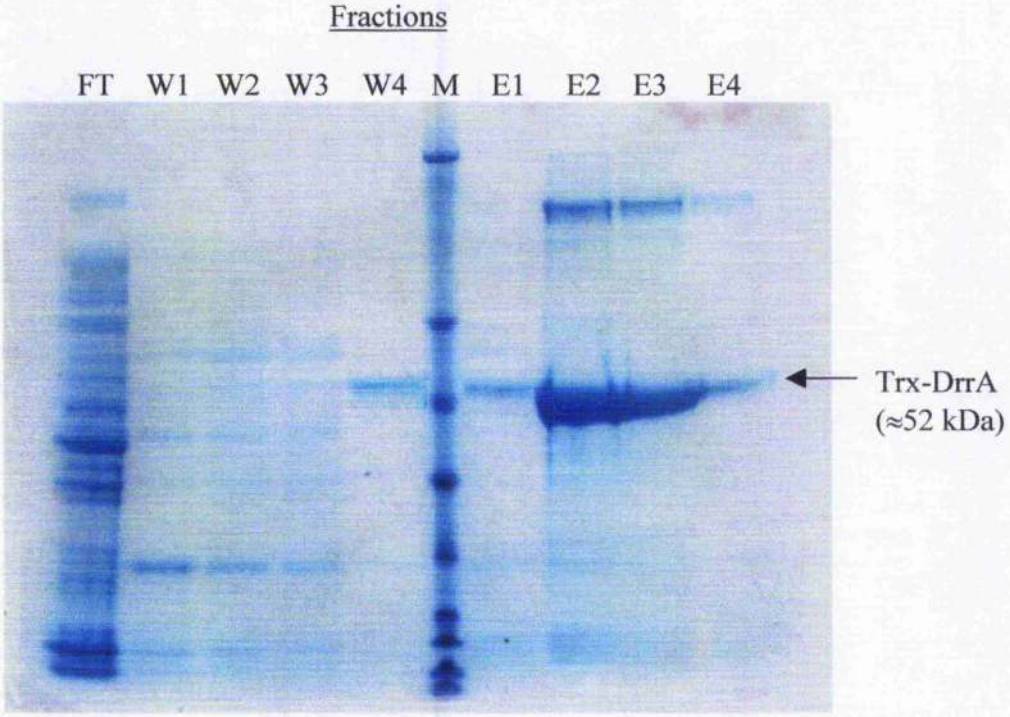
Significant quantities of Trx-DrrA are released from the IMAC resin using the 300mM imidazole elution buffer (lanes E1-E4). However, despite extensive washing of the IMAC resin during the early stages of the procedure (lanes W1-W3), a large number of contaminating proteins are seen to co-elute with Trx-DrrA. The contamination is particularly evident in the first two elution fractions where the majority of Trx-DrrA is found.

Figure 5.5 illustrates that the low temperature growth and induction protocol employed was sufficient to generate acceptable quantities of soluble Trx-DrrA for further experiments. However, both biochemical studies and crystallisation trials require the availability of protein samples having purity levels in the region of 90-95%. The IMAC purification of Trx-DrrA shown in figure 5.5 yielded protein with a purity of little more than 60%, well below the level required for any meaningful analysis of biochemical activity.

A number of variations on the basic gravity-fed column chromatography procedure were tested as a means of enhancing Trx-DrrA purity. These variations included processing of increased quantities of cell lysate in a single purification experiment so as to saturate the IMAC resin with specifically bound Trx-DrrA, thereby reducing the availability of binding sites for contaminating proteins. Alternative host strains for the pTrx-DrrA expression plasmid were also tested in order to identify a strain capable of generating higher yields per litre of culture. Whilst some minor improvements in yield were observed when *E. coli* LMG194 was used as the expression host, this variable had little effect on the purity of Trx-DrrA. Since the major problem with the procedure lay in the co-elution of contaminating proteins, a number of variations in the column washing protocols were adopted. These included washing the IMAC resin with both increased volumes of wash buffer and the use of more stringent buffers (containing higher concentrations of imidazole). Figure 5.6 illustrates how all of these factors combined led to significant increases in both the purity and yield of Trx-DrrA obtained by this methodology. Despite the observed improvements, figure 5.6 also illustrates the limitations of this particular form of IMAC. Lane W4, a sample obtained whilst washing the IMAC resin with a buffer containing 75mM imidazole, contains a low concentration of Trx-DrrA and few if any contaminants. This suggests that an imidazole concentration of 75mM was sufficient to cause the leaching of Trx-DrrA from the resin, whilst not completely removing contaminating proteins (see lanes E2 and E3).

The experiments described above served as an important 'proof-of-principle' for the purification of soluble Trx-DrrA by IMAC, yet had failed to produce protein of sufficient purity for further analysis. Given the problems encountered, an alternative purification procedure was adopted. This modified procedure, described in the next section, eventually led to the generation of highly purified Trx-DrrA.

Figure 5. 6 **Improved yields of Trx-DrrA generated by a modification of the basic IMAC procedure**



Coomassie blue stained SDS-PAGE gel showing purification of Trx-DrrA by gravity flow IMAC chromatography. A number of modifications to the basic protocol were used in attempts to increase the purity and yield of Trx-DrrA. The above analysis shows protein purified from a 6 litre culture of LMG194/pTrx-DrrA cells grown and induced at 17 °C.

FT = Flow-through, W1-W4 = Wash fractions (see below), M = Molecular size markers, E1-E4 = Elution fractions

Basic buffer : 50mM Tris-HCl (pH 8.0), 300mM NaCl, 10% glycerol (v/v)

W1 = Buffer + 25mM Imidazole (20ml)

W2 = Buffer + 40mM Imidazole (20ml)

W3 = Buffer + 50mM Imidazole (20ml)

W4 = Buffer + 75mM Imidazole (20ml)

E1-E4 = Buffer + 300mM Imidazole (4 x 1.5ml)

5. 5 Effective purification of soluble Trx-DrrA using an enhanced chromatography procedure

Up to this point all attempts to purify soluble Trx-DrrA had been performed using Qiagen's 'Ni-NTA-agarose' IMAC resin. The purification procedure itself had involved two distinct phases, an initial bulk binding phase followed by gravity flow column chromatography. In the following series of experiments an alternative procedure was adopted. Purification was performed using a pre-packed IMAC matrix purchased from a commercial source, and all binding, washing and elution phases were performed under positive pressure provided by a peristaltic pump.

1ml 'HiTrap Chelating HP' columns were purchased from Amersham Pharmacia Biotech and were loaded with Ni^{2+} ions according to the manufacturer's instructions (see chapter 2). A single 1ml column, loaded with Ni^{2+}, served as the chromatography matrix in all of the following experiments. The major difference between this matrix and the Ni-NTA matrix used previously is the nature of the ligand group to which Ni^{2+} ions are bound. In the case of the HiTrap columns Ni^{2+} ions are bound to a Sepharose matrix through an *iminodiacetic acid* (IDA) ligand, rather than being bound to agarose beads via a *nitrilotriacetic acid* (NTA) ligand. The experiments described below show how this relatively small difference led to vastly improved yields and purity levels during the preparation of Trx-DrrA.

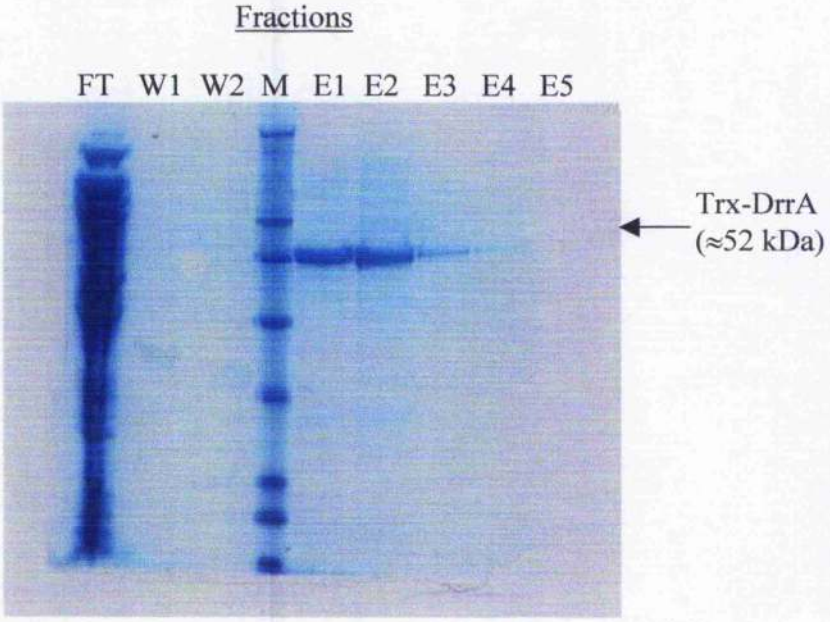
A trial experiment using the Ni-IDA matrix was performed according to the following protocol. 1 litre culture of *E. coli* LMG194/pTrx-DrrA was grown using the low temperature growth and induction procedure described in section 5.3. These cells were resuspended in 50ml of an ice-cold buffer containing 50mM Tris-HCl (pH 8.0), 400mM NaCl, 10% glycerol and EDTA-free protease inhibitors. The cells were disrupted and a cleared lysate containing soluble Trx-DrrA was prepared by centrifugation. The elevated NaCl concentration used in all buffers was an attempt to reduce non-specific interactions between the column matrix and proteins present in the cleared lysate. Trx-DrrA was immobilised by pumping the cleared lysate (approx. 45ml) through a pre-equilibrated IMAC column at a constant flow-rate of 0.8ml/minute. This and all subsequent stages of the procedure were performed at 4 °C in a cold-room. After the completion of the binding phase unbound protein was removed from the column matrix by passing 25ml of

resuspension buffer through the column at the same flow-rate. Non-specifically bound contaminants were removed from the resin by washing with 75ml of the same buffer supplemented with 60mM imidazole. The final stage of the procedure involved elution of the bound Trx-DrrA using buffer supplemented with 250mM imidazole. 10ml of this elution buffer was pumped through the column at a flow rate of 0.8ml/minute and 1.5 ml fractions collected. The results of the experiment described above were analysed by SDS-PAGE as shown in figure 5.7.

The use of a pre-packed Ni-IDA column and a peristaltic pump clearly resulted in a vastly improved preparation of soluble Trx-DrrA. Little background contamination was evident in the elution fractions and the protein was estimated to be approximately 90% pure, ideal for subsequent biochemical assays. The experiment illustrated in figure 5.7 was performed using a cleared lysate prepared from just 1 litre of cell culture.

Following the success of the initial purification trials, larger scale purification experiments were initiated in order to provide the necessary quantities of Trx-DrrA required for subsequent studies. The purification protocols developed previously were simply applied to cleared lysates obtained from larger cell cultures. Figure 5.8 shows the large scale purification of Trx-DrrA from a 3 litre culture of LMG194/pTrx-DrrA cells. Exactly the same IMAC binding, washing and elution steps as had been performed in the smaller scale experiments were used here also. Figure 5.8 shows that the purification methodology developed was robust, reproducible and able to yield Trx-DrrA of the required purity (greater than 95%) for further studies. In addition, the protein purified in this manner proved to be stable at 4 °C for up to a week, and could be dialysed without risk of precipitation.

Figure 5. 7 **Enhanced purification of Trx-DrrA using a Ni-IDA chromatography matrix**

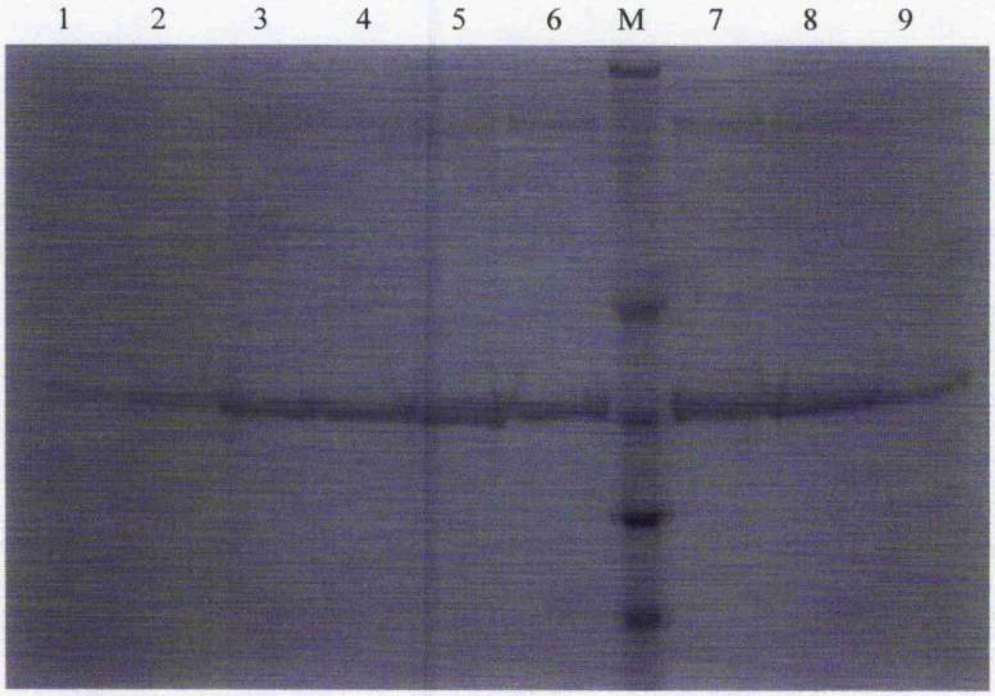


A Coomassie blue stained SDS-PAGE gel showing IMAC purification of soluble Trx-DrrA. Purified protein was isolated from a 1 litre culture of LMG194/pTrx-DrrA grown and induced as described in section 5.3. Column chromatography was performed using a pre-packed 1ml Ni-IDA-Sepharose column loaded and developed at a flow rate of 0.8ml/minute. Trx-DrrA prepared by this method was estimated to be approximately 90% pure. This represented a significant improvement in the quality of the Trx-DrrA protein preparation when compared with previous gravity-flow IMAC experiments.

Column Buffer : 50mM Tris-HCl (pH 8.0), 400mM NaCl, 10% glycerol (v/v)

W1 = Buffer + 0mM Imidazole (25ml)
W2 = Buffer + 60mM Imidazole (75ml)
E1-E5 = Buffer + 300mM Imidazole (5 x 1.5ml)

Figure 5. 8 **Large-scale purification of soluble Trx-DrrA**



A Coomassie blue stained SDS-PAGE gel showing purified soluble Trx-DrrA. The above protein was purified from 3 litres of LMG194/pTrx-DrrA cells grown and induced at 17 °C. Cleared lysate from these cells was applied to a 1ml HiTrap Chelating IMAC column at a constant flow rate of 0.8ml/minute using a peristaltic pump. Wash and elution buffers (see below) were applied to the column in the same fashion. Trx-DrrA purified by this technique was estimated to be greater than 95% pure, suitable for further biochemical analysis. Protein samples were stored at -80°C.

Purification protocol: -

Column Buffer	50mM Tris-HCl (pH 8.0), 400mM NaCl, 10% glycerol (v/v)
Wash 1	Buffer + 0mM Imidazole (25ml)
Wash 2	Buffer + 60mM Imidazole (75ml)
Elutions	Buffer + 300mM Imidazole (15ml, collected in 1.5ml fractions)

5. 6 Discussion

Recombinant protein over-expression and purification is a complex process, the success of which is influenced by a wide range of variables. Factors such as the source of the recombinant protein, the expression host, and the nature of the expression system employed may all have a significant role in determining the outcome of experiments. All previous attempts to express a soluble form of recombinant *M. tuberculosis* DrrA using the pET-based expression system had failed due to the formation of inclusion bodies. However, as the data in this chapter demonstrates, such problems can be overcome by careful modulation of variables known to affect protein expression and solubility.

The eventual production of soluble *M. tuberculosis* DrrA was achieved through introduction of multiple, sequential changes to expression and purification protocols. The exact degree to which each variable influenced the eventual outcome is difficult to assess. However, modifications to the expression system, growth conditions, bacterial strains and chromatography techniques were all shown to contribute to the improved results. Perhaps the greatest improvements in DrrA expression were derived from the use of an alternative expression vector. The pBAD-ThioFusion system not only provided high levels of expression (as the pET vector had done), but also had a significant influence on the sub-cellular localisation of the protein. Whereas cells carrying the pDrrA plasmid had expressed the protein in the form of insoluble aggregates, pTrx-DrrA led to accumulation of the fusion protein in the *E. coli* cytosol. The use of fusion proteins is becoming an increasingly common strategy when attempting to improve the solubility of ABC transporter NBD's. The technique has been applied to proteins from both eukaryotic and prokaryotic sources, most commonly employing MBP as the fusion partner (e.g. see Gartner *et al.*, 2003; Honisch and Zumft, 2003). The data presented here suggests that *E. coli* thioredoxin may be an equally effective fusion partner for some bacterial NBD's. The proximity of the structurally important Walker A nucleotide-binding motif to the N-terminal of DrrA was a major factor in the choice of thioredoxin as the fusion partner. It was considered possible that the presence of a high molecular weight protein, such as MBP, near to the nucleotide-binding fold might influence the biochemical properties of the ATPase. This may be even more of a consideration in the light of structural data suggesting that important functional interactions between ABC transporter NBD's are mediated by juxtaposition of two nucleotide-binding folds (Chang,

G., 2003). Determining whether or not this is a valid concern will require the availability of high resolution structural data for a NBD fusion protein.

The quantity of soluble Trx-DrrA produced was strongly influenced by the growth temperature of expression cultures. Rapid growth at 37 °C was shown to promote the formation of insoluble aggregates similar to those produced by the pET system, whilst growth at 17 °C generated primarily soluble, cytosolic protein. Clearly the speed at which heterologous proteins are synthesised in *E. coli* can have a major effect on the efficiency of protein folding, especially when so many of the cell's resources are being used to this end.

Having established conditions suitable for the expression of soluble Trx-DrrA, further analysis required development of an efficient purification protocol. Previous experiments using C-terminal His-Tagged DrrA had shown that IMAC was an acceptable means of purification. However, further iterative alterations to the basic IMAC protocol were required to identify optimal purification parameters. Initial experiments using a combination of batch-binding and gravity flow column chromatography were only partially successful in the purification of Trx-DrrA. Whilst acceptable levels of Trx-DrrA were recovered using this technique, the purity of the preparations was deemed to be insufficient for subsequent biochemical analysis. By switching to an alternative metal affinity resin and using larger quantities of starting material the problem of protein impurity was virtually eliminated (see figure 5.8). Once again a relatively small change in experimental methodology had produced a significant improvement in results.

The expression and purification protocols described here were used to produce Trx-DrrA for a range of biochemical experiments described in the next chapter. Yields of the highly purified protein were estimated to be 2-3mg/litre of culture, and the protein could be stored at -80 °C for extended periods of time without significant loss of function or aggregation.

Chapter 6

Biochemical analysis of the Trx-DrrA fusion protein

Sequence analysis of the *M. tuberculosis drr*ABC operon (presented in chapter 3) strongly suggests that DrrA constitutes the NBD of an ABC transporter complex. Having established a suitable means of over-expressing and purifying the protein using *E. coli* as the host organism, the next stage of analysis was to examine the biochemical properties of the protein.

A significant body of experimental data has accrued regarding the biochemical properties of ABC transporter NBD's. Much of this data relates to the medically important eukaryotic ABC transporters, CFTR and P-glycoprotein, and to model bacterial systems such as the nutrient uptake transporters. Despite this, a definitive picture of NBD behaviour has yet to emerge. Whilst certain fundamental properties of these proteins remain constant, for example their ability to bind and hydrolyse ATP, other properties show a marked degree of variation. Variables identified include the sensitivity of ATPase activity to a range of inhibitors, whether or not activity is stimulated by transport substrates, and the proposed multimeric state of these proteins. Differences in the experimental systems used by different groups to over-express, purify and assay the activity of NBD's no doubt contribute to this variability.

The data presented in this chapter show the biochemical activity of the Trx-DrrA fusion protein to be broadly compatible with a functional role as the NBD of a mycobacterial ABC transporter. A partial characterisation of the protein was achieved through the use of traditional biochemical assays to determine features of the protein such as its enzymatic activity and substrate preference. Efforts to describe more detailed features of Trx-DrrA behaviour using sophisticated techniques such as stopped-flow fluorescence spectroscopy were essentially unsuccessful. Possible reasons for the failure of these experiments are described at the end of the chapter.

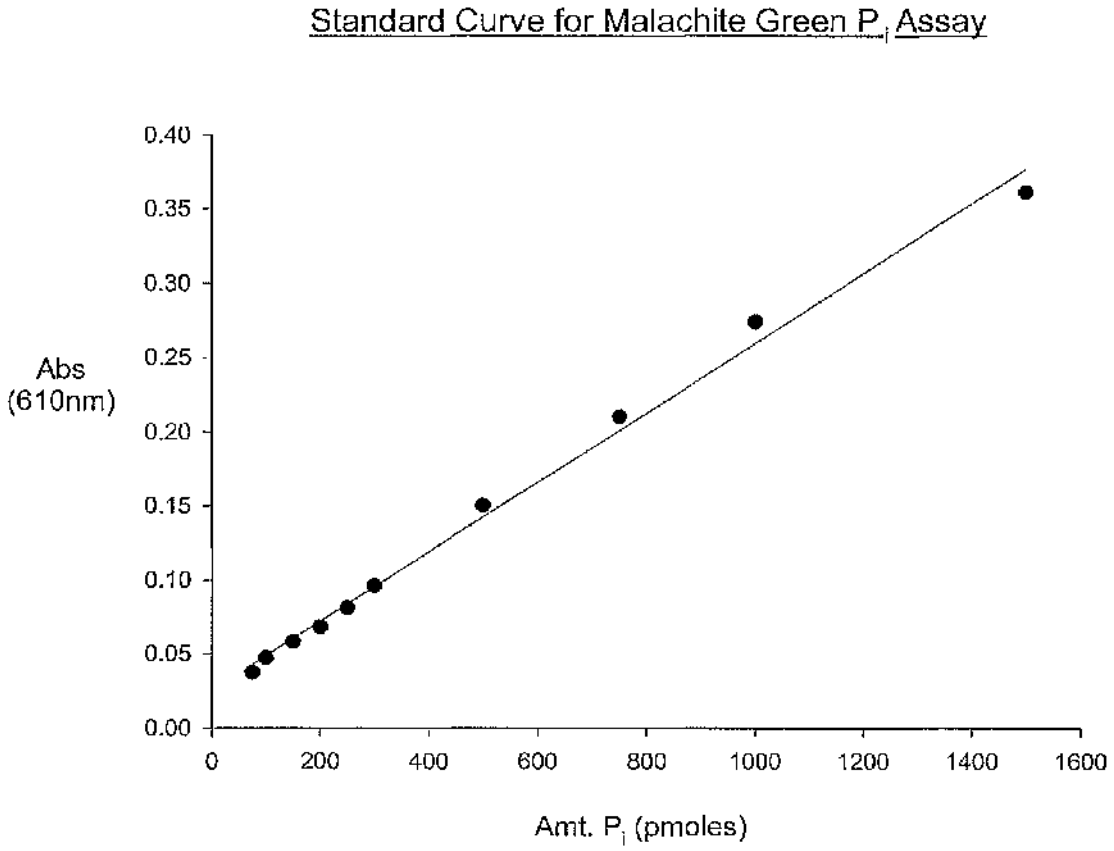
6.1 Trx-DrrA exhibits cation dependent ATPase activity

The ability to bind and hydrolyse ATP is a fundamental property of all ABC transporter NBD proteins identified to date; the phosphate bond energy released being used to drive the substrate translocation process. Many ABC transporter NBD's have been shown to retain their ATP-binding/hydrolysing properties even when expressed and assayed in isolation, i.e. in the absence of associated membrane-spanning components of the transporter complex (e.g. see Nikaido *et al.*, 1997). The high level of sequence homology between *M. tuberculosis* DrrA and other NBD proteins suggested that this protein should demonstrate a similar ATPase activity. In fact, establishing such activity was considered to be a vital first step in demonstrating not only the efficacy of the expression and purification protocol developed, but also in confirming the hypothesis that DrrA does indeed constitute the NBD of a novel mycobacterial ABC transporter.

The ATPase activity of recombinant Trx-DrrA was initially assayed using a modification of the protocol developed by Harder and colleagues. The technique relies upon the use of a chromogenic reagent in order to detect inorganic phosphate (P_i) released into the assay medium through the activity of the ATPase enzyme. The detection reagent, a mixture of malachite green and acidified ammonium molybdate, reacts quantitatively with P_i to form a green coloured complex absorbing strongly at 610nm (see chapter 2 and Harder *et al.*, 1994). Performed in 96-well microtitre plates, the assay provides excellent sensitivity and linearity of response when testing samples containing a range of P_i concentrations from 1 to 30 μM (50 to 1500 picomoles of P_i). Standard curves were prepared in parallel with each ATPase assay in order to quantify the amount of P_i released. Figure 6.1 shows the data generated for a typical standard curve.

To date, the ATPase activity of virtually all ABC transporter NBD's has been shown to be dependent upon the presence of a divalent cation, most usually magnesium. Indeed, the function of the highly conserved aspartate residue in the Walker B ATP-binding motif is believed to be co-ordination of the divalent cation. The presence of this cation is thought to be absolutely vital to the mechanism of ATP hydrolysis and a number of ABC transporter NBD's show dramatically reduced ATPase activities when the conserved aspartate is removed by site directed mutagenesis (e.g. see Zhou and Rosen, 1999).

Figure 6. 1



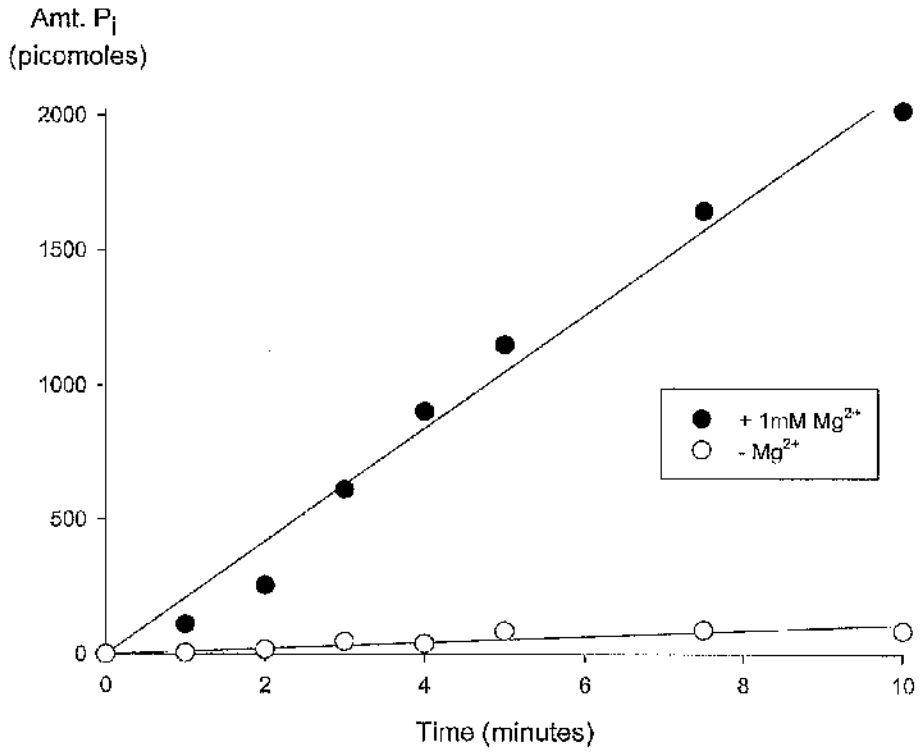
The standard curve above shows the response of the malachite green P_i assay over the range 50 to 1500 picomoles of P_i. Standard samples containing various amounts of phosphate were prepared by diluting a 1mM stock of NaH_2PO_4 into assay buffer (50mM MOPS, 100mM NaCl; pH 7.4). A 45µl aliquot of each standard sample was then mixed with 5µl of 500mM EDTA in a separate well of a 96-well microtitre plate. The assay was developed by the addition of 100µl of malachite green detection reagent to each well and the absorbance of samples at 610nm measured in a plate reader.

Detection reagent: 3 parts 0.045% w/v malachite green
1 part 4.2% w/v ammonium molybdate in 6N HCl
0.01% v/v Tween 20

The first ATPase assays performed were designed to identify whether or not the Trx-DrrA fusion protein was capable of ATP hydrolysis, and also whether such activity was dependent upon the presence of Mg^{2+} ions. Two identical reaction mixtures, detailed in figure 6.2, were prepared and incubated at 37 °C for 10 minutes in a thermostatically controlled heating block. After this initial equilibration period the reaction in one of the two tubes was initiated by the addition of $MgCl_2$ to a final concentration of 1mM. A similar volume of water was added to the second reaction mixture (the negative control). Over the ensuing ten minutes 45µl samples of each reaction mixture were withdrawn at various time points and mixed with 5µl of 500mM EDTA in a separate well of a microtitre plate. The role of the high concentration EDTA was to rapidly chelate Mg^{2+} ions, thereby immediately halting cation-dependent ATP hydrolysis in each sample. After a complete set of samples had been collected for both the experimental and control tubes, the phosphate content of each withdrawn sample was analysed by the addition of 100µl of malachite green detection reagent. The results were analysed immediately after the addition of the detection reagent by measuring absorbance at 610nm. This was done as quickly as possible in order to reduce the interfering effect of acid-hydrolysis of ATP caused by the low pH of the detection reagent.

The results of these initial experiments are shown in figure 6.2. The Trx-DrrA fusion protein clearly shows ATPase activity, as measured by an increase in P_i content of the assay mixture over time. A similar level of ATPase activity is absent from the control sample indicating that activity is, as expected, dependent upon the presence of a divalent cation. Whilst a small increase in phosphate content is observed in the negative control reaction over the course of the experiment, it is many times lower than that observed for the magnesium-containing sample. This small increase can perhaps be attributed to spontaneous ATP hydrolysis occurring over the total 20 minute period of the experiment. The assay was performed under 'steady-state' conditions, i.e. the concentration of substrate (5mM ATP) was far in excess of that of the enzyme (1.2µM, based upon a calculated M.W. of 52 kDa for Trx-DrrA). Under such conditions the enzyme would be expected to display linear ATPase activity. Although some deviation from linearity was observed in the earliest time-points, overall the activity was a good fit to a straight line ($R^2 = 0.98$).

Figure 6. 2 **Magnesium-dependent ATP hydrolysis by Trx-DrrA**



Two identical reaction mixtures (see below) were incubated at 37 °C for ten minutes prior to the addition of 1mM $MgCl_2$ (filled circles), or H_2O (open circles). 45 µl samples were withdrawn from the reaction mixtures at the time points indicated and ATPase activity quenched by mixing with 5µl of 500mM EDTA. Samples were analysed for P_i content by measuring absorbance at 610nm after the addition of 100µl of malachite green detection reagent. Absolute quantities of P_i in each sample were calculated by comparison with a standard curve.

Reaction conditions

Assay buffer = 50mM MOPS (pH 7.4), 100mM NaCl

[ATP] = 5mM

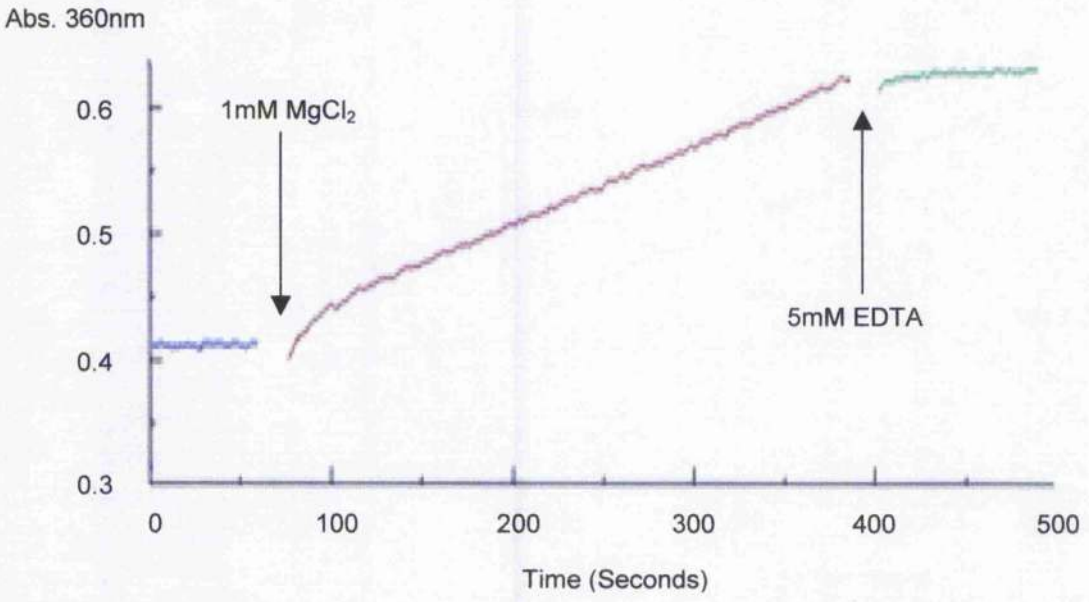
[Trx-DrrA] = 1.2µM

[$MgCl_2$] = 1mM (filled circles), 0mM (open circles)

The gradient of the activity curve indicated that P_i was being released into the reaction mixture at a rate of 209 picomoles/min/45µl sample analysed. By subtracting the apparent rate of spontaneous ATP hydrolysis observed in the control reaction (11 picomoles/min/45µl) a corrected figure of 198 picomoles/min/45µl was derived. ATPase data is more commonly presented in terms of activity/minute/mg protein. In order to allow effective comparison of Trx-DrrA with other NBD proteins the activity data was converted to these alternative units using the calculated M.W of the protein and the fact that the assay was carried out in a total volume of 800µl. On this basis, the ATPase activity of Trx-DrrA, under the conditions described, was found to be 56.41 nmoles P_i/minute/mg protein. The only other mycobacterial NBD protein for which ATPase activity data is available is PstB, the ATPase domain of the phosphate-specific transporter (Sarin *et al.*, 2001). The activity of this protein was found to be 122 nmoles P_i/min/mg, approximately twice that observed for Trx-DrrA. The significance of this and other comparisons are considered in the discussion section at the end of this chapter.

A second means of measuring the ATPase activity of Trx-DrrA was also tested in order to confirm the results of the malachite green P_i release assay. This second assay, the EnzChek Phosphate Assay system (Molecular Probes, see chapter 2), had the advantage that P_i accumulation could be monitored in real-time by spectroscopic means. This feature of the assay allowed both the stimulatory effect of Mg^{2+} ions and the quenching effect of EDTA to be shown directly. It also provided a further demonstration that the steady-state hydrolysis of ATP by Trx-DrrA was indeed linear, and that the deviations from linearity observed in the early time-point samples of the malachite green P_i assay were artefactual. Figure 6.3 shows that virtually no ATP hydrolysis is observed in the sample prior to the addition of 1mM $MgCl_2$ (0-60sec, blue line). Thereafter, the P_i content of the sample increased linearly over the period 77-387sec (red line) as measured by the increase in absorbance of the sample at 360nm. Finally, EDTA was shown to completely halt the ATPase activity of Trx-DrrA when added to the reaction mixture at a final concentration of 5mM (green line). The data produced by this assay were not directly quantified because of the need to use large quantities of the expensive assay reagents in order to generate P_i standard curves.

Figure 6. 3 **Direct observation of ATP hydrolysis by Trx-DrrA using a continuous spectroscopic assay**



The EnzCheck Phosphate Assay system (Molecular Probes) was used to monitor the continuous hydrolysis of ATP by Trx-DrrA in a spectroscopic assay. The assay mixture was prepared according to the manufacturers directions and the generation of P_i monitored as an increase in absorbance of the sample at 360nm (see below and chapter 2). ATPase activity was induced by the addition of 1mM $MgCl_2$ to the sample cuvette, left hand arrow, and stopped by the addition of the chelator EDTA, right hand arrow. Between these points the rate of ATP hydrolysis by Trx-DrrA was essentially linear and entirely dependent upon the availability of free Mg^{2+} ions. The assay was performed at 22 °C in a 1.5ml cuvette with a 1cm path length. Increase in absorbance of the experimental sample was calculated with respect to an identical cuvette containing dH_2O.

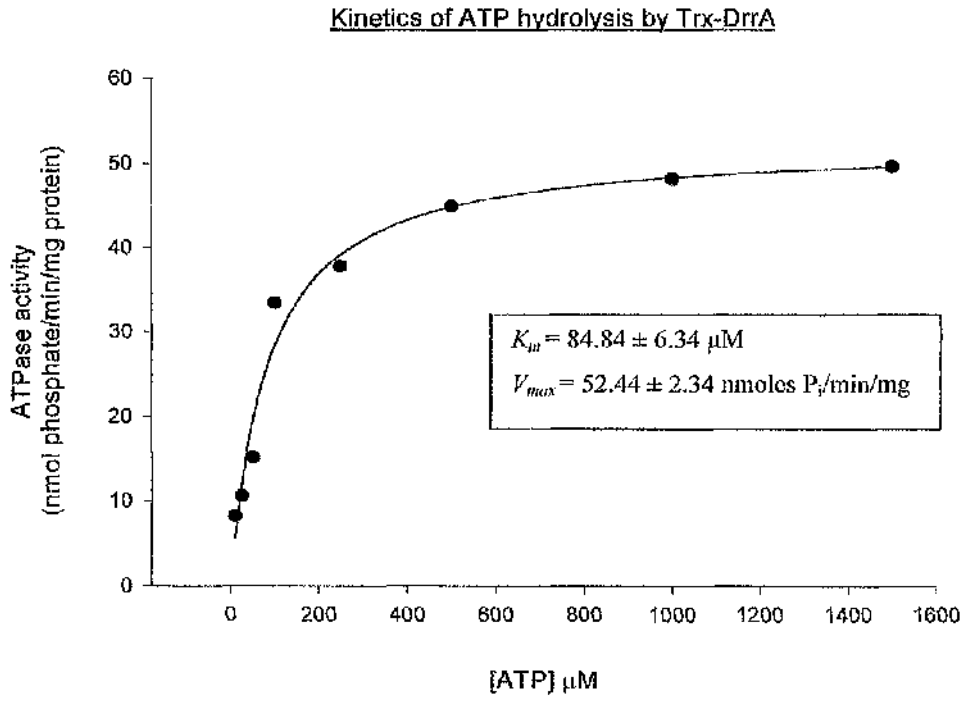
Assay Component	Volume (total 1ml)
MOPS (pH 7.4), 100mM NaCl	720µl
PNP	10µl
ATP (100mM stock)	10µl
MESG	200µl
Trx-DrrA (300µg/ml)	50µl
$MgCl_2$ (100mM stock)	10µl

6. 2 Dependence of Trx-DrrA ATPase activity on ATP concentration – Determination of V_{max} and K_m

Having established that the purified Trx-DrrA fusion protein was indeed a functional ATPase, a series of experiments were performed in order to determine some of the basic kinetic parameters of the enzyme's activity, thereby allowing meaningful comparison with other NBD proteins. The malachite green P_i release assay was used to determine the rate-dependence of ATP hydrolysis upon ATP concentration (under steady-state conditions). ATPase activity data was collected over a 150-fold range of ATP concentrations from 10μM to 1.5mM. The Trx-DrrA concentration used in all assays was 1μM (52μg/ml). Figure 6.4 shows that the rate of ATP hydrolysis varied in a hyperbolic manner with increasing ATP concentration, indicative of simple Michaelis-Menten type kinetics. Non-linear regression analysis (Sigma Plot 4.01, SPSS) showed that the data was a good fit to a single hyperbolic function indicating that Trx-DrrA had a V_{max} of 52.4 ± 2.3 nmoles P_i/min/mg and a K_m of 84.8 ± 6.3 μM.

Whilst regression analysis of this kind is undoubtedly the best way of accurately determining kinetic parameters from untransformed rate data, indirect, graphical means are also available. The same rate data illustrated in figure 6.4 were re-plotted in a straight line form according to the method of Hanes and Woolf. The values of the gradient and x-intercept were used to calculate V_{max} and K_m respectively (see figure 6.5). Using this graphical means Trx-DrrA was shown to have a V_{max} of 52.4 nmoles P_i/min/mg and a K_m of 85.5μM, both figures showing excellent correlation with those derived by non-linear regression. This calculated V_{max} value of 52.4 ± 2.3 nmoles P_i/min/mg for Trx-DrrA also shows good correlation with the earlier ATPase data described in section 6.1. This previous assay, performed using a very high ATP concentration of 5mM (≈ 60 times K_m), suggested an activity of 56.4 nmoles P_i/minute/mg protein, a figure that is in good overall agreement with the calculated value described here. Comparison of these data with those obtained for other prokaryotic ABC transporter NBD's is discussed in section 6.11 at the end of this chapter.

Figure 6. 4 **ATPase activity of Trx-DrrA as a function of [ATP]**



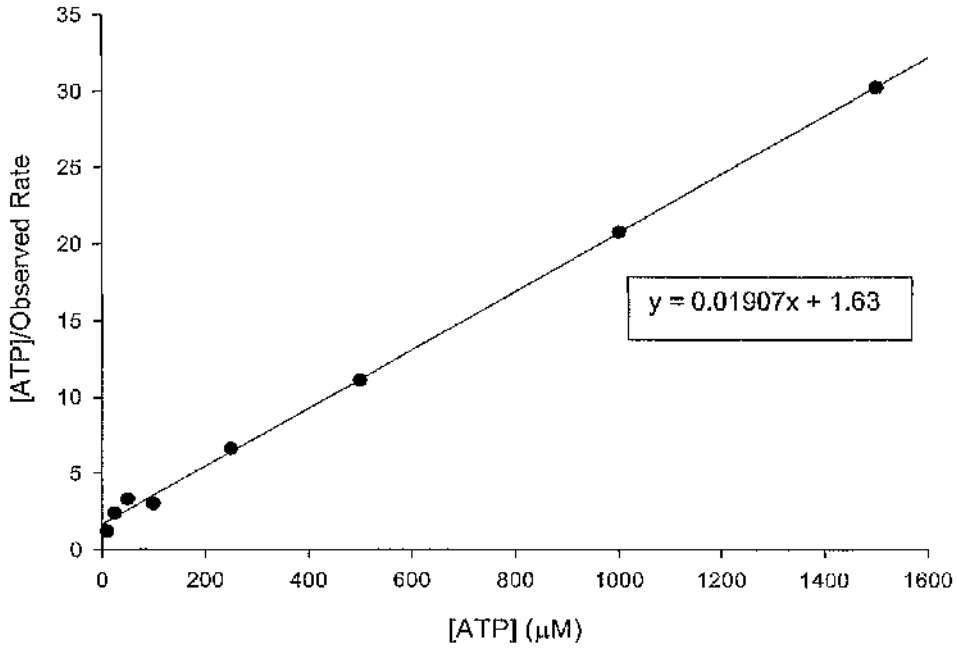
The malachite green P_i release assay was used to determine the rate of ATP hydrolysis by 1μM (52 μg/ml) Trx-DrrA in the presence of various ATP concentrations ranging from 10μM to 1500μM. ATPase assays were performed at 37 °C in a thermostatically controlled heating block and initiated by the addition of $MgCl_2$ to a final concentration of 1mM. All assays were conducted in a total volume of 800μl with 45μl samples analysed for P_i content at 1 minute time intervals over a time course of 10 minutes. A P_i standard curve was used to convert P_i accumulation data to rate data in the format nmoles P_i/min/mg protein.

The rate data were fitted to a single hyperbolic (Michaelis-Menten) function of the type:

$$y = \frac{a \, . \, x}{b + x}$$

where the co-efficients a and b correspond to kinetic parameters V_{max} and K_m, and x and y refer to the substrate concentration and measured rates respectively. The values for V_{max} and K_m returned by the Sigma Plot 4.01 non-linear regression software are shown on the graph.

Figure 6. 5 **Hanes-Woolf determination of V_{max} and K_m for Trx-DrrA**



The ATPase activity data shown in figure 6.4 were re-plotted according to Hanes and Woolf (above). The data were an excellent fit to a straight line with the equation: - $y = 0.01907x + 1.63$ ($R^2 = 0.998$). The Hanes-Woolf plot is a re-arrangement of the hyperbolic Michaelis-Menten plot and can be used to determine the kinetic parameters V_{max} and K_m by graphical means. The gradient of the straight line fit corresponds to the value $1/V_{max}$, whilst solving the linear equation for $y = 0$ provides the $-K_m$ value.

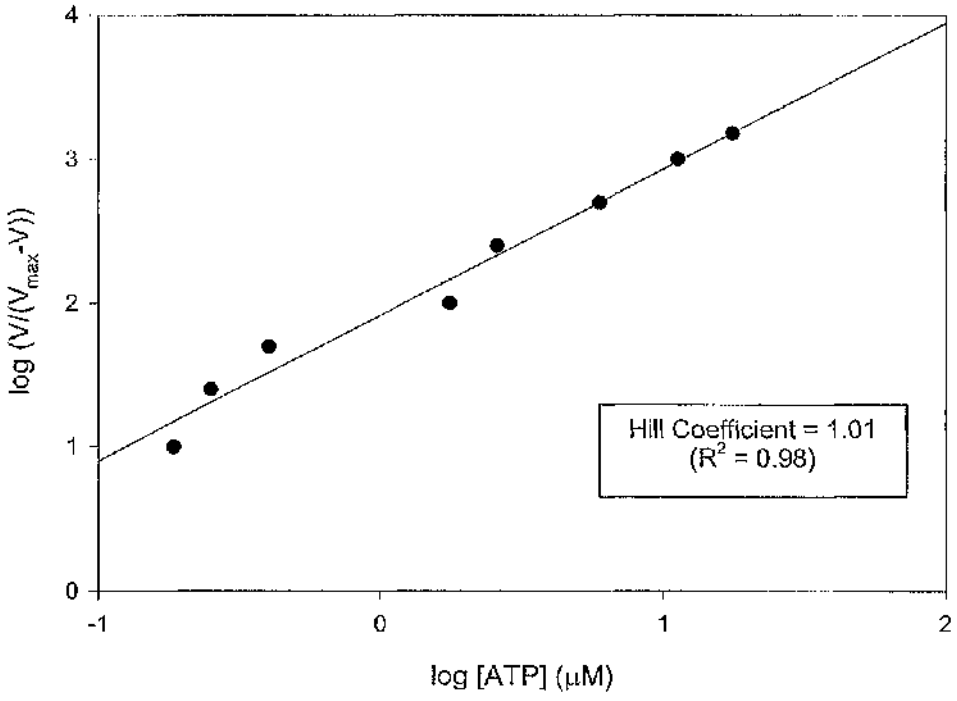
Evaluated by this means, the ATPase activity of Trx-DrrA was shown to have a V_{max} of 52.4 nmoles P_i/min/mg of protein and a K_m of 85.4μM. Both of these figures are in excellent agreement with those calculated for the same parameters by non-linear regression analysis of the untransformed data.

6. 3 The Trx-DrrA fusion protein does not display positive cooperativity with regard to ATP hydrolysis

The fact that Trx-DrrA appeared to hydrolyse ATP according to simple Michaelis-Menten kinetics suggested the absence of any physical or mechanistic interaction between protein sub-units during the ATP hydrolysis cycle. Indeed, a Hill plot of the kinetic data indicated that individual Trx-DrrA molecules behave completely independently of one another (Hill coefficient = 1.01, see figure 6.6). However, interaction between NBD's has become an increasingly central feature of models regarding ABC transporter function (e.g. see Locher *et al.*, 2002). It was considered that up to this point all ATPase activity had been measured at low enzyme concentrations, conditions which reduce the likelihood of protein/protein interactions, particularly if the affinity of sub-units for one another is weak. In order to investigate whether protein concentration might be a factor in the formation of Trx-DrrA dimers (or higher order multimers), a series of ATPase assays were conducted using increasing concentrations of protein whilst maintaining a constant ATP concentration. Protein concentration in the ATPase assays was varied over a 50-fold range from 1 to 50 μM. This upper limit was largely determined by the ability to produce solutions of Trx-DrrA up to a maximum concentration of approximately 3mg/ml (before incurring protein precipitation and loss of activity). The EnzChek phosphate assay system, rather than malachite green P_i assay, was used to measure ATPase activity for a number of reasons. Firstly, the continuous assay retains linearity of response over a wider range of P_i concentrations (0-100nmoles), and secondly, the rate of reaction could be inferred directly from the continuously monitored absorbance at 360nm with out the need to transform the data.

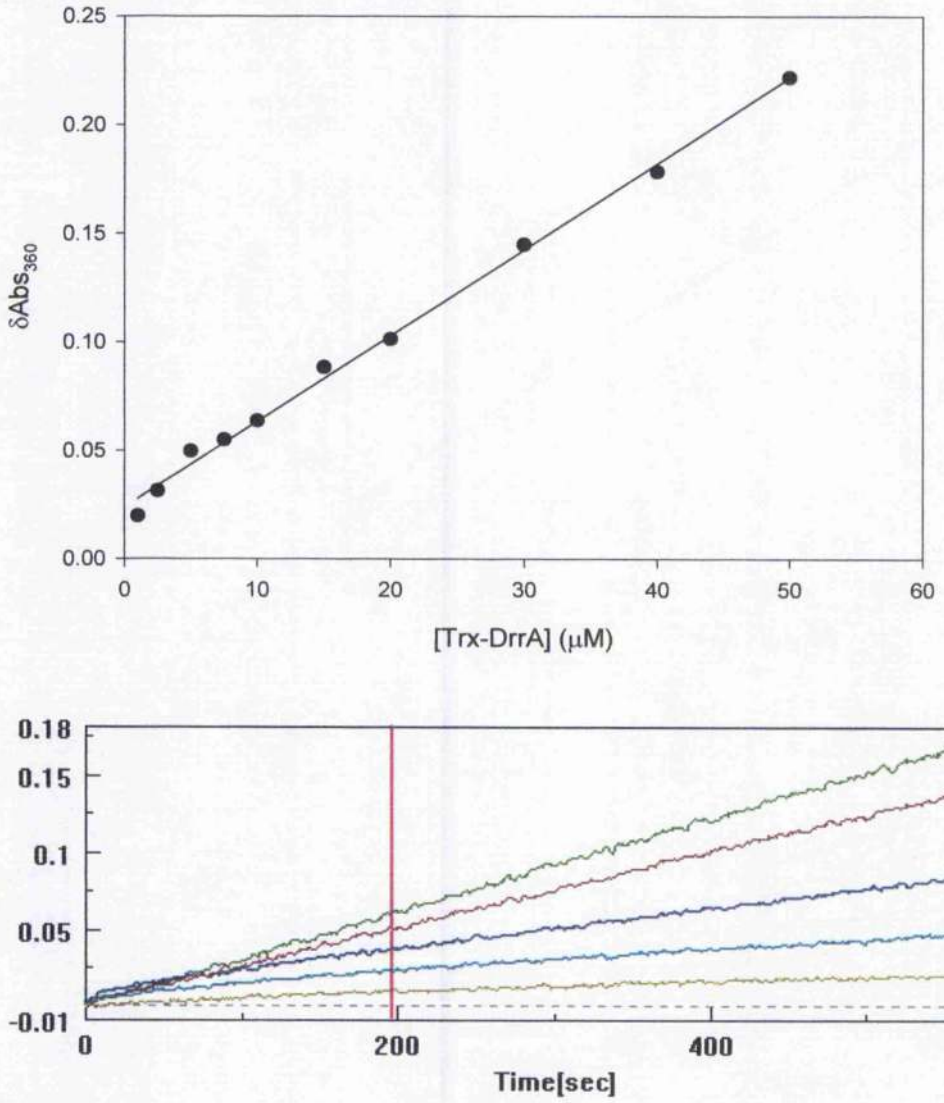
The results of these experiments are illustrated in figure 6.7. The plot of ATPase activity vs. Trx-DrrA concentration is essentially linear, once again suggesting that Trx-DrrA molecules hydrolyse ATP independently of one another, i.e. ATP hydrolysis by Trx-DrrA is a non-cooperative process. These results initially appear to be at odds with a significant amount of experimental data concerning the behaviour of both eukaryotic and prokaryotic ABC transporters. For example, the $MalFGK_2$ and $HisQMP_2$ transporters, from *E. coli* and *S. typhimurium* respectively, were both shown to display

Figure 6. 6 **Trx-DrrA hydrolyses ATP in a non-cooperative manner**



The Hill plot of Trx-DrrA ATPase activity shown above indicates that ATP hydrolysis occurs with a Hill coefficient of 1, i.e. it is a non-cooperative process. This data suggests that each Trx-DrrA molecule binds and hydrolyses ATP independently of all others and that increasing concentrations of ATP do not drive the formation of Trx-DrrA dimers or multimers that behave differently from the monomeric species.

Figure 6. 7 **Dependence of ATPase activity on protein concentration**



The rate of ATP hydrolysis by Trx-DrrA was shown to vary linearly with increasing protein concentration (top). Assays were performed using the EnzChek phosphate assay system in which the hydrolysis of ATP (and hence the accumulation of P_i) is monitored by a change in absorbance at 360nm (bottom). All assays were pre-incubated at room temperature with 1mM ATP before the reaction was initiated by the addition of 1mM $MgCl_2$. The fact that the top plot does not pass through the origin suggests a small amount of spontaneous ATP hydrolysis occurred over the 10 minute time-course of the assay.

positive cooperativity with regard to ATP hydrolysis when the activity of the intact transporter complex was measured (Davidson *et al.*, 1996; Liu *et al.*, 1997). Cooperative interaction between NBD's has also been demonstrated for the human multidrug ABC transporter P-glycoprotein/MDR1 (Qu and Sharom, 2001). However, in some cases, the behaviour of purified NBD's assayed in dilute solution appears to vary from that displayed by the same proteins when they are present as part of an intact (four domain) ABC transporter complex. For example, neither the MalK nor HisP NBD's show strong evidence of cooperative ATP hydrolysis when their activity is assayed in solution (Nikaido *et al.*, 1997; Morbach *et al.*, 1993). In response to this apparent contradiction Nikaido and colleagues proposed that removal of HisP from the ABC transporter complex may result in structural rearrangements that affect the enzymatic properties of the protein.

Only recently has direct evidence of ATP-driven dimerisation of isolated NBD's been presented. Moody and colleagues used very high resolution gel filtration chromatography to demonstrate that the soluble *M. jannaschii* ATP-binding cassette, MJ0796, forms stable dimers specifically in response to ATP-binding (Moody *et al.*, 2002). The formation of NBD dimers was shown to be highly dependent upon the ionic strength of the assay medium, a factor which suggests electrostatic interactions are important mediators of NBD-NBD association. The authors go on to propose a model of ABC transporter function in which ATP-driven association of NBD's is a vital first stage in coupling the event of ATP binding to the structural changes in the transporter complex necessary for substrate translocation. Upon ATP hydrolysis subtle electrostatic and structural effects result in dissociation of the dimer and the transporter returns to its initial state.

The biochemical behaviour of the Trx-DrrA fusion protein appears to mirror that of the isolated HisP and MalK NBD's, rather than that of MJ0796. However, the data presented here does not preclude the possibility that the native mycobacterial DrrA NBD behaves differently when it is part of a complete ABC transporter. Indeed, it illustrates rather well that subtle effects such as enzymatic cooperativity may be difficult to identify when the components of multi-protein complexes are treated in isolation. Whilst in some instances simple biochemical data may hint at the underlying mechanism of an enzymatic reaction, in others a more sophisticated analysis may be necessary. Increasingly, both biochemical

and structural data are beginning to suggest that ABC transporter function is highly dependent upon interactions between the various domains of the complex. As such, a complete understanding of their function may require researchers to adopt a 'systems-approach' to their analysis, rather than more traditional techniques.

6. 4 The Walker A structural motif in Trx-DrrA is vital for efficient ATP hydrolysis

The presence of the Walker A and B nucleotide-binding motifs in the mycobacterial DrrA protein is one of the key indicators that the enzyme is an ATPase, and possibly the energy transducing sub-unit of a novel mycobacterial ABC transporter. Crystallographic studies have shown that the amino acid residues of the Walker A motif in particular mediate vital hydrogen bonding interactions between the NBD and associated nucleotide (e.g. see Hung *et al.*, 1998). Mutations to residues within this motif are usually very poorly tolerated and can lead to complete loss of catalytic activity of ABC transporter NBD proteins (e.g. Saveanu *et al.*, 2001). In order to ascertain whether the Walker A residues of *M. tuberculosis* DrrA are similarly vital in the observed ATPase activity of protein, two different amino acid mutations were introduced into the motif by site directed mutagenesis (see chapter 2). The mutations introduced at key positions in the motif are illustrated below (most highly conserved residues are bold): -

Mutant 1 (G47S)	42- **G** P N G A **S** **K** T -49
	↑
Native DrrA Walker A motif	42- **G** P N G A **G** **K** T -49
	↓
Mutant 2 (K48E)	42- **G** P N G A **G** **E** T -49

Each mutation introduced was designed to significantly alter the chemical nature of the amino acid present at a given location in the motif. In the first mutant the small, apolar, glycine residue at position 47 is changed to the larger hydroxyl-containing amino acid serine. The second mutation introduced changes the positively charged lysine residue at position 48 to a negatively charged glutamate.

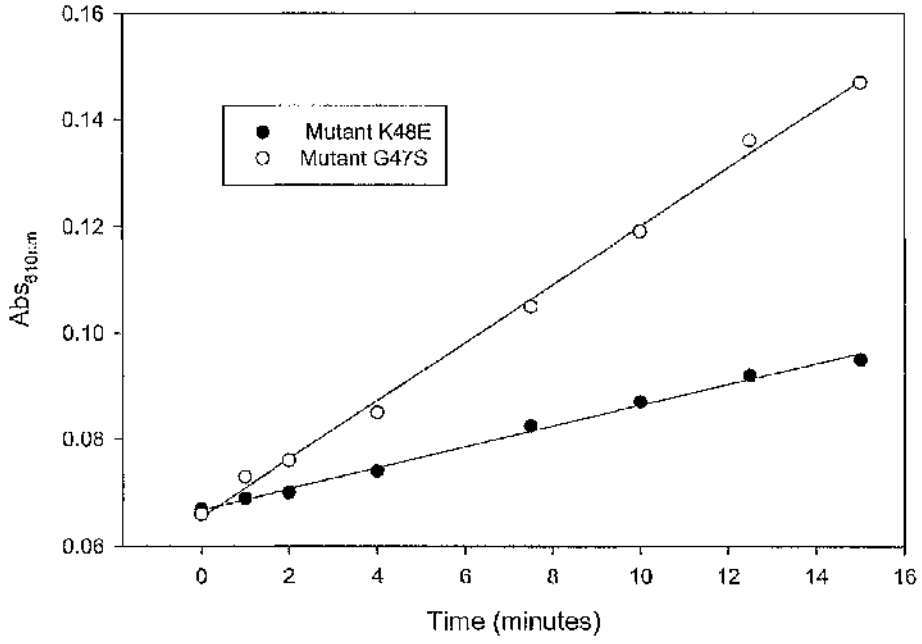
native Trx-DrrA fusion protein. The mutations appeared to have no significant effect on the expression levels, solubility or purity of the protein.

The biochemical effect of the Walker A mutations was assessed in terms of the ATPase activity displayed by the mutant proteins. Under conditions similar to those used for the initial characterisation of the native Trx-DrrA protein, i.e. 1mM ATP, 1mM $MgCl_2$, 1 M protein, no cation stimulated ATPase activity was detectable for either mutant using the malachite green P_i release assay. However, when the ATP concentration in the assay was increased to 5mM, a limited degree of Mg^{2+} stimulated activity could be detected for both proteins. The results of these assays and an analysis of the data are presented in figure 6.8.

Similarly to the native protein, both the G47S and K48E mutants were shown to hydrolyse ATP in a linear manner, ATPase activity being dependent upon the presence of Mg^{2+} ions. When the rate of ATP hydrolysis by each mutant protein was calculated in terms of a specific activity, significant differences emerged both between the mutants, and in comparison with native Trx-DrrA. The K48E mutant showed the lowest level of activity, 2.86 nmoles P_i/min/mg. This represents just 5% of the activity of the native Trx-DrrA protein assayed under the same conditions. The G47S mutant was a little more active, displaying a specific activity of 7.96 nmoles P_i/min/mg. Nonetheless, this activity still represented just 14% of that of the native protein. The very low activity of the two mutant proteins precluded a rigorous analysis of the kinetic properties of these enzymes e.g. determination of K_m and V_{max} values. However, it is apparent from these empirical data that both the G47S and K48E mutations result in dramatically reduced enzymatic activity.

The experiments provided important confirmation that the ATPase activity of Trx-DrrA was at least 95% attributable to the mycobacterial DrrA moiety of the fusion protein. Furthermore, ATP-binding appears to be dependent on the same nucleotide/protein interactions as occur in other well-characterised ABC transporter NBD's. The dramatic decreases in ATPase activity of the mutant proteins almost certainly reflect a reduced affinity for ATP caused by structural distortion or physico-chemical alteration of the nucleotide binding fold.

Figure 6. 8 **Mutations in the Walker A motif of Trx-DrrA result in reduced ATPase activity**



Protein	Specific Activity	% Activity
Wild-Type	56.4 nmoles P_i/min/mg	100*
Mutant K48E	2.86 nmoles P_i/min/mg	5
Mutant G47S	7.96 nmoles P_i/min/mg	14
		* see section 6.1

The malachite green P_i release assay was used to calculate the ATPase activity of mutant Trx-DrrA proteins. Assays were performed at 37 °C in a total volume of 800µl. Each assay contained 1µM mutant protein, 5mM ATP and was initiated by the addition of $MgCl_2$ to a final concentration of 1mM. 45µl samples of the assay mixtures were withdrawn at the time-points indicated and mixed with 5µl of 500mM EDTA to quench the reaction. Phosphate content of each sample was analysed by Abs_{610nm} after mixing with 100µl of malachite green detection reagent. Reaction rates measured in terms of δAbs_{610} (see figure) were converted to specific activities using a P_i standard curve. Neither protein exhibited activity in the absence of Mg^{2+} (data not shown).

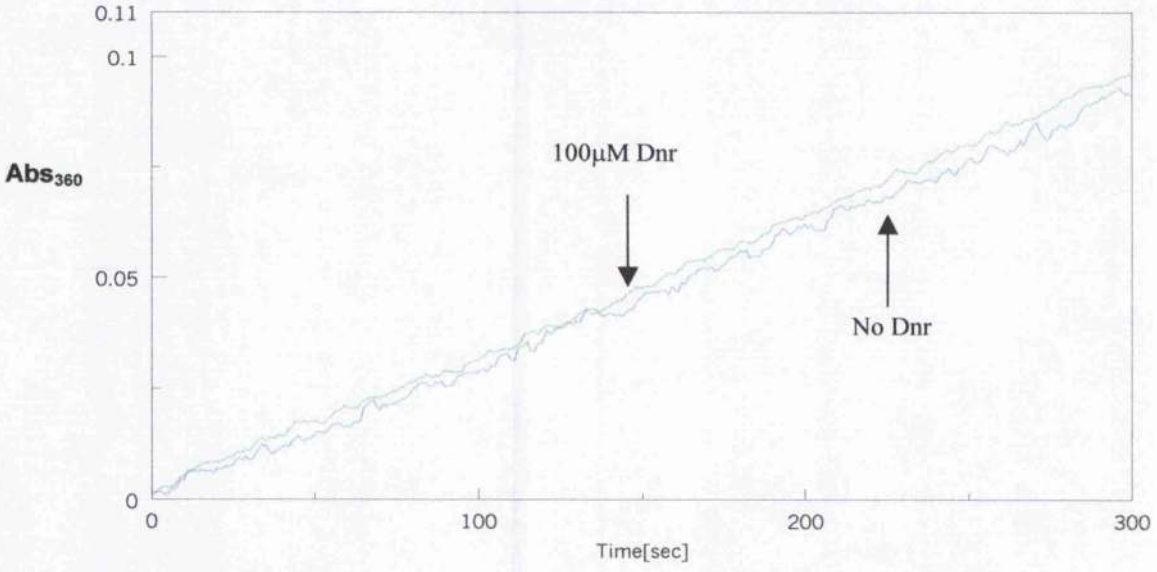
6. 5 Neither Doxorubicin nor Daunorubicin activate the rate of ATP hydrolysis by the Trx-DrrA fusion protein

A number of ABC transporters exhibit significantly enhanced ATPase activities when exposed to their transport substrates, e.g. mammalian P-glycoproteins, and the ArsAB arsenite/antimonite efflux transporter of *E. coli* (Scarborough, 1995; Kaur, P., 1999). However, as with so many other features of ABC transporter research, this property is by no means universal and the mechanism by which it occurs is a matter of some controversy. For example the ArsA ATPase was shown by X-ray crystallography to possess three metalloid binding sites, occupancy of which can be correlated with increases in the rate of ATP binding and hydrolysis by this protein (Zhou *et al.*, 2000). On the other hand, only scant evidence exists to suggest that the NBD's of P-glycoproteins possess similar substrate binding sites. To add further complexity to the picture, neither free histidine nor the liganded soluble receptor HisJ, affect the ATPase activity of purified HisP, whilst spermidine, the substrate of the PotA ABC transporter, actively inhibits the ATPase activity of this NBD (Nikaido *et al.*, 1997; Kashiwagi *et al.*, 2002). Despite this apparent complexity of regulation, which may involve the presence of substrate binding sites in either the NBD or MSD components of a transporter, it was decided to analyse whether or not the ATPase activity of the Trx-DrrA fusion protein was affected by the presence of the potential transport substrates doxorubicin and daunorubicin. An obvious stimulatory effect caused by the presence of these antibiotic molecules might be interpreted as an indication that the DrrABC transporter is indeed an antibiotic efflux system, and that *M. tuberculosis* retains some of the antibiotic transport capabilities widely distributed amongst other members of the actinomycetaeles.

ATPase assays were performed using the EnzChek phosphate assay system. 40μM Trx-DrrA was pre-incubated in the presence of 1mM ATP and doxorubicin or daunorubicin at varying concentrations between 0 and 100μM. After a five minute incubation period ATPase activity was initiated by the addition of 1mM $MgCl_2$. ATPase activity was monitored spectroscopically by a change in absorbance of the assay mixture at 360nm over a period of five minutes.

Figure 6.9 shows the results of two such ATPase assays, one conducted in the absence of antibiotic (blue line), and the second in the presence of 100μM daunorubicin (green line).

Figure 6. 9 **The ATPase activity of Trx-DrrA is unaffected by anthracycline antibiotics**



The ATPase activity of Trx-DrrA monitored in the presence and absence of the potential transport substrate daunorubicin (Dnr). ATPase assays were performed at room temperature using 40μM Trx-DrrA (approx. 2mg/ml) and 1mM ATP. Assays were prepared in a total volume of 1ml according to the manufacturers instructions (see chapter 2). Dnr was added to the assay by microdilution of a concentrated stock solution prepared in DMSO. After a five minute pre-incubation period ATP hydrolysis was initiated by the addition of $MgCl_2$ to a final concentration of 1mM. The rate of ATP hydrolysis was monitored by the increase in absorbance of the assay mixture at 360nm.

The gradients of the two activity curves above were used as a direct measure of ATPase activity. In the first assay, conducted in the absence of daunorubicin, the ATPase activity corresponded to a δAbs_{360} of 0.0182 units/minute (blue line). The second assay, conducted in the presence of 100μM daunorubicin, showed a δAbs_{360} of 0.0192 units/minute (green line). The difference between these two figures represents that a less than 5% increase in ATPase activity occurring in response to the presence of the antibiotic. This very small figure suggests that anthracycline antibiotics have little or no effect on the ATPase activity of Trx-DrrA.

Clearly, the presence of the antibiotic has no significant effect on the rate of ATP hydrolysis by Trx-DrrA, even at this highest concentration tested. One possible interpretation of these results is that the Trx-DrrA fusion protein does *not* possess a binding site(s) for doxorubicin/daunorubicin. Alternatively the results may simply suggest that any interaction between the NBD and the anthracycline antibiotics has no stimulatory effect on the rate of ATP binding or hydrolysis. However, two interesting counterpoints to these data are have been presented in the literature. In her initial studies on the *S. peucetius* DrrAB antibiotic efflux system Kaur was able to demonstrate that 35μM doxorubicin significantly enhanced the binding of radiolabelled ATP to DrrA (Kaur, P., 1997). In this experimental system radiolabelling was performed using whole cell membranes from a recombinant *E. coli* strain expressing both the DrrA and B proteins, rather than using the purified NBD. Kaur suggests that the data can be interpreted in two alternative ways. The first, simplest, interpretation is that the NBD protein DrrA is an allosteric enzyme with binding sites for doxorubicin that, when occupied, directly enhance ATP binding. The second interpretation proposes that binding of doxorubicin to the MSD component of the transporter, DrrB, enhances ATP binding to DrrA through conformational changes transmitted from one domain to the other.

A second, and perhaps more significant contradiction to the ATPase activity data presented here was published by Choudhuri and colleagues in 2002. This group, also working on the *M. tuberculosis* Drr proteins, developed an artificial operon for simultaneous recombinant expression of both DrrA and DrrB in *E.coli* (Choudhuri *et al.*, 2002). Much of the research published by this group is discussed in the next chapter which considers the possible physiological roles of the *M. tuberculosis* DrrABC system. However, in a series of experiments essentially similar to those performed by Kaur, this group were able to demonstrate that ATP binding to DrrA was again enhanced by both doxorubicin and daunorubicin (40μM) when binding activity was measured in whole cell membranes. Choudhuri uses these, and other data presented in the same work, to argue that the mycobacterial DrrA and B proteins together constitute a functional daunorubicin/doxorubicin efflux transporter. It should be noted however that they chose to express only two of the three genes present in the transcriptionally linked *M. tuberculosis drr*ABC operon.

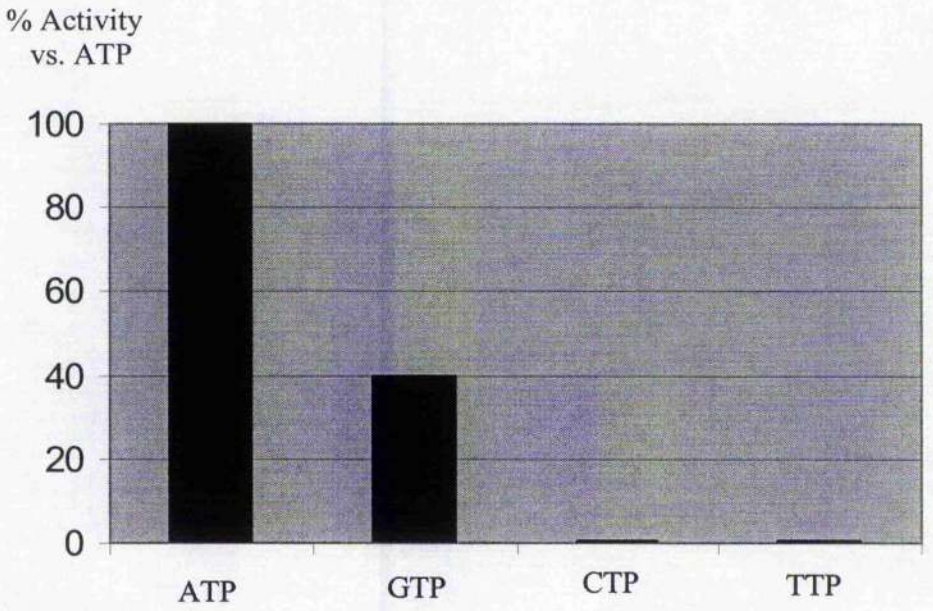
Nonetheless, both the experiments of Kaur and Choudhuri further suggest the potential importance of inter-domain interactions between the various components of prokaryotic ABC transporters. The fact that the ATPase activity of the soluble Trx-DrrA fusion protein is independent of daunorubicin concentration would seem to suggest that any enhancement of ATP-binding must be the product of antibiotic binding to a component of the transporter other than DrrA.

6. 6 Trx-DrrA displays a marked preference for purine nucleotides

The ability to bind substrates other than Mg^{2+}ATP has been demonstrated for a number of ABC transporter NBD's (e.g. see Buche *et al.*, 1997). Whilst this ability is unlikely to be of physiological significance, it is indicative of a degree of conformational flexibility within the NBD core. In order to ascertain whether Trx-DrrA possesses a similar degree of conformational flexibility the malachite green P_i release assay was used to test alternative nucleotide-triphosphates as potential substrates. GTP, CTP and TTP were all tested at 1mM concentrations and their rate of hydrolysis compared with that of ATP (see figure 6.10).

The results indicated that Trx-DrrA *was* capable of hydrolysing the purine nucleotide GTP, whilst displaying virtually no activity towards the pyrimidine substrates CTP and TTP. The rate of hydrolysis of GTP was approximately 40% of that observed for the same concentration of ATP, suggestive of a lower affinity for this substrate. A similar ability to bind GTP was demonstrated for the *S. peucetius* DrrA protein (Kaur, P. 1997). The only other mycobacterial NBD protein for which similar data is available is PstB, the NBD subunit of the phosphate specific uptake transporter (Sarin *et al.*, 2001). This protein was able to hydrolyse not only ATP and GTP, but also the pyrimidine CTP. In fact the affinity of this protein for CTP was higher than that for GTP (K_m[CTP] = 130μM, K_m[GTP] = 211μM). Generally speaking there is very little consistency in the literature regarding the ability of different ABC transporter NBD's to bind and hydrolyse alternative NTP's. Differences appear to be case specific although GTP is by far the most common alternative substrate. Whilst the F_1-type ATP-binding core of all NBD proteins is thought to adopt a highly conserved structure, it is possible that minor differences in the amino acid residues in and around the ATP-binding fold affect the ability of different NBD's to utilise alternative substrates.

Figure 6. 10 Hydrolysis of alternative NTP substrates by Trx-DrrA



Substrate	Specific activity*	% activity vs. ATP
ATP	42.08	100
GTP	16.66	39.6
CTP	0.29	0.7
TTP	0.35	0.8

* nmoles P_i /min/mg Trx-DrrA

NTPase assays were performed at 37 °C in a total volume of 800µl using 1µM Trx-DrrA and 1mM NTP. Reactions were initiated by the addition of 1mM $MgCl_2$. 45µl samples were withdrawn from each reaction at 1 minute intervals over a 10 minute period and analysed for P_i content as described in chapter 2 and elsewhere in this chapter. Specific activities with regard to each substrate were calculated and normalised by comparison to ATP. Whilst ATP is clearly the preferred substrate, Trx-DrrA is also able to bind and hydrolyse the structurally similar molecule GTP. The fact that neither CTP nor TTP are effective substrates indicates there is only a limited degree of conformational flexibility in the ATP-binding core of the mycobacterial DrrA protein.

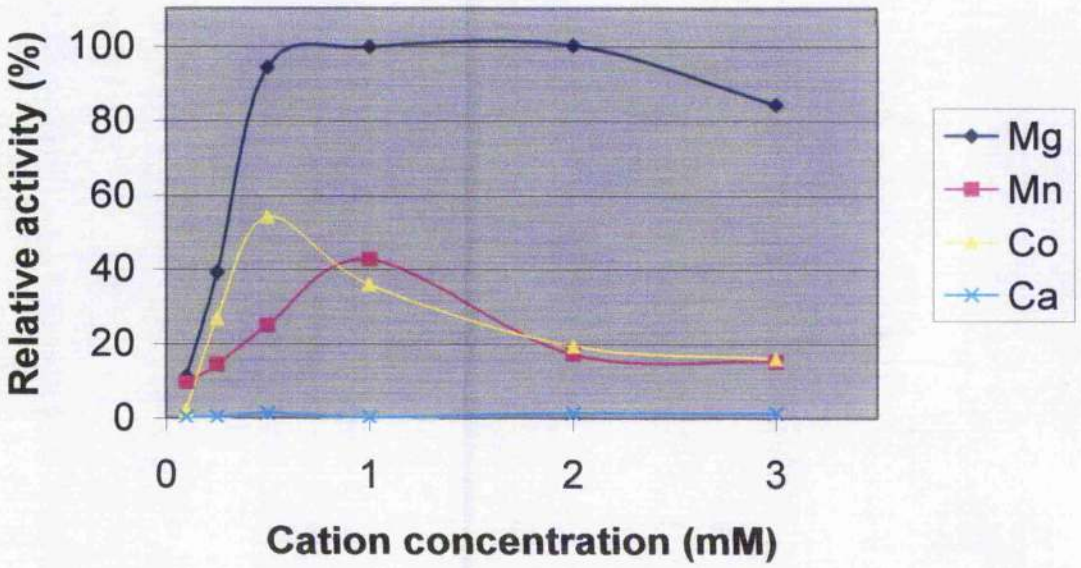
Conformational flexibility of secondary structural elements in the region of the ATP-binding fold may also have a significant effect. These experiments suggest that the Trx-DrrA NBD has only a limited ability to bind and utilise alternative substrates.

6. 7 Both Mn^{2+} and Co^{2+} ions will support ATP hydrolysis by Trx-DrrA

The data presented in section 6.1 demonstrated that the catalytic activity of Trx-DrrA was entirely dependent upon the presence of Mg^{2+} ions. In order to ascertain whether alternative divalent cations might substitute for Mg^{2+} in the reaction mechanism of Trx-DrrA a series of ATPase assays were performed in the presence of different salts. The same series of experiments were also used to determine the rate dependency of the ATPase reaction upon the concentration of the various ions tested. The dichloride salts of manganese, cobalt, zinc and calcium were all tested for their ability to support ATP hydrolysis. The concentration of each salt was varied over a range from 100μM to 3mM and rate data presented as percentage of the maximum rate of ATPase activity supported by Mg^{2+} (see figure 6.11). All ATPase assays were conducted at 37 °C using the malachite green P_i release assay and constant concentrations of Trx-DrrA and ATP of 1μM and 1mM respectively.

The data shown in figure 6.11 clearly indicate that Mg^{2+} ions support the highest level of ATPase activity across all concentrations tested. The optimal concentration for Mg^{2+} ions appears to lie in a broad range between 0.5mM and 2mM. Only at the highest concentrations tested do Mg^{2+} become inhibitory to the ATPase activity of Trx-DrrA. On the basis of this data 1mM $MgCl_2$ was used in all further ATPase assays. The next best stimulator of ATPase activity was cobalt dichloride. Co^{2+} ions produced 53% maximal ATPase activity at a concentration of 0.5mM with stimulatory activity dropping-off sharply above this concentration. Manganese ions were also able to support ATP hydrolysis by Trx-DrrA, a maximal stimulation of 42% occurring at a concentration of 1mM. Neither Zn^{2+} nor Ca^{2+} ions were able to support ATPase activity.
The physical and mechanistic implications of the fact that a range of divalent cations can support ATP hydrolysis by ABC transporter NBD's are unclear. Structural and biochemical data suggest that the divalent cation is liganded by the highly conserved aspartate residue of the Walker B motif (e.g. see Hung *et al.*, 1998). The presence of a divalent cation at this site is clearly central to the catalytic mechanism of ABC

Figure 6. 11 **Cation dependency of ATP hydrolysis by Trx-DrrA**



The divalent cations Mg^{2+}, Mn^{2+}, Co^{2+}, Ca^{2+} and Zn^{2+} were all tested for their ability to stimulate the ATPase activity of Trx-DrrA. All assays were performed at 37 °C in a total volume of 800µl. The ATP concentration in each assay was constant at 1mM and Trx-DrrA concentration was maintained at 1µM. ATPase activity was initiated by the addition of dichloride metal salts at the concentrations indicated. The malachite green P_i release assay was used to calculate the rate of ATP hydrolysis in each experiment and all data is presented as a percentage of the maximal rate of ATP hydrolysis observed.

Mg^{2+} ions were the most effective stimulators of ATP hydrolysis (100% at 1mM), whilst both Mn^{2+} and Co^{2+} produced a moderate stimulatory effect at concentrations between 0.5 and 1mM. Neither Zn^{2+} nor Ca^{2+} produced any stimulatory effect at any of the concentrations tested (data for Zn^{2+} not shown as it overlaps with that of Ca^{2+}).

transporters since mutations to the Walker B aspartate are poorly tolerated and frequently lead to complete loss of ATPase activity. The behaviour of Trx-DrrA with respect to the different cations tested is broadly similar to that observed for model NBD proteins such as HisP. Optimal concentrations of cation lie in the range 1-2mM with Mg^{2+}, Mn^{2+} and Co^{2+} ions all supporting ATP hydrolysis to a greater or lesser extent (Nikaido *et al.*, 1997). It should perhaps be noted that the mycobacterial PstB NBD exhibits some highly unusual properties with regard to its use of cations. Mg^{2+} ions over a range of concentration 10μM to 100mM were reported to have absolutely no effect on ATPase activity of the enzyme. ATPase activity could even be detected in the complete absence of Mg^{2+} ions. Even more strangely Mn^{2+} ions appeared to actively inhibit the ATPase activity of the enzyme. At the same time, mutation of residue Asp118, the Walker B aspartic acid, to lysine, dramatically decreased the observed ATPase activity (Sarin *et al.*, 2001). These two groups of data are very difficult to reconcile with one another and with what is known about the activity of ABC transporter NBD's in general. The highly unusual behaviour of this mycobacterial protein awaits clarification.

6. 8 Intrinsic tryptophan fluorescence studies of Trx-DrrA

Tryptophan fluorescence spectroscopy has proven to be a powerful tool for investigating the properties of both ABC transporter NBD's and complete transporter complexes (see Borges-Walmsley *et al.*, 2003). The technique relies upon the observation that the fluorescence properties of individual tryptophan residues within a protein are uniquely responsive to even small changes in their microenvironment. Consequently, conformational changes in a protein, induced by events such as ligand binding, can frequently be correlated with a change in the measured tryptophan fluorescence. Collected under steady-state conditions, protein fluorescence data can be used to calculate important biochemical parameters such as the equilibrium dissociation constant of a particular protein-substrate interaction. The *rate of change* of tryptophan fluorescence linked to a particular event can also be used to determine similarly important reaction parameters. Walmsley and colleagues have successfully used a combination of these techniques to develop a detailed kinetic mechanism describing the activity of the of the *E. coli* ArsA ATPase. Conformational changes associated with ATP binding, metalloid binding, ATP hydrolysis and product release were all detected through changes in the tryptophan fluorescence of the protein (Walmsley *et al.*, 2001).

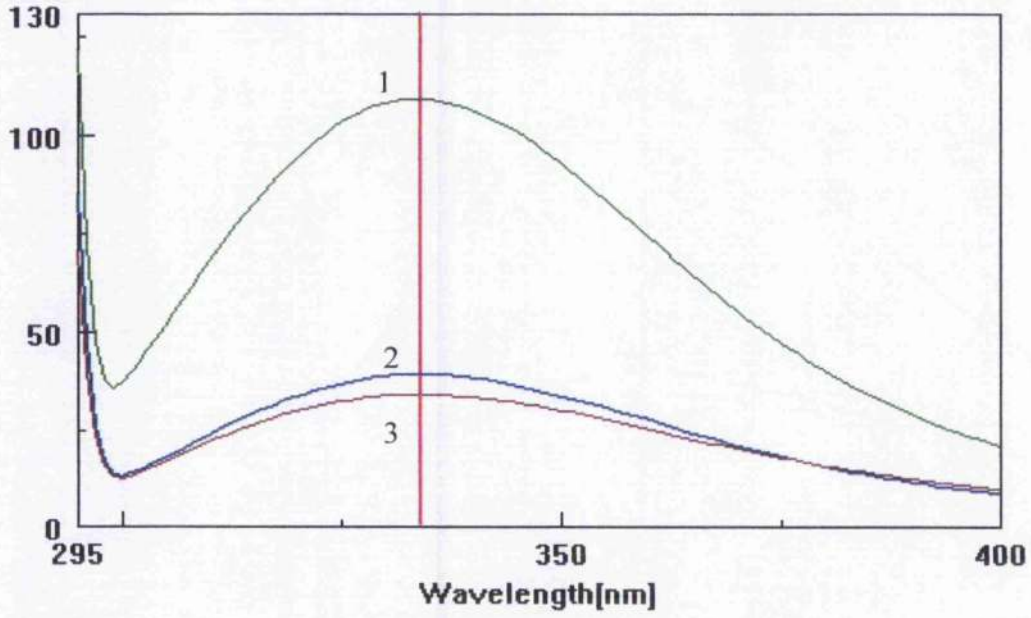
In order to ascertain whether similar conformational changes might be detected during the ATP binding and hydrolysis cycle of Trx-DrrA, steady-state tryptophan emission spectra were collected for the protein both in the presence and absence of nucleotide. The resulting spectra are shown in figure 6.12.

In the absence of nucleotides Trx-DrrA exhibits an emission maximum of 109 Units at a wavelength of 334nm. The addition of 1mM ATP to the protein resulted in a large quench in the measured fluorescence (64% decrease), and the subsequent addition of 1mM $MgCl_2$ to the mixture caused a further quench of 4%. Neither addition had any effect on the λ_{max} value which remained constant at 334nm. These data initially suggested that one or more of the four Trp residues in Trx-DrrA was reporting a nucleotide induced conformational change in protein structure. However, the adenine ring of ATP is known to interfere with fluorescence measurements by absorbing photons, thereby reducing the intensity of both the excitation beam and of the observed emission spectrum. This phenomenon, known as an *inner filter effect*, is clearly an important consideration when measuring the steady-state fluorescence of ATP-binding proteins. After a series of careful controls in which the magnitude of the inner filter effect was calculated for different concentrations of ATP, it was determined that the observed 64% decrease in Trx-DrrA fluorescence was entirely attributable to an inner filter effect. Given the magnitude of the inner filter effect it was considered that any small changes in the tryptophan fluorescence of Trx-DrrA, occurring as a result of specific interactions with ATP, were unlikely to be detected by this technique.

6. 9 Construction of single tryptophan Trx-DrrA mutants

In order to enhance the detectability of any fluorescent signal associated with ATP binding to Trx-DrrA a site-directed mutagenesis approach was adopted. Firstly, it was known that Trx-DrrA contained a total of four Trp residues, Trp30 and Trp33 in the thioredoxin moiety of the fusion protein, and Trp110 and Trp180 in the DrrA protein. It was considered that neither of the tryptophan residues in the thioredoxin moiety was likely to be involved in the binding and hydrolysis of ATP by the fusion protein. As such, whilst these two residues were undoubtedly contributing to the overall fluorescence of the protein, their lack of response in any binding assay would merely produce a background fluorescence that would lower the overall signal-to-noise ratio. As a result of these considerations it was decided that both Trp30 and Trp33 could be

Figure 6. 12 **Tryptophan fluorescence emission spectra of Trx-DrrA**



The fluorescence emission spectrum of 2.5μM Trx-DrrA measured at room temperature in a buffer of 50mM MOPS, 100mM NaCl, pH 7.4. Tryptophan fluorescence was excited at a wavelength of 285nm and emission data measured between 295 and 400nm. Data were collected on a Jasco FP-750 spectrofluorimeter with emission and excitation bandwidths set at 5nm. The response of the detection photomultiplier tube was set to medium.

1) 2.5μM Trx-DrrA	λ_{max} = 334 nm	Peak height = 109 Units
2) 2.5μM Trx-DrrA + 1mM ATP	λ_{max} = 334 nm	Peak height = 39.1 Units
3) 2.5μM Trx-DrrA + 1mM ATP + 1 mM $MgCl_2$	λ_{max} = 334 nm	Peak height = 33.9 Units

safely mutated to alternative residues without affecting any possible ATP binding signal. This was achieved by using oligonucleotide directed mutagenesis; the single primer 'Trx2W' being used to simultaneously mutate tryptophan residues 30 and 33 to glycine residues (see chapter 2). A comparison of the primary sequence of DrrA with those of structurally defined NBD's was used to determine the likely location of tryptophans 110 and 180. Trp110 is in a region of the protein between Walker A and the ABC transporter signature-sequence motif. In structurally characterised NBD's this region is generally poorly conserved and contributes to the transitional region between the F_1-type ATP-binding core of the protein and the ABC transporter-specific α-helical domain. Tryptophan180 on the other hand lies slightly downstream of the Walker B sequence, a region of the protein believed to be involved in cation coordination. Using the published structures of HisP and the MJ0796 NBD's as a guide, Trp180 is likely to be at some distance from the ATP binding site in another weakly conserved region of the NBD. Given the likely presence of the two DrrA Trp residues in catalytically inactive regions of the protein, the decision was made to mutate these residues to non-fluorescent glycines (using oligonucleotides W110G and W180G, see chapter 2). The combined result of these mutational experiments was the construction of a tryptophan-free Trx-DrrA molecule in which all four tryptophan residues were replaced by glycines (see figure 6.13). DNA sequencing of the pTrx-DrrA expression plasmid was used to confirm the successful introduction of all four mutations.

The tryptophan-free Trx-DrrA molecule was used as a low fluorescence background into which two alternative mutations were introduced. Each mutant protein was designed to have a fluorescently reactive Trp residue in a structurally and/or mechanistically active region of the protein. The structures of HisP and MalK both contain aromatic residues at positions close to the N-terminal of the NBD. These residues, Y16 in HisP and W13 in MalK, are thought to contribute to an important aromatic-stacking interaction between the protein and the adenine ring of a bound ATP molecule, thereby stabilising ATP binding (Hung *et al.*, 1998; Walter *et al.*, 1992). The equivalent of this residue in *M. tuberculosis* DrrA is tyrosine17. It was hoped that mutating this residue to tryptophan would result in the generation of a Trx-DrrA molecule that was spectroscopically active in response to ATP binding. Indeed, working with MalK, Schneider and colleagues were able to demonstrate a 37% decrease in protein fluorescence in response to the addition of ATP, possibly a direct result of the stacking interaction between W13 and ATP

(Schneider *et al.*, 1994). A second single-tryptophan Trx-DrrA mutant was also produced. The so-called ABC signature-sequence motif is one of the most highly conserved regions of all ABC transporter NBD's. It generally conforms to the consensus 'LSGGQ' and is found in the ABC transporter-specific α-helical sub-domain of the NBD. A number of crystallographic NBD dimers suggest this motif is part of a composite ATP binding site formed by juxtaposition of one NBD's Walker A motif with the signature-sequence motif of a second molecule. In *M. tuberculosis* DrrA the signature-sequence motif is degenerated to the sequence 'YSGGM'. Mutation of the motif's initial tyrosine residue (Y141) to a potentially fluorescently active tryptophan residue is particularly attractive because this region of the protein may well be intimately involved in both dimerisation of the NBD and the mechanism of ATP hydrolysis.

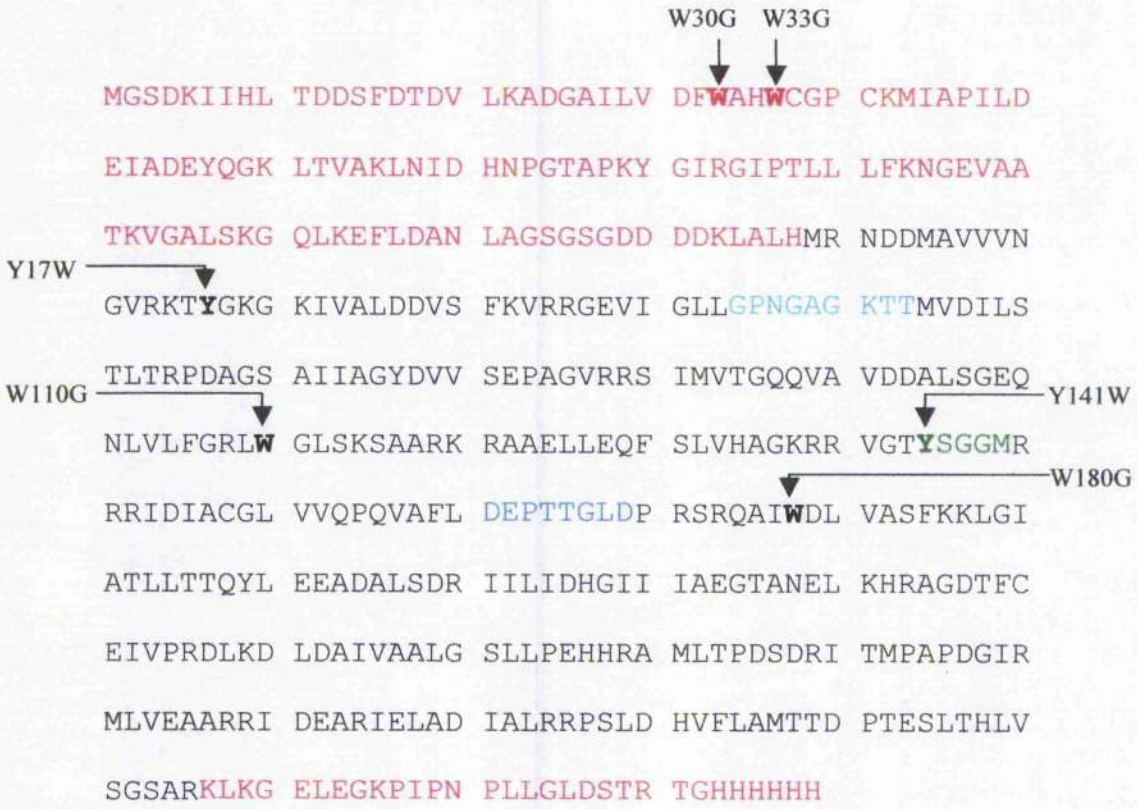
The two single-tryptophan Trx-DrrA mutants Y141W and Y17W were produced by oligonucleotide directed mutagenesis of a plasmid encoding the tryptophan free molecule (see chapter 2). Both molecules were overexpressed in similar quantities to the wild-type protein and demonstrated essentially identical ATPase activity (data not shown). The positions of all relevant mutations are shown in figure 6.13.

6. 10 Stopped-flow fluorescence spectroscopy of ATP binding to single-tryptophan mutant Trx-DrrA proteins

Stopped-flow fluorescence spectroscopy allows the very rapid detection of changes in protein fluorescence in response to ligand binding. Unlike measurement of changes under steady-state conditions, the use of the stopped-flow system negates inner filter effects because any reduction in signal due to the presence of ATP occurs prior to the onset of data collection. A second advantage of the system is that even very small and rapid changes in fluorescence that may be missed under steady-state conditions can be detected and time-resolved to reveal mechanistic details of the interaction (e.g. see Walmsley *et al.*, 2001).

Fluorescent signals produced by mixing the two mutant proteins with 1mM ATP and 1mM $MgCl_2$ in the stopped flow instrument are shown in figures 6.14. All data were collected at 22 °C in an Applied Photophysics stopped-flow device (see chapter 2).

Figure 6. 13 Positions of Trx-DrrA tryptophan mutations



Positions of mutations introduced into pTrx-DrrA by oligonucleotide directed mutagenesis (section 6.9 and chapter 2). Mutated residues shown in bold.

Red text = Vector encoded sequence
Turquoise text = DrrA Walker A motif
Green text = DrrA ABC transporter signature-sequence
Blue text = DrrA Walker B motif

Light at 285nm was selected for specific excitation of tryptophan fluorescence using two serially connected monochromators. Changes in protein fluorescence above 320 nm were measured using a 320nm cut-off filter covering the window of the detection photomultiplier tube. All changes in fluorescence were measured relative to an initial 4 Volt signal to allow percentage changes in protein fluorescence to be calculated.

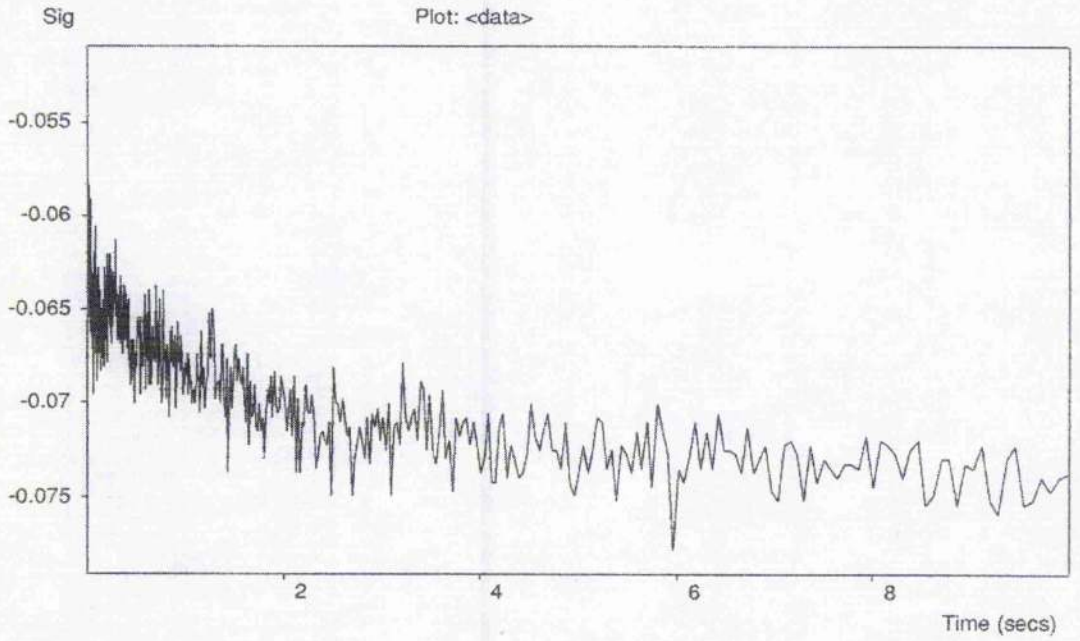
As the data in figure 6.14 show, the ATP binding signal for both proteins was extremely low, a change of much less than 0.1% of the total protein fluorescence in both cases. Whilst minor differences in quality and size of signal can be seen for the two proteins, notably on the time-scales over which they occur, little useful information can be derived from such data. The major problem encountered during data collection was a consistently low signal-to-noise ratio. Neither the magnitude nor rate of changes in tryptophan fluorescence could be shown to vary systematically with changes in the concentration of ligand (ATP). As such, the small decreases in protein fluorescence observed proved impossible to unambiguously assign as being the direct result of protein binding to ATP.

In an attempt to circumvent the problems of the low signal-to-noise ratio observed in the tryptophan fluorescence experiments an alternative approach was adopted. MANT-ATP is a fluorescent analogue of ATP whose fluorescence above 420nm increases in direct response to being bound to proteins. Using the serial monochromators MANT-ATP fluorescence was specifically excited at 255nm in the stopped-flow mixing chamber. Emission data were collected above 420nm using an appropriate cut-off filter. However, as the data in figure 6.15 shows, very little signal was observed upon mixing 100μM MANT-ATP with Trx-DrrA (10μM) in the stopped-flow device. A similar lack of signal was observed when the same mixing experiments were performed in a standard spectrofluorimeter under steady-state conditions (data not shown).

The failure of the stopped-flow experiments to generate usable data is difficult to reconcile with the known ATP-binding properties of Trx-DrrA (as demonstrated by its ATPase activity). Possible explanations for these observations are put forward in section 6.11.

Figure 6. 14 **Changes in protein fluorescence of Trx-DrrA(Y17W) and Trx-DrrA(Y141W) associated with ATP binding**

Y17W



Y141W

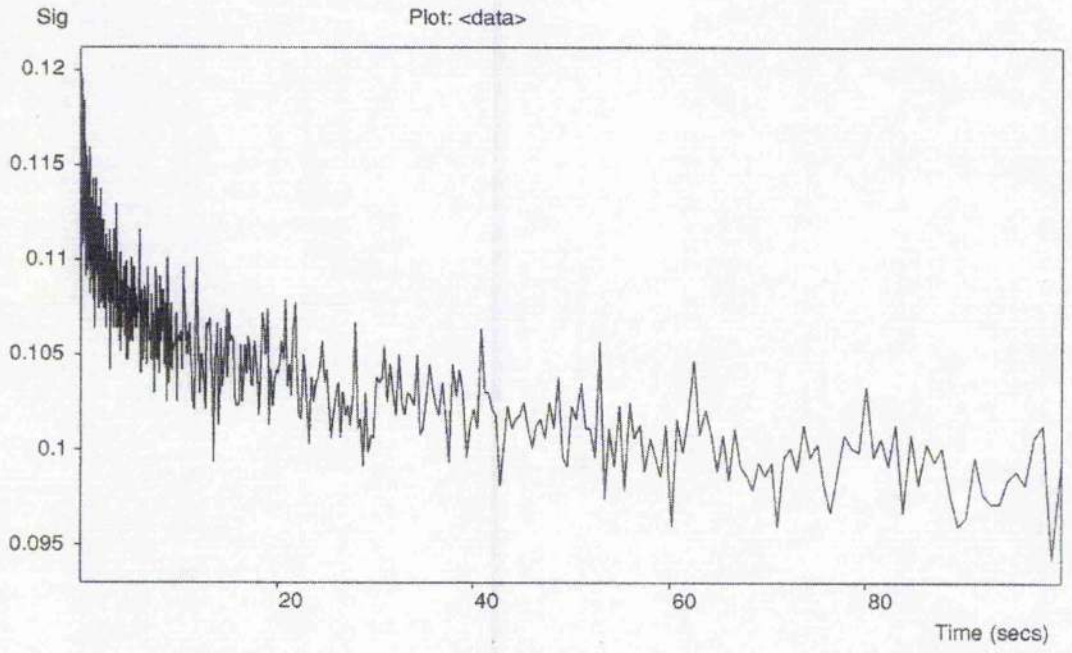
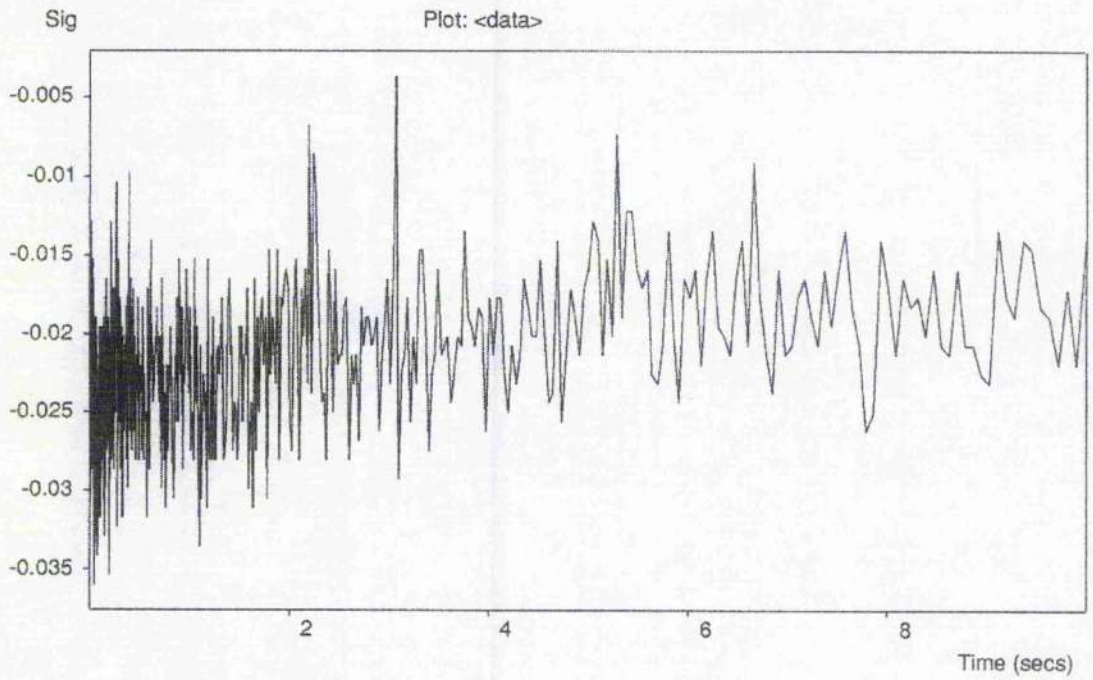


Figure 6. 15 **Changes in MANT-ATP fluorescence upon mixing with Trx-DrrA**



Stopped-flow fluorescence trace generated by mixing 10μM Trx-DrrA (wild type) with 100μM MANT-ATP + 1mM $MgCl_2$ (mixing chamber concentrations). MANT-ATP fluorescence was excited at 255nm and emission data collected above 420nm (λ_{max} MANT-ATP = 448nm). All reagents were prepared at twice the mixing chamber concentration in a buffer of 50mM MOPS (pH 7.4), 100mM NaCl.

6. 11 Discussion

Many of the data presented in this chapter support a working hypothesis in which the *M. tuberculosis* DrrA protein comprises the ATP-binding sub-unit of an ABC transporter. Expressed as a soluble fusion protein in the absence of its cognate MSD's DrrB and DrrC, DrrA was shown to retain a number of biochemical characteristics typical of ABC transporter NBD proteins.

Firstly, Trx-DrrA displayed a measurable catalytic activity consistent with a role in energising transmembrane transport. The protein was shown to be a cation dependent ATPase capable of binding and hydrolysing ATP in a similar manner to the NBD proteins of well-characterised prokaryotic ABC transporters. Whilst Choudhuri and colleagues have previously shown the native DrrA protein is able to bind ATP (Choudhuri *et al.*, 2002), the data presented here are the first to show functional ATP hydrolysis. The calculated V_{max} of ATP hydrolysis by Trx-DrrA was somewhat lower than that described for other prokaryotic NBD proteins. However, the affinity of the protein for ATP, as measured by the K_m value, lies in a very similar range to those same NBD's (see table 6.16 below).

Table 6. 16

Protein	K_m (μM)	V_{max}*	Reference
Trx-DrrA	84.8	52	This Study
HisP (*S. typhimurium*)	205	500	Nikaido *et al.*, 1997
PstB (*M. tuberculosis*)	71.5	122	Sarin *et al.*, 2001
MalK (*E. coli*)	23.9	322	Morbach *et al.*, 1993

* nmoles P_i/min/mg

On the basis of the kinetic analyses performed we were not able to find any direct evidence of cooperativity with regard to ABC hydrolysis, Trx-DrrA appearing to behave as a typical Michaelis-Menten type enzyme. Whilst such cooperativity *has* been demonstrated for a few isolated NBD proteins, e.g. Mdl1p from *S. cervevisiae* (Janas *et al.*, 2003), it is by no means a universal property of all NBD's assayed. For example, neither the MalK nor HisP proteins display cooperative ATP hydrolysis when assayed in

isolation, yet appear to behave cooperatively when part of an intact transporter complex (Nikaido *et al.*, 1997). This raises the important point that NBD behaviour assayed in isolation may not necessarily be a true reflection of their behaviour when forming part of a complete ABC transporter complex. A great deal of data now suggest that ATP hydrolysis is performed by NBD dimers which adopt a similar structure to the Rad50cd dimer crystallised by Hopfner and colleagues (Hopfner *et al.*, 2000). Excellent evidence in support of such a model was provided by Janas *et al.* in their examination of the Mdl1p NBD. They successfully demonstrated the existence of stable NBD dimers trapped at various stages in the ATP hydrolysis cycle (Janas *et al.*, 2003). The formation of such dimers, in which a composite ATP-binding site is formed by juxtaposition of the Walker A region of one NBD with the 'LSGGQ' motif of a second NBD, is further supported by the work of Davidson and Fetsch. This group used photocleavage experiments to demonstrate the proximity of these two structural motifs during ATP hydrolysis by MalK (Fetsch and Davidson, 2002). The relatively low ATPase activity of Trx-DrrA, and the absence of cooperativity, may be perhaps explained in terms of a reduced ability of the protein to form ATPase competent dimers. The presence of the thioredoxin moiety near to the N-terminal of DrrA, where the Walker A sequence is located, is potentially a steric hindrance to the formation of dimers. The reduced ATPase activity of the Trx-DrrA Walker mutants G47S and K48E clearly demonstrates the importance of this particular region of the protein to the overall ATPase mechanism. An alternative explanation for the low activity and absence of cooperativity is that DrrA monomers have a very weak affinity for one another. Under such circumstances dimer formation may require very high localised concentrations of the NBD monomers, these conditions only being satisfied in the complete transporter complex.

The failure to demonstrate antibiotic stimulated ATPase activity for Trx-DrrA provides another example of the potential benefits of working with complete ABC transporter complexes. Choudhuri and colleagues successfully demonstrated enhanced ATP binding to native *M. tuberculosis* DrrA in the presence of doxorubicin, one of the potential antibiotic substrates of the transporter. This work was however performed using membrane vesicles containing both DrrA *and* one of the putative MSD components of the transporter, DrrB. Combination of that data with the absence of substrate stimulated ATPase activity presented here suggests that any transporter-substrate interactions take place via the MSD. Under such circumstances enhancement of ATP binding must

presumably be achieved via conformational changes transmitted from one domain to the other.

Fluorescence spectroscopy has proven to be a powerful tool in the characterisation of ABC transporter behaviour. The combined use of intrinsic protein fluorescence and external fluorescent probes has allowed researchers to gather important data regarding reaction mechanisms, substrate binding, transport rates and the inter-domain interactions of ABC transporters (e.g. see Borges-Walmsley *et al.*, 2003). Unfortunately, the Trx-DrrA fusion protein proved remarkably resistant to investigation using these techniques. The major problem encountered in both the steady state and transient binding studies using Trx-DrrA was that the protein exhibited a very low signal-to-noise ratio during the investigations. Despite adopting a rational approach to the construction of single-tryptophan mutant proteins, this problem was never sufficiently overcome to allow the generation of useful data.

A number of scenarios can be put forward to explain the failure of these experiments. Firstly, it remains possible that no significant structural rearrangements take place during the ATP-binding and hydrolysis cycle of DrrA. This however seems unlikely given the functional role of ABC transporter NBD's, i.e. the transduction of energy from ATP hydrolysis into the gross conformational changes of the transporter required to drive transmembrane transport. Additionally, differences in the crystal structures of ATP and ADP bound forms of NBD proteins suggest the existence of specific nucleotide-induced structural alterations (Yuan *et al.*, 2001). Moreover, structural changes induced in the *S. typhimurium* MalK protein upon ATP binding have been correlated with a change in proteolytic cleavage pattern, once again suggestive of significant conformational rearrangements (Schneider *et al.*, 1994). A second, more plausible, explanation for the lack of a strong signal upon ATP binding to Trx-DrrA relates to the presence of the thioredoxin moiety and the low ATPase activity of the protein. If NBD dimers are indeed the functional form giving rise to ATPase activity, it may be that only a small proportion of molecules in a given situation are able dimerise due to the steric effects of the thioredoxin moiety. Under such circumstances most molecules would not contribute to the observed ATPase activity (or give rise to a fluorescent signal). Any observed binding signal or activity may be attributable to a much smaller number of active ATPase dimers.

This interpretation is supported by the absence of a strong signal for the binding of the external fluorescence probe MANT-ATP to Trx-DrrA (see section 6.10).

On the whole it seems likely that the *M. tuberculosis* DrrA protein is indeed the NBD of an ABC transporter, both sequence data and biochemical data supporting this interpretation. However, a more complete understanding of its catalytic behaviour will probably require the overexpression and purification of a native form of the protein lacking an N-terminal fusion partner. To this end, the publication by Choudhuri and colleagues of a system for producing native DrrA in *E. coli* is of particular interest, especially given that the same system may potentially be adapted to co-express the MSD's of the transporter.

Chapter 7

Investigating the physiological function of the *M. tuberculosis* DrrABC system

Previous chapters have described attempts to over-express and purify the *M. tuberculosis* DrrABC proteins as individual entities. However, the fact that all three genes of the *drr* system are present in a single, transcriptionally-linked operon suggests that the physiological roles of the Drr proteins are closely tied to one another at a functional level as well as a genetic one. Biochemical evidence presented in the previous chapters, such as the membrane localisation of DrrB and the catalytic activity of DrrA, support a model of DrrABC function in which DrrA constitutes the energy transducing sub-unit of an ABC transporter (the NBD), whilst DrrB contributes to a transmembrane translocation pathway. Expression of the third gene, *drr*C, could not be unambiguously demonstrated in *E. coli*, however its presence in the same operon strongly suggests that it too is part of the same transporter complex. The hydrophobic nature of the DrrC protein, and its homology to DrrB (see chapter 3), is an indication that this protein may well be a second transmembrane component of the transporter. Given the fact that all ABC transporters characterised to date have a minimum of four functional domains, the *M. tuberculosis* DrrABC system seems most likely to operate as 'A_2BC' complex similar to the $MalK_2FG$ and $HisP_2MQ$ transporters of *E. coli* and *S. typhimurium* respectively.

The experiments described in this chapter were aimed at establishing simultaneous co-expression of all three *M. tuberculosis drr* genes in the heterologous host *E.coli*. In an attempt to identify possible physiological roles of the DrrABC transporter, use was made of mutant *E. coli* strains defective in certain transmembrane transport capabilities. Throughout this work a number of experimental difficulties were encountered, primarily relating to the large evolutionary distance between *E. coli* and *M. tuberculosis*. Much of the work in this chapter was performed before the publication of an 'artificial operon' system designed for co-expression of the DrrA and B proteins (Choudhuri *et al.*, 2002). A number of interesting parallels can be drawn between the work presented here and the results of Choudhuri and colleagues. These comparisons also demonstrate how differences in experimental systems used to measure ABC transporter activity can have a significant bearing on the results obtained.

7. 1 Heterologous co-expression of DrrA, B and C in *E. coli* using the pET-based expression construct pDrrABC

Initial attempts to achieve co-expression of the Drr proteins were performed using the pDrrABC expression construct described in chapter 3. This plasmid, based upon the pET-33 vector, contained the entire 2.7Kb *M. tuberculosis drr*ABC operon cloned between the *Nde*I and *Hind*III restrictions sites of the vector MCS. Whereas expression of the individual Drr proteins had used pET-21 based plasmids which encode a single C-terminal His-Tag, pET-33 allowed for tagging of two of three genes present in the DrrABC operon. DrrA, the first gene of the operon, was fused to a vector-encoded N-terminal His-Tag sequence, whilst DrrC, the final gene, was fused to a similar tag at its C-terminal end. A schematic representation of the pDrrABC construct is shown in figure 7.1. As is the case with all pET-series vectors, transcription of the inserted operon was to be driven by a bacteriophage T7 promoter located upstream of the cloned sequence. In previous studies the DrrB protein had been shown to be prone to aggregation when expressed and purified in isolation, whilst no conclusive evidence for successful over-expression of DrrC had ever been obtained. These results suggested that the membrane bound components of the transporter were particularly unstable and/or sensitive to degradation within the cell. It was hoped that co-expression of all three genes might lead to improved stability of these proteins. Similar co-expression studies have in the past facilitated the purification of complete ABC transporter complexes and allowed for demonstration of ATP-dependent transport activity in cell-free systems (Ames *et al.*, 2001).

Expression trials were conducted by transforming the pDrrABC plasmid construct into the *E. coli* host strain BL21(DE3), the DE3 lysogen providing the IPTG inducible T7 RNA polymerase necessary to drive transcription of the recombinant DNA sequence. Recombinant colonies were selected on solid LB-agar medium containing kanamycin at 50μg/ml. A single recombinant colony was used to inoculate 100ml of fresh liquid LB medium containing kanamycin at the same concentration. The liquid culture was grown at 30 °C with rotary shaking at approximately 200rpm. It was considered that a reduced growth temperature might help to improve the stability of the DrrB and C proteins whilst reducing inclusion body formation by the aggregation prone DrrA.

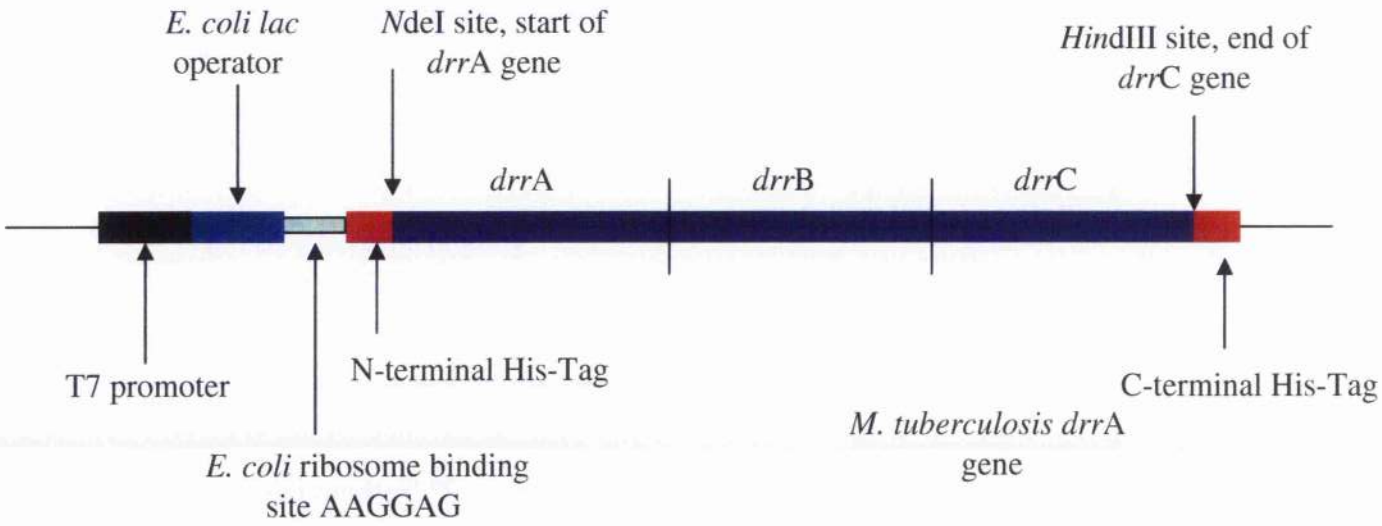
Expression of the *drr*ABC operon was induced by the addition of 1mM IPTG to the liquid culture when the cells reached an OD_{600} of 0.6 (mid log phase). After induction the cells were grown for a further 3 hours at 30 °C before being harvested by centrifugation (6000rpm, Beckman JA-10 rotor).

Both structural and over-expression studies suggest that, in a complete four-domain ABC transporter complex, the NBD's are peripherally associated with the cytoplasmic membrane through specific interactions with the MSD's. As such, successful over-expression of the *drr* operon might be expected to result in the presence of a membrane-localised complex containing all three Drr proteins. To determine whether or not this was the case total cell membranes were isolated from the expression culture as described in chapter 2. All membrane bound proteins were solubilised in a buffer containing 8M urea and then separated by SDS-PAGE. Given the low expression levels of the DrrB and DrrC proteins encountered in previous studies, the decision was made to probe for the presence of the membrane bound complex by Western blot. The SDS-PAGE separated proteins were transferred to a PVDF membrane by electro-blotting and probed using an anti-His_6 monoclonal antibody to detect the His-tagged DrrA and DrrC proteins. The results of this blot along with experimental details of its preparation are shown in figure 7.2.

The results of the Western blot were initially encouraging. A strong signal was observed at approximately 40 kDa, a size consistent with His-tagged DrrA. The presence of significant quantities of DrrA being associated with the cell membrane was also an indication of successful expression of the membrane bound DrrB and C proteins. Two much smaller signals were also observed at a size just below 30 kDa. Either of these signals might possibly be attributed to smaller quantities of His-tagged DrrC being present in the *E. coli* membrane.

The need to confirm the identity of the proteins giving rise to positive Western blot signals led to a number of attempts to purify the proteins concerned by IMAC chromatography. The small signals observed below 30 kDa, putatively assigned to DrrC, suggested that expression levels of this protein were very low. This factor, combined with the known instability of purified MSD's, dictated that purification experiments be performed under denaturing conditions so as to limit yield losses due to proteolysis or

Figure 7. 1 **The pDrrABC expression construct**



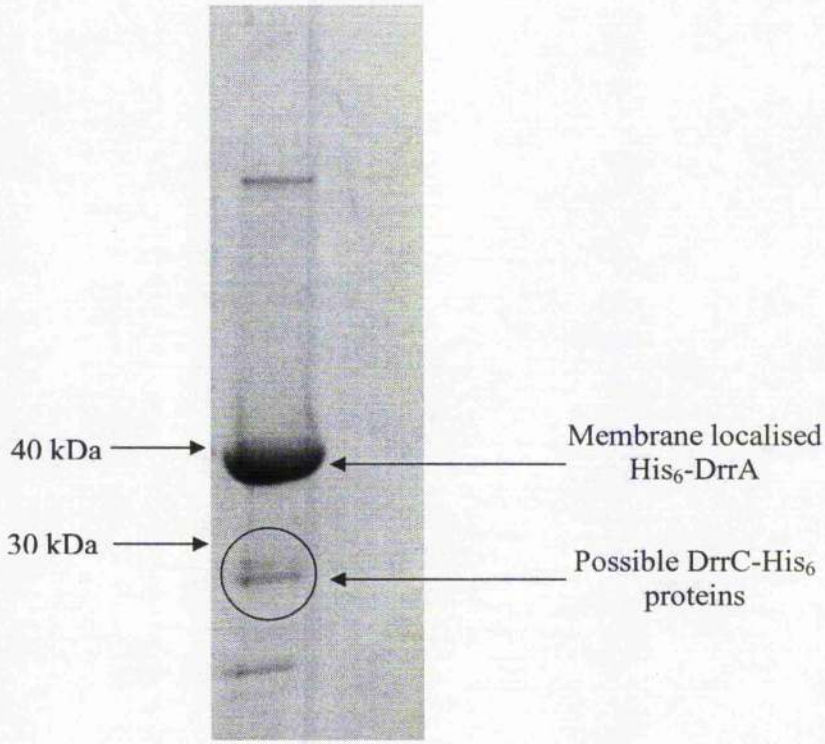
Additional pET-33 vector encoded features include: -

Transcriptional start immediately after T7 promotor

Constitutively expressed kanamycin resistance marker

Bacteriophage T7 transcriptional terminator downstream of C-terminal His-Tag

Figure 7. 2 **Western blot of total cell membrane protein isolated from a pDrrABC expression culture**



Membrane associated proteins from a 100ml pDrrABC expression culture were extracted into 5ml of the following buffer: -

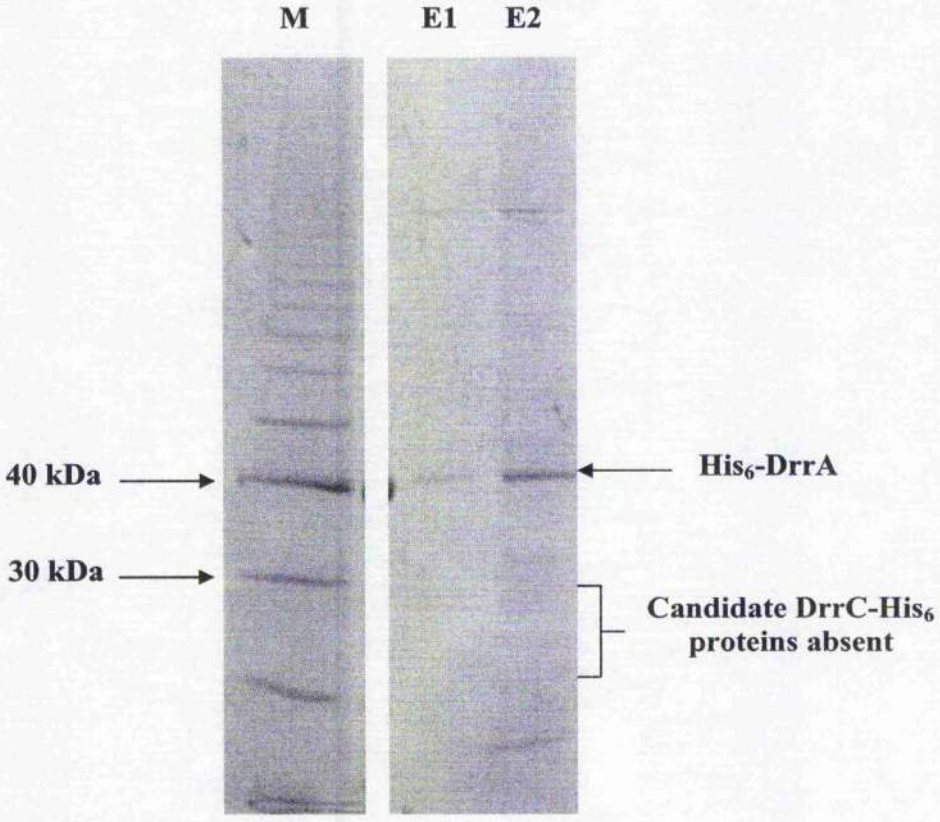
50mM Tris-HCl (pH 8.0), 300mM NaCl, 8M Urea.

The high urea concentration of the buffer resulted in disruption of the lipid membranes and solubilisation of associated proteins. 30µl of the urea extract was mixed with 10µl of SDS-PAGE sample buffer and the proteins separated by electrophoresis on a 12% polyacrylamide gel. Separated proteins were transferred to PVDF membrane by electro-blotting and the resulting blot probed with an anti-His_6 monoclonal antibody. The blot was developed using the ImmuneStar chemiluminescence detection system (see chapter 2).

aggregation. A number of purification conditions were tested; typically using total cell membranes isolated from a 1 litre expression cultures. Membrane proteins were extracted and solubilised in 8M urea buffer and the resulting preparations applied to Ni-NTA spin columns. The column resin was washed extensively according to the manufactures recommended protocols before bound protein was finally eluted in urea buffer supplemented with 600mM imidazole (see chapter 2). The results of one such purification experiment are shown in figure 7.3.

A 40kDa protein, believed to be His-tagged DrrA, was consistently isolated from membrane extracts prepared from cells containing the pDrrABC plasmid. However, neither of the approximately 30 kDa proteins identified in Western blots of the same membranes could be purified from these extracts. A number of explanations for these apparently conflicting results can be suggested. Firstly, the failure to purify any His-tagged DrrC may simply be a reflection of its extreme instability both within the cell and once it is removed from the membrane environment. A second and perhaps more likely explanation is that the 30 kDa Western blot signals are experimental artefacts, and that little or no DrrC expression occurred using this expression construct. This latter interpretation is supported by the earlier results presented in chapter 4 where apparently positive Western blot signals attributed to DrrC were later confirmed as being the result of a histidine-rich *E. coli* protein. The failure of the pDrrABC expression construct to express all three mycobacterial genes might also be explained in terms of genetic differences between *E. coli* and *M. tuberculosis*. A vector-encoded T7 promoter drives transcription of the *drr*ABC operon from pDrrABC via an IPTG induced RNA polymerase. However, *translation* of the mRNA signal is entirely dependent on the host cell ribosomal machinery. The overlapping nature of the *drr*ABC operon means that translation of the *drr*B and *drr*C genes would require *E.* coli ribosomes to recognise binding sites encoded within the mycobacterial DNA. The failure of *E. coli* ribosomes to recognise mycobacterial binding sites would simply result in the ribosomes 'falling-off' or failing to bind the mRNA transcript, thereby leading to loss of expression. Translation of *drr*A would be unaffected by this problem since an optimised *E. coli* ribosome binding site is encoded in the vector sequence upstream of *drr*A.

Figure 7. 4 **Attempted co-purification of the DrrA and DrrC proteins by IMAC chromatography under denaturing conditions**



Total membrane proteins from a 1 litre expression culture of BL21(DE3)/pDrrABC were solubilised in 5ml of a buffer containing 50mM Tris-HCl (pH 8.0), 300mM NaCl, 8M urea. The resulting extract was applied to a Ni-NTA spin column and unbound protein removed by washing with 5x 800µl of the same buffer. Non-specifically bound proteins were removed by washing the column resin with 5x 800µl of buffer supplemented with 50mM imidazole. Finally, His_6-tagged proteins were eluted from the resin by washing twice with 400µl of buffer containing 600mM imidazole. Eluted proteins were separated by SDS-PAGE and visualised by Coomassie blue staining (Lanes E1 and E2).

The results show elution of the 40 kDa N-terminal His-tagged DrrA only. No His-tagged proteins with molecular weights less than 30 kDa were purified by this method. These results suggested that there was little or no expression of C-terminal His-tagged DrrC in these cultures. All of the expression and purification experiments described here were performed before the publication of Choudhuri and colleagues. This group encountered an exactly analogous series of problems during their attempts to express just two of the *M. tuberculosis drr* genes, A and B. Expressed independently of one another using pET-based constructs, DrrA was found exclusively as inclusion bodies, and the DrrB protein localised to the membrane yet was impossible to extract due to its instability. These results are very similar to those described in chapter 4 of this work. Additionally, expression of the *drr*A and B genes in tandem with one another resulted in the detectable expression of only the *drr*A gene product. To circumvent the co-expression problem Choudhuri constructed an artificial operon in which the *drr*A and B genes were separated from their native overlapping structure and an artificial *E. coli* ribosome binding site incorporated between the end of the *drr*A gene and the beginning of *drr*B. Only when this construct was used were they able to demonstrate the stable, membrane associated expression of both the DrrA and B proteins (Choudhuri *et al.*, 2002).

The experiments of Choudhuri and colleagues are very interesting in that they not only describe a potentially successful route to the heterologous co-expression of all three *M. tuberculosis drr* genes, but also further demonstrate the interdependence of ABC transporter sub-domains on one another for stability. Additionally their work illustrates that differences in the genetics of *E. coli* and *M. tuberculosis* can be responsible for the failure of heterologous expression.

7. 2 Antibiotic sensitive *E. coli* bearing the pDrrABC plasmid do not exhibit increased resistance to anthracycline antibiotics

Despite failure to co-purify the components of a membrane bound DrrABC complex, a number of further experiments were performed to investigate whether co-expression of the DrrABC proteins might be detected on a phenotypic basis. Sequence homologies between the DrrABC proteins and components of the *S. peucetius* doxorubicin/ daunorubicin efflux transporter are one indication that the two transporters may share a similar function. In order to identify whether cells carrying the pDrrABC plasmid

exhibited increased resistance to the anthracycline antibiotics, a host strain sensitive to their action was required. Standard laboratory *E. coli* strains are intrinsically resistant to both doxorubicin and daunorubicin by virtue of their possessing a number of chromosomally encoded multidrug efflux pumps. Foremost among these is the AcrB multidrug resistance pump, a member of the RND family of secondary transmembrane transporters (Ma *et al.*, 1995). Two further accessory proteins are vital to the function of this pump, the AcrA membrane fusion protein and the TolC outer membrane channel (Elkins and Nikaido, 2003). *E. coli* strains deleted for any of these three proteins exhibit massively increased susceptibility to a wide range of antibiotics. In a systematic study of the contributions different efflux systems make to the intrinsic drug resistance of *E. coli*, Sulavik and colleagues report that both Δ*acr*AB and Δ*tol*C knockout strains exhibit doxorubicin resistance levels of below 1μg/ml. This compares with a resistance level of >80μg/ml for wild-type strains (Sulavik *et al.*, 2001).

In order to enable a direct comparison with the *S. peucetius* DrrAB transporter, a Δ*acr*A *E.coli* strain called 'N43' was chosen as the host organism for the pDrrABC plasmid. This strain had been successfully used by Kaur to demonstrate the antibiotic transport capabilities of the *S. peucetius* doxorubicin efflux pump (Kaur, 1997). *E.coli* N43 does not possess a chromosomal copy of the DE3 lysogen required to drive transcription from pET-series plasmids. This was provided by co-transformation of the mutant strain with a second compatible plasmid called pT7 (Aldema *et al.*, 1996). This plasmid encodes an IPTG inducible T7 RNA polymerase as well as an ampicillin resistance marker to enable selection of doubly transformed cells. Both the pDrrABC and pT7 plasmids were transformed into chemically competent N43 and recombinant colonies were selected on solid LB-agar medium containing 25μg/ml kanamycin and 50μg/ml ampicillin.

The antibiotic susceptibilities of both the parent N43 strain and the transformed strain were determined in terms of a *Minimum Inhibitory Concentration* (MIC). An MIC is defined as the lowest concentration of antibiotic that causes a particular bacterial strain to cease growing. MIC's were determined by the broth microdilution method recommended by the National Committee for Clinical Laboratory Standards (see chapter 2). Antibiotic susceptibilities with regard to a range of structurally unrelated compounds were tested. The results of these experiments are shown in table 7.5.

Table 7.5 MIC Determinations

	Strain MIC (μg/ml)		
Drug	**N43**	**N43/pDrrABC (-)***	**N43/pDrrABC (+)***
Doxorubicin	4	4	4
Daunorubicin	4	4	4
Erythromycin	4	4	2
Rhodamine-6G	16	8	8
Ethidium Bromide	80	40	40
SDS	100	100	50

* +/- 1mM IPTG

The data presented above represent average values from between 4 and 6 independent experiments with each drug.

The MIC data obtained suggested that the presence of the pDrrABC plasmid had no positive effect on levels of antibiotic resistance. In some cases, e.g. the dyes rhodamine-6G and ethidium bromide, the plasmid even appeared to cause a lowering of the MIC. Addition of 1mM IPTG to the cultures had no effect. The MIC's of daunorubicin and doxorubicin were both in the region of 4μg/ml, a value consistent with that reported by Kaur for *E. coli* N43. These data would tend to support the idea that there is little or no expression of a complete/functional DrrABC transporter from the pDrrABC plasmid.

At this point it is worth comparing these data with those obtained by Choudhuri and colleagues. Using their artificial operon system for co-expression of the *drr*A and B genes, this group report apparent increases in resistance towards both doxorubicin and daunorubicin. The MIC for doxorubicin rose from 20μg/ml to 40μg/ml, while the MIC for daunorubicin increased from 20μg/ml to 60μg/ml. Co-expression of DrrA and B also enhanced resistance towards ethidium bromide (where the MIC increased from 20μg/ml to 60μg/ml). It should be noted however that these experiments were performed using a standard *E. coli* host strain, BL21(DE3), and that the MIC figures reported are all

somewhat lower than those generally accepted for laboratory strains of *E. coli* (Kaur, 1997; Sulavik *et al.*, 2001).

7. 3 The pDrrABC plasmid does not complement a temperature sensitive mutation in the *E. coli* MsbA ABC transporter

The MsbA ABC transporter is a homodimeric half-transporter responsible for export of a major lipid component of the *E. coli* outer membrane, 'Lipid A'. In 2001 Chang and Roth produced a high-resolution crystal structure for the MsbA homodimer making it the first ABC transporter for which such information was available (Chang and Roth, 2001). In a further attempt to identify a specific transport function associated with the *M.tuberculosis drr*ABC genes, use was made of a temperature sensitive MsbA mutant. This strain, identified by Doerrler, grows normally at 30 °C but is unable to do so at the elevated temperature of 44 °C. The biochemical basis of this growth defect was identified as a failure to target MsbA to the cytoplasmic membrane at the higher temperature (Doerrler *et al.*, 2001). Although *E. coli* and *M. tuberculosis* are evolutionarily diverse bacteria with very different cellular structures, the broad substrate specificity of some ABC transporters towards lipophilic molecules suggested that the DrrABC transporter might be able to complement the temperature sensitive defect in Lipid A export.

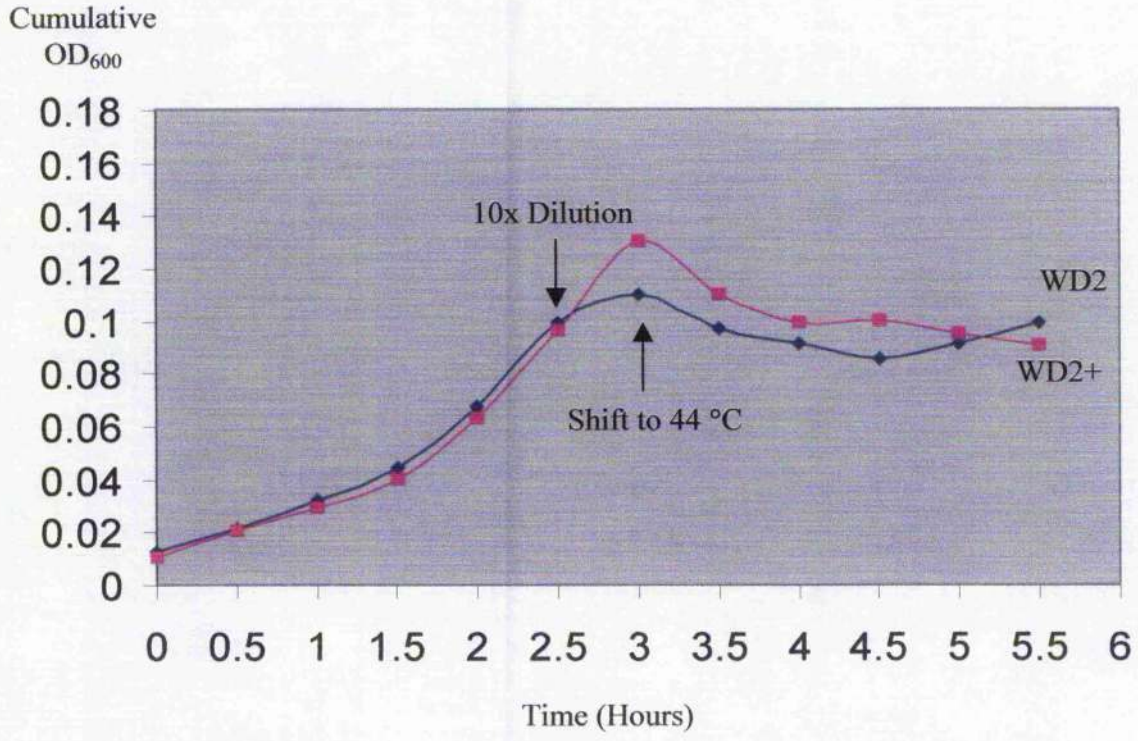
The temperature sensitive mutant strain, WD2, was co-transformed with the pDrrABC and pT7 plasmids as described in the previous section. Transformants were selected on solid LB-agar medium containing 25μg/ml kanamycin and 50μg/ml ampicillin. The temperature sensitive nature of the WD2 strain required that transformation plates be incubated at 30 °C rather than the normal 37 °C.

A single colony of the doubly transformed WD2 strain (WD2+) was used to inoculate 10ml of fresh LB medium supplemented with both kanamycin and ampicillin. At the same time an untransformed WD2 colony was used to prepare a control culture (LB only). Both liquid cultures were incubated at 30 °C with rapid rotary shaking. Growth of the cultures was monitored by an increase in OD_{600} with measurements taken at 30 minute intervals. The cultures were allowed to grow for 2.5 hours at 30°C in order to establish that their growth was normal at the permissive temperature. At this point each

culture was diluted 10x into fresh medium pre-warmed to 30 °C and 1mM IPTG was added to the WD2+ strain. After a further 30 minutes of growth at 30 °C both liquid cultures were transferred to a 44 °C incubator and the effect on growth monitored by changes in OD_{600}. The growth curves obtained for both cultures are shown in figure 7.6.

As the data show, the pDrrABC plasmid had no effect on the growth of WD2 at either the permissive temperature or the elevated temperature. Both the test and control strains stopped growing almost immediately after they were warmed to 44 °C. The same experiment was repeated a number of times with essentially no difference in the results observed. These data can only be interpreted as either failure of the DrrABC transporter to support Lipid A export, or, more likely in the light of similar results, a failure of the pDrrABC expression construct to produce all three mycobacterial proteins.

Figure 7. 6 **Growth of *E. coli* lipid efflux mutants WD2 and WD2+**



E. coli strains WD2 (blue line) and WD2+ (pink line) were grown at 30 °C in 10ml liquid cultures. Growth of the cultures was monitored by increases in OD_{600} (y-axis). After 2.5 hours each culture was diluted 10x into fresh, pre-warmed medium (in order to prevent the emergence of revertant organisms), and 1mM IPTG was added to the WD2+ culture to induce expression from the pDrrABC plasmid. At 3 hours the cultures were shifted to 44 °C and the effect on growth was monitored for a further 3 hours. All OD figures after 2.5 hours are corrected for the dilution effect.

Both cultures ceased growing almost immediately after the shift to 44 °C indicative of incorrect targeting of the *E. coli* MsbA Lipid A transporter. The presence of the pDrrABC plasmid in the WD2+ organisms clearly had no effect on growth rate of the culture either before or after the shift to 44 °C. These data suggest either a lack of expression from the pDrrABC plasmid, or an inability of the mycobacterial DrrABC proteins to support *E. coli* Lipid A efflux.

7. 4 Discussion

The various negative results reported in this chapter can largely be attributed to failure of the pDrrABC expression construct. They do however demonstrate some of the wider difficulties associated with ABC transporter research and with heterologous expression systems in general. ABC transporters are distributed throughout all biological kingdoms and their activity has been correlated with a range of processes having a direct impact on human health. The emerging phenomenon of bacterial multidrug resistance, the failure of anticancer chemotherapy and human genetic diseases have all been linked to the function of members of the ABC transporter superfamily. An ability to combat these problems will increasingly rely on detailed structural and functional data relating to ABC transporter activity. Such data can only be generated if sufficient source material is available for scientists to work with, and, as the work presented here and elsewhere has shown, generating such source material is far from routine. Researchers are still heavily reliant on *E. coli* as a host for recombinant protein expression, yet this organism is by no means ideal for the expression of proteins from evolutionary diverse sources such as eukaryotes or the mycobacteria. Moreover, ABC transporters are by their very nature multi-domain proteins, the function and stability of the various domains being closely tied to one another. If researchers are to continue to rely on *E. coli* as a vehicle for recombinant expression then improved systems for co-expression and correct targeting of multi-protein complexes must be developed.

The work of Choudhuri and colleagues is an example of a modified expression system that proved to be very effective for the co-expression of the *M. tuberculosis drr*A and B genes (Choudhuri *et* al., 2002). Their experiments clearly establish successful expression of both the mycobacterial *drr*A and *drr*B genes in *E. coli* through the use of an artificial operon system. Moreover, the DrrA and B proteins produced were shown to form a complex associated with the cytoplasmic membrane, an architecture entirely consistent with the hypothesis that these proteins constitute components of an ABC transporter complex.

Building on the strength of their expression data, Choudhuri and colleagues presented additional experimental results that suggested a possible *functional* role for the DrrA and B proteins. Firstly, the accumulation of [^{14}C]doxorubicin in *E. coli* cells co-expressing DrrA and B was shown to be significantly lower than in control strains. Moreover,

accumulation of [^{14}C]doxorubicin was shown to revert to control levels when cells expressing DrrAB were exposed to sub-lethal concentrations of *reserpine*, a known inhibitor of ABC transporter function (Choudhuri, 2002). These data would seem to indicate that the mycobacterial DrrA and B proteins form a functional ABC transporter complex very similar to the *S. peucetius* DrrAB transporter described by Kaur i.e. an efflux ABC transporter capable of extruding anthracycline antibiotics (Kaur, 1997).

A second series of experiments described in the same paper lend further weight to the hypothesis that the DrrAB complex is a functional efflux transporter. The probe molecule BCECF-AM is a non-fluorescent, neutral compound able to diffuse freely across cellular membranes. Once inside a cell BCECF-AM is rapidly hydrolysed to a free acid (BCECF) that is fluorescent and no longer able to cross the cytoplasmic membrane. The change in fluorescence of a cellular preparation to which BCECF-AM has been added can therefore be used as a direct measure of the intracellular accumulation of BCECF. When BCECF-AM was added to *E. coli* cells co-expressing the mycobacterial DrrA and B proteins, BCECF accumulation was markedly reduced compared with control cells. Once again the ABC transporter inhibitor reserpine was shown to reduce the magnitude of this effect. Choudhuri and colleagues interpret these data as an indication that BCECF-AM is a substrate of the DrrAB transporter complex. This raises the interesting possibility that the DrrAB transporter is capable of mediating the efflux of at least two distinct substrate types i.e. cationic, hydrophobic molecules such as doxorubicin *and* neutral compounds such as BCECF-AM.

The reduced accumulation of [^{14}C]doxorubicin, reduced accumulation of BCECF and increased antibiotic resistance exhibited by *E. coli* cells co-expressing DrrA and B are all strong indications that these two mycobacterial proteins combine to generate a functional ABC efflux transporter. Given the four-domain architecture of all known ABC transporters, Choudhuri's observations imply the formation of a $DrrA_2B_2$ tetramer exactly analogous to that proposed by Kaur for the *S. peucetius* doxorubicin transporter. However, such a model does not take into account the presence of the *drr*C gene in the same operon as *drr*A and B, and nor does it correlate with observations of Azad and colleagues suggesting that the DrrC protein plays a vital role in the export of mycobacterial cell wall lipids (Azad *et al.*, 1997).

The two groups of experimental data regarding the physiological function of the *M. tuberculosis* DrrABC transport system are somewhat difficult to reconcile. Whilst it remains possible that the transport complex may be able to adopt two or more different stoichiometries, this phenomenon has not previously been described amongst members of the ABC transporter superfamily.

A more complete understanding of the physiological role of the DrrAB(C) transport system could ideally be achieved by studying the expression of the genes and proteins in their native host, *M. tuberculosis*. However, given the difficulties associated with the growth, handling and genetic manipulation of this organism, it seems likely that researchers will remain reliant on *E. coli* as a vehicle for expression and functional studies. Against this experimental background, the co-expression and assay systems reported by Choudhuri and colleagues represent significant advances. Similar systems may eventually allow for simultaneous co-expression of all three *M. tuberculosis drr* genes and the development of new assays to assess the true function of this transporter.

Chapter 8

Final Discussion

More than 50 years after the discovery of the first effective anti-tubercular agents, pulmonary tuberculosis remains the leading cause of mortality due to a bacterial pathogen. Around 2-3 million deaths annually can be attributed to *M. tuberculosis* infection, a situation that is likely to worsen over the next decade as a result of the global HIV/AIDS pandemic. Once considered to have been virtually eradicated in the rich, industrialised nations of the West, TB has re-emerged in new and deadlier drug-resistant forms. The increased availability of air travel and a growing number of displaced people have allowed the bacterium to spread rapidly and with no respect for geographical borders. Together these various factors have forced national governments and the scientific community to focus significant efforts on the control and treatment of tuberculosis. Many of the current problems associated with TB treatment are related to an over-reliance on a limited number of chemotherapeutic agents. The four drugs forming the cornerstone of most modern anti-tubercular therapy were all first used in the middle of the last century. Worse still, incorrect usage of these drugs has resulted in the emergence of new MDR strains that have the potential to become untreatable. Halting the spread of TB clearly requires the development of new and effective TB treatments and the introduction of a prophylactic vaccine.

The publication of the complete *M. tuberculosis* genome sequence was a major milestone in TB research (Cole *et al.*, 1998). The genetic sequence of the organism essentially represents a detailed map of every possible drug target and antigen that might be used to combat the spread of the disease. However, this genetic data is not, in itself, enough to ensure the eradication of TB when even some of the most basic facets of mycobacterial physiology remain poorly understood. How are the organisms able to avoid immune clearance and survive within the hostile intracellular environment? Which factors affect mycobacterial dormancy and re-activation of latent infection? Why does *M. tuberculosis* display such a high level of intrinsic drug resistance? Answers to these questions, and many more facing the TB research community, will require a sustained international effort involving scientists in every field from basic microbiology to X-ray crystallography. The work presented in this thesis aimed to address just one tiny aspect of the *M. tuberculosis* puzzle; what is the role of the *drr*ABC operon?

The *M. tuberculosis* genome sequencing project revealed the existence of as many as 37 potential ABC transporter systems (Braibant *et al.*, 2000). These versatile, multi-domain transmembrane transport systems have been identified in all kingdoms of life, from archea through to man, so their presence in the *M. tuberculosis* genome is far from surprising. However, at the outset of these investigations little or nothing was known about the role that ABC transporters play in the biology of *M. tuberculosis*. Much of the interest in ABC transporters derives from their many physiological roles. ABC transporters from various organisms have been shown to be involved in nutrient acquisition, virulence factor export, antibiotic resistance and cellular assembly (see chapter 1.2). Any or all of these processes may yet prove to be important determinants of *M. tuberculosis* pathogenicity.

Of all the potential ABC transporters encoded in the *M. tuberculosis* genome the *drr*ABC system was revealed to be particularly interesting on a number of levels. Bioinformatic tools such as primary sequence analysis were used to show that the proteins encoded in *drr*ABC operon bear significant homology to an antibiotic efflux transporter from the evolutionarily related actinomycete *Streptomyces peucetius*. Active efflux of antibiotics is not known to be one of the major drug resistance mechanisms employed by *M. tuberculosis*, yet the possibility remains that it may be a contributory factor. A second potential function of the *drr*ABC genes was identified by the presence of a poorly conserved but narrowly distributed amino acid sequence known as the 'ABC-2' motif. Found in both the DrrB and C proteins, ABC-2 is restricted to the transmembrane components of ABC transporters involved in the assembly of extracellular structures such as the cell wall and glycocalix (Reizer *et al.*, 1992). Could it be that the *drr*ABC system is somehow involved in transport of molecules contributing to the complex and unique mycobacterial cell wall? This latter hypothesis was lent a great deal of credence by the work of Camacho and colleagues who demonstrated that genetic disruption of the *drr*ABC operon resulted in decreased transport rates for one particular class of complex lipids (Camacho *et al.*, 1999). Perhaps even more importantly, this transport deficiency could be correlated with decreased virulence of the mutant organisms in a mouse model of pulmonary tuberculosis (Cox *et al.*, 1999).

Both potential roles of the *drr*ABC system marked this ABC transporter out for further study. On one hand *drr*ABC is in itself an attractive drug target because of a possible role in cell wall assembly, on the other hand an improved understanding of the basic mechanisms of mycobacterial drug resistance may help us to design effective new chemotheraputics. However useful bioinformatic approaches are for identifying members of the ABC transporter superfamily and identifying *potential* functions, it is at the biochemical level where the most valuable information is to be gained. To this end most of the work presented thesis was aimed at overexpressing and purifying the protein products of the *M. tuberculosis drr*ABC operon.

From the earliest days of recombinant DNA technology and the development of techniques for the overexpression and purification of proteins, researchers have relied heavily on the use of particular model organisms. In the case of prokaryotic protein expression *E. coli* is invariably the organism of choice as no other bacterium is so well characterised at a genetic, biochemical or physiological level. Given the extensive availability of genetic systems for protein overexpression in *E. coli*, this organism was selected as a heterologous host for studies of the *M.* tuberculosis DrrABC proteins. Whilst the use of a rapidly-growing Gram-negative organism as a means of producing proteins native to a slow-growing, Gram-positive pathogen is clearly far from ideal, few other options were available because of the many problems associated with working with live *M. tuberculosis*.

ABC transporters are highly modular in nature and consist of a minimum of four functional domains, two ATP-binding/hydrolysing domains and two membrane-spanning domains (Higgins *et al.*, 1986). The exact mode of expression of these domains varies enormously between different members of the ABC transporter superfamily but amongst the prokaryotic systems a different gene usually encodes each domain. In the case of the *drr*ABC operon, sequence analysis indicated that *drr*A gene encoded the nucleotide-binding domains of the transporter, whilst *drr*B and C each encoded membrane-spanning domains. Throughout the history of biochemical research it has been common practice to investigate the properties of individual components of a system before analysing how the various components contribute to the whole. Such an approach has been highly successful in studies of the ABC transporter family, particularly in understanding how the various domains contribute to substrate translocation (Davidson *et al.*, 1996). For this reason the

same approach was initially adopted with regard to the DrrABC proteins. Each protein was treated as an individual entity to be expressed in isolation using a separate expression construct. In the case of the DrrA protein this technique initially yielded huge quantities of recombinant protein (using the pET-based construct pDrrA). However, as is so often the case with recombinant overexpression systems, the vast majority of the DrrA produced was insoluble, inactive inclusion bodies. The lack of biological function associated with this form of the protein prevented further biochemical analysis of native DrrA. Various attempts were made to refold the inclusion body aggregates into an active conformation. Unfortunately these efforts ultimately proved abortive. Inclusion body formation is in fact a rather common phenomenon when overexpressing the NBD proteins of ABC transporters. Yuan and colleagues have suggested that certain structural features of the NBD promote non-physiological interactions between individual molecules, a process that ultimately results in aggregation and insolubility (Yuan *et al.*, 2001).

An entirely different set of problems were encountered when attempting to overexpress the MSD's, DrrB and DrrC. In many respects these problems were typical of those facing all researchers working with integral membrane proteins and are mostly related to low levels of expression and protein instability. DrrB was successfully expressed and purified as a C-terminal His-tagged protein but proved to be highly unstable once removed from the membrane environment. This instability primarily manifested itself as low yields during the purification procedure (even when large culture volumes were used for protein preparation), and a tendency towards spontaneous aggregation and precipitation. DrrC on the other hand proved impossible to express using a similar pET-based expression system. Although a candidate protein was isolated, N-terminal amino acid sequencing identified it as a histidine-rich protein native to *E. coli*, SlyD. The fact that the protein structure databases contain far fewer examples of integral membrane protein structures than of soluble proteins is an indication of just how difficult working with these proteins is. After almost forty years of ABC transporter research and despite huge advances in detergent chemistry, only three complete ABC transporter structures are available, two of these being variants of the same transporter from different organisms (Chang and Roth, 2001; Locher *et al.*, 2002; Chang, 2003). A number of groups have suggested that MSD stability is only a property of complete ABC transporter complexes, i.e. the various domains are highly dependent upon one another, and possibly also the membrane itself,

to maintain a stable conformation. Both Kaur and Russell and Choudhuri and colleages have reported similar MSD instability in their attempts to overexpress the DrrB proteins of *S. peucetius* and *M. tuberculosis* respectively (Kaur and Russell, 1998; Choudhuri *et al.*, 2002).

Having failed to express DrrA as a native protein, an alternative mode of expression was sought. By expressing the mycobacterial protein as a translational fusion to *E. coli* thioredoxin a stable soluble DrrA was eventually produced in useful quantities. Similar techniques have been used extensively in the heterologous overexpression of both prokaryotic and eukaryotic ABC transporter NBD's. The Trx-DrrA fusion protein was shown to exhibit cation dependent ATPase activity, a feature consistent with the proposed role of DrrA as the energy transducing sub-unit of a mycobacterial ABC transporter. ATPase activity was shown to be dependent upon the structural integrity of the 'Walker A' nucleotide-binding motif. Two mutant proteins in which this motif was disrupted by site-directed mutagenesis both exhibited dramatically reduced levels of ATP hydrolysis. This data is the first to show that the *M. tuberculosis* DrrA protein is a functional ATPase and only the second *M. tuberculosis* ABC transporter NBD to have been characterised in this way (the other being the PstB NBD, Sarin *et al.*, 2002).

The availability of significant quantities of Trx-DrrA allowed the ATPase activity of the protein to be investigated in greater detail. Trx-DrrA was shown to have K_m and V_{max} parameters broadly similar to those of other isolated NBD's although the maximal rate of ATP hydrolysis was somewhat lower than might have been expected. Trx-DrrA was shown to exhibit a degree of substrate promiscuity with regard to both nucleotide triphosphates and cation usage. Such promiscuity is very common amongst isolated bacterial NBD's and, although unlikely to be of physiological relevance, is an indication that Trx-DrrA behaves as a typical NBD protein. A growing body of biochemical and structural data suggest that the ATPase competent form of ABC transporter NBD's is a dimer similar in structure to the Rad50cd dimer crystallised by Hopfner and colleagues (Hopfner *et al.*, 2000; Moody *et al.*, 2002). In fact, a number of researchers have proposed that ATP-dependent NBD dimerisation may be a key first stage in driving the substrate translocation process. Despite several attempts to identify Trx-DrrA dimer formation, no biochemical data in support of such a model was obtained. In all respects Trx-DrrA behaved as a simple Michaelis-Menten type enzyme hydrolysing ATP with a

Hill coefficient of 1. This data is consistent with that produced for a number of NBD proteins and is an indication that the behaviour of isolated NBD's may be different from that which they exhibit when forming part of an intact ABC transporter complex. A second interpretation of this data is that the low ATPase activity of Trx-DrrA may be a reflection of a reduced ability to form ATPase active dimers as a direct consequence of steric hindrance by the thioredoxin moiety. In order to ascertain this we intend to test an alternative expression system in which the thioredoxin molecule is fused at the C-terminal end of DrrA, rather than the N-terminal end where the Walker A nucleotide-binding motif is located. The possibility that the DrrA fusion protein displays non-physiological activity when assayed in isolation is a strong argument in favour of treating ABC transporters as complex, interdependent multi-component systems. A complete understanding of the ABC transporter translocation mechanism will undoubtedly require such an approach and begs the question as to how much useful information can be derived from investigating the behaviour of individual components.

Having argued that ABC transporter activity should be treated as a multi-domain phenomenon, one aspect of NBD activity that *is* open to investigation in isolation is the mechanism of ATP hydrolysis. Intrinsic tryptophan fluorescence spectroscopy is a powerful tool for detecting even minor changes in protein structure that may be linked to the catalytic cycle of an enzyme (Walmsley *et al.*, 2001). Unfortunately the ATPase activity of Trx-DrrA proved impossible to study by this means despite extensive attempts to optimise any signal through the construction of single-tryptophan mutant proteins. It was considered that the lack of a strong fluorescence signal during ATP-binding/hydrolysis might be a further manifestation of the low activity of the protein. In order to address this question, and if time allows, we intend to produce further mutant proteins with tryptophan residues inserted at different points within the primary sequence. Such an approach may eventually lead to the generation of a spectroscopically active protein. It will also be interesting to see if moving the thioredoxin moiety of the protein from one end of the molecule to the other has any effect on the fluorescence properties of the protein. The ATPase activity of Trx-DrrA was shown to be totally unaffected by the presence of the potential DrrABC substrates doxorubicin and daunorubicin. This would seem to argue against the presence of an antibiotic binding site or sites anywhere on the DrrA molecule. Whilst many ABC transporters exhibit substrate modulated activity, this is yet another example of a property that may largely be confined to the intact ABC

transporter complex. Under such circumstances, and with a few notable exceptions such as the *E. coli* ArsA protein, ABC transporters should not be regarded as typical allosteric enzymes. It is more likely that the ATPase activity of the NBD is modulated by conformational changes brought about through interaction with the substrate-bound MSD's of the transporter, rather than by direct interaction with the transport substrate itself.

Clearly demonstration of interdomain interactions requires the ability to co-express more than one domain of an ABC transporter simultaneously in the same organism. We attemped to achieve this goal by using a very simple system in which the entire *drr*ABC operon was cloned into a single expression vector. Unfortunately these efforts were abortive due to differences in the mycobacterial and *E. coli* translational machinery. The data presented in chapter 7 seems to indicate that *E. coli* ribosomes are unable to interact with the embedded ribosome binding sites present in a transcriptionally linked *M. tuberculosis* operon. As such, only the first gene in the operon, *drr*A, was correctly translated and expressed, whist there was no evidence for expression of the downstream genes *drr*B and *drr*C. In 2002 Choudhuri and colleagues, also recognising the potential importance of the *M. tuberculosis drr* system, developed a system for co-expression of both the DrrA and B proteins in *E. coli* (Choudhuri *et al.*, 2002). By separating the overlapping genes, and inserting an *E. coli* ribosome-binding site upstream of *drr*B, this group were able to effect simultaneous co-expression of both DrrA and DrrB. Whilst this group acknowledged the probable role of the DrrA and B proteins in mycobacterial lipid export, their work focussed largely on the role of the proteins as possible mediators of antibiotic resistance. It is rather unclear as to why they chose to express only two of the three *drr* genes but present an argument that DrrA and B alone are able to function as doxorubicin/daunorubicin efflux system. Some of the data presented should perhaps be regarded with a degree of scepticism since the use of an intrinsically drug resistant *E. coli* strain to demonstrate antibiotic resistance is not necessarily an entirely appropriate experimental system.

The work of Choudhuri does however serve an important proof-of-principle. It seems entirely plausible that extending their artificial operon system to include not only the *drr*A and B genes, but also *drr*C, may provide a route to successful co-expression of all three components of the *M. tuberculosis* ABC transporter. Under such circumstances it

may eventually be possible to co-purify the entire ABC transporter complex and perhaps reconstitute it into a cell free system. Similar experiments have allowed extensive characterisation of the transport capability and catalytic activity of a number of ABC transporters, most notably the $MalFGK_2$ maltose permease of *E. coli* (Davidson *et al.*, 1996). Another possible advantage of such a co-expression system is a likely improvement in stability of the DrrB and C MSD's.

In conclusion, much remains to be learned about the *M. tuberculosis* DrrABC transporter and the role of mycobacterial ABC transporters in general. The studies described in this work represent only a tiny advance in our knowledge of these systems yet hint that much more remains to be discovered. The multi-domain nature of ABC transporters seems to lie at the very heart of their activity but also contributes to the difficulties associated with studying them. After forty years of ABC transporter research the publication of three complete, atomic-resolution, ABC transporter structures within the last three years should serve as a major encouragement to all researchers working in this field.

References

Abramson, J. *et al.* "Structure and mechanism of the lactose permease of *Escherichia coli*." Science 301.5633 2003): 610-5.

Ainsa, J. A. *et al.* "Aminoglycoside 2'-N-acetyltransferase genes are universally present in mycobacteria: characterization of the aac(2')-Ic gene from *Mycobacterium tuberculosis* and the *aac*(2')-Id gene from *Mycobacterium smegmatis*." Mol.Microbiol. 24.2 (1997): 431-41.

al Shawi, M. K., I. L. Urbatsch, and A. E. Senior. "Covalent inhibitors of P-glycoprotein ATPase activity." J.Biol.Chem. 269.12 (1994): 8986-92.

Aldema, M. L. *et al.* "Purification of the *Tn10*-specified tetracycline efflux antiporter TetA in a native state as a polyhistidine fusion protein." Mol.Microbiol. 19.1 (1996): 187-95.

Allikmets, R. *et al.* "Mutation of a putative mitochondrial iron transporter gene (ABC7) in X-linked sideroblastic anemia and ataxia (XLSA/A)." Hum.Mol.Genet. 8.5 (1999): 743-49.

Altschul, S. F. *et al.* "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res. 25.17 (1997): 3389-402.

Ames, G. F. "Bacterial periplasmic transport systems: structure, mechanism, and evolution." Annu.Rev.Biochem. 55 (1986): 397-425.

Ames, G. F. *et al.* "Purification and characterization of the membrane-bound complex of an ABC transporter, the histidine permease." J.Bioenerg.Biomembr. 33.2 (2001): 79-92.

Antoniou, A. N. *et al.* "Interactions formed by individually expressed TAP1 and TAP2 polypeptide subunits." Immunology 106.2 (2002): 182-89.

Aparicio, G. *et al.* "Characterization of the ATPase activity of the N-terminal nucleotide binding domain of an ABC transporter involved in oleandomycin secretion by *Streptomyces antibioticus*." FEMS Microbiol.Lett. 141.2-3 (1996): 157-62.

Aubourg, P. *et al.* "Adrenoleukodystrophy gene: unexpected homology to a protein involved in peroxisome biogenesis." Biochimie 75.3-4 (1993): 293-302.

Azad, A. K. *et al.* "Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria." J.Biol.Chem. 272.27 (1997): 16741-45.

Azzaria, M., E. Schurr, and P. Gros. "Discrete mutations introduced in the predicted nucleotide-binding sites of the *mdr*1 gene abolish its ability to confer multidrug resistance." Mol.Cell Biol. 9.12 (1989): 5289-97.

Bakos, E. *et al.* "Functional multidrug resistance protein (MRP1) lacking the N-terminal transmembrane domain." J.Biol.Chem. 273.48 (1998): 32167-75.

Baltz, R. H. "Genetic manipulation of antibiotic-producing Streptomyces." Trends Microbiol. 6.2 (1998): 76-83.

Bauer, B. E., H. Wolfger, and K. Kuchler. "Inventory and function of yeast ABC proteins: about sex, stress, pleiotropic drug and heavy metal resistance." Biochim.Biophys.Acta 1461.2 (1999): 217-36.

Baulard, A. R., G. S. Besra, and P. J. Brennan. "The Cell-Wall Core of *Mycobacterium*: Structure, Biogenesis and Genetics." Ed. C. Ratledge and J. Dale. Blackwell Science, 2000. 240-58.

Bavoil, P., M. Hofnung, and H. Nikaido. "Identification of a cytoplasmic membrane-associated component of the maltose transport system of Escherichia coli." J.Biol.Chem. 255.18 (1980): 8366-69.

Berger, E. A. "Different mechanisms of energy coupling for the active transport of proline and glutamine in Escherichia coli." Proc.Natl.Acad.Sci.U.S.A 70.5 (1973): 1514-18.

Berger, E. A. and L. A. Heppel. "Different mechanisms of energy coupling for the shock-sensitive and shock-resistant amino acid permeases of *Escherichia coli*." J.Biol.Chem. 249.24 (1974): 7747-55.

Berkower, C. and S. Michaelis. "Mutational analysis of the yeast a-factor transporter STE6, a member of the ATP binding cassette (ABC) protein superfamily." EMBO J. 10.12 (1991): 3777-85.

Bermudez, L. E. and J. Goodman. "*Mycobacterium tuberculosis* invades and replicates within type II alveolar cells." Infect.Immun. 64.4 (1996): 1400-06.

Binet, R. *et al.* "Protein secretion by Gram-negative bacterial ABC exporters--a review." Gene 192.1 (1997): 7-11.

Bishop, L. *et al.* "Reconstitution of a bacterial periplasmic permease in proteoliposomes and demonstration of ATP hydrolysis concomitant with transport." Proc.Natl.Acad.Sci.U.S.A 86.18 (1989): 6953-57.

Blattner, F. R. *et al.* "The complete genome sequence of *Escherichia coli K-12*." Science 277.5331 (1997): 1453-74.

Bonington, A. *et al.* "Use of Roche AMPLICOR *Mycobacterium tuberculosis* PCR in early diagnosis of tuberculous meningitis." J.Clin.Microbiol. 36.5 (1998): 1251-54.

Borges-Walmsley, M. I., K. S. McKeegan, and A. R. Walmsley. "The structure and function of efflux pumps that confer resistance to drugs." Biochem.J. Pt (2003).

Borst, P., N. Zelcer, and A. van Helvoort. "ABC transporters in lipid transport." Biochim.Biophys.Acta 1486.1 (2000): 128-44.

Braibant, M., P. Gilot, and J. Content. " The ATP binding cassette (ABC) transport systems of *Mycobacterium tuberculosis*." FEMS Microbiol.Rev. 24.4 (2000): 449-67.

Brown, T. J., E. G. Power, and G. L. French. "Evaluation of three commercial detection systems for *Mycobacterium tuberculosis* where clinical diagnosis is difficult." J.Clin.Pathol. 52.3 (1999): 193-97.

Buche, A., C. Mendez, and J. A. Salas. "Interaction between ATP, oleandomycin and the OleB ATP-binding cassette transporter of *Streptomyces antibioticus* involved in oleandomycin secretion." Biochem.J. 321 (Pt 1) (1997): 139-44.

Camacho, L. R. *et al.* "Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis." Mol.Microbiol. 34.2 (1999): 257-67.

Cantwell, M. F. *et al.* "Epidemiology of tuberculosis in the United States, 1985 through 1992." JAMA 272.7 (1994): 535-39.

Chami, M. *et al.* "Three-dimensional structure by cryo-electron microscopy of YvcC, an homodimeric ATP-binding cassette transporter from *Bacillus subtilis*." J.Mol.Biol. 315.5 (2002): 1075-85.

Chang, G. "Structure of MsbA from *Vibrio cholera*: a multidrug resistance ABC transporter homolog in a closed conformation." J.Mol.Biol. 330.2 (2003): 419-30.

Chang, G. and C. B. Roth. "Structure of MsbA from *E. coli*: a homolog of the multidrug resistance ATP binding cassette (ABC) transporters." Science 293.5536 (2001): 1793-800.

Choudhuri, B. S. *et al.* "Overexpression and functional characterization of an ABC (ATP-binding cassette) transporter encoded by the genes *drr*A and *drr*B of *Mycobacterium tuberculosis*." Biochem.J. 367.Pt 1 (2002): 279-85.

Cole, S. T. "*Mycobacterium tuberculosis*: drug-resistance mechanisms." Trends Microbiol. 2.10 (1994): 411-15.

Cole, S. T. *et al.* "Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence." Nature 393.6685 (1998): 537-44.

Contreras, M. *et al.* "Topology of ATP-binding domain of adrenoleukodystrophy gene product in peroxisomes." Arch.Biochem.Biophys. 334.2 (1996): 369-79.

Couture, M. *et al.* "A cooperative oxygen-binding hemoglobin from *Mycobacterium tuberculosis*." Proc.Natl.Acad.Sci.U.S.A 96.20 (1999): 11223-28.

Cox, J. S. *et al.* "Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice." Nature 402.6757 (1999): 79-83.

Dassa, E. and P. Bouige. "The ABC of ABCS: a phylogenetic and functional classification of ABC systems in living organisms." Res.Microbiol. 152.3-4 (2001): 211-29.

Davidson, A. L., S. S. Laghaeian, and D. E. Mannering. "The maltose transport system of *Escherichia coli* displays positive cooperativity in ATP hydrolysis." J.Biol.Chem. 271.9 (1996): 4858-63.

Davidson, A. L. and S. Sharma. "Mutation of a single MalK subunit severely impairs maltose transport activity in *Escherichia coli*." J.Bacteriol. 179.17 (1997): 5458-64.

Davidson, A. L., H. A. Shuman, and H. Nikaido. "Mechanism of maltose transport in *Escherichia coli*: transmembrane signaling by periplasmic binding proteins." Proc.Natl.Acad.Sci.U.S.A 89.6 (1992): 2360-64.

Dean, M., Y. Hamon, and G. Chimini. "The human ATP-binding cassette (ABC) transporter superfamily." J.Lipid Res. 42.7 (2001): 1007-17.

Diederichs, K. *et al.* "Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon *Thermococcus litoralis*." EMBO J. 19.22 (2000): 5951-61.

Doerrler, W. T., M. C. Reedy, and C. R. Raetz. "An *Escherichia coli* mutant defective in lipid export." J.Biol.Chem. 276.15 (2001): 11461-64.

Doolittle, R. F. *et al.* "Domainal evolution of a prokaryotic DNA repair protein and its relationship to active-transport proteins." Nature 323.6087 (1986): 451-53.

Duffieux, F. *et al.* "Nucleotide-binding domain 1 of cystic fibrosis transmembrane conductance regulator production of a suitable protein for structural studies." Eur.J.Biochem. 267.17 (2000): 5306-12.

Dye, C. *et al.* "Worldwide incidence of multidrug-resistant tuberculosis." J.Infect.Dis. 185.8 (2002): 1197-202.

Elkins, C. A. and H. Nikaido. "3D structure of AcrB: the archetypal multidrug efflux transporter of *Escherichia coli* likely captures substrates from periplasm." Drug Resist.Updat. 6.1 (2003): 9-13.

Eytan, G. D., R. Regev, and Y. G. Assaraf. "Functional reconstitution of P-glycoprotein reveals an apparent near stoichiometric drug transport to ATP hydrolysis." J.Biol.Chem. 271.6 (1996): 3172-78.

Farr, B. M. "Diagnostic tests for healthcare epidemiology." Curr.Opin.Infect.Dis. 14.4 (2001): 443-47.

Fernandez-Moreno, M. A. *et al.* "A silent ABC transporter isolated from *Streptomyces rochei* F20 induces multidrug resistance." J.Bacteriol. 180.16 (1998): 4017-23.

Fetsch, E. E. and A. L. Davidson. "Vanadate-catalyzed photocleavage of the signature motif of an ATP-binding cassette (ABC) transporter." Proc.Natl.Acad.Sci.U.S.A 99.15 (2002): 9685-90.

Finken, M. *et al.* "Molecular basis of streptomycin resistance in *Mycobacterium tuberculosis*: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot." Mol.Microbiol. 9.6 (1993): 1239-46.

Finzi, A., J. Cloutier, and E. A. Cohen. "Two-step purification of His-tagged Nef protein in native condition using heparin and immobilized metal ion affinity chromatographies." J.Virol.Methods 111.1 (2003): 69-73.

Friedrich, M. J., L. C. de Veaux, and R. J. Kadner. "Nucleotide sequence of the *btu*CED genes involved in vitamin B12 transport in *Escherichia coli* and homology with components of periplasmic-binding-protein-dependent transport systems." J.Bacteriol. 167.3 (1986): 928-34.

Froshauer, S. *et al.* "Genetic analysis of the membrane insertion and topology of MalF, a cytoplasmic membrane protein of *Escherichia coli*." J.Mol.Biol. 200.3 (1988): 501-11.

Gadsby, D. C. and A. C. Nairn. "Control of CFTR channel gating by phosphorylation and nucleotide hydrolysis." Physiol Rev. 79.1 Suppl (1999): S77-S107.

Gartner, J. *et al.* "Functional characterization of the adrenoleukodystrophy protein (ALDP) and disease pathogenesis." Endocr.Res. 28.4 (2002): 741-48.

Gentschev, I. and W. Goebel. "Topological and functional studies on HlyB of *Escherichia coli*." Mol.Gen.Genet. 232.1 (1992): 40-48.

Gerlach, J. H. *et al.* "Homology between P-glycoprotein and a bacterial haemolysin transport protein suggests a model for multidrug resistance." Nature 324.6096 (1986): 485-89.

Gilson, E. *et al.* "Extensive homology between membrane-associated components of histidine and maltose transport systems of *Salmonella typhimurium* and *Escherichia coli*." J.Biol.Chem. 257.17 (1982): 9915-18.

Gilson, E., H. Nikaido, and M. Hofnung. "Sequence of the malK gene in *E.coli* K12." Nucleic Acids Res. 10.22 (1982): 7449-58.

Goebel, W. and J. Hedgpeth. "Cloning and functional characterization of the plasmid-encoded hemolysin determinant of *Escherichia coli*." J.Bacteriol. 151.3 (1982): 1290-98.

Gray, C. H. *et al.* "The *pfr*1 gene from the human pathogenic fungus *Paracoccidioides brasiliensis* encodes a half-ABC transporter that is transcribed in response to treatment with fluconazole." Yeast 20.10 (2003): 865-80.

Guzman, L. M. *et al.* "Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter." J.Bacteriol. 177.14 (1995): 4121-30.

Hance, A. J. *et al.* "Detection and identification of mycobacteria by amplification of mycobacterial DNA." Mol.Microbiol. 3.7 (1989): 843-49.

Harder, K. W. *et al.* "Characterization and kinetic analysis of the intracellular domain of human protein tyrosine phosphatase beta (HPTP beta) using synthetic phosphopeptides." Biochem.J. 298 (Pt 2) (1994): 395-401.

Heifets, L. *et al.* "Two liquid medium systems, mycobacteria growth indicator tube and MB redox tube, for *Mycobacterium tuberculosis* isolation from sputum specimens." J.Clin.Microbiol. 38.3 (2000): 1227-30.

Heppel, L. A. The effect of osmotic shock on release of bacterial proteins and on active transport. J.Gen.Physiol. 54, 95s. 1969.

Higgins, C. F. "ABC transporters: from microorganisms to man." Annu.Rev.Cell Biol. 8 (1992): 67-113.

Higgins, C. F. *et al.* "A family of closely related ATP-binding subunits from prokaryotic and eukaryotic cells." Bioessays 8.4 (1988): 111-16.

Higgins, C. F. *et al.* "Complete nucleotide sequence and identification of membrane components of the histidine transport operon of *S. typhimurium*." Nature 298.5876 (1982): 723-27.

Higgins, C. F. *et al.* "A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria." Nature 323.6087 (1986): 448-50.

Higgins, C. F. *et al.* "Nucleotide binding by membrane components of bacterial periplasmic binding protein-dependent transport systems." EMBO J. 4.4 (1985): 1033-39.

Hiles, I. D. *et al.* "Molecular characterization of the oligopeptide permease of *Salmonella typhimurium*." J.Mol.Biol. 195.1 (1987): 125-42.

Hobson, A. C., R. Weatherwax, and G. F. Ames. "ATP-binding sites in the membrane components of histidine permease, a periplasmic transport system." Proc.Natl.Acad.Sci.U.S.A 81.23 (1984): 7333-37.

Holland, I. B., B. Kenny, and M. Blight. "Haemolysin secretion from *E. coli*." Biochimie 72.2-3 (1990): 131-41.

Honisch, U. and W. G. Zumft. "Operon structure and regulation of the nos gene region of *Pseudomonas stutzeri*, encoding an ABC-Type ATPase for maturation of nitrous oxide reductase." J.Bacteriol. 185.6 (2003): 1895-902.

Hopfner, K. P. *et al.* "Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily." Cell 101.7 (2000): 789-800.

Hottenrott, S. *et al.* "The *Escherichia coli* SlyD is a metal ion-regulated peptidyl-prolyl cis/trans-isomerase." J.Biol.Chem. 272.25 (1997): 15697-701.

Howard, E. M. and P. D. Roepe. "Purified human MDR 1 modulates membrane potential in reconstituted proteoliposomes." Biochemistry 42.12 (2003): 3544-55.

Hsu, C. M. and B. P. Rosen. "Characterization of the catalytic subunit of an anion pump." J.Biol.Chem. 264.29 (1989): 17349-54.

Huang, Y. *et al.* "Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*." Science 301.5633 (2003): 616-20.

Hung, L. W. *et al.* "Crystal structure of the ATP-binding subunit of an ABC transporter." Nature 396.6712 (1998): 703-07.

Hunke, S., S. Dose, and E. Schneider. "Vanadate and bafilomycin A1 are potent inhibitors of the ATPase activity of the reconstituted bacterial ATP-binding cassette transporter for maltose (MalFGK2)." Biochem.Biophys.Res.Commun. 216.2 (1995): 589-94.

Hunke, S. *et al.* "ATP modulates subunit-subunit interactions in an ATP-binding cassette transporter (MalFGK2) determined by site-directed chemical cross-linking." J.Biol.Chem. 275.20 (2000): 15526-34.

Hyde, S. C. *et al.* "Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport." Nature 346.6282 (1990): 362-65.

IseZaki, M. *et al.* "A putative ATP-binding cassette transporter YbdA involved in sporulation of *Bacillus subtilis*." FEMS Microbiol.Lett. 204.2 (2001): 239-45.

Janas, E. *et al.* "The ATP hydrolysis cycle of the nucleotide-binding domain of the mitochondrial ATP-binding cassette transporter Mdl1p." J.Biol.Chem. 278.29 (2003): 26862-69.

Jones, P. M. and A. M. George. "A new structural model for P-glycoprotein." J.Membr.Biol. 166.2 (1998): 133-47.

Jones, P. M. and A. M. George. "Subunit interactions in ABC transporters: towards a functional architecture." FEMS Microbiol.Lett. 179.2 (1999): 187-202.

Karpowich, N. *et al.* "Crystal structures of the MJ1267 ATP binding cassette reveal an induced-fit effect at the ATPase active site of an ABC transporter." Structure.(Camb.) 9.7 (2001): 571-86.

Kashiwagi, K. *et al.* "The ATPase activity and the functional domain of PotA, a component of the spermidine-preferential uptake system in *Escherichia coli*." J.Biol.Chem. 277.27 (2002): 24212-19.

Katila, M. L., P. Katila, and R. Erkinjuntti-Pekkanen. "Accelerated detection and identification of mycobacteria with MGIT 960 and COBAS AMPLICOR systems." J.Clin.Microbiol. 38.3 (2000): 960-64.

Kaur, P. "Expression and characterization of DrrA and DrrB proteins of *Streptomyces peucetius* in *Escherichia coli*: DrrA is an ATP binding protein." J.Bacteriol. 179.3 (1997): 569-75.

Kaur, P. "The anion-stimulated ATPase ArsA shows unisite and multisite catalytic activity." J.Biol.Chem. 274.36 (1999): 25849-54.

Kaur, P. and J. Russell. "Biochemical coupling between the DrrA and DrrB proteins of the doxorubicin efflux pump of *Streptomyces peucetius*." J.Biol.Chem. 273.28 (1998): 17933-39.

Kennedy, K. A. and B. Traxler. "MalK forms a dimer independent of its assembly into the MalFGK2 ATP-binding cassette transporter of *Escherichia coli*." J.Biol.Chem. 274.10 (1999): 6259-64.

Kerppola, R. E. *et al.* "The membrane-bound proteins of periplasmic permeases form a complex. Identification of the histidine permease HisQMP complex." J.Biol.Chem. 266.15 (1991): 9857-65.

Kispal, G. *et al.* "The ABC transporter Atm1p is required for mitochondrial iron homeostasis." FEBS Lett. 418.3 (1997): 346-50.

Kolaczkowski, M. *et al.* "Anticancer drugs, ionophoric peptides, and steroids as substrates of the yeast multidrug transporter Pdr5p." J.Biol.Chem. 271.49 (1996): 31543-48.

Koronakis, V., E. Koronakis, and C. Hughes. "Isolation and analysis of the C-terminal signal directing export of *Escherichia coli* hemolysin protein across both bacterial membranes." EMBO J. 8.2 (1989): 595-605.

Koronakis, V. and C. Hughes. "Bacterial signal peptide-independent protein export: HlyB-directed secretion of hemolysin." Semin.Cell Biol. 4.1 (1993): 7-15.

Lankat-Buttgereit, B. and R. Tampe. "The transporter associated with antigen processing: function and implications in human diseases." Physiol Rev. 82.1 (2002): 187-204.

Lazarevic, V. and D. Karamata. "The *tag*GH operon of *Bacillus subtilis* 168 encodes a two-component ABC transporter involved in the metabolism of two wall teichoic acids." Mol.Microbiol. 16.2 (1995): 345-55.

Lill, R. and G. Kispal. "Mitochondrial ABC transporters." Res.Microbiol. 152.3-4 (2001): 331-40.

Linton, K. J. and C. F. Higgins. "The *Escherichia coli* ATP-binding cassette (ABC) proteins." Mol.Microbiol. 28.1 (1998): 5-13.

Lippincott, J. and B. Traxler. "MalFGK complex assembly and transport and regulatory characteristics of MalK insertion mutants." J.Bacteriol. 179.4 (1997): 1337-43.

Liu, C. E. and G. F. Ames. "Characterization of transport through the periplasmic histidine permease using proteoliposomes reconstituted by dialysis." J.Biol.Chem. 272.2 (1997): 859-66.

Liu, C. E., P. Q. Liu, and G. F. Ames. "Characterization of the adenosine triphosphatase activity of the periplasmic histidine permease, a traffic ATPase (ABC transporter)." J.Biol.Chem. 272.35 (1997): 21883-91.

Liu, J., C. E. Barry, and H. Nikaido. "Cell Wall: Physical Structure and Permeability." Ed. C. Ratledge and J. Dale. Blackwell Science, 2000. 220-39.

Liu, L. X. *et al.* "Homo- and heterodimerization of peroxisomal ATP-binding cassette half-transporters." J.Biol.Chem. 274.46 (1999): 32738-43.

Liu, Y. C., T. S. Huang, and W. K. Huang. "Comparison of a nonradiometric liquid-medium method (MB REDOX) with the BACTEC system for growth and identification of mycobacteria in clinical specimens." J.Clin.Microbiol. 37.12 (1999): 4048-50.

Locher, K. P., A. T. Lee, and D. C. Rees. "The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism." Science 296.5570 (2002): 1091-98.

Loo, T. W. and D. M. Clarke. "Reconstitution of drug-stimulated ATPase activity following co-expression of each half of human P-glycoprotein as separate polypeptides." J.Biol.Chem. 269.10 (1994): 7750-55.

Ma, D. *et al.* "Genes *acr*A and *acr*B encode a stress-induced efflux system of *Escherichia coli*." Mol.Microbiol. 16.1 (1995): 45-55.

Maupin-Furlow, J. A. *et al.* "Genetic analysis of the *mod*ABCD (molybdate transport) operon of *Escherichia coli*." J.Bacteriol. 177.17 (1995): 4851-56.

McKeegan, K. S., M. I. Borges-Walmsley, and A. R. Walmsley. "The structure and function of drug pumps: an update." Trends Microbiol. 11.1 (2003): 21-9.

McNeil, M. R. and Brennan, P. J. "Structure, function and biogenesis of the cell envelope of mycobacteria in relation to bacterial physiology, pathogenesis and drug resistance; some thoughts and possibilities arising from recent structural information." Res.Microbiol. 8, 451-63. (1991).

Mdluli, K. *et al.* "Mechanisms involved in the intrinsic isoniazid resistance of *Mycobacterium avium*." Mol.Microbiol. 27.6 (1998): 1223-33.

Mendez, C. and J. A. Salas. "The role of ABC transporters in antibiotic-producing organisms: drug secretion and resistance mechanisms." Res.Microbiol. 152.3-4 (2001): 341-50.

Minnikin, D. E. "Lipids: Complex Lipids, Their Chemistry, Biosynthesis, and Roles." Ed. C. Ratledge and J. Stanford. London: Accademic Press, 1982. 95-184.

Miroux, B. and J. E. Walker. "Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels." J.Mol.Biol. 260.3 (1996): 289-98.

Mitterauer, T. *et al.* "Metal-dependent nucleotide binding to the *Escherichia coli* rotamase SlyD." Biochem.J. 342 (Pt 1) (1999): 33-39.

Moody, J. E. *et al.* "Cooperative, ATP-dependent association of the nucleotide binding cassettes during the catalytic cycle of ATP-binding cassette transporters." J.Biol.Chem. 277.24 (2002): 21111-14.

Morbach, S., S. Tebbe, and E. Schneider. "The ATP-binding cassette (ABC) transporter for maltose/maltodextrins of *Salmonella typhimurium*. Characterization of the ATPase activity associated with the purified MalK subunit." J.Biol.Chem. 268.25 (1993): 18617-21.

Mori, H. and K. Ito. "The Sec protein-translocation pathway." Trends Microbiol. 9.10 (2001): 494-500.

Morita, Y. *et al.* "NorM of *Vibrio parahaemolyticus* is an Na^+-driven multidrug efflux pump." J.Bacteriol. 182.23 (2000): 6694-7.

Mourez, M., M. Hofnung, and E. Dassa. "Subunit interactions in ABC transporters: a conserved sequence in hydrophobic membrane proteins of periplasmic permeases defines an important site of interaction with the ATPase subunits." EMBO J. 16.11 (1997): 3066-77.

Mukherjee, S., A. Shukla, and P. Guptasarma. "Single-step purification of a protein-folding catalyst, the SlyD peptidyl prolyl isomerase (PPI), from cytoplasmic extracts of *Escherichia coli*." Biotechnol.Appl.Biochem. 37.Pt 2 (2003): 183-86.

Murakami, S. *et al.* "Crystal structure of bacterial multidrug efflux transporter AcrB." Nature 419.6907 (2002): 587-93.

Nikaido, K., P. Q. Liu, and G. F. Ames. "Purification and characterization of HisP, the ATP-binding subunit of a traffic ATPase (ABC transporter), the histidine permease of *Salmonella typhimurium*. Solubility, dimerization, and ATPase activity." J.Biol.Chem. 272.44 (1997): 27745-52.

Pao, S. S., I. T. Paulsen, and M. H. Saier, Jr. "Major facilitator superfamily." Microbiol.Mol.Biol.Rev. 62.1 (1998): 1-34.

Patzlaff, J. S., T. van der Heide, and B. Poolman. "The ATP/substrate stoichiometry of the ATP-binding cassette (ABC) transporter OpuA." J.Biol.Chem. 278.32 (2003): 29546-51.

Paulsen, I. T. *et al.* "Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes." J.Mol.Biol. 301.1 (2000): 75-100.

Pavelka, M. S., Jr., S. F. Hayes, and R. P. Silver. "Characterization of KpsT, the ATP-binding component of the ABC-transporter involved with the export of capsular polysialic acid in *Escherichia coli* K1." J.Biol.Chem. 269.31 (1994): 20149-58.

Pearce, S. R. *et al.* "Membrane topology of the integral membrane components, OppB and OppC, of the oligopeptide permease of *Salmonella typhimurium*." Mol.Microbiol. 6.1 (1992): 47-57.

Poolman, B. and W. N. Konings. "Secondary solute transport in bacteria." Biochim.Biophys.Acta 1183.1 (1993): 5-39.

Porath, J. *et al.* "Metal chelate affinity chromatography, a new approach to protein fractionation." Nature 258.5536 (1975): 598-99.

Qu, Q. and F. J. Sharom. "FRET analysis indicates that the two ATPase active sites of the P-glycoprotein multidrug transporter are closely associated." Biochemistry 40.5 (2001): 1413-22.

Qu, Q. and F. J. Sharom. "Proximity of bound Hoechst 33342 to the ATPase catalytic sites places the drug binding site of P-glycoprotein within the cytoplasmic membrane leaflet." Biochemistry 41.14 (2002): 4744-52.

Raibaud, O. *et al.* "Structure of the *mal*B region in *Escherichia coli* K12. 1. Genetic map of the *mal*K-*lam*B operon." Mol.Gen.Genet. 174.3 (1979): 241-48.

Rainwater, D. L. and P. E. Kolattukudy. "Fatty acid biosynthesis in *Mycobacterium tuberculosis* var. bovis Bacillus Calmette-Guerin. Purification and characterization of a novel fatty acid synthase, mycocerosic acid synthase, which elongates n-fatty acyl-CoA with methylmalonyl-CoA." J.Biol.Chem. 260.1 (1985): 616-23.

Raviglione, M. C., D. E. Snider, Jr., and A. Kochi. "Global epidemiology of tuberculosis. Morbidity and mortality of a worldwide epidemic." JAMA 273.3 (1995): 220-26.

Reizer, J., A. Reizer, and M. H. Saier, Jr. "A new subfamily of bacterial ABC-type transport systems catalyzing export of drugs and carbohydrates." Protein Sci. 1.10 (1992): 1326-32.

Rigg, G. P., B. Barrett, and I. S. Roberts. "The localization of KpsC, S and T, and KfiA, C and D proteins involved in the biosynthesis of the *Escherichia coli* K5 capsular polysaccharide: evidence for a membrane-bound complex." Microbiology 144 (Pt 10) (1998): 2905-14.

Riordan, J. R. *et al.* "Amplification of P-glycoprotein genes in multidrug-resistant mammalian cell lines." Nature 316.6031 (1985): 817-19.

Riordan, J. R. *et al.* "Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA." Science 245.4922 (1989): 1066-73.

Riska, P. F. *et al.* "Rapid film-based determination of antibiotic susceptibilities of *Mycobacterium tuberculosis* strains by using a luciferase reporter phage and the Bronx Box." J.Clin.Microbiol. 37.4 (1999): 1144-49.

Roach, T. I. *et al.* "Macrophage activation: lipoarabinomannan from avirulent and virulent strains of *Mycobacterium tuberculosis* differentially induces the early genes *c-fos*, *KC*, *JE*, and tumor necrosis factor-alpha." J.Immunol. 150.5 (1993): 1886-96.

Rogl, H. *et al.* "Refolding of *Escherichia coli* produced membrane protein inclusion bodies immobilised by nickel chelating chromatography." FEBS Lett. 432.1-2 (1998): 21-26.

Romsicki, Y. and F. J. Sharom. "Phospholipid flippase activity of the reconstituted P-glycoprotein multidrug transporter." Biochemistry 40.23 (2001): 6937-47.

Rosenberg, M. F. *et al.* "Structure of the multidrug resistance P-glycoprotein to 2.5 nm resolution determined by electron microscopy and image analysis." J.Biol.Chem. 272.16 (1997): 10685-94.

Rosenberg, M. F. *et al.* "The structure of the multidrug resistance protein 1 (MRP1/ABCC1). crystallization and single-particle analysis." J.Biol.Chem. 276.19 (2001): 16076-82.

Rouse, D. A. *et al.* "Characterization of the *kat*G and *inh*A genes of isoniazid-resistant clinical isolates of *Mycobacterium tuberculosis*." Antimicrob.Agents Chemother. 39.11 (1995): 2472-77.

Rusch-Gerdes, S. *et al.* "Multicenter evaluation of the mycobacteria growth indicator tube for testing susceptibility of *Mycobacterium tuberculosis* to first-line drugs." J.Clin.Microbiol. 37.1 (1999): 45-48.

Saier, M. H. Jr. *et al.* "The major facilitator superfamily." J.Mol.Microbiol.Biotechnol. 1.2 (1999): 257-79.

Saier, M. H. Jr. "A functional-phylogenetic classification system for transmembrane solute transporters." Microbiol.Mol.Biol.Rev. 64.2 (2000): 354-411.

Saraste, M., P. R. Sibbald, and A. Wittinghofer. "The P-loop--a common motif in ATP- and GTP-binding proteins." Trends Biochem.Sci. 15.11 (1990): 430-34.

Sarin, J. *et al.* "B-subunit of phosphate-specific transporter from *Mycobacterium tuberculosis* is a thermostable ATPase." J.Biol.Chem. 276.48 (2001): 44590-97.

Sasaki, S. *et al.* "*Mycobacterium leprae* and leprosy: a compendium." Microbiol.Immunol. 45.11 (2001): 729-36.

Saunders, B. M., A. A. Frank, and I. M. Orme. "Granuloma formation is required to contain bacillus growth and delay mortality in mice chronically infected with *Mycobacterium tuberculosis*." Immunology 98.3 (1999): 324-28.

Saurin, W. and E. Dassa. "Sequence relationships between integral inner membrane proteins of binding protein-dependent transport systems: evolution by recurrent gene duplications." Protein Sci. 3.2 (1994): 325-44.

Saurin, W., M. Hofnung, and E. Dassa. "Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters." J.Mol.Evol. 48.1 (1999): 22-41.

Saveanu, L., S. Daniel, and P. M. Van Endert. "Distinct functions of the ATP binding cassettes of transporters associated with antigen processing: a mutational analysis of Walker A and B sequences." J.Biol.Chem. 276.25 (2001): 22107-13.

Scarborough, G. A. "Drug-stimulated ATPase activity of the human P-glycoprotein." J.Bioenerg.Biomembr. 27.1 (1995): 37-41.

Scarparo, C. *et al.* "Comparison of enhanced *Mycobacterium tuberculosis* amplified direct test with COBAS AMPLICOR *Mycobacterium tuberculosis* assay for direct detection of *Mycobacterium tuberculosis* complex in respiratory and extrapulmonary specimens." J.Clin.Microbiol. 38.4 (2000): 1559-62.

Schneider, E. and S. Hunke. "ATP-binding-cassette (ABC) transport systems: functional and structural aspects of the ATP-hydrolyzing subunits/domains." FEMS Microbiol.Rev. 22.1 (1998): 1-20.

Schneider, E., S. Wilken, and R. Schmid. "Nucleotide-induced conformational changes of MalK, a bacterial ATP binding cassette transporter protein." J.Biol.Chem. 269.32 (1994): 20456-61.

Selwyn, P. A. *et al.* "A prospective study of the risk of tuberculosis among intravenous drug users with human immunodeficiency virus infection." N.Engl.J.Med. 320.9 (1989): 545-50.

Senior, A. E., M. K. al Shawi, and I. L. Urbatsch. "The catalytic cycle of P-glycoprotein." FEBS Lett. 377.3 (1995): 285-89.

Shapiro, A. B. and V. Ling. "Extraction of Hoechst 33342 from the cytoplasmic leaflet of the plasma membrane by P-glycoprotein." Eur.J.Biochem. 250.1 (1997): 122-29.

Sharom, F. J. *et al.* "Characterization of the ATPase activity of P-glycoprotein from multidrug-resistant Chinese hamster ovary cells," Biochem.J. 308 (Pt 2) (1995): 381-90.

Shyamala, V. *et al.* "Structure-function analysis of the histidine permease and comparison with cystic fibrosis mutations." J.Biol.Chem. 266.28 (1991): 18714-19.

Silhavy, T. J. *et al.* "Structure of the *mal*B region in *Escherichia coli* K12. II. Genetic map of the *mal*E,F,G operon." Mol.Gen.Genet. 174.3 (1979): 249-59.

Smith, P. C. *et al.* "ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer." Mol.Cell 10.1 (2002): 139-49.

Spaink, H. P., A. H. Wijfjes, and B. J. Lugtenberg. "*Rhizobium* NodI and NodJ proteins play a role in the efficiency of secretion of lipochitin oligosaccharides." J.Bacteriol. 177.21 (1995): 6276-81.

Sreevatsan, S. *et al.* "Ethambutol resistance in *Mycobacterium tuberculosis*: critical role of *emb*B mutations." Antimicrob.Agents Chemother. 41.8 (1997): 1677-81.

Sudre, P., G. ten Dam, and A. Kochi. "Tuberculosis: a global overview of the situation today." Bull.World Health Organ 70.2 (1992): 149-59.

Sulavik, M. C. *et al.* "Antibiotic susceptibility profiles of *Escherichia coli* strains lacking multidrug efflux pump genes." Antimicrob.Agents Chemother. 45.4 (2001): 1126-36.

Tattevin, P. *et al.* "The validity of medical history, classic symptoms, and chest radiographs in predicting pulmonary tuberculosis: derivation of a pulmonary tuberculosis prediction model." Chest 115.5 (1999): 1248-53.

Telenti, A. *et al.* "The *emb* operon, a gene cluster of *Mycobacterium tuberculosis* involved in resistance to ethambutol." Nat.Med. 3.5 (1997): 567-70.

Thiagalingam, S. and L. Grossman. "Both ATPase sites of *Escherichia coli* UvrA have functional roles in nucleotide excision repair." J.Biol.Chem. 266.17 (1991): 11395-403.

Tomii, K. and M. Kanehisa. "A comparative analysis of ABC transporters in complete microbial genomes." Genome Res. 8.10 (1998): 1048-59.

Urbatsch, I. L., M. K. al Shawi, and A. E. Senior. "Characterization of the ATPase activity of purified Chinese hamster P-glycoprotein." Biochemistry 33.23 (1994): 7069-76.

Urbatsch, I. L. *et al.* "P-glycoprotein is stably inhibited by vanadate-induced trapping of nucleotide at a single catalytic site." J.Biol.Chem. 270.33 (1995): 19383-90.

van Helvoort, A. *et al.* "MDR1 P-glycoprotein is a lipid translocase of broad specificity, while MDR3 P-glycoprotein specifically translocates phosphatidylcholine." Cell 87.3 (1996): 507-17.

van Veen, H. W. "Towards the molecular mechanism of prokaryotic and eukaryotic multidrug transporters." Semin.Cell Dev.Biol. 12.3 (2001): 239-45.

van Veen, H. W. *et al.* "A bacterial antibiotic-resistance gene that complements the human multidrug-resistance P-glycoprotein gene." Nature 391.6664 (1998): 291-95.

van Veen, H. W. and W. N. Konings. "Multidrug transporters from bacteria to man: similarities in structure and function." Semin.Cancer Biol. 8.3 (1997): 183-91.

van Veen, H. W. and W. N. Konings. "The ABC family of multidrug transporters in microorganisms." Biochim.Biophys.Acta 1365.1-2 (1998): 31-36.

van Veen, H. W. *et al.* "The homodimeric ATP-binding cassette transporter LmrA mediates multidrug transport by an alternating two-site (two-cylinder engine) mechanism." EMBO J. 19.11 (2000): 2503-14.

van Veen, H. W. *et al.* "Multidrug resistance mediated by a bacterial homolog of the human multidrug transporter MDR1." Proc.Natl.Acad.Sci.U.S.A 93.20 (1996): 10668-72.

von Reyn, C. F., R. Waddell, and K. Pallanygo. "Extrapulmonary tuberculosis." Lancet 358.9286 (2001): 1010-11.

Walker, J. E. *et al.* "Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold." EMBO J. 1.8 (1982): 945-51.

Walmsley, A. R. *et al.* "A kinetic model for the action of a resistance efflux pump." J.Biol.Chem. 276.9 (2001): 6378-91.

Walter, C., Bentrup K. Honer zu, and E. Schneider. "Large scale purification, nucleotide binding properties, and ATPase activity of the MalK subunit of *Salmonella typhimurium* maltose transport complex." J.Biol.Chem. 267.13 (1992): 8863-69.

Wang, C. *et al.* "Expression and purification of the first nucleotide-binding domain and linker region of human multidrug resistance gene product: comparison of fusions to glutathione S-transforase, thioredoxin and maltose-binding protein." Biochem.J. 338 (Pt 1) (1999): 77-81.

Welsh, M. J. and A. E. Smith. "Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis." Cell 73.7 (1993): 1251-54.

WHO Communicable Diseases Division, Geneva. Global Tuberculosis Control Report 2000. 2000.

WHO Communicable Diseases Division, Geneva. Tuberculosis Fact Sheet 104 (Revised August 2002). 2002.

Woods, G. L. "Molecular methods in the detection and identification of mycobacterial infections." Arch.Pathol.Lab Med. 123.11 (1999): 1002-06.

Wooff, E. *et al.* "Functional genomics reveals the sole sulphate transporter of the *Mycobacterium tuberculosis* complex and its relevance to the acquisition of sulphur in vivo." Mol.Microbiol. 43.3 (2002): 653-63.

Yuan, Y. R. *et al.* "The crystal structure of the MJ0796 ATP-binding cassette. Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter." J.Biol.Chem. 276.34 (2001): 32313-21.

Zhou, T. *et al.* "Structure of the ArsA ATPase: the catalytic subunit of a heavy metal resistance pump." EMBO J. 19.17 (2000): 4838-45.

Zhou, T. and B. P. Rosen. "Asp45 is a Mg^{2+} ligand in the ArsA ATPase." J.Biol.Chem. 274.20 (1999): 13854-58.

Zhou, Z. *et al.* "Function of *Escherichia coli* MsbA, an essential ABC family transporter, in lipid A and phospholipid biosynthesis." J.Biol.Chem. 273.20 (1998): 12466-75.

Zurawski, C. A. *et al.* "Pneumonia and bacteremia due to *Mycobacterium celatum* masquerading as *Mycobacterium xenopi* in patients with AIDS: an underdiagnosed problem?" Clin.Infect.Dis. 24.2 (1997): 140-43.