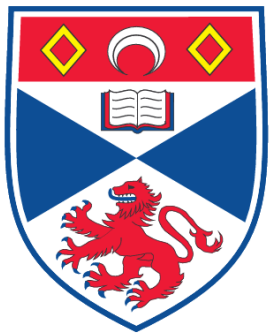
INVESTIGATION OF THE TRANSCRIPTIONAL RESPONSE OF *SULFOLOBUS SOLFATARICUS* TO DAMAGING AGENTS

Stacey Munro

A Thesis Submitted for the Degree of PhD
at the
University of St. Andrews



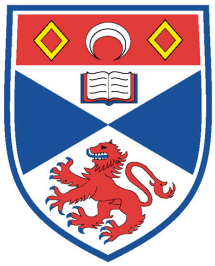
2009

INVESTIGATION OF THE TRANSCRIPTIONAL RESPONSE OF *SULFOLOBUS SOLFATARICUS* TO DAMAGING AGENTS

STACEY MUNRO

A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

2009



University
of
St Andrews

Table of Contents

Table of Contents

Figures and Tables

Abbreviations

ADP	Adenosine diphosphate
AP site	Apurinic or apyrimidinic site
AsnC	Asparagine-dependent activator of asparagines synthase
ATP	Adenosine triphosphate
BER	Base excision repair
bp	Base pair
BRE	Transcription factor B recognition element
BSA	Bovine serum albumin
CAT	Catalase
Cdc6	Cell division control protein
cDNA	Complementary DNA
ChIP	Chromatin Immunoprecipitation
CP values	Crossing point values
Cy Dye	Cyanine Dye
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
Dps	DNA protection protein from starved cells
dsDNA	Double strand DNA
DTT	Dithiothreitol
DtxR	Diphtheria toxin repressor
EDTA	Ethylenediamine tetraacetic acid
EMSA	Electrophoretic mobility shift assay
FEN-1	Flap Endonuclease-1
GG-NER	Global genome NER
IPTG	Isopropyl β-D-thiogalactopyranoside
KO	Knockout
Lrp	Leucine-responsive regulatory protein
MALDI-ToF	Matrix-associated laser desorption ionization-time of flight
MDR-1	Metal dependent repressor-1
MMC	Mitomycin C
MMS	Methyl methane sulfonate

NER	Nucleotide excision repair
NRAMP	Natural resistance-associated macrophage protein
PBL 2025	Paul Blum's *Sulfolobus solfataricus* strain.
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
RNAPII	RNA polymerase II
rNTP	Ribosomal nucleotide triphosphate
ROS	Reactive oxygen species
RT	Room temperature
RT real time PCR	Reverse transcription real time PCR
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-Polyacrylamide gel electrophoresis
SOD	Superoxide dismutase
SSB	Single stranded DNA binding protein
SSC	Sodium chloride sodium citrate
ssDNA	Single stranded DNA
Sso	*Sulfolobus solfataricus*
TAF	Transcription accessory factors
TBE	Tris-borate-EDTA buffer
TBP	TATA box binding protein
TC-NER	Transcription coupled NER
TE	Tris-EDTA Buffer
TFB	Transcription factor B
TFIIB	Transcription factor II B
UV	Ultraviolet radiation
WT	Wild type

Declaration

I, Stacey Munro, hereby certify that this thesis, which is approximately 40'000 words in length, has been written by me, that it is the record of work carried out by me, and that it has not been submitted in any previous application for a higher degree.

Date …….. Signature of candidate ……

I was admitted as a research student in October 2005 and as a candidate for the degree of PhD in October 2006; the higher study for which this is a record was carried out in the University of St Andrews between 2005 and 2008.

Date …….. Signature of candidate ……

I herby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date …….. Signature of supervisor ……..

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for the use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. We have obtained any third-party copyright permission that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

Access to printed copy and electronic publication of thesis through the University of St Andrews.

Date …… Signature of candidate……..…….

Signature of supervisor…….……..

Abstract

It is vital for the survival of an organism that it can repair damage to its DNA. Exogenous and endogenous sources of damage are dealt with by a variety of repair pathways that have evolved to repair specific types of damage. Organisms in the archaeal domain, the third domain of life, contain homologues of many of the eukaryotic repair proteins, however little is known about how damage is detected in the archaeal domain.

Microarray studies in the archaeal species *Sulfolobus solfataricus* determined a number of genes whose expression was effected by UV radiation (work by Dr D Götz). The change in expression of nine of these genes was confirmed by RT real time PCR. The expression of these genes was then investigated after exposure to different damaging agents, Mitomycin C, Methyl methane sulfonate, Phleomycin and Hydrogen peroxide. The expression of two genes, transcription factor *tfb-3* and cell division control gene *cdc6-2,* was up regulated in all damage conditions.

There was a huge induction of the *dps-like* gene (*sso2079*) after hydrogen peroxide damage. Transcription from this genes promoter was shown to be strong *in vitro* (work by Dr S Paytubi) suggesting a repressor was controlling the gene *in vivo*. A palindromic repeat in the promoter of the *dps-like* gene was used to 'fish' for a transcriptional repressor and the Sso2273 protein, a homologue of the diphtheria toxin repressor (DtxR) from *Corynebacterium diphtheria,* was identified as a possible repressor.

Sso2273 was expressed and purified, and its crystal structure solved, its paralogue, Sso0669, was also expressed and purified. Electrophoretic mobility shift assays showed that the Sso2273 protein does not bind DNA, and had no effect on transcription from any promoter used in *in vitro* transcription assays. However Sso0669 appeared to inhibit transcription, although the inhibition was not sequence specific.

A knockout strain of *S. solfataricus* PBL2025 missing the *sso2273* gene was produced and used in microarray experiments in an attempt to determine the role of Sso2273 within the cell. The absence of Sso2273 appeared to have no effect on the expression of the *dps-like* gene, however strong repression of an operon containing genes involved in Sulphur assimilation was observed.

Acknowledgements

I have found my PhD a challenging but enjoyable experience and I would like to thank all the people involved.

Thank you to my supervisor Professor Malcolm White for his support, advice and guidance throughout my PhD. Thanks also go to the Mass spec service and the SSPF lab, especially Dr Stephen McMahon for their contributions to this project. I would also like to thank Dr Sonja Albers and her lab at the University of Groningen for all their help with the knockout and their hospitality while I was in Holland, and Professor Rolf Bernander and his lab, especially Erik Karlsson, at the University of Uppsala for all their help with the microarrays, and for their hospitality while I was in Sweden, and the BBSRC (and Professor White) for funding these trips.

But the biggest thanks has to go to past and present member of the White and Coote labs for making the lab such a supportive, friendly and fun place to work. Thanks to Dr Dorothee Götz for her help and advice on the UV damage, and microarray experiments and special thanks to Dr Sonia Paytubi whose help and advice in so many aspects of my work was invaluable (Pregunta!).

And last, but by no means least, a huge thank-you to my family and my boyfriend Kieran Emptage for their support and encouragement throughout my PhD.

Chapter One

Introduction

1.1 Archaea: The third domain of life

It was first proposed that archaea be designated as a separate domain in 1977, when Carl Woese and George Fox used ribosomal RNA sequence analysis to show that this group of organisms were different enough to be considered a third domain of life, distinct from bacteria and eukaryotes (Woese and Fox 1977).

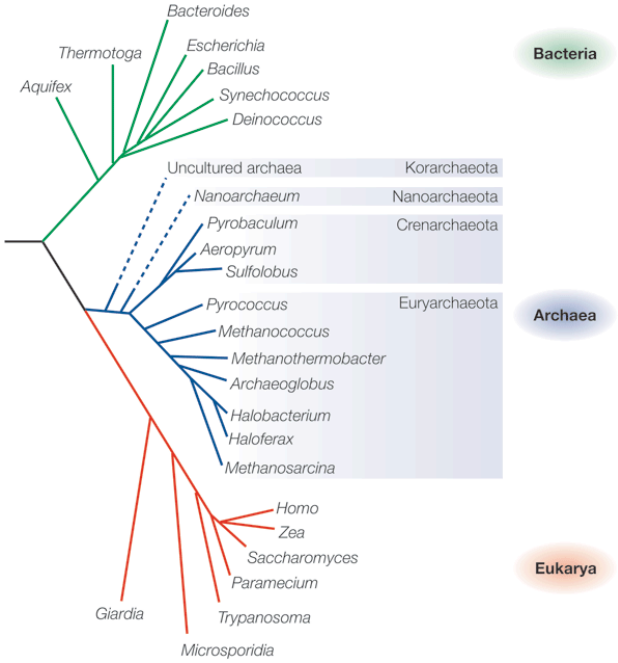


Figure 1.1 Phylogenetic tree showing the three domains of life
Analysis of ribosomal RNA sequences led to a tree of life with three domains; bacteria, archaea and eukarya (Allers and Mevarech 2005).

The idea caused much contention, previously organisms had been grouped together according to their physical characteristics, Woese argued that "molecular sequencing can reveal evolutionary relationships in a way and to an extent that classical phenotypic criteria and even molecular function cannot" (Woese, Kandler *et al.* 1990). The phylogenetic tree in Figure 1.1 shows the archaeal domain split into four kingdoms, euryarchaeota, crenarchaeota, nanoarchaeota and korarchaeota. The euryarchaeota contain a diverse mix of halophiles, methanogens, thermophiles and psychrophiles. The crenarchaeota contain thermophiles, hyperthermophiles and initially marine and environmental mesophiles were placed in this kingdom as well. However analysis of the first sequenced mesophilic crenarchaea, *Cenarchaeum symbiosum*, has shown that mesophilic crenarchaea are different from hyperthermophilic crenarchaea and it has been proposed that they be considered a separate domain called the Thaumarcheaota (Brochier, Boussau et al. 2008). It was initially believed that the archaea represented an ancient group of organisms

inhabiting niche extreme environments, but work by Fuhrman (Fuhrman, McCallum *et al*. 1992) and DeLong (Delong 1992) showed that crenarchaea (or Thaumarchaea) actually constitute the dominant fraction of bacterioplankton in the ocean, highlighting their importance in global systems. A number of features of the euryarchaea suggest they are more eukaryotic in nature than the crenarchaea, for example the euryarchaea possess eukaryotic-like histones for DNA packing, these proteins are missing from all crenarchaea. The family D polymerase, FtsZ cell-division protein, replication protein A (RPA) and a full length XPF endonuclease are all present in euryarchaea, but missing from crenarchaea suggesting differences in methods of chromosome packing and replication between the two domains (White 2003; Brochier, Gribaldo *et al*. 2005).

The nanoarchaeota kingdom was suggested after the discovery of *Nanoarchaeum equitans,* a hyperthermophilic organism, only 400 nm in diameter, possessing the smallest cellular genome ever sequenced (480 kb). *N. equitans* grows on the surface of a larger archaea, of the *Ignicoccus* genus, and cannot be cultivated alone suggesting a parasitic or symbiotic relationship between the two organisms. When *N. equitans* was isolated from a submarine hot vent its 16S sequence showed that it could not be grouped with either the euryarchaeota or crenarchaeota, and it was believe to represent a new kingdom (Huber, Hohn *et al*. 2002). However analysis using concatenated ribosomal protein sequences from 25 archaeal genomes showed that *N.equitans* was in fact a fast-evolving, very divergent euryarchaeon showing strong similarity to *Thermococcales* (Brochier, Gribaldo *et al*. 2005).

The fourth kingdom, Korarchaeota, was first predicted by 16S rRNA analysis back in 1996 (Barns, Delwiche *et al*. 1996), but it was not until recently that a member of this kingdom, *Candidatus* Korarchaeum crytofilum, was isolated from one of the hot springs at Yellow Stone Park (Elkins, Podar et al. 2008). It has an ultra thin filamentous morphology and analysis of its genome sequence suggests it represents a more ancestral form of the archaea.

Archaea are a chimeric mix of bacterial, eukaryal and some uniquely archaeal features. Their morphology is similar to that of bacteria, they are single celled

organisms, and their DNA is usually contained in a single circular chromosome. Their metabolic processes are also similar to those of bacteria, however many of their cellular processes, such as DNA replication (Kelman and White 2005), recombination (Allers and Mevarech 2005), repair (White 2003) and transcription (Bell and Jackson 2001) are more similar to eukaryotic systems. Although the archaea possess homologues of proteins from eukaryotic systems, they are often stripped down and streamlined versions, possessing far fewer proteins and providing a simplified system in which to study these complex processes. But archaea are more than just a mix of eukaryotic and bacterial features, two thirds of the genes found in archaea have no homologues in bacteria or eukaryotes. For example the only organisms to possess reverse gyrase are hyperthermophiles. Reverse gyrase is a DNA topoisomerase that induces positive supercoiling, which affords the DNA more thermal stability (Napoli, Valenti *et al.* 2005). Some of the hyperthermophiles are bacteria however it is believed that they obtained the gene by lateral transfer from hyperthermophilic archaea (Brochier and Forterre 2006).

The archaea also possess a unique cell membrane structure, a number of chemical differences distinguishes the archaeal cell membrane from its bacterial or eukaryotic counterparts. For example, the glycerol molecules used to make the phospholipids of the archaeal cell membrane are in the L-isoform, not the D-isoform as in bacteria. The linkages used to attach side chains to the glycerol in archaea are ether linkages, rather that the ester linkages used in bacteria. In bacteria and eukaryotes the side chains are fatty acids, while in archaea isoprene molecules are used, one benefit of using these molecules is that they can join together forming transmembrane phospholipids, providing more structural stability to the membrane, making it more resistant to oxidation and high temperature (Van der Vossenberg, Driessen *et al.* 1998), see Figure 1.2.

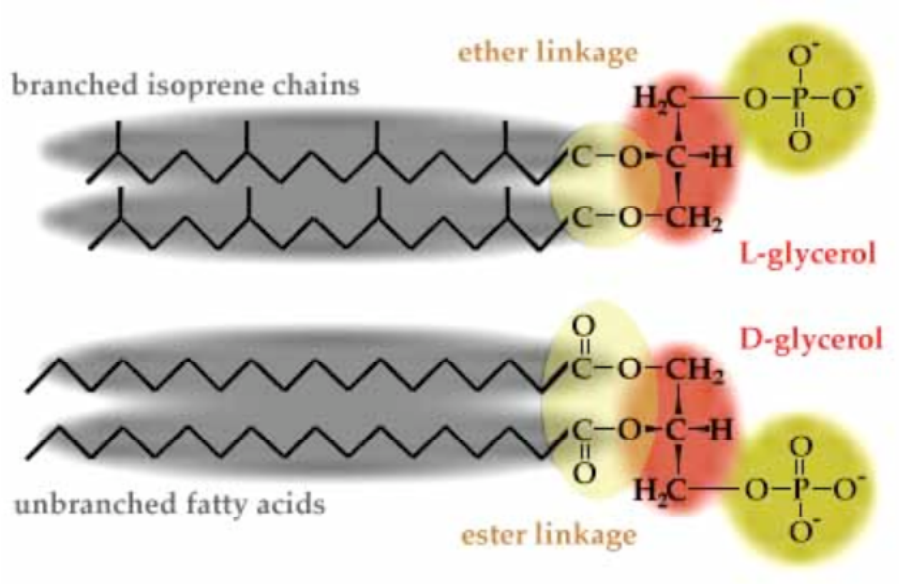


Figure 1.2 Archaeal cell membrane structure
Archaeal cell membranes are chemically different from those of bacteria and eukaryotes, they contain the L-isoform of glycerol, their side chains are attached with ether, instead of ester, linkages and their side chains are made up of isoprene chains instead of fatty acids. Image taken from www.ucmp.berkley.edu/archaea/archaeamm.html.

1.2 *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius*

The archaeons *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius* are hyperthermophiles that grow optimally at temperature of 75 - 80 °C and pH 2 - 4 (She, Singh *et al*. 2001; Chen, Brugger *et al*. 2005), conditions that would prove fatal for most organisms, yet studies in *S. acidocaldarius* show its mutation rate is no higher than that of mesophilic *E. coli* (Grogan, Carver *et al*. 2001). This suggests that these archaea have highly effective protection and repair mechanisms, in order to deal with the risks posed by the conditions they inhabit. The thermo tolerance of the proteins of these organisms makes them a useful tool for investigating the processes of transcription and repair; it has also made them of great interest to the biotechnology industry as a means of providing more robust enzymes (Podar and Reysenbach 2006).

1.3 DNA damage

DNA is under constant attack from endogenous and exogenous agents. As well as damage from external sources such as ultraviolet (UV) light, heat and chemicals, the cell has to deal with mutations caused by reactive oxygen species produced as by-products of its own cellular metabolism (Bertram and Hass 2008). No single repair

system can deal with the array of different types of damage the cell experiences, and so a number of damage specific repair pathways have evolved, see Figure 1.3.

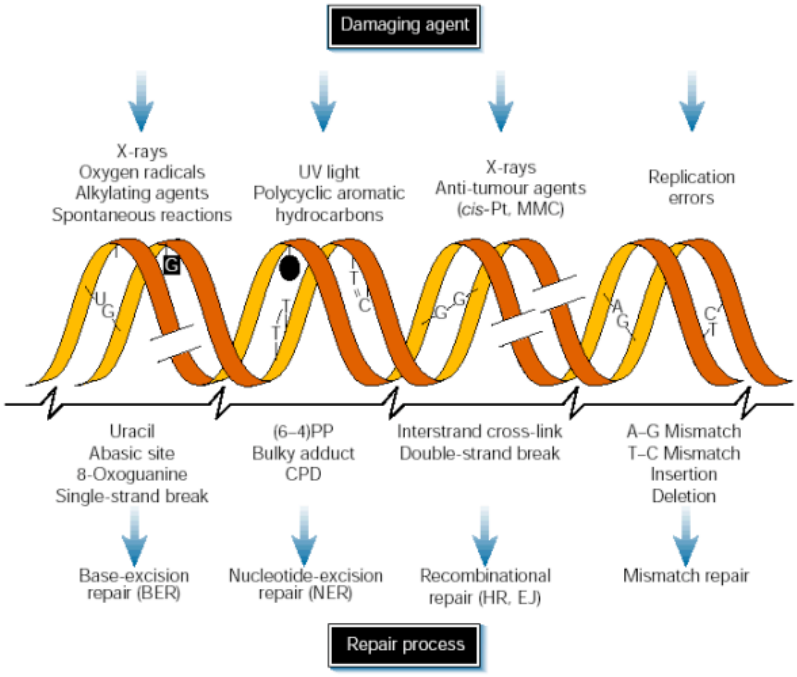


Figure 1.3 DNA damaging agents and the subsequent repair processes
Different types of DNA damaging agents are shown along the top of the image, the resulting damage is shown in the middle, while the specific repair pathways are shown underneath. Image taken from Hoeijmaker 2001.

1.3.1 Ultraviolet (UV) radiation

All organisms exposed to the sun's rays have to cope with the damaging effects of UV radiation. The most common types of damage caused by UV radiation are cyclobutane pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4 PP) (Pfeifer, You *et al*. 2005) see Figure 1.4.

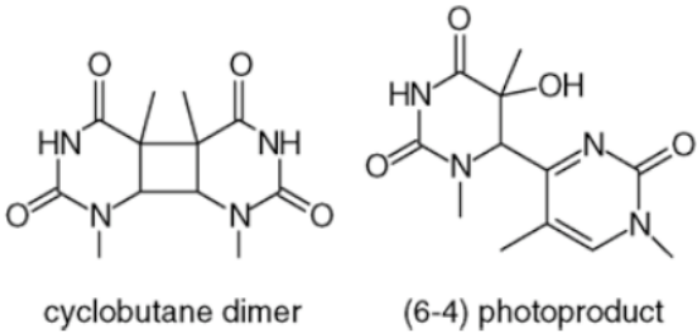


Figure 1.4 Cyclobutane pyrimidine dimer (CPD) and 6,4-photoproduct
Cyclobutane pyrimidine dimers and (6,4) photoproducts are the most common types of damage caused by UV irradiation (Cadet, Sage et al. 2005).

The fastest way to repair these bulky lesions is direct reversal with proteins from the photolyase/cryptochrome family. These proteins are between 55 and 75 kDa in weight and contain two non-covalently bound chromophores; the first is flavin (FADH) and the second is usually folate (MTHF) or deazaflavin (8-HDH) (Christmann, Tomicic *et al*. 2003). The enzymes work by binding to UV damage and using visible light as an energy source to convert the damage back to its original form (Menck 2002). Photolyases are found in all domains of life, but are missing from higher mammals, including humans. Proteins homologous to photolyase have been identified in humans, however they no longer function in repair, but are responsible for regulation of circadian rhythms (Thompson and Sancar 2002). Photolyases are present in archaea, but they are by no means ubiquitous. They are present in the three *Sulfolobus* species and a couple of methanogen and halophiles but are missing from the majority of archaeal genomes (Kelman and White 2005). The crystal structure of the photolyase from *Sulfolobus tokodaii* has been solved and both chromophores are flavin molecules (Fujihashi, Numoto et al. 2007). Normally the two chromophores perform different functions, the FAD catalyses the repair of CPDs while the other chromophore harvests photon energy. The fact that both of the chromophores in the *S. tokodaii* photolyases are FAD molecules suggests a novel light-harvesting process in this organism. In situations where visible light is not available, or in organisms that do not possess photolyase, an alternative pathway performs repair; the nucleotide excision repair (NER) pathway

1.3.1.1 Repair of UV damage by nucleotide excision repair (NER)

In bacteria and eukaryotes repair of damage in DNA strands that are being actively transcribed is faster than repair of damage in non-transcribed strands. The faster transcription-coupled repair (TC-NER) is initiated by the stalling of RNA polymerase (RNAP) at the site of damage. In eukaryotes additional proteins CSA and CSB are recruited to the site of damage and are believed to displace the stalled RNAP allowing repair proteins access to the damage (Li and Bockrath 1995). Damage in parts of the genome that are not being actively transcribed is dealt with by global genome repair (GG-NER). This type of NER is initiated by the binding of damage recognition proteins, in eukaryotes XPC-hHR23B, DDB1 and DDB2 are responsible for damage

recognition. After the initial damage recognition step the NER pathway is the same in GG and TC-NER. The TFIIH complex is recruited to the site, and the damaged DNA is unwound by DNA helicases XPB (3'-5') and XPD (5'-3'). The DNA around the damage site is then cut by endonucleases, ERCC1-XPF (5' of the damage) and XPG (3' of the damage), and removed. The open complex is stabilised by RPA and the remaining strand used as a template for synthesis of new DNA, that is then sealed by DNA ligase (Bell, Cairns et al. 1999; Hoeijmakers 2001; Christmann, Tomicic et al. 2003; Fleck and Nielsen 2004) see Figure 1.5.

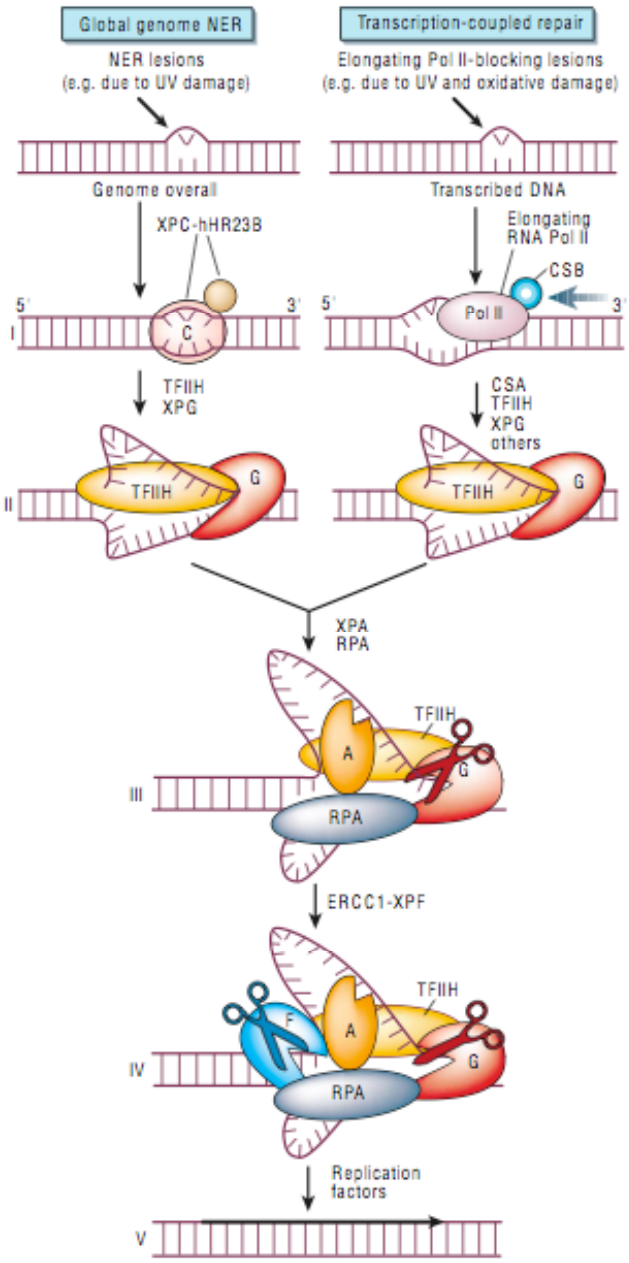


Figure 1.5 Nucleotide excision repair in eukaryotes
UV damage forms a bulky lesion. In GG-NER damage is recognised by XPC-hHR23B. In TC-NER stalling of the RNAP initiates repair, CSA-CSB proteins displace the RNAP allowing the other repair proteins access to the damage. The TFIIH complex is recruited to the site of damage, helicases XPB and XPD, contained within the TFIIH complex, unwind the DNA around the site of damage. Nucleases, ERCC1-XPF and XPG, cut the DNA either side of the damage and the strand containing the damage is released. The complementary strand is used as a template for DNA polymerase to synthesis new DNA. The new DNA is sealed with DNA ligase, leaving a repaired, intact DNA strand. RPA protects the template strand from degradation while repair is performed, image taken from Hoeijmakers 2001.

In bacteria NER is performed by the ATP-dependent UvrABC system. When damage occurs on a non-transcribed strand (GG-NER) the bulky lesion is detected by a UvrA dimer. After binding the damage site the UvrA dimer recruits UvrB, this protein is involved in verifying the presence of damage. The recruitment of UvrB triggers hydrolysis of the ATP bound to the UvrA dimer and results in the monomerization and releases of UvrA from the DNA, leaving the UvrB-DNA complex, which is recognised by UvrC. UvrC cuts the DNA four or five bases 3' and eight bases 5' of the damage site. Once the DNA is cut UvrC dissociates from the DNA allowing UvrD (DNA helicase II) to release the damaged DNA, DNA polymerase I fills the gap and removes UvrB leaving DNA ligase to complete the process by sealing the repair (Truglio, Croteau *et al*. 2006). If damage occurs on a strand being actively transcribed the stalling of RNAP initiates the repair, the transcription repair coupling factor (TRCF) sometimes called Mfd, releases the stalled RNAP and recruits UvrA to the damage site, after the initial damage recognition the process is the same as for global genome NER, see Figure 1.6.

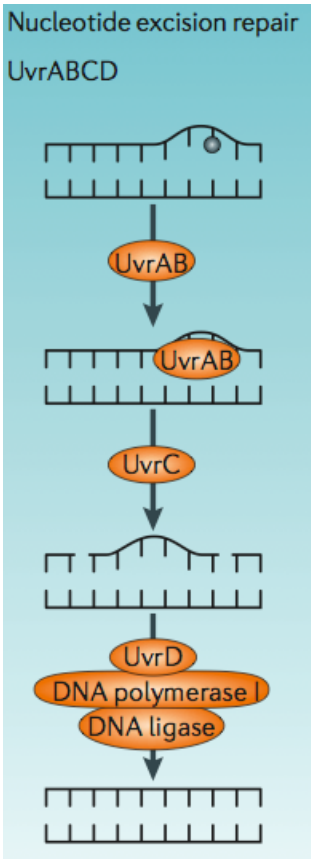


Figure 1.6 Nucleotide excision repair in bacteria

The UvrABC complex is responsible for NER is bacteria. UvrA binds the damaged DNA recruiting UvrB to the site. UvrB confirms the presence of damage before recruiting UvrC to the site. UvrC excises the damaged piece of DNA, which is removed by UvrD. The intact strand is used as a template by DNA polymerase to replicate new DNA, and the repair is sealed by DNA ligase (Kang and Blaster 2006).

Whether the archaea possess an NER pathway is unclear. Many archaeal species possess homologues of the eukaryotic NER proteins, however the damage recognition proteins (XPA, XPC) are absent from all archaeal genomes sequenced so far. Some species, such as *Halobacterium salinarum* and *Methanosarcina mazei,* also possess homologues of the bacterial UvrABC proteins (Grogan 2004), although their presence is believed to be due to lateral gene transfer events (White 2003). The absence of the damage recognition proteins in archaea, suggests a novel method of damage detection. The single-stranded DNA binding protein (SSB) has been suggested as a possible damage recognition protein (Cubeddu and White 2005).

Sulfolobus solfataricus possesses a photolyase and so can perform rapid repair of UV damage by this light dependent process (Grogan 1997). As in all hyperthermophiles no homologues of the damage recognition proteins XPC and XPA are found, however homologues of XPB, XPD, XPF and XPG (Fen-1 in *S. solfataricus*) are present (Grogan 2004). These proteins have other functions in the cells and may not necessarily function in an NER pathway. However repair of damage in the dark, (ie repair without photolyase) has been observed (Salerno, Napoli et al. 2003). Two independent studies in *S. solfataricus* detected no transcription-coupled repair, as removal of damage from transcribed and non-transcribed strands was performed at the same rate. The rate of GG-NER in *S. solfataricus* is as fast as the rate of TC-NER in *E. coli*. This suggests there is no need for a faster TC-NER pathway in *S. solfataricus*, as the GG-NER is fast enough already (Brune, Werner *et al.* 2006; Dorazi, Gotz *et al.* 2007).

1.3.2 Oxidative stress

Over 2 billion years ago the evolution of O_2-producing cyanobacteria led to the accumulation of oxygen in the Earth's atmosphere (Dismukes, Klimov *et al.* 2001). This allowed some organisms to evolve a new way of producing energy; aerobic metabolism. Considering that aerobic metabolism is around 18 times more efficient than anaerobic metabolism (Lodish, Berk *et al.* 1999), it seems logical that organisms would utilize this form of energy production. However there is a catch, the oxygen needed for aerobic metabolism has the potential to produce extremely damaging

molecules called reactive oxygen species (ROS), See Table 1.1. In order to safely utilise oxygen for aerobic metabolism the cell must have a means of protecting itself against the damaging effects of ROS.

Name	Notation	Description
Superoxide anion	$O_2^{\bullet-}$	One-electron reduction state of O_2. Formed by photo-oxidation, electron transfer reactions.
Hydrogen Peroxide	H_2O_2	Formed from $O_2^{\bullet-}$ by dismutation, used by the cell in its defence against pathogens
Hydroxyl Radical	$OH^{\bullet}$	Formed by the Fenton Reaction and decomposition of peroxynitrite. Extremely reactive, will attack most cellular components
Ozone	O_3	Formed by UV irradiation or electrical discharge in the stratosphere

Table 1.1 Reactive oxygen species involved in oxidative stress.
Information in the table is taken from the following (Pedone, Bartolucci et al. 2004; Scandalios 2005; Valko, Rhodes et al. 2006).

Aerobic metabolism is not the only source of ROS, they are also produced by various exogenous sources, such as UV light, ionizing radiation, cigarette smoke and air pollution. Single and double strand breaks, abasic sites, DNA cross-linking and oxidation of DNA bases can all be caused by ROS. The most commonly produced base lesion, is 8-hydroxyguanine (8-OHD), see Figure 1.7, which is often used as a measure of oxidative damage.

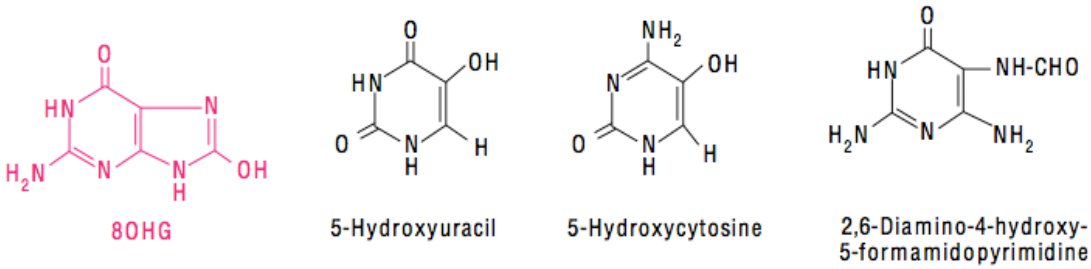


Figure 1.7 DNA bases damaged by oxidative stress
Damaged DNA bases produced by oxidative stress, 8-hydroxyguanine shown in red is the most commonly produced (Wiseman and Halliwell 1996).

Oxidative stress is used to describe the imbalance in the levels of damaging ROS, and protecting antioxidants in the cell. The cell has an arsenal of antioxidant molecules such as Ascorbate (Vitamin C), β Carotene, α Tocopherol (Vitamin E), Thioredoxin and Glutathione that will donate an electron to the ROS and so neutralise it. There are also a number of enzymes that protect the cell in a variety of other ways (see Table 1.2).

Name	Function
Superoxide Dismutase (SOD)	Catalyzes the dismutation of $O_2^{\bullet -}$ to O_2 and H_2O_2
Catalase (CAT)	Promotes the conversion of H_2O_2 to H_20 and O_2
Peroxiredoxins	Scavenges peroxides using the thioredoxin reductase (Tr) system as an electron donor.
Dps (DNA protection protein from starved cells)	Physically shields DNA by binding to it, sequesters components of the Fenton reaction (H_20_2 and ferrous iron) and so reduces the production of hydroxyl radicals
Thioredoxin reductase	Reduces oxidised thioredoxin to its active form using NADPH as an electron donor
Glutathione reductase	Reduces oxidised glutathione to it active form using NADPH as an electron donor

Table 1.2 Enzymes and proteins involved in protecting the cell against oxidative stress Information in the table is taken from the following (Pedone, Bartolucci *et al*. 2004; Scandalios 2005; Limauro, Pedone *et al*. 2006; Valko, Rhodes *et al*. 2006; Limauro, Pedone *et al*. 2008).

Superoxide dismutase and catalase work together as the cells first line of defence against oxidative stress. The antioxidants keep ROS at bay by donating electrons, and once oxidised are recycled by reductases (Valko, Rhodes *et al*. 2006). The Dps protein acts in a number of ways to reduce the effects of oxidative stress. It physically protects the DNA by shielding it from ROS and also removes hydrogen peroxide and ferrous iron, components of the Fenton reaction, reducing the production of extremely damaging hydroxyl radicals (Zhao, Ceci *et al*. 2002). Dps is discussed in more detail in Chapter 4.

1.3.2.1 Repair of oxidative stress damage by base excision repair (BER)

Oxidised bases, apurinic and apyrimidinic (AP) sites and single-strand breaks induced by oxidative stress are predominantly repaired by the base excision repair (BER) pathway (Fleck and Nielsen 2004). The first step in BER involves glycosylases that recognise specific types of damage. Glycosylases are split into two groups; the monofunctional and the bifunctional. Monofunctional glycosylases scan along the DNA searching for damaged bases by compressing the DNA backbone, damaged bases are flipped out and accommodated in a specific binding pocket in the protein where the base is cleaved at the N-glycosidic bond, resulting in an abasic site. The abasic site is then cut by an AP-endonuclease leaving a 3'-hydroxyl group and a 5'-deoxyribose phosphate. Deoxyribophosphodiesterase (dRPase) removes the 5'-

deoxyribose phosphate and the gap is filled by DNA polymerase and sealed by DNA ligase (Demple and Harrison 1994; Fortini, Pascucci *et al*. 2003), see Figure 1.8.

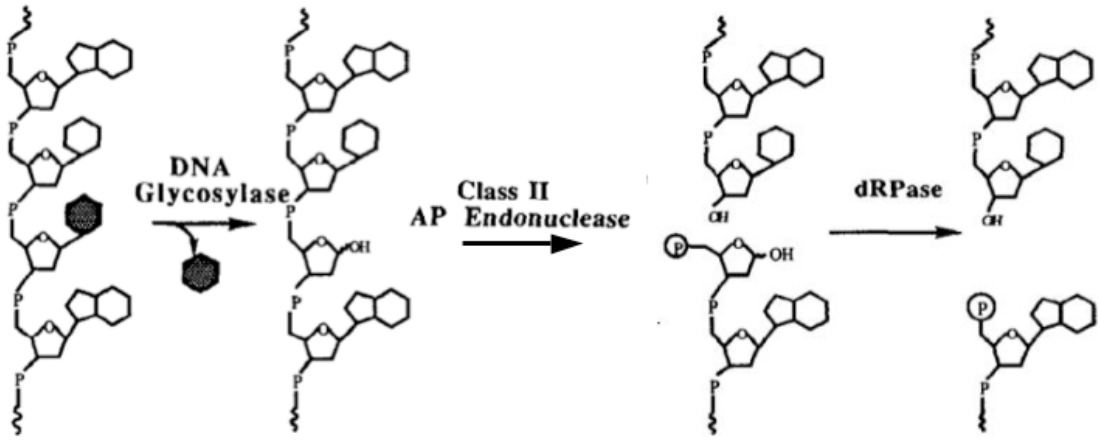


Figure 1.8 Base excision repair by monofunctional glycosylases
The damaged base is recognised by its specific glycosylase and removed from the DNA backbone leaving an abasic site. AP endonuclease cuts the abasic site and dRPase removes the 5' deoxyribose phosphate. The gap is filled by DNA polymerase and sealed by DNA ligase (Demple and Harrison 1994).

The process is slightly more complicated with bifunctional glycosylases due to their 3' AP lyase activity. Once the damaged base is in the binding pocket of the glycosylase the active site residue forms a transient Schiff base with the deoxyribose, after hydrolysis of the damaged base. The phosphodiester bond is then cleaved by the AP lyase using a β-elimination reaction, resulting in a single-stand break, see Figure 1.9. How the break is processed to allow repair to be completed by DNA polymerase and DNA ligase is not fully understood (Hazra, Izumi et al. 2003; Huffman 2006).

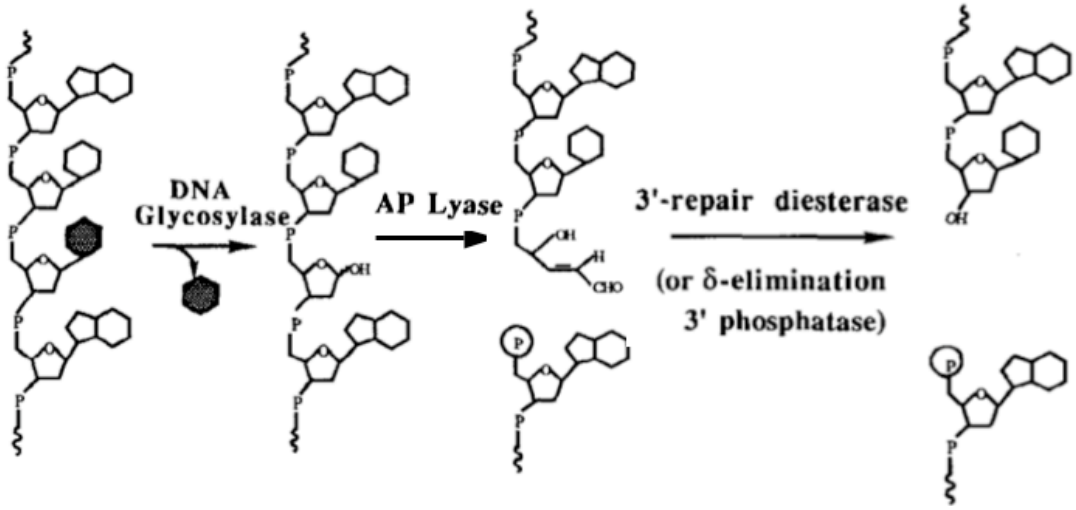


Figure 1.9 Base excision repair by bifunctional glycosylases
The AP lyase activity of the bifunctional glycosylases results in a single strand break produced by a β-elimination reaction. Theses types of glycosylases commonly repair DNA damage induced by oxidative stress. (Demple and Harrison 1994)

Glycosylases responsible for repairing bases damaged by oxidative stress are bifunctional and fall into two sub-groups; those that repair oxidised pyrimidines and those that repair oxidised purine. The endonuclease III (endoIII or Nth) and endonuclease VIII (endo VIII or Nei) families repair pyrimidine lesions such as 5-hydroxyuracil and 5-hydroxycytosine. The 8-oxyguanine-DNA glycosylase (OGG1), MutM and MutY glycosylases recognise the oxidation products of purines such as 8-oxoguanine and formamidopyrimidines. Oxidised bases occur regularly and spontaneously and it is extremely beneficial for the cell to have a large number of glycosylases, with overlapping substrates, to repair this damage.

1.3.2.2 Repair of oxidative stress damage by homologous recombination (HR) and non-homologous end rejoining (NHEJ)

DNA double strand (ds) breaks can result from exposure to ionizing radiation, X-rays, free radicals, and the physical strain experienced when a replicating DNA polymerase encounters a single-strand break. These lesions can have dire consequences for the cell if they are not repaired quickly and effectively. The two main processes for repair of ds breaks are homologous recombination and non-homologous end joining.

The process of homologous recombination is highly conserved in bacteria and eukaryotes; the proteins discussed here are from the human system. Because HR uses

a homologous strand as a template the ends of the break are regressed to aid strand invasion. The complex Rad50/Mre11/Nbs1 uses 5'-3' exonuclease activity to expose the 3' ends of the break. RPA coats the resulting single-stranded DNA and along with Rad52 aids assembly of the Rad51 nucleoprotein filament. The Srs2 helicase ensures no untimely or inappropriate formation of the nucleoprotein filament occurs, such as Rad51 binding ssDNA at stalled replication forks, as this can be detrimental to the cell, inducing strong cell cycle arrest or cell death (Macris and Sung 2005; Malik and Symington 2008). Once a homologous strand has been located Rad51 catalyses strand exchange, the damaged DNA molecule invades the undamaged DNA duplex displacing one strand as a D-loop. DNA polymerase uses the undamaged strand to repair the damage, and the ends are ligated by DNA ligase. The DNA is now in the form of a Holliday junction which is resolved by a specific resolving enzyme to leave two intact DNA molecules, (D'Amours and Jackson 2002; Jackson 2002; Lavin 2007) see Figure 1.10.

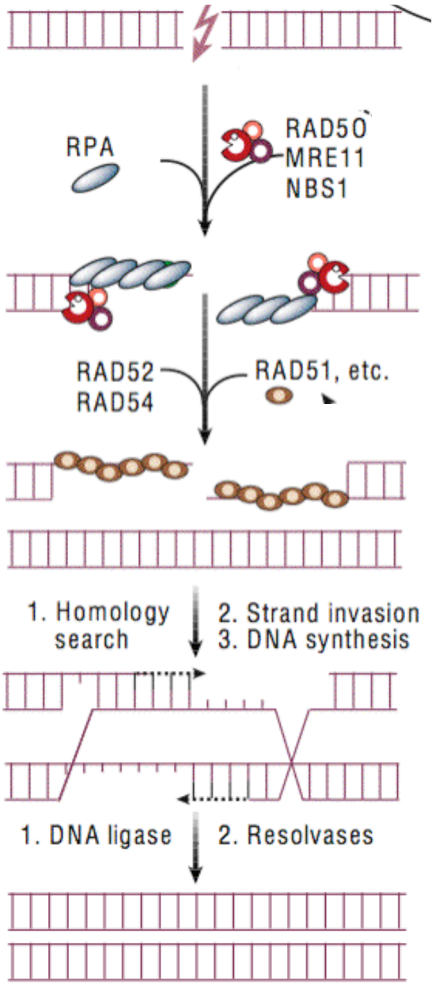


Figure 1.10 Repair of a double strand break by homologous recombination
Once a double strand break has been detected the ends are regressed by the Rad50-Mre11-Nbs1 complex.
The resulting single-stranded DNA is coated with RPA, which aids Rad52 in forming the Rad51 nucleoprotein filament.
An undamaged homologous strand of DNA is used as a template for repair and the resulting Holliday junction is resolved by resolving enzymes leaving two intact DNA duplexes, figure amended from Hoeijmakers 2001.

Non-homologous end joining does not use another, undamaged, DNA molecule as a template for repair, the fact that some processing of the DNA ends must occur before they can be religated makes NHEJ a more error prone process than HR. The central protein in this repair pathway is Ku, a heterodimer made up of two subunits, Ku70 and Ku80. Ku forms an open ring-type structure that can thread onto the broken ends of the DNA. Once bound to the DNA ends the Ku protein recruits the DNA-dependent protein kinase catalytic subunit (DNA-PKca) and activates its protein kinase activity. DNA-PKcs main phosphorylation target appears to be itself, as auotphosphorlyation is needed for the remodelling of the DNA-PKcs complex to make the DNA ends accessible for ligation. In eukaryotes there are a number of proteins involved in the processing of the broken ends, the Artemis nuclease is believed to process broken ends left after damage with ionizing radiation. Y-type polymerases add bases at the broken ends of the DNA to produce complementary ends for ligation. Polynucleotide kinase is also recruited to the damage site to remove 3'-phosphates and add 5'-phosphates to facilitate ligation. The Rad50-Mre11-Nbs1 complex and protein kinases such as ATM are also required for ds break repair. The long flexible, coiled arm of the Rad50 proteins is predicted to be involved in holding the two ends together and the ATM protein facilitates accumulation of the Rad50-Mre11-Nbs1 complex at the damage site by altering the chromatin environment around the damage. Once processing of the ends is complete XRCC4-ligase IV complex ligates the ends, (Hoeijmakers 2001; D'Amours and Jackson 2002; Jackson 2002; Cuadrado, Martinez-Pastor et al. 2006) see Figure 1.11. Homologues of the Ku proteins have been found in bacteria, they are structurally similar to their eukaryotic counterparts despite relatively low sequence similarity (Pitcher, Brissett et al. 2007).

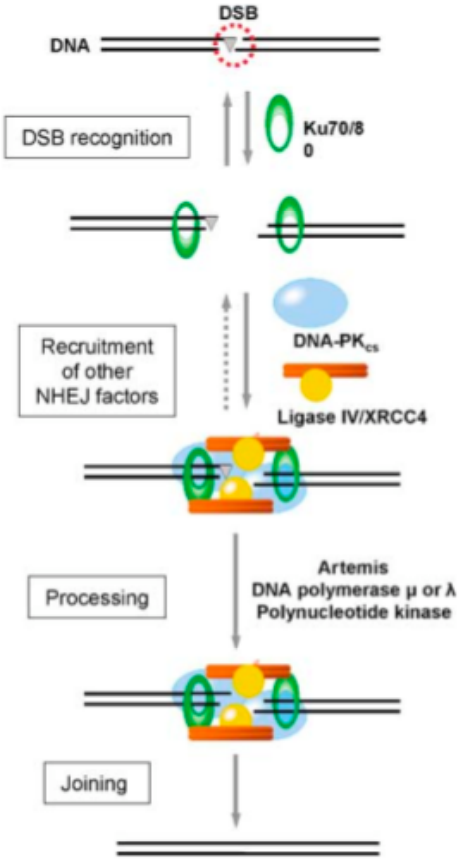


Figure 1.11 Repair of a double strand break by non-homologous end rejoining

Once a double strand break is detected the Ku protein threads onto the ends of the DNA and recruits the DNA-PKcs. Depending of the damage, processing of the DNA ends is performed by Artemis, Rad50-Mre11-Nbs1 and Y family polymerase. XRCC4-ligaseIV joins the ends of the DNA. The process is error prone due to the reprocessing of the broken ends of the DNA. (Van Gent and Van der Burg 2007).

Archaea possess homologues of Rad50, Mre11 and Rad51 (called RadA in archaea). The function of these proteins in archaea is not fully understood, however Mre11 and RadA have been shown to be recruited to DNA after damage with gamma radiation in *S. acidocaldarius* (Quaiser, Florence *et al*. 2008). They are also found bound to DNA before and after gamma irradiation in *Pyroccocus abyssi* suggesting the repair system is preassembled and continually available to repair any damage that occurs (Jolivet, Matsunaga *et al*. 2003). RadA, along with other genes believe to be involved in homologous recombination, are up regulated in *Halobacterium NRC-1* in response to UV irradiation, suggesting homologous recombination plays a role in repairing UV damage in this organism (McCready, Muller *et al*. 2005). The *rad50* and *mre11* genes are clustered with two other genes in the *S. acidocaldarius* genome; *nurA* and *herA*. NurA exhibits both a single-stranded endonucleases activity and a 5'-3' exonuclease activity on single and double-stranded DNA and HerA exhibits helicase activity and can unwind DNA from both a 3' and a 5' single–stranded overhang. The fact that these gene are transcribed together and exhibit functions vital for processing of ds breaks suggested that they played a role in HR in archaea (Quaiser, Florence *et al*.

2008) and subsequent work from Tanya Paull's lab showed all four proteins work cooperatively to create 3'ssDNA for strand invasion in the archaeon *P. furiosus* (Hopkin and Paull 2008)

1.4 Detection of damage

To prevent mutations being passed on to future generations, the cell must be able to detect damage, initiate repair pathways and slow down replication to allow time for the repair to be completed (Early, Drury *et al.* 2004). It has been established that archaea have a transcriptional response to DNA damage by UV light (Salerno, Napoli *et al.* 2003; Frols, Gordon *et al.* 2007; Gotz, Paytubi *et al.* 2007), gamma radiation (Williams, Lowe *et al.* 2007) and oxidative stress (see Chapter 6) but the means by which the damage is detected is still unclear. Detection of damage in the three domains of life is discussed in the next few sections.

1.4.1 Bacterial DNA damage detection

In *E. coli* the job of damage detection and initiation of repair is performed by the SOS response. UV radiation, chemicals such as Mitomycin C and Methyl methane sulfonate and stalled replication forks can all induce the SOS response. The SOS response is initiated when RecA (recombinase A) binds to single-stranded DNA (a product of many types of damage) and becomes active. Levels of RecA in the cell are high, it has a strong affinity for single-stranded DNA and binds with a ratio of one RecA proteins to every three DNA bases. The active form of RecA (RecA*) interacts with the LexA (locus for X-ray sensitivity A) repressor stimulating its auto-cleavage activity. LexA binds some promoters more weakly than others, so some genes maybe induced by minor damage while other are only induced after a strong and persistent damage signal. The cleaved LexA can no longer bind the "SOS box" (palindromic binding site of LexA) and so the repression of more than 40 genes involved in the damage response is relieved (Janion 2008; Justice, Hunstad et al. 2008) see Figure 1.12.

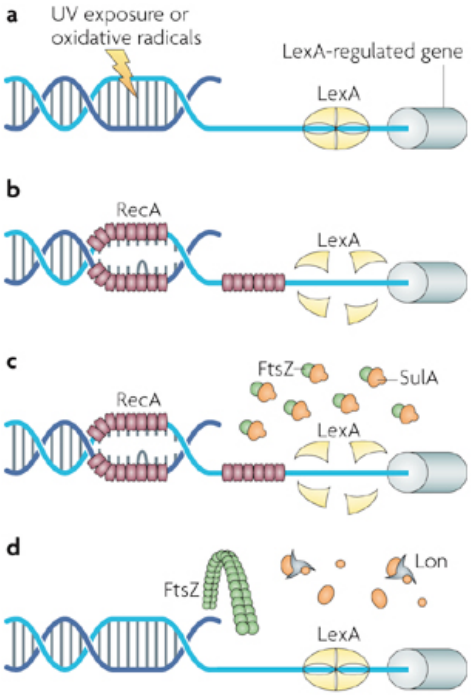


Figure 1.12 The SOS response in bacteria

(a) UV light or oxidative stress causes damage.
(b) The resulting single-stranded DNA is bound by RecA. This binding changes RecA into its active form.
(c) Active RecA* degrades the LexA repressor.
(d) One of the genes repressed by LexA is the cell-division inhibitor, SulA. When released SulA prevents cell division by inhibiting polymerization of cell division proteins. This prevents the transmission of mutant DNA to new daughter cells.
Image taken from Justice, Hunstad *et al* 2008.

RecA* also plays a role in the assembly of the SOS-induced mutagenic DNA polymerase V (PolV). PolV performs translesion synthesis across damage allowing DNA replication to occur despite the presence of damage, a process that often leads to mutations, these mutations presumably facilitate bacterial evolution in times of stress (Butala, Zgur-Bertok et al. 2008).

One of the genes induced by the SOS response is SulA, the protein coded for by this gene inhibits cell division by binding to cell division proteins FtsZ and preventing their polymerization, and the formation of the Z-ring required for cell division (Dajkovic, Mukherjee et al. 2008). The SulA gene is not repressed again until later in the SOS response, inhibition of the cell division allows the cell time to perform the necessary repair (Justice, Hunstad *et al*. 2008).

1.4.2 Eukaryotic DNA damage detection

No SOS type response has been shown in eukaryotes, however, Replication protein A (RPA), a single stranded DNA binding protein, has been shown to accumulate at the site of damage and recruits other repair proteins (Zou, Liu *et al*. 2006). RPA is a single-stranded DNA binding protein and is involved in a large number of cellular functions, such as DNA replication, recombination, cell cycle and DNA repair. RPA

is involved in all major repair pathways (BER, NER, MMR) interacting with various repair proteins such as XPA and XPG in NER, uracil-DNA glycosylase in BER, Rad51, Rad52 and Rad54 in HR. The role of RPA in NHEJ is less clear but its presence stimulates the rate and extent of rejoining (Binz, Sheehan et al. 2004; Fanning, Klimovich et al. 2006; Zou, Liu et al. 2006).

Homologues of RPA are conserved throughout all domains, the human RPA is a heterotrimer, each subunit consists of at least one oligo-saccharide/oligonucleotide binding fold (OB-fold), a C-terminal α-helix domain for protein interaction, and RPA32 has an unstructured N-terminal phosphorylation domain (Binz, Sheehan et al. 2004). After damage RPA is hyperphosphorylated by protein kinases ATM and ATR, which are early signalling molecules known to initiate transduction cascades at sites of single (ATR) and double strand (ATM) breaks. ATM and ATR then phosphorylate numerous different substrates. Some of ATM and ATR substrates, such as Chk1 and Chk2, are checkpoint kinases that initiate cell cycle arrest and allow time for repair of damage (Cuadrado, Martinez-Pastor et al. 2006; Ashwell and Zabludoff 2008). In this way RPA and other damage detection proteins act as damage sensors, effecting transducers (protein kinases such as ATM and ATR) to elicit kinase cascades that results in DNA repair, cell cycle arrest and even apoptosis if the damage is too severe to repair (Jackson 2002), see Figure 1.13.

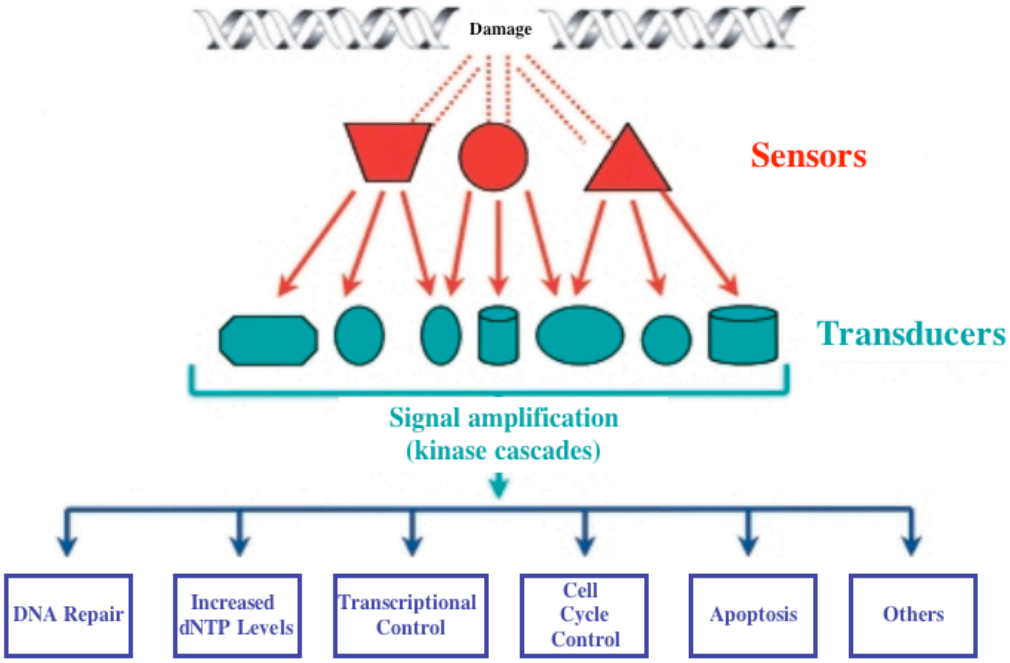


Figure 1.13 Damage detection and signalling pathways
DNA damage is detected by sensors, that signal transducers, which then amplify the signal through different kinase cascades leading to repair, cell cycle regulation and apoptosis, image taken from Jackson 2002.

1.4.3 Archaeal DNA damage detection

Homologues of the SOS LexA proteins are absent from the majority of archaeal genomes, as are the NER damage detection proteins XPC and XPA. Some euryarchaeal species, such as *Pyrococcus*, possess a heterotrimeric RPA with four OB-folds, similar to eukaryotic RPA, but these are absent from the crenarchaea. The crenarchaeal species *S. solfataricus* does possess a single-stranded DNA binding protein (SSB), which is a monomer with a single OB-fold similar to those found in eukaryal RPA and a flexible C-terminal tail similar to that found on bacterial SSB. The C-terminal tail is predicted to interact with other proteins and it has been proposed that SSB may act as a damage recognition protein in *Sulfolobus* (Cubeddu and White 2005). SSB has been shown to melt DNA containing a mismatch or lesion *in vitro*, and its flexible C-terminal tail can interact with RNA polymerase (Richard, Bell *et al*. 2004), reverse gyrase (Napoli, Valenti *et al*. 2005) and the helicase XPB1 (Cubeddu and White 2005). Microarray studies have shown that there is early up regulation of SSB after UV irradiation, suggesting a role in UV repair (Gotz, Paytubi *et al*. 2007). Further study into SSB will hopefully elucidate what role it plays in damage recognition and repair.

1.5 Transcription initiation and control

One consequence of DNA damage is the induction or reduction of the transcription of numerous genes. Transcription initiation in the three domains of life is discussed in this section, and the regulation of transcription in the subsequent section.

1.5.1 Transcription initiation in bacteria

Bacteria contain only one RNA polymerase (RNAP) which comprises of subunits $\beta\beta'\alpha_2\omega$. To recognise a promoter RNAP must first bind a sigma factor and form the holoenzyme. There are several sigma factors in bacteria, which allow binding to different promoters. *E. coli* contains one main sigma factor, σ^{70}, which allows recognition of most promoters, however there are also sigma factors that accumulate in response to specific stress, such as heat shock (Missiakas and Raina 1998). Sigma factors have four domains joined by linkers and are responsible for the following; aiding the RNAP in recognising specific promoter sequences, positioning the RNAP at the promoter and facilitating unwinding of the DNA at the transcription start site (Browning and Busby 2004).

Bacterial promoters contain four sequence elements that are involved in transcription initiation, see Figure 1.14. The –10 and –35 hexamers are located –10 and –35 bases upstream of the transcription start site respectively. The –10 hexamer is recognized by domain 2 of the RNAP sigma subunit, while the –35 hexamer is recognized by domain 4 of the RNAP sigma subunit. Also needed for promoter recognition are the extended –10 and UP element sequences. The extended –10 element is located 3-4 bases upstream of the –10 hexamer and is bound by domain 3 of the RNAP sigma factor. The UP element is located upstream of the –35 hexamer and is bound by the C-terminal domain of the RNAP α subunit (Browning and Busby 2004), see Figure 1.14.

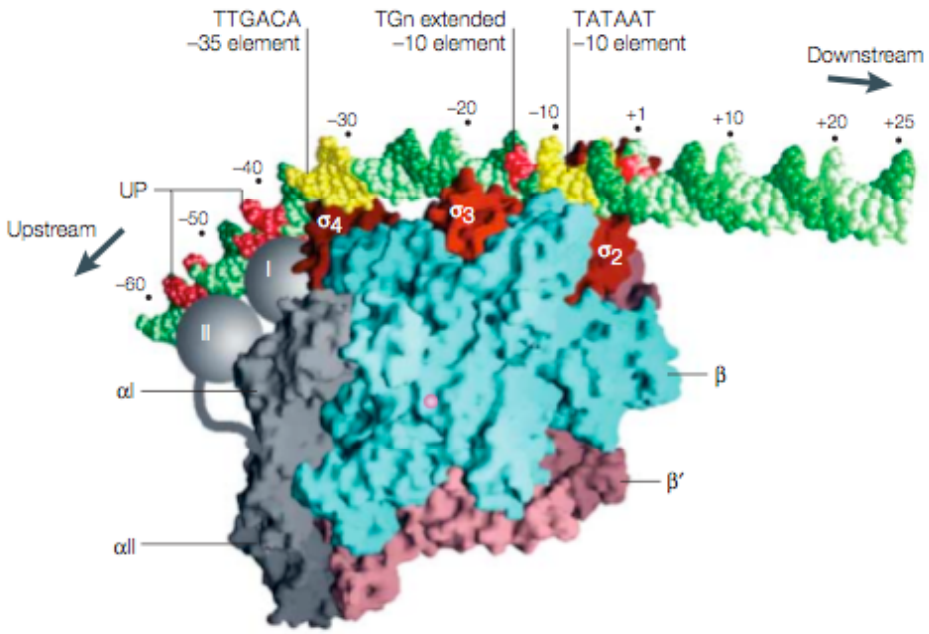


Figure 1.14 Bacterial RNA polymerase binding to a promoter
The crystal structure of bacterial RNAP and promoter DNA shows the contacts between the sigma subunit and the promoter elements. The DNA strand is shown in green with the –10 and –35 elements in yellow, and the extended –10 and UP element in pink. The RNAP subunits β and β' are shown in blue and pink respectively, while the α subunits are shown in grey. The sigma subunits are coloured red. The pink sphere in the middle of the blue β subunit represents the magnesium ion in the active site of RNAP, image taken from Browning and Busby 2004.

Transcription is initiated as follows; RNAP binds a sigma factor forming the holoenzyme. The 4 regions of the promoter (-10, -35, extended –10 and UP element) are bound by the holoenzyme. A section of the DNA from around -10 to +2 is unwound forming the open complex. The non-template strand is bound by domain 2 of the RNAP sigma subunit and the template strand moves to the active site of RNAP so RNA synthesis can begin (Gourse, Ross *et al*. 2000).

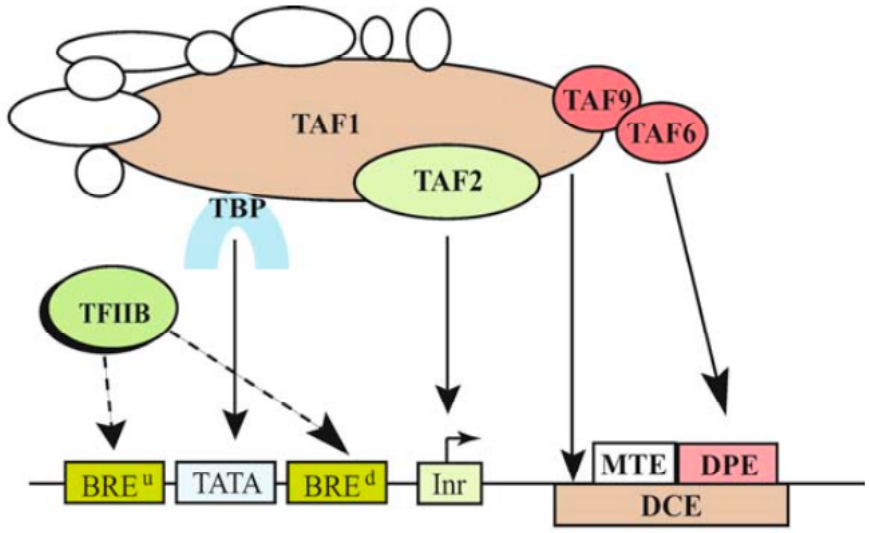
1.5.2 Transcription initiation in eukaryotes

Transcription initiation in eukaryotes is more complex than in bacteria. Eukaryotes possess four RNAP (I to IV) each with a different function. Transcription initiation with RNAP II (involved in production of mRNA) requires six (some multi protein) transcription factors (see Table 1.4) and numerous transcription accessory factors (TAF).

Factor	Function
TFIIA	Stabilizes TBP-TATA complex
TFIIB	Start site selection, stabilizes TBP-TATA complex, RNAP II/ TFIIF recruitment
TFIID	Core promoter-binding factor, co-activator, protein kinase, histone acetyltransferase
TFIIE	Recruits TFIIH, involved in promoter clearance
TFIIF	Binds RNAP II and facilitates its recruitment to the promoter, recruits TFIIE and TFIIH, functions with TFIIB and RNAP II for start site selection, facilitates RNAP II promoter escape, enhances efficiency of RNAP II elongation.
TFIIH	ATPase activity for transcription initiation and promoter clearance, helicase activity for promoter opening, kinase activity for activating RNAP II, ubiquitin ligase activity

Table 1.3 Eukaryotic Transcription Factors
Information in the table is taken from Thomas and Chiang 2006.

The eukaryotic promoters are also more complex than their bacterial counterparts, and possess seven core promoter elements required for the correct assembly and orientation of the preinitiaiton complex (PIC), see Figure 1.15.



Core Promoter Element	Position	Bound Protein
BRE^{u} (TFIIB recognition element up stream of the TATA box)	-38 to –32	TFIIB
TATA (A/T rich sequence)	-31 to –24	TBP
BRE^{d} (TFIIB recognition element down stream of the TATA box)	-23 to –17	TFIIB
Inr (Initiator, pyrimidine rich sequence)	-2 to +5	TAF1/TAF2
MTE (Motif ten element)	+18 to +29	n.a
DPE (Down stream promoter element)	+28 to +34	TAF6/TAF9
DCE (Down stream core element)	3 sub-elements (+6 to +11), (+16 to +21), (+30 to +34)	TAF1

Figure 1.15 Core promoter elements in eukaryotes
The core promoter elements and the transcription factors that bind to them are shown in the top image. The table indicates the position of the core elements in the promoter. Not all promoters contain all of the core promoter elements, leading to a huge amount of diversity in eukaryotic promoters, image taken from Thomas and Chiang 2006.

Transcription initiation with RNAP II occurs in a step-wise manner; first TFIID binds to the promoter region, then TFIIA and TFIIB help stabilise the TFIID-promoter complex, TFIIF and RNAP II are then recruited to the site. Once a stable TFIID-TFIIA-TFIIB-RNAP II/TFIIF-promoter complex has been formed TFIIE is recruited followed by TFIIH. Once the PIC has formed RNA synthesis can begin (Hahn 2004; Kadonaga 2004; Thomas and Chiang 2006).

1.5.3 Transcription initiation in archaea

Transcription in archaea can be initiated with a minimal set of proteins homologous to eukaryotic transcription proteins. The recognition elements found in archaeal promoters are also similar to those found in eukaryotes, such as the TATA box and the BRE (Qureshi and Jackson 1998). Transcription is initiated by the TATA box binding protein TBP, binding upstream of the transcription start site at the TATA box. This is followed by the binding of the transcription factor B protein (TFB, homologue of the eukaryotic TFIIB) which binds to TBP and makes additional contacts with the DNA at the BRE, the binding of TFB also determines the direction of transcription (Bell, Kosa *et al*. 1999). These two proteins recruit RNAP to the transcription start site, at which point the promoter undergoes conformational change unwinding a short section of the DNA allowing the active site of RNA polymerase to bind the DNA and start transcription, see Figure 1.16 (Bartlett 2005).

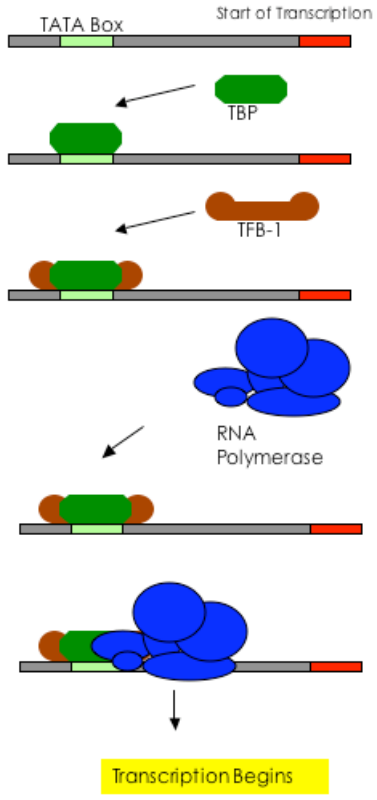


Figure 1.16 Archaeal transcription initiation

Transcription initiation begins with binding of the TATA box binding protein (TBP) to the TATA box. This is followed by the binding of Transcription factor B (TFB) to TBP and the TFB recognition element (BRE) up stream of the TATA box. These factors recruit RNA polymerase to the start site and transcription begins.

1.6 Basal transcription proteins in archaea

1.6.1 TATA box binding protein (TBP)

The first step in transcription initiations is the binding of TBP to the TATA box in the promoter. TBP is a dimer in solution and forms a saddle-like structure, the protein contacts eight base pairs of DNA causing distortion and partial unwinding. This distortion is due to the projection of bulky phenylalanine residues into the minor groove of the DNA (Soppa 1999) , see Figure 1.17.

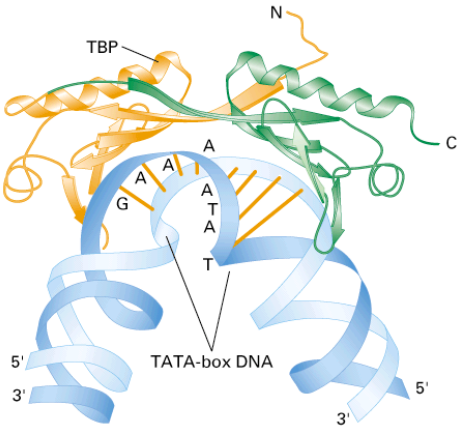


Figure 1.17 Binding of TATA box binding protein (TBP) to DNA
The binding of the TATA box binding protein (TBP) causes bending of the DNA by inserting phenylalanine residues into the minor grooves of the DNA.
Image taken from http://departments.oxy.edu/biology/Stillman/bi221/110300/rna_polymerases.htm.

The archaeal TBP is a homologue of the eukaryotic TBP, which is part of the TFIID complex along with other TBP-associated factors. The N-terminal domain of the eukaryotic TBP is missing from archaeal TBPs, it is not well conserved in eukaryotes but in invertebrates it is involved in transcription performed by RNAP III (Hickey, Conway de Macario *et al*. 2002). The structure of the TBP from the hyperthermophilic archaeon *Pyrococcus woesei* has been solved and unsurprisingly, considering its habitat, it displays significantly higher thermostability and salt tolerance than eukaryotic TBPs (DeDecker, O'Brien *et al*. 1996).

Most of the available archaeal genomes show only one TBP gene, however there are exceptions; *Halobacterium NRC-1* encodes six TBPs, and seven TFBs (Baliga, Goo *et al*. 2000). It has been suggested that these multiple transcription factors may act in different combinations to control transcription initiation, in a similar manner to bacterial sigma factors (Baliga, Goo *et al*. 2000).

1.6.2 Transcription factor B (TFB)

The second protein to bind at the promoter is TFB, the homologue of eukaryotic TFIIB, TFB binds to the C-terminal domain of TBP and makes sequence specific contact with the DNA upstream of the TATA-box at a sequence called the TFB recognition element (BRE) (Qureshi and Jackson 1998), see Figure 1.18. This binding determines the orientation of transcription (Bell, Kosa *et al*. 1999) and the DNA-TBP-TFB complex is able to recruit RNAP to the site.

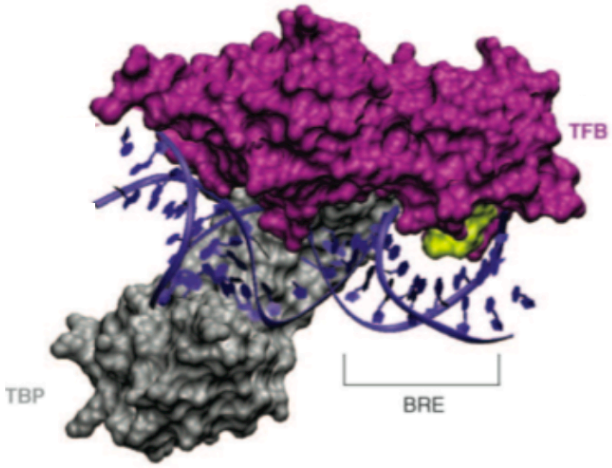


Figure 1.18 DNA-TBP-TFB complex

Structure showing DNA in purple, TBP in grey and TFB in pink. The BRE recognition helix-turn-helix of TFB is shown in yellow, image taken from Bartlett 2005.

TFB comprises of a C-terminal domain containing repeated sequences that form the globular TBP/BRE binding domain, joined by a flexible linker to the N-terminal domain. The N-terminal domain of archaeal TFBs contains a zinc ribbon and B-finger sequence. In *S. solfataricus* TFB has been shown to make contact with the K subunit of RNAP, this interaction is stronger with the zinc ribbon alone, compared to the full length TFB and the TBP-DNA-TFB complex is actually stabilized if the zinc ribbon is absent. However, recruitment of RNAP to the promoter is impaired if the zinc ribbon is mutated (Magill, Jackson *et al*. 2001). In *M. jannaschii* TFB missing the zinc ribbon was still able to support promoter specific transcription, however if the zinc ribbon and the B-finger were deleted a substantial reduction in transcription was observed. In *Saccharomyces cerevisiae* the B-finger penetrates deep into the active site of RNAP II and is believed to stimulate abortive transcription. The B-finger of *M. jannaschii* TFB is believed to interact with RNAP in a similar way (Werner and Weinzierl 2005).

Archaeal genomes possess at least one TFB but many contain more; *Pyrococcus horikoshii* contains two, *S. solfataricus* has three while *Halobacterium NRC-1* contains seven. The reason for these multiple transcription factors is not known, but one of the TFBs in both *S. solfataricus* and *P. horikoshii* is missing the B-finger sequence, and it has been suggested that this truncated TFB may function as a transcriptional regulator, competing with the other TFBs for binding of the RNAP (Soppa 1999; Gotz, Paytubi *et al*. 2007). As mentioned before it has been suggested that the multiple transcription factors may control gene expression in a similar way to sigma factors in bacteria (Baliga, Goo *et al*. 2000).

1.6.3 RNA polymerase (RNAP)

RNA polymerases are highly conserved throughout all domains of life. The archaea possess only one RNAP, its 12 subunits are homologues of eukaryotic RNAP II, and the polymerases are similar in composition and architecture, see Figure 1.19. The 12 subunits can be grouped in terms of function; A', A", B' and B" are involved in catalysis, L, N, D, P are involved in assembly and F, E, H, K have auxiliary functions (Werner 2007).

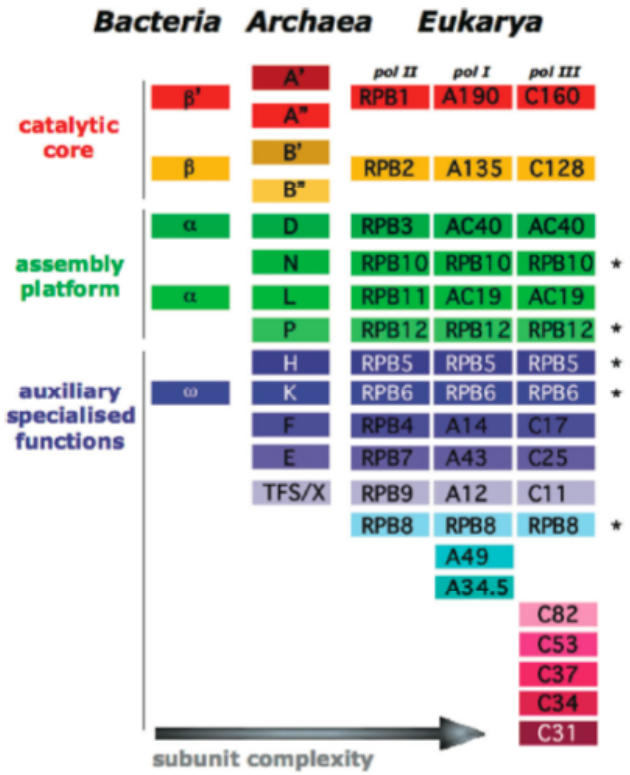


Figure 1.19 RNA polymerase subunit composition in the three domains of life
Subunits are organised by function rather than size, and subunits of similar function are shown in the same colour. Subunits marked with an asterisk are conserved in all three eukaryotic RNAPs. Image taken from Werner 2007.

Archaeal RNAP is around 370 kDa and roughly 70 % of its weight is made up by the four catalytic subunits (A', A'', B', B''), the active site lies between these subunits. The assembly platform (L, N, D, P) allows assembly of the two catalytic subunits (Hirata, Klein *et al*. 2008). Subunit H forms the lower 'jaw' and interacts with downstream DNA, and the F/E 'stalk' subunit interacts with newly synthesized RNA transcript and is predicted to interacts with TFE (Werner and Weinzierl 2005) see Figure 1.21.

RNAP cannot bind DNA directly, and needs TBP and TFB for transcription initiation. Three region of RNAP interacts with TFB; the docking domain interacts with the zinc finger, the TBP-DNA and TFB-RNAP complexes are linked by the TFB core domain and the TFB B-finger domain inserts deep into the active centre of the polymerase (Werner 2007).

1.6.4 TFE

All archaeal genomes sequenced so far contain TFE, a homologue of the α-subunit of the eukaryal transcription factor TFIIE. Although transcription can be initiated with only TBP, TFB and RNAP, TFE has been shown to stimulate transcription from weak promoters (Bell, Brinkman *et al*. 2001). TFE is believed to play a role in stabilizing the initiation complex by influencing the positioning of the F and E subunits of the RNAP (Werner and Weinzierl 2005), see Figure 1.20.

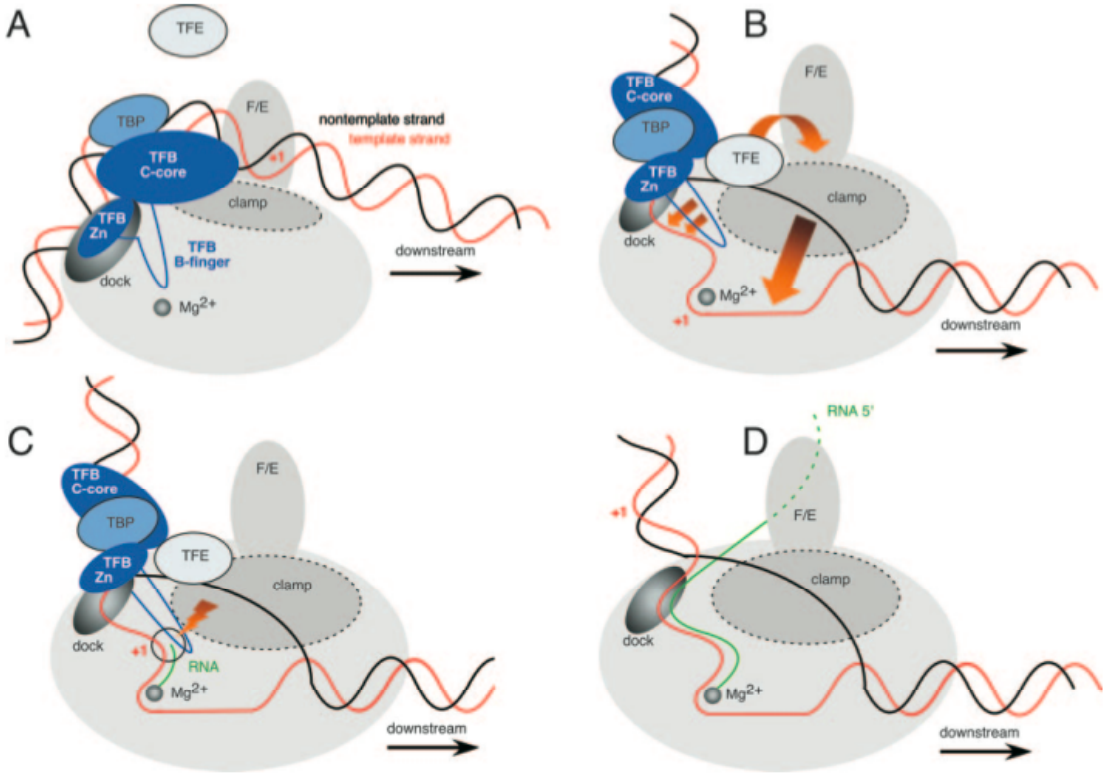


Figure 1.20 Role of TFE in transcription initiation
(A) The DNA-TBP-TFB complex recruits RNAP to the promoter. RNAP makes only superficial contacts with the DNA at this point. (B) During open complex formation the melted DNA strand is inserted into the active site of RNAP where it is stabilized by the B-finger of TFB. The complex is further stabilised by TFE, which influences the position of the RNAP subunits F and E, and possibly interacts with the non-template strand. (C) The TFB B-finger stimulates abortive transcription, possibly by aiding the formation of a DNA-RNA-rNTP configuration (indicated by the orange lightening flash). (D) The basal transcription proteins are released from the promoter and the elongation phase begins. Image taken from Werner and Weinzierl 2005.

1.7 Transcriptional regulation

1.7.1 Transcriptional regulation in bacteria

The key protein regulating transcription in bacteria is RNA polymerase. This multi-subunit protein is responsible for all transcription in combination with sigma factors. Different sigma factors target the RNAP to different promoters, for example binding of the heat shock sigma factor (σ^H) results in the transcription of genes involved in the heat shock response (Missiakas and Raina 1998).

Transcriptional regulators, proteins that bind the promoter and elicit either an inhibitory or stimulatory effect are also involved in the regulation of transcription. Transcription factors allow the cell to alter gene expression in response to their

environment and they are tightly regulated. Factors such as the concentration of small ligands, structural changes caused by modifications such as phosphorylation or binding of a regulatory protein to the transcription factor function to control their action. The transcriptional regulators have been grouped into families based on their sequence similarity. One of these families, the Lrp/AsnC family is found extensively in bacteria and archaea, but not in eukaryotes. Their effect can be global (in the case of Lrp) or specific (in the case of AsnC, which is an asparagine-dependent activator of asparagine synthase). They have a molecular weight of around 15 kDa and comprise of an N-terminal helix-turn-helix DNA-binding domain, joined by a linker to a C-terminal regulation of amino acid synthesis (RAM) domain, responsible for co-factor binding and oligomerization. The *E. coli* Lrp regulates up to 75 genes, it can act as a repressor or activator and is most commonly modulated by leucine (Brinkman, Ettema *et al*. 2003).

Another family of regulators with homologues in the archaea is the Diphtheria toxin repressor (DtxR) family. DtxR from *Corynebacterium diphtheria* is an iron activated transcription factor involved in the regulation of the diphtheria toxin and genes involved in iron homeostasis (Brune, Werner et al. 2006). DtxR forms a dimer, each monomer comprises of an N-terminal helix-turn-helix DNA-binding domain, a dimerization domain containing two metal ion-binding sites and a flexible C-terminal domain (White, Ding *et al*. 1998). DNA binding sites for the proteins are palindromic repeat sequences (Schmitt, Twiddy et al. 1992) and two dimers of DtxR bind either side of the DNA, see Figure 1.21.

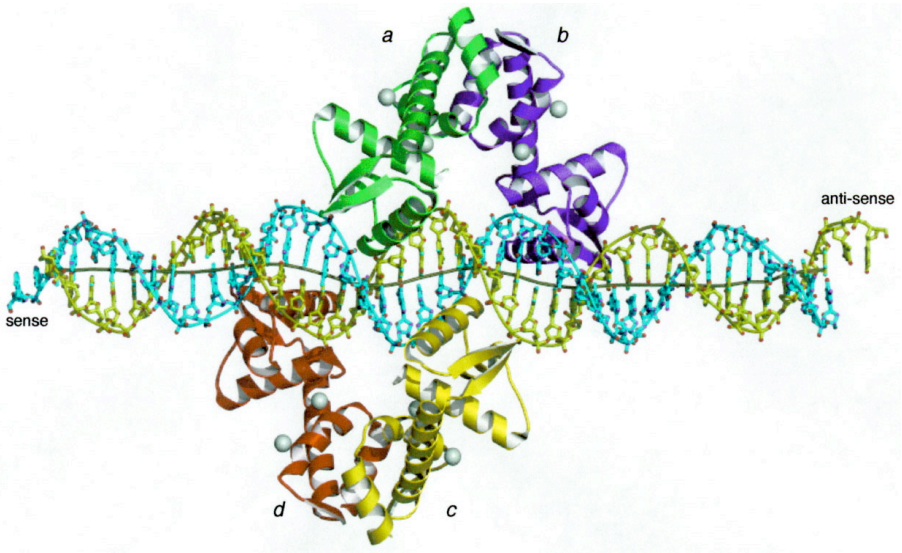


Figure 1.21 Two DtxR dimers binding DNA
Two dimers of DtxR binding palindromic repeats on opposite faces of the DNA (White, Ding et al. 1998).

Both metal binding sites in the protein must be occupied for it to adopt the correct DNA–binding conformation. *In vivo* the metal occupying the sites is ferrous iron, however *in vitro* experiments have shown that other divalent metal can also occupy the sites (Spiering, Ringe *et al*. 2003; Rangachari, Marin *et al*. 2005). When ferrous iron is not available to occupy the binding site the protein cannot adopt the correct DNA binding conformation and so repression of transcription is relieved. DtxR is predicted to regulate transcription of the Dps protection protein in *C. diphtheria*, a protein involved in the cells response to oxidative stress (Yellaboina, Ranjan *et al*. 2004).

1.7.2 Transcriptional regulation in eukaryotes

Gene transcription in eukaryotes is a multi-step process involving a large number of proteins and numerous levels of control. *In vivo* a cell's genome is carefully packed away by histones, not only so that it fits within the cell, but also to protect it from damage. For a gene to be transcribed the transcription proteins first need access to the DNA. Modification of histones, by acetylation, phosphorylation, methylation, sumoylation and ubiquitinylation plays a vital role in gene regulation. Sequences within the DNA act as targets for activator proteins and the binding of these proteins results in recruitment of histone and chromatin modifying enzymes that make the

gene accessible to general transcription factors, see Figure 1.22. Activator sequences are then bound by activator proteins, initiating recruitment of RNAP. RNAP on its own cannot bind the promoter, and acts in concert with a large number of accessory factors to initiate transcription. These accessory factors are specific for *cis*-regulatory sites within the DNA and allow the RNAP to be targeted to specific promoters.

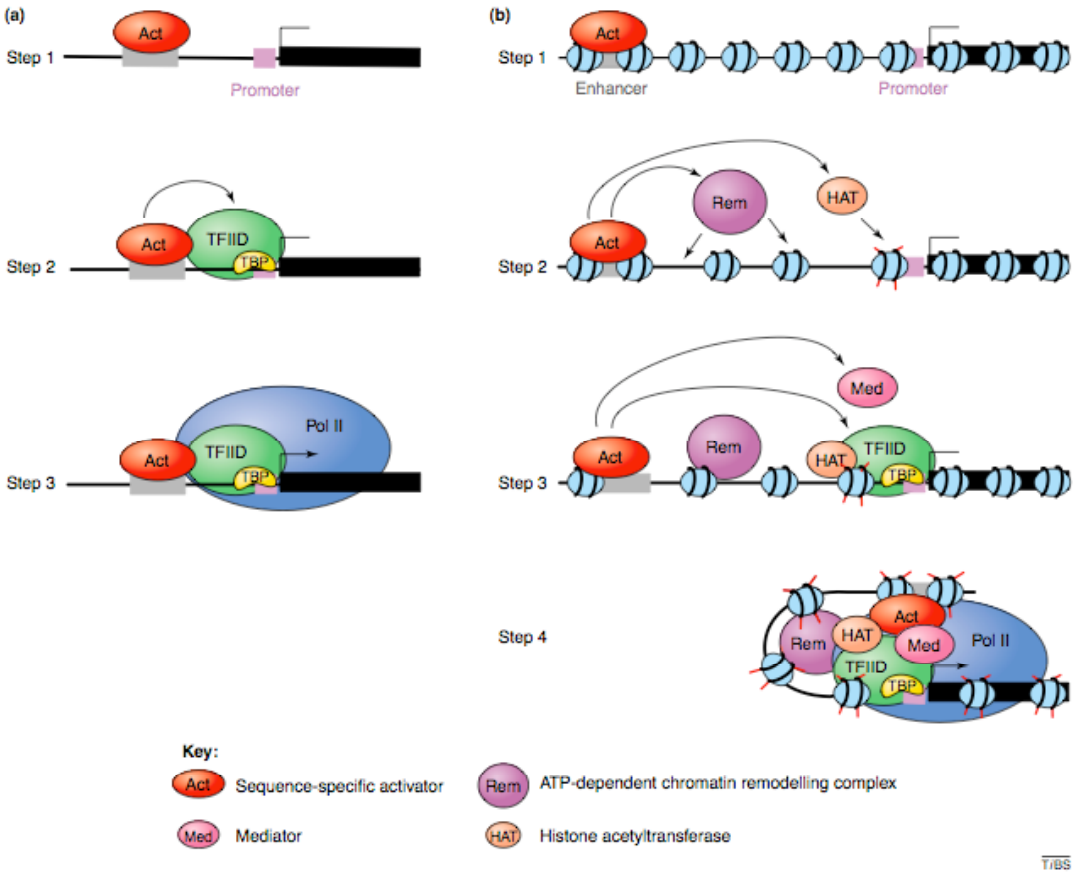


Figure 1.22 Transcription initiation of 'naked' and histone bound DNA
(a) Shows the process of transcription initiations from 'naked' DNA, activation is initiated by the binding of a sequence specific activator to the enhancer sequence up stream of the promoter, which recruits TFIID. After interactions of other transcription factors RNAP is recruited and transcription can begin. In (b) transcription is initiated from DNA bound by histones. An activator binds a distant enhancer sequence, leading to recruitment of chromatin (Rem) and histone (HAT) modifying enzymes. The transcription proteins can now access the promoter and mediator complexes (Med) provide a link between the activator bound enhancer element and the basal transcription machinery at the promoter. Transcription can proceed once direct contact between the enhancer and promoter is established by looping of the DNA (Szutorisz, Dillon et al. 2005).

General transcription factors are important in regulation of gene expression, recruiting the basal transcription proteins to specific promoters. Binding sites for enhancer proteins can be as much as 2 kilobases upstream of the transcription start site. For

transcription to begin contact between the proteins at the promoter and the enhancer must be established. It has been proposed that this could occur either by the RNAP tracking along the DNA or the DNA looping back on itself to make the contact (Szutorisz, Dillon et al. 2005).

1.7.3 Transcriptional regulation in archaea

In archaea the cellular processes of transcription, translation, replication, recombination and repair are more similar to those of eukaryotic systems than bacterial ones, so it is surprising that the majority of archaeal transcriptional activators and repressors discovered so far are bacterial homologues (Dussurget, Timm *et al*. 1999; Bell 2005).

Homologues of the Lrp/AsnC family make up the majority of transcription regulators found so far in the archaea. Many self-regulate by binding their own promoter. An example of a self regulating Lrp-like transcriptional regulator is the LrpA proteins from *P. furiosus*. Multiple copies of this protein binds its own promoter and inhibit transcription, possibly by distorting the DNA (Leonard, Smits *et al*. 2001), see Figure 1.23.

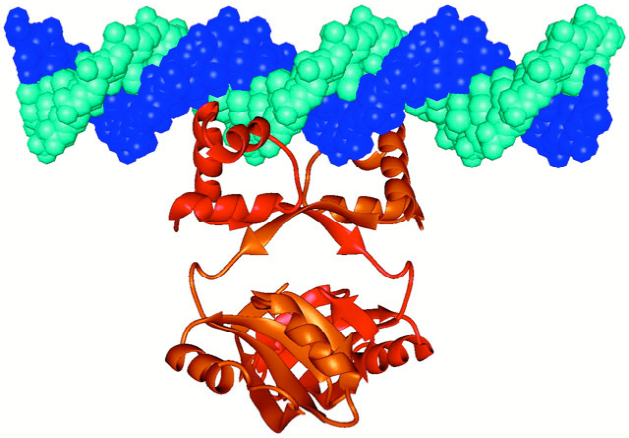


Figure 1.23 LrpA regulator protein from *Pyrococcus furiosus* bound to DNA
LrpA forms a dimer and the two recognition helices access adjacent turns of the major groove, figure taken from Leonard, Smits *et al*. 2001.

A number of the known archaeal transcriptional regulators are discussed in Table 1.4.

Species	Protein Name	Activator or Repressor	Action	Reference
Sulfolobus solfataricus	Ss-LrpB	Activator at low conc, repressor at high conc	Forms 3 distinct complexes depending on whether 1, 2 or 3 of the binding sites in its own promoter/operator region are occupied. When all 3 sites are bound a bend of 148° occurs in the DNA	(Peeters, Thia-Toong et al. 2004)
Sulfolobus solfataricus	Lrs14	Repressor	Exerts negative autoregulation by binding the promoter region and preventing access to TBP and TFB	(Napoli, van der Oost et al. 1999)
Sulfolobus solfataricus	LysM	Activator	Is involoved in the regulation of Lysine. Binds up stream of the BRE in the weak *lysW* promoter, enhancing transcription, with the help of an as yet unidentified co-activator. Lysine binding to LysM decreases its affinity for DNA	(Brinkman, Bell et al. 2002)
Pyrococcus furiosus	LrpA	Repressor	Binds to a 46 base pair sequence overlapping the transcription start site of its own promoter, and so inhibits recruitment of RNAP	(Brinkman, Dahlke et al. 2000; Leonard, Smits et al. 2001)
Pyrococcus furiosus	Phr	Repressor	Regulator of the heat shock response. Represses transcription form its own promoter, Hsp20 heat shock protein and a AAA+ ATPase. Acts by preventing RNAP from binding the TBP/TFB complex, by binding a 29 base pair sequence overlapping the transcription start site.	(Vierke, Engelmann et al. 2003)
Methanocaldococcus jannaschii	Ptr2	Activator	Stimulates transcription by facilitating recruitment of TBP to the TAT box	(Ouhammouch, Langham et al. 2005)
Archaeoglobus fulgidus	MDR-1	Repressor	Binds to its own promoter in a metal dependent manner, blocking RNAP recruitment	(Bell, Kosa et al. 1999)
Haloferax mediterranei	GvpE	Activator	Activates transcription from 14 *mc-gvp* genes that produce gas vesicles. GvpE recognises a 20 base pair nucleotide sequence up stream and adjacent to the BRE	(Zimmermann and Pfeifer 2003)
Haloferax mediterranei	GvpD	Repressor	Represses transcrption from 14 mc-gvp genes that produce gas vesicles. The mode of action is not fully understood.	(Scheuch, Marschaus et al. 2008)

Not all regulators in archaea are homologues of the Lrp/AsnC family. The metal dependent repressor (MDR1) from *Archaeoglobus fulgidus* is homologous to proteins of the DtxR family. MDR-1 binds to the operator sequences in its own promoter, and like DtxR, needs metal for DNA-binding, it represses transcription by preventing recruitment of RNAP to the promoter (Bell, Cairns *et al*. 1999).

There are some examples of eukaryotic-like regulators in archaea. The Tfx protein (a transcriptional repressor from *Methanothermobacter thermautotrophicus*) contains a possible acidic activation domain similar to those found in eukaryotic regulators (Hochheimer, Hedderich et al. 1999). The GvpE transcriptional activator from *Haloferax mediterranei* belongs to the leucine zipper class of eukaryotic transcriptional regulators and activates transcription of genes involved in the formation of gas vesicles (Zimmermann and Pfeifer 2003).

A putative transcriptional repressor (Sso2273), homologous to the DtxR family was expressed and characterised for this thesis. This protein was isolated after investigation of the transcriptional response of a set of genes (shown to be affected by UV radiation) to different kinds of DNA damage. In particular the response of the Dps protein to hydrogen peroxide damage was looked at and Sso2273 was isolated, using pull down assays, as a possible transcriptional repressor for the *dps* gene. The function of the *sso2273* gene was explored using gene knockout and microarray technologies.

Chapter Two

Materials and Methods

2.1 Growth of *Sulfolobus solfataricus* cultures

Sulfolobus solfataricus P2 (DSMZ 1617) frozen pellets (provided by Dr Dorothee Götz) were grown aerobically at 80 °C with vigorous shaking in *Sulfolobus* media (Brock TD 1972) supplemented with yeast extract (1 g/l) and adjusted to pH 3 with H_2SO_4.

2.2 Extraction of DNA from *S. solfataricus* cultures

DNA was extracted from aliquots of growing *S. solfataricus* culture using Qiagen Dneasy Blood and Tissue kit following the protocol for gram-positive bacteria.

2.3 Extraction of RNA from *S. solfataricus* cultures (Qiagen Kit Method)

RNA was extracted from aliquots of growing *S. solfataricus* culture using Qiagen Rneasy Tissue kit following the protocol for gram-positive bacteria. The lysozyme step was replaced with a proteinase K incubation. Cell pellet was resuspended in 100 µl of TE buffer supplemented with 2.5 µl of 20 mg/ml proteinase K and samples were incubated at 65 °C for an hour. After extraction, RNA was run on a 1% agarose:TBE gel to check for DNA contamination.

2.3.1 Extraction of RNA from *S. solfataricus* cultures (Phenol:Chloroform method)

50 ml aliquots of growing *S. solfataricus* culture were pelleted by centrifugation at 4,000 x g at 4 °C. Pellets were resuspended in Lysis buffer (4 M Guanidine isothiocyanate, 25 mM Sodium citrate pH 7, 0.5 % Sarcosyl, 100 mM β Mercaptoethanol) and incubated on ice for 15 min. 0.2 M Sodium acetate, 500 µl water saturated phenol and 100 µl 24:1 (v/v) chloroform:isoamylalcohol was added to the samples and they were incubated on ice for 15 min before being centrifuged at 11,000 x g for 15 min. The water phase was transferred to a new tube and 500 µl 24:1 (v/v) chloroform isoamylalcohol added. Samples were again incubated on ice for 15 min then centrifuged. The water phase was transferred to new tubes and precipitated overnight at –20 °C with an equal volume of isopropanol. Samples were centrifuged and the pellet washed with 75% ethanol. Pellet was speedvac dried and resuspended in 169 µl of RNase free water. Samples were DNase treated (Qiagen) and cleaned

using the phenol:chloroform method as before. The pellet was speedvac dried and resuspended in 50 µl RNA storage solution (Ambion). The quality of the RNA was checked on a 2 % agarose:TBE gel and the concentration determined by measuring the UV absorbance at 260 nm.

2.4 DNA damage treatments of *Sulfolobus* cultures

2.4.1 UV irradiation of *Sulfolobus* cultures

Sulfolobus cultures were grown to early exponential phase (OD_{600} approx 0.2). Cultures were irradiated in 50 ml volumes in tissue culture plates and returned to the incubator immediately after irradiation. Cultures were exposed to 200 J/m_2 UV light (254 nm) using a Stratalinker® UV crosslinker 1800 (Stratagene). A control culture was removed from the incubator at the same time so any effects caused by cold shock could be taken into account. Samples of 1 ml were taken at 30, 60, 90 and 120 min after irradiation.

2.4.2 Mitomycin C, Methyl methane sulfonate, Phleomycin and Hydrogen peroxide damage

Initially, cultures were treated with a series of concentrations of the damaging agent and growth curves used to determine a concentration that would produce damage, but also allow the cells to recover. The final concentrations used were as follows: Mitomycin C 2 µM, Methyl Methane Sulfonate 300 µM, Phleomycin 200 µM, Hydrogen Peroxide 1 µM (RT real time experiments) or 5 µM (microarray experiments).

2.5 Reverse transcription real time PCR with *Sulfolobus* RNA

Reverse transcription real time PCR was performed using BioRad iScript One-Step RT PCR kit with SYBR green. First the amplification efficiency for each primer set was determined using serial dilutions of genomic DNA. All experiments were performed on a BioRad iCycler and samples for each time point were performed in triplicate. Crossing point (CP) values of treated and control samples were used to determine the ratio of gene expression, using the formula from (Pfaffl 2001). Primer sequences for each gene are shown in Appendix A.

2.6 Determining gene expression using the Pfaffl equation

An equation from an article by Pfaffl *et al* (2001) was used to determine the change in gene expression after damage. The equation uses the real time PCR efficiencies, determined by amplifying a serial dilution of DNA, and the crossing point values of the control and treated sample, to work out the relative expression ratio for each gene.

$$E = 10^{[-1/slope]} \qquad \text{Expression Ratio} = \frac{(E_{target})^{\Delta CP_{target}(control - sample)}}{(E_{ref})^{\Delta CP_{ref}(control - sample)}}$$

Figure 2.1 Equations for reaction efficiency and expression ratio determination for real time PCR.
The equation on the left is used to determining PCR efficiency (E) using the value for the slope produced from a real time PCR amplification of a serial dilution of *S. solfataricus* DNA. The equation on the right is used to determine the expression ratio using the PCR efficiency of the target and reference genes (E_{target}, E_{ref}) and the difference in crossing point (CP) value between the control and treated samples, which is determined by (control – sample).

2.7 Pull down assays using Biotin labelled DNA

Sulfolobus solfataricus cultures were grown at 80 °C to an OD_{600} of 0.4. The culture was divided in two. Half was treated with 1 μM Hydrogen peroxide and the other half was left as an untreated control culture. Both cultures were placed back in the incubator for an hour.

Cultures were centrifuged at 4 °C, 6,000 x g for 30 min using a Beckman JLA 8.1000 rotor. Cell pellets were resuspended in 2 ml of Binding buffer (20 mM Mes (pH 6.5), 50 mM NaCl, 5 mM EDTA, 1 mM DTT). Cells were sonicated on ice and then centrifuged at 4 °C, 20,000 x g for 30 min. The supernatant was filtered using 45 μm Tuffyrn Filters.

A doubled stranded oligo 45 bases in length, representing the upstream region of the *dps* (DNA protection protein from starved cells) promoter, was formed by annealing two complementary single stranded oligos (Operon). One oligo possessed a biotin label (see Appendix A for sequence), equal amounts of the two complementary oligos were placed in a water bath at 95 °C for 2 min. The water bath was turned off and the oligos allowed to cool and anneal.

700 pmols of *dps* promoter oligo, or T6 SSV1 single stranded oligo (see Appendix A for sequence), were added to 0.6 mgs of Promega Magnasphere magnetic beads (which had been washed in 0.5 x SSC) and incubated for 10 min at room temperature. Beads were then washed 3 times with 0.1 x SSC. Beads were incubated with 20 mg of cell lysate (H_2O_2 treated or Control) at 50 °C for 1 hr on a shaking platform.

Beads were captured with a magnet and the supernatant removed. The beads were washed 5 times with Binding buffer containing 150 mM NaCl. Bound proteins were eluted using 300 µl of Binding buffer with 250 mM NaCl, then 300 µl of Binding buffer with 500 mM NaCl and finally 300 µl of Binding buffer with 1 M NaCl.

Eluted proteins were Trichloroacetic acid (TCA) precipitated as follows. Samples were incubated on ice for 30 min with 150 µl of 45 % TCA. Samples were centrifuged for 15 min at 13,000 x g in the cold room, and supernatant discarded. 500 µl of cold acetone was added and samples centrifuge for 15 min at 13,000 x g in the cold room. Supernatant discarded and pellet dried for 5 min in the speed vac. The white pellet produced was resuspended in 30 µl of protein loading buffer and placed at 95 °C for 5 min. Samples were analysed by SDS-PAGE, gel stained with SYPRO Ruby Stain and visualised using Image Gauge FLA5000 (Fuji). Bands of interest were cut out and identified by the University of St Andrews Mass Spectrometry service.

2.8 Cloning and expression of *sso2273* and *sso0669*

2.8.1 Cloning procedure and vectors

The gene sequence of *sso2273* and *sso0669* was obtained from the *Sulfolobus solfataricus* P2 complete genome sequencing project database (http://www-archbac.u-psud.fr/projects/sulfolobus/). The genes were amplified from genomic DNA using Taq polymerase (Promega). The oligonucleotides used were designed with *Bam*HI/*Nco*I restriction sites on the forward primers and a *Sal*I restriction site on the reverse primers for ligation into a pET28c plasmid (Novagen). See Appendix A for sequences.

2.8.2 Over expression in *Escherichia coli*

Sso2273 and Sso0669 were expressed in *Escherichia coli* BL21 Rosetta cells (Novagen) and cultures grown in Luria Bertani (LB) medium containing 35 μg/ml Kanamycin. For some experiments cells were also grown in LB media containing $ZnCl_2$. 50 μM $ZnCl_2$ was added at the same time as the IPTG was added.

Cultures were grown in a shaking incubator at 37 °C to an OD_{600} of 0.9 – 1.0. Once this OD_{600} was reached protein expression was induced with 0.5 mM of IPTG and cultures were grown at 37 °C for a further 3 hrs.

2.9 Purification of Sso2273 and Sso0669

2.9.1 Sso2273 purification

After 3 hrs induction cells were pelleted by centrifugation at 4 °C, 8,000 x g for 20 min. Pellets were resuspended in 10 ml of Lysis buffer (50 mM Tris pH 7.5, 50 mM NaCl, 1 mM DTT, 1 EDTA-free Protease inhibitor tablet (Roche) per 50 mls of buffer) and sonicated on ice 4 x 2 min. The sonicated cells were centrifuged at 4 °C, 20,000 x g for 20 min. Supernatant was placed at 70 °C for 15 min, to precipitate any *E.coli* proteins, and then centrifuged again at 20,000 x g for 20 min. The resulting supernatant was filtered through an Aerodisc 0.45μm syringe filter.

Filtered lysate was diluted 1:1 with Buffer A (20 mM Mes pH 6, 1 mM DTT) and loaded onto a Heparin column (HiTrap™ 5 ml HP, Amersham Biosciences) and eluted over a 100 ml linear gradient comprising 0-1 M NaCl. Fractions corresponding to peaks were collected and analysed by SDS-PAGE (Invitrogen). Fractions containing proteins of the same molecular weight as Sso2273 (15kDa) were pooled and concentrated (Vivaspin columns). Concentrated protein was loaded onto a HiLoad® 26/60 Superdex® 200 size exclusion column (Amersham Biosciences) equilibrated with Gel Filtration buffer (20 mM Mes pH 6, 150 mM NaCl, 1 mM DTT). Fractions corresponding to peaks were checked by SDS-PAGE (Invitrogen) and fractions containing the protein were pooled and concentrated using Vivaspin 5 kDa cut off spin filters. The identity of the protein was confirmed by the Mass Spectrometry service at the University of St Andrews.

2.9.2 Sso0669 purification

Cell Lysate was obtained following the same protocol as for Sso2273, see section 2.9.1. Filtered lysate was diluted 1:1 with Q Column buffer A (40 mM Tris pH 8, 1 mM DTT) and loaded onto a Q column (HiTrap™ 5 ml Q HP, Amersham Biosciences). Protein Sso0669 was eluted over a 100 ml linear gradient comprising 0-1 M NaCl. Fractions corresponding to peaks were collected and analysed by SDS-PAGE (Invitrogen). Fractions containing proteins of the same molecular weight as Sso0669 (24 kDa) were pooled and concentrated using Vivaspin 5 kDA cut off columns. Concentrated protein was loaded onto a HiLoad® 26/60 Superdex® 200 size exclusion column (Amersham Biosciences) equilibrated with Gel Filtration buffer (40 mM Tris pH 7.5, 20 mM MgCl, 2 mM DTT). Fractions corresponding to peaks were checked by SDS-PAGE (Invitrogen) and fractions containing the protein were pooled and concentrated using Vivaspin 5 kDa cut off spin filters. The identity of the protein was confirmed by the Mass Spectrometry service at the University of St Andrews.

2.9.3 Determination of protein concentration

Protein concentrations were determined using absorbance at 280 nm, taking into account the protein extinction coefficient (from ProtParam http://ca.expasy.org/tools/protparam.html) or using a Bradford Assay (Bio-Rad). Bradford reagent was used as per manufacturers instructions, and protein concentration calculated using a BSA standard curve.

2.9.4 Analytical gel filtration

A HiPrep™ 16/60 Sephacryl™ S-300 HR column (GE Healthcare) was equilibrated with 50 mM Tris pH 7.5, 100 mM KCl. 1 ml of Sso2273 at concentration 3 mg/ml, that had been grown with $ZnCl_2$ in the growth media (see section 2.8.2) was run through the column at 1 ml/min, the absorbance trace at 280nm was recorded. The same procedure was repeated with Sso2273 that had been grown up without $ZnCl_2$, and protein that had been reconstituted with $NiCl_2$ and $MnCl_2$.

A standard curve had been produced previously using proteins of known molecular weight; blue dextran (2000 kDa), beta-amylase (200 kDa), alcohol dehydrogenase (150 kDa), albumin (66 kDa), carbonic anhydrase (29 kDa) and cytochrome c (12.4 kDa). The standard curve was used to determine whether Sso2273 formed a dimer upon the addition of metal.

2.10 Bathophenanthroline assay for determination of protein Iron content

100 µl of 50 µM Sso2273 was mixed with 30 µl of concentrated HCl and heated for 15 min at 100 °C. A negative control made with buffer was also treated. Samples were centrifuged for 2 min at 13,000 x g, 100 µl of the resulting supernatant was transferred to 2 ml tubes and 1.3 mls of 500 mM Tris-HCl (pH 8.5) was added. 100 µl of freshly prepared 5 % ascorbic acid was added and samples vortexed, 400 µl 0.1 % bathophenanthroline was added and samples vortexed. Finally samples were incubated for an hour at room temperature and the absorbance measured at 535 nm.

2.11 Metal reconstitiution

Sso2273 was reconstituted with divalent metal salts ($Fe(NH_4)_2SO_4$, $MnCl_2$, $NiCl_2$, $CoCl_2$). Purified Sso2273 in various concentrations was added to 1 ml of Reconstitution buffer (10 mM Mes pH 6, 50 mM NaCl) plus 5 mM metal salt. Reaction mixtures were incubated at room temperature for an hour then applied to PD-10 desalt columns (GE Healthcare) that had been pre-equilibrated with 10 mM Tris-HCl pH 6. The flow-through was discarded; 3 mls of 10 mM Tris-HCl pH 6 was applied to the columns and collected. The metal reconstituted protein was concentrated from 3 mls to 500 µl using a 5 kDa cut off Vivaspin column. Concentration of the metal reconstituted proteins was determined using a spectrophotometer.

2.12 Inductively coupled plasma – optical emission spectrometry (ICP-OES)

ICP-OES was performed on a Perkin-Elmer Optima 5330 ICP-OES machine at the Chemistry Department of the University of Edinburgh. Samples were purified and metal reconstituted as previously described. Standards for Fe and Mn of 10 ppm, 5 ppm, 1 ppm, 0.1 ppm were prepared to use for the standard curve. 60 µM of native

Sso2273 and Fe and Mn reconstituted Sso2273 were analysed on the machine and compared to the standard curve. Buffer that flowed through the concentrators at the end of the metal reconstitution was used as a blank.

2.13 DNA interactions

2.13.1 Electrophoretic mobility shift assays (EMSA)

Increasing concentrations of Sso2273 were incubated with two different [γ-^{32}P] ATP-labelled oligonucleotides, one corresponding to a 45 bp sequence up-stream of the *dps* gene and a second that contains the same bases but scrambled, to check for sequence specific binding. 10 nM labelled oligonucleotide and different concentrations of Sso2273 were incubated in Binding buffer (20 mM Tris pH 7.5, 50 mM NaCl, 2 mM DTT, 0.002 % Triton X, 0.1 mg/ml BSA) at room temperature for 30 min, control reactions without protein were also set up. Finally 3 µl of loading dye (15 % Ficoll, 0.25 % bromophenol blue) was added and the sample loaded onto an 8% acrylamide:TBE gel and run at 130 V for 3 hrs in 1 x TBE buffer. Gels were exposed to photoscreen and imaged using Fuji Image Gauge software.

2.13.2 Primer extension transcription assays

A PCR 2.1 TOPO vector (Invitrogen) plasmid containing a section of the *dps* promoter region was digested with *Xho*I to produce linear plasmid DNA. 75 ng of linear plasmid DNA was incubated in 2 x Transcription buffer (40 mM Tris-HCl pH 8, 20 mM $MgCl_2$, 440 mM KCl, 4 mM DTT) with various concentrations of Sso2273 or Sso0669, at 65 °C for 10 min. TBP (80 nM), TFB-1 (80 nM) and RNAP (40 nM) were then added and the samples incubated at 70 °C for a 10 min. rNTPs (200 µM each) were then added and the samples incubated for a further 20 min. Placing the tubes on ice stopped the reaction.

The RNA was then used as a template for the primer extension reaction. A [γ-^{32}P]ATP-labelled reverse oligonucleotide specific to the sequence downstream of the *dps* promoter was used as a primer for the reverse reaction.

Once the primer and RNA had been hybridized by incubation at 70 °C for 5 min RevertAid H Minus M-MuLV Reverse Transcriptase 5 x Buffer (Fermentas) was added along with dNTPs (25 mM each) and 4 u of Rnasin (Promega) and the samples incubated at 37 °C for 5 mins. 200 u of Fermentas RevertAid H Minus M-MuLV Reverse Trancriptase was added and the samples incubated at 42 °C for an hour. Loading dye was added and the samples were run on a 12 % Acrylamide, 7 M Urea gel, at 95 watts for 1 hour 45 min. Once run, the gel it was placed in a cassette with a photoscreen and left overnight. Photoscreen was imaged using Fuji Image Gauge software.

2.14 *nramp* PCR with *S. solfataricus* DNA

According to the *Sulfolobus* Genome website the *nramp* gene is disrupted by a 1065 base transposon which is inserted 360 bases into the gene. To determine if the lab's strain of *S. solfataricus* contained this transposon, primers were designed for the first half of the *nramp* gene (*sso2076*), and the second half (*sso2078*). Two primers were also designed for the internal transposase (*sso2077*) that interrupts the *nramp* gene.

The primers were used in PCR reactions with *S.solfaraticus* DNA, 1 x GoTaq Flexi Buffer, 0.2 μM dNTPs, 2 mM $MgCl_2$, 2 μM of each primer, 0.5 u Go Taq flexi (Qiagen). PCR program 95 °C for 30 sec 1 cycle; 95 °C 1 min, 55°C 1 min, 72 °C 1 min for 30 cycles; 72 °C for 10mins, 1 cycle; 4°C 10mins 1 cycle.

Amplification was also attempted with KOD Polymerase (Novagen) and Pfu Polymerase (Promega). Reaction mixtures were taken from the protocols supplied with the polymerases and the same PCR program was used.

2.15 Construction of knockout plasmids for *sso2273*

2.15.1 PCR with knockout primers

Promega GoTaq was used to amplify Sso DNA in a PCR reaction with the following knockout primers. *Kpn*ISso2273f and *Nco*ISso2273r, *Bam*HISso2273f and *Not*ISso2273r, *Kpn*ISso0669f and *Nco*ISso0669r, *Bam*HISso0669f and *Not*ISso0669r. Sequences for all primers can be found in Appendix A.

The recognition sequences for each of the restriction enzymes was obtained from the New England Biolabs website and were incorporated into the primers. The reaction conditions used were are follows; 1x Go Taq Buffer, 0.4 mM each dNTPs, 0.4 mM of forward and reverse primers, 1 u Go Taq in a 50 µl reaction. Amplification program 95 °C 1 min, 95 °C 30 secs, 47 °C 30 secs, 72 °C 1 min; 72 °C 5 min, 4 °C hold.

2.15.2 Cloning of pET 2268 and knockout PCR products

The pET 2268 plasmid (supplied by Dr Sonja Albers of the University of Groningen) was digested with *Kpn*I and *Nco*I FastDigest enzymes from Fermentas for an hour at 37 °C then purified from a gel using Qiagen Gel purification kit.

PCR products amplified with *Kpn*ISso2273f & *Nco*ISso2273r, *Kpn*ISso0669f & *Nco*ISso0669r were also digested with *Kpn*I and *Nco*I FastDigest enzymes from Fermentas after being cleaned up using Qiagen gel purification kit (reaction mixture clean up protocol). The PCR products were also cleaned after digestion.

The digested plasmid and PCR products were ligated using Fermentas T4 DNA Ligase, ligation reaction was left overnight at room temperature. Around half the ligation reaction was transformed into *E.coli* TOP 10 competent cells (Invitrogen) and cells plated onto agar plates containing 100 µg/ml ampicillin.

The colonies were grown up and a mini prep (Qiagen) performed. The plasmid DNA was digested with the *Kpn*I and *Nco*I, and the digested sample run on a gel to check for the insert. Plasmids containing the insert were sent to the Sequencing service in Dundee. Once the insert was confirmed to be correct, the plasmid was digested with *Bam*HI and *Not*I FastDigest enzymes, again for one hour at room temperature. Cleaned digested PCR products from primer set *Bam*HISso2273f and *Not*ISso2273r, *Bam*HISso0669f and *Not*ISso0669r were ligated into the plasmid and the colonies checked by sequencing as before.

The plasmids, one containing the upstream and downstream regions of gene *sso2273*, and one containing the upstream and down stream regions of gene *sso0669,* were sent to Dr Sonja Albers at the University of Groningen.

2.15.3 Electroporation of PBL2025 cultures

A 50 ml culture of *S. solfataricus* PBL2025 at an OD_{600} between 0.2 and 0.3 was centrifuged at 4 °C, 4000 x g for 20 min. The cell pellet was resuspended in 50 ml 20 mM cold Sucrose, and centrifugation repeated. The cell pellet was resuspended in 10 ml 20 mM cold Sucrose and centrifugation step repeated. Cells were resuspended in 20 mM cold Sucrose at approximately 10^{10} cells/ml. 50 µl of cells and 100-300 ng of DNA were used for each electroporation. Electroporation conditions were 1.5 kV, 25 µF, 400 W with 2-mm cuvettes (Thermotron) and a Genepulser II Electroporator (Bio-Rad). After electroporation 1 ml of demineralised water was added and the cells transferred to eppendorf tubes and left on ice for 1-3 min. Cells were placed at 75 °C for 10 min, then inoculated into pre-warmed (0.4%) lactose growth media.

2.15.4 Plating and selection of *sso2273* knockouts

Once cells reached an OD_{600} of 0.15 (8-14 days) 1 ml was transferred to 50 ml (0.4%) lactose growth media. Once this culture had reached an OD_{600} of 0.1 cells were plated on lactose minimal media plates. Once colonies were visible the plates were sprayed with X-gal. The X-gal was dissolved, 25 mg/ml, in dimethylformamide. Plates were sprayed with a solution containing 5 mg/ml X-gal diluted in demineralised water. Selected blue colonies were then picked for PCR analysis to check the knockout had been successful.

2.16 Chromatin immunoprecipiation assay

2.16.1 Formaldehyde crosslinking and sonication

100 mls of *Sulfolobus solfataricus* was grown to mid log phase before the addition of 1 % formaldehyde. The culture was then allowed to cool for 20 min. 125 mM Glycine was added to quench the reaction. Cells were harvested by centrifugation and washed with PBS containing 300 mM NaCl, then resuspended and lysed in 1 ml TBSTT (20 mM Tris.Cl [pH 7.5], 150 mM NaCl, 0.1 % Tween 20, 0.1% Trition X-100, protease

inhibitors). The resuspended cells were sonicated on ice 4 x 30 sec and then centrifuged to collect the fragmented DNA.

2.16.2 Pre-blocking and binding of antibodies to Dynabeads

30 µl of Dynabeads®Protein G (Invitrogen) were washed twice with TBSTT. 10 µl of serum (either Sso2273, Sso Alba or pre-immune antibodies) and 30 µl of TBSTT were added to the beads and the mix was placed in a bench top shaker at room temperature for an hour.

Beads were pelleted with a magnet and washed once with TBSTT/BSA (10 µg), once with TBSTT/Calf thymus DNA (10 µg) and once with TBSTT. Beads were finally pelleted with the magnet.

2.16.3 Immunoprecipitation

200 µl of extract was added to the beads, they were left on a tilt table overnight at 4 °C.

2.16.4 Washing and reversal of cross-links

Beads were pelleted with the magnet and washed with 200 µl TBSTT 6 x 5 min shaking. Beads were then washed once with TBSTT containing 300 mM NaCl and once with TBSTT containing 0.5 % Tween-20 and 0.5 % Triton X-100.

Immune complexes were disrupted by resuspending beads in 100 µl 20mM Tris [pH 7.5], 10mM EDTA, 0.5 % SDS and heated to 65 °C for 30 min.

The resulting supernatant was mixed with 100 µl of 20mM Tris [pH 7.5] and 10 µg/ml proteinase K and left at 65 °C for 6 hrs then 37 °C for 12 hrs. DNA was recovered by phenol:chloroform extraction and ethanol precipitation.

2.17 Microarray

RNA for microarray experiments was extracted from *S. solfataricus* culture at an OD_{600} of 0.1 using the phenol:chloroform extraction method (see 2.3.1)

2.17.1 cDNA synthesis and Cye-dye labelling

cDNA was prepared using random nonomer primers (Operon). Each cDNA reaction (20 µl) contained 10 µg RNA, 5 µg primer, 0.5mM dATP, 0.5 mM dGTP, 0.5 mM dCTP, 0.1 mM dTTP, 0.4 mM aminoallyl-dUTP (Sigma), 42 mM DTT and 400 u Superscript II reverse transcriptase (Invitrogen).

The RNA was incubated with the random nonomers at 70 °C for 15 min and cooled rapidly on ice. Nucleotides and reverse transcriptase were added to the primed RNA and cDNA synthesis was carried out at 42 °C for 2 hrs. RNA was degraded by the addition of 2 µl of 200 mM EDTA and 3 µl of 1 M NaOH, the reaction was incubated at 70 °C for 15 min, and the reaction neutralised by the addition of 3 µl 1 M HCL. The cDNA was purified using a MinElute PCR purification kit (Qiagen). The cDNA was eluted in 20 µl 0.1 M $NaHCO_3$ and labelled with either Cy3 or Cy5 monoreactive dye (GE Healthcare) in the dark for 90 min. Samples to be co-hybridised were mixed and purified using Qiagen MinElute columns, The labelled cDNA was eluted in 20 µl Elution Buffer (10 mM Tris pH 8.5).

2.17.2 Microarray design

DNA Microarray containing 2488 gene-specific tags (GSTs) for *Sulfolobus solfataricus* were produced as previously described (Andersson, Bernander et al. 2005)

2.17.3 Microarray hybridization

Microarray slides were pre-hybridized with 50 ml pre-hybridisation solution (5 x SSC, 0.1 % SDS) containing 10 mg/ml of BSA at 42 °C. After incubating for one hour, the slides were washed three times in distilled water. Slides were dipped in isopropanol and dried using a slide centrifuge.

The labelled cDNA was mixed with 57 µl Hybridization mix (6.5 x SSC, 0.16 % SDS, 66 % formamide) followed by the addition of 10 µg of tRNA and 10 µg of herring sperm DNA (Ambion). Samples were denatured at 95 °C for 2 min then placed on ice.

The array slide was placed in a hybridization chamber (TeleChem International Sunnydale, CA, USA) and a cover slip (LifterSlip, ERIE Scientific Company) placed over the printed area. The sample was applied to the slide and the hybridization chamber sealed with the screw on top. Slides were incubated at 42 °C for 16-20 hrs.

Coverslips were removed by dipping the slide in Wash I (2 x SSC, 0.1 % SDS, 2mM DTT). Slides were then washed in Wash I at 42 °C for 5 min. Slides were washed in Wash II (0.1 % SSC, 0.1 % SDS, 2mM DTT) in a Coplin jar on a shaking platform for 10 min at room temperature, then washed 5 x in Wash III (0.1 % SSC, 2 mM DTT) and slides dried by centrifugation. Slides were scanned using a GenePix Personal 4100A Microarray scanner (Molecular Devices Corp, Sunnydale, USA).

2.17.4 Microarray data analysis

Data analysis was carried out using GenePix Pro 5.1 (Molecular Devices) and the web-based BioArray Software Environment (BASE). Normalization within arrays was performed to remove dye specific effects and to compensate for differences in arrays. Data from duplicate spots were merged, as was data from dye swap and duplicate experiments. The top, up and down regulated, genes were grouped according to their ratio values.

Chapter 3

Real Time PCR Detection Of Changes In Gene Expression After Damage With Ultraviolet Radiation

3.1 Introduction

It is vital for the survival of an organism that it can protect its genome against damage, and repairs any damage it does experience. The fact that *Sulfolobus* species inhabit conditions of high temperature and low pH, both of which are damaging to DNA, would suggest that they are exposed to a higher degree of damage than that experienced by mesophilic organisms. However the rate of mutation in *S. acidocaldarius* is no higher than in mesophilic *Escherichia coli* (Grogan, Carver *et al*. 2001) suggesting the protection and repair mechanisms in *S. acidocaldarius* are highly effective. *Sulfolobus* species live in hot springs and so are exposed to ultraviolet (UV) radiation from the rays of the sun. Several types of DNA damage are produced after exposure to UV, the most common are cyclobutane pyrimidine dimers (CPDs) and 6-4 Photoproducts (6-4 PP) (Pfeifer, You *et al*. 2005). These types of damage are repaired in one of two ways. The most efficient is the light repair pathway, which uses a light activated enzyme, called Photolyase, to reverse the damage and return the bases to their original form. Photolyase is present in all three domains of life, but at some point in evolution the enzyme was lost from placental mammals and so is not present in mice or humans (Menck 2002). Repair of UV damage in the species missing photolyase, is performed by the second repair pathway, nucleotide excision repair (NER). This pathway is utilised by all organisms if repair must be performed in the dark. In bacteria and eukaryotes damage in transcribed strands is repaired faster than damage in non-transcribed strands, therefore transcription coupled nucleotide excision repair (TC-NER) is performed faster that global genome nucleotide excision repair (GG-NER).

In bacteria NER is performed by a multi enzyme complex called UvrABC endonuclease (Truglio, Croteau *et al*. 2006). The UvrABC complex is found in a number of mesophilic archaea, and is believed to be the result of lateral transfer events (White 2003). The majority of archaeal species contain homologues for the eukaryal NER proteins (Kelman and White 2005).

S. solfataricus contains homologues of the helicases, XPD and XPB, and the endonuclesases XPF and FEN-1 (a homologue of XPG), but is missing the

recognition proteins, XPA and XPC. Not all archaeal species possess all of these homologues, and the fact that all these proteins have other functions in the cell could mean that they are not in fact used in an NER pathway, but only perform their alternate functions.

The repair pathways in archaea are still poorly understood, in particular how damage is detected, and repair initiated. A homologue of the LexA protein involved in the bacterial SOS response is missing from thermophilic archaea (Robb, Antranikian *et al*. 2007). To help elucidate the *S. solfataricus* response to UV irradiation microarray studies were carried out.

Microarray technology provides a means of investigating the transcriptional response to UV damage across the whole genome. They provide a massive amount of information relatively inexpensively and quickly. With the use of extremely accurate robotic machinery, PCR products that represent a small segment of every gene in the genome are spotted onto a glass slide (Andersson, Bernander *et al*. 2005). These slides are then 'washed' with cDNA that has been produced from the RNA extracted from untreated control cells and irradiated cells. cDNA from the control and irradiated cells is labelled with two different Cy dyes. Where there has been up regulation of a gene in response to DNA damage more cDNA molecules of that gene are present in the treated sample compared to the control, and by comparing the fluorescence in the two different conditions any changes in expression can be determined and genes that are up or down regulated identified. This technology was used to investigate the effect of UV radiation on the regulation of genes in *Sulfolobus solfataricus* at four time points (30, 60, 90 and 120 minutes). Dr Dorothee Götz performed the microrarray experiments; treatment of the cells, RNA extraction, cDNA synthesis and Cy dye labelling was performed at the University of St Andrews. The labelled samples were then sent to the University of Uppsala, to Professor Rolf Bernander's lab where the arrays were hybridised, washed and scanned. The microarrays produced a huge amount of data and although experimental and biological replicates were performed the changes in expression of a number of the most interesting genes was checked by reverse transcription (RT) real time PCR.

Real time PCR allows the user to follow the production of DNA throughout the course of the reaction. In traditional PCR the end products of the PCR are quantified by running them on an agarose gel, this often does not show the differences in the levels of starting material. The real time PCR technique is sensitive, relatively simple and easily reproducible (Bustin 2000; Dussault and Pouliot 2006).

SYBR green is the simplest detection method for real time PCR. The fact that SYBR green fluoresces 13 times more strongly when bound to dsDNA than when bound to ssDNA (Vitzthum, Geiger *et al*. 1999), allows the levels of dsDNA to be measured throughout the PCR by recording the emission at 521nm (the wavelength at which SYBR green fluoresces). The sensitivity of real time PCR means that care must be taken when designing primers and setting up the experiment. SYBR green will readily bind to any double stranded DNA and so primer dimers will give a false positive signal, and even slight contamination will be detected due to the sensitivity of the technique. Melt curves for each RT PCR can be checked to ensure no contamination is present. Reverse transcriptase real time PCR (RT-PCR), where the initial step is the production of cDNA from mRNA, provides a simple way to look at the relative levels of gene transcript, and by comparing control and treated samples the effects of a damaging agent on the expression levels of specific genes can be investigated. RT Real Time PCR was chosen as the technique to provide confirmation of the changes in gene expression that were seen by microarray analysis of *S. solfataricus* after UV damage.

3.2 Microarray analysis of *Sulfolobus solfataricus* response to UV damage

Microarray experiments had been performed by Dr Dorothee Götz (Gotz, Paytubi *et al*. 2007). Figure 3.1 shows the experimental set up used for the treatment of the cells. An *S. solfataricus* culture was grown to early log phase; half the culture was damaged with 200 J/m^2 of UV radiation, and the other half left as an undamaged control. Samples were taken at four time points (30, 60, 90, 120 minutes) after UV treatment and RNA extracted.

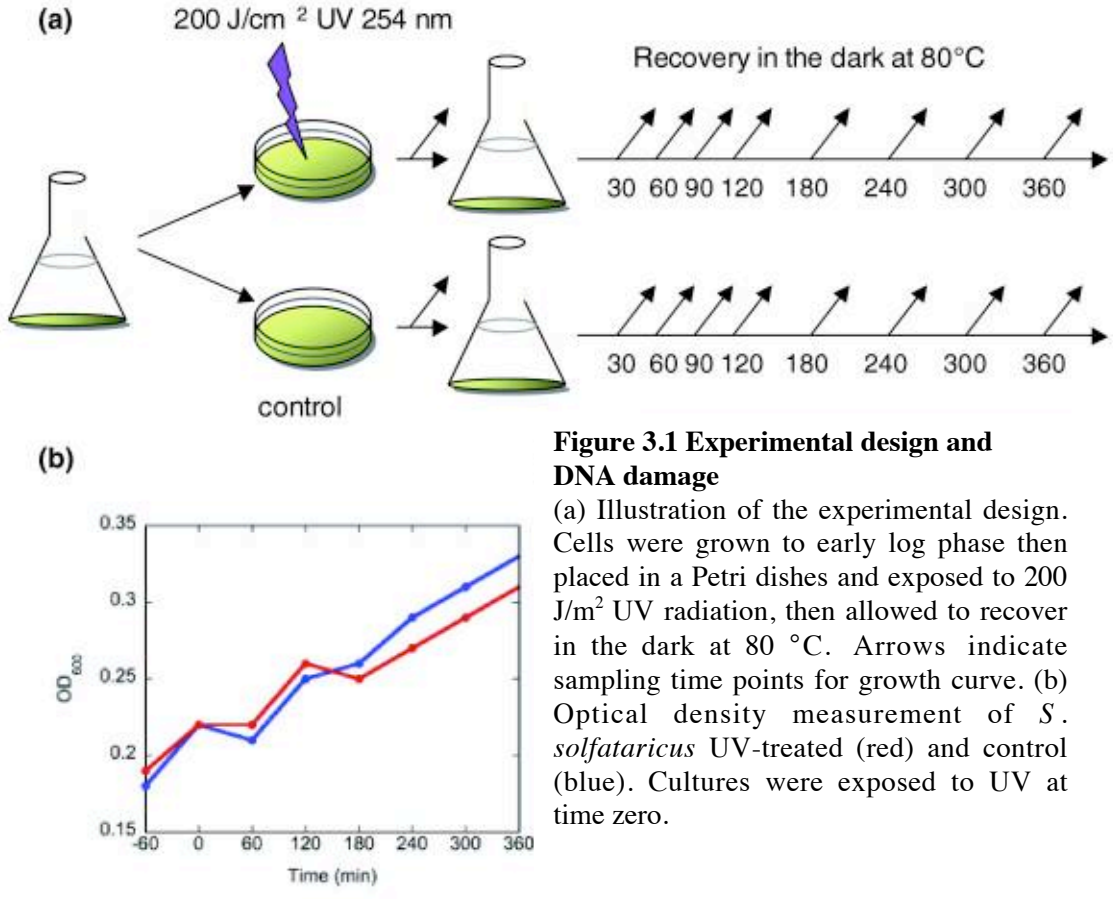


Figure 3.1 Experimental design and DNA damage
(a) Illustration of the experimental design. Cells were grown to early log phase then placed in a Petri dishes and exposed to 200 J/m^2 UV radiation, then allowed to recover in the dark at 80 °C. Arrows indicate sampling time points for growth curve. (b) Optical density measurement of *S. solfataricus* UV-treated (red) and control (blue). Cultures were exposed to UV at time zero.

RNA was then converted into cDNA and labelled with Cy dye (either Cy3 or Cy5). The labelled cDNA was hybridized to the arrays, and after numerous washes the arrays were scanned and analysed. A huge amount of data was produced from the microarray experiments but genes of particular interest are shown in Table 3.1. It was decided that RT real time PCR would be used to confirm the changes in expression seen in these genes.

			Time After Damage (min)			
Sso	Description	Function	30	60	90	120
0771	Cdc6-2	Cell division control	1.99	7.66	10.16	6.9
0446	TFB1	Transcription factor	0.97	1.29	1.75	1.70
0946	TFB2	Transcription factor	0.61	0.33	0.42	0.46
0280	TFB3	Transcriptional factor	1.11	2.82	8.26	8.03
0959	XPB1	Helicase	1.00	0.99	0.98	0.95
1459	DNA polymerase II	DNA synthesis	1.28	1.76	4.63	2.76
2079	DNA protection protein from starved cells	DNA Protection	0.95	1.69	4.67	5.60
2364	Single Stranded Binding protein	Binds ssDNA	2.28	1.24	1.07	0.95
0121	Hypothetical protein	Secretion or pili formation	1.07	2.00	7.65	9.96
0961 (Ref)	Hypothetical protein	?	1.02	1.08	1.18	1.25
2506 (Ref)	Tetracyclin Repressor (tetR)	Putative transcriptional repressor	0.98	1.03	1.17	1.02

Table 3.1 Genes of interest from microarray experiments
The ratios of gene expression were obtained by comparing the intensity of Cy dye labelled cDNA hybridised to the spot in control and irradiated conditions. If the fluorescence intensity is higher from the Cy dye used for the control culture gene transcript is more abundant in control conditions. However if the fluorescence intensity is higher from the Cy dye used for the irradiated cultures the gene transcript is more abundant in damage conditions. Two reference genes (ref), whose expression did not change after UV damage, were chosen to control for any loading differences, or differences in the quality of the RNA.
The ratios are colour coded to indicate the degree of up or down regulation. Red - >3, Orange-2-3, Light Orange – 1.5 – 2, Yellow – 0.5 – 1.5, Green - <0.5.

3.3 Determining gene expression by RT real time PCR using the Pfaffl method

The experimental design shown in Figure 3.1 was used to obtain RNA for use in RT real time PCR experiments. Primers were designed for each of the genes of interest and a PCR with a serial dilution of DNA was performed to produce a standard curve (see Figure 3.2), which allowed the PCR efficiency (E) with each primer set to be determined using the slope of the standard curve, see (a) in Figure 3.3.

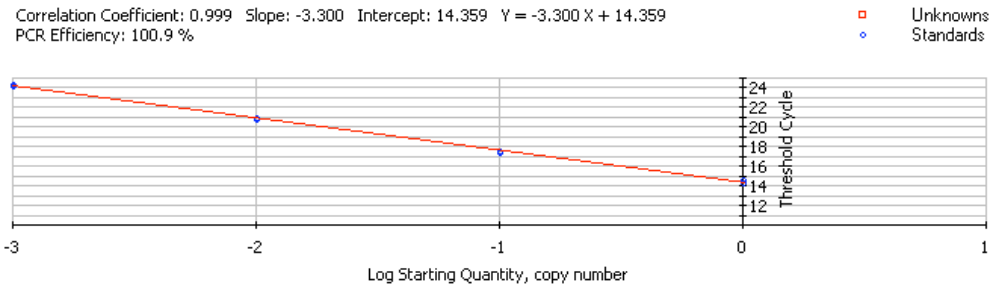


Figure 3.2 Standard curve used to determine PCR efficiency of real time reaction
Serial dilutions of DNA were PCR amplified to obtain a standard curve, which could be used to determine the efficiency of the PCR reaction with different primer sets. The threshold cycle on the y-axis indicates the cycle at which the intensity of SYBR green fluorescence crosses a pre set threshold. The threshold is defined by the real time PCR software using the background intensity of the first 10 cycles. 100% PCR efficiency indicates a doubling of DNA every cycle.

Primers were also designed for two reference genes, these genes (*sso0961* and *sso2506*) were chosen because their expression levels in the microarray had remained constant. The reference genes allow normalisation of any differences in the amounts or quality of RNA in the sample (Radonic, Thulke *et al*. 2004).

RNA from the control and treated cultures was used in RT real time PCR reactions with each primer set, samples were run in triplicate. The average crossing point (CP) value for the control and treated samples was used in equation (b) in Figure 3.3 to determine the ratio of expression. The CP value is the point at which the intensity of the SYBR green passes a predetermined threshold. The software of the real time machine sets this threshold; it does this by reading the levels of fluorescein in the samples. Fluorescein is added to the reagents and used as an internal control to account for any differences in loading, the background intensity of SYBR green in the first 10 cycles is also taken into account. The equation used to determine the gene expression ratio was taken from (Pfaffl 2001), it uses the PCR efficiency, and the delta CP values for the gene of interest (target gene) and the reference gene to work out the expression ratio.

(a)

$$E = 10^{[-1/slope]}$$

(b)

$$\text{Expression Ratio} = \frac{(E_{target})^{\Delta CP_{target}(control - sample)}}{(E_{ref})^{\Delta CP_{ref}(control - sample)}}$$

Figure 3.3 Equation for determining gene expression ratios
Equation (a) is used to determine PCR efficiency (E) using the value for the slope produced from a real time PCR amplification of a serial dilution of *Sulfolobus* DNA. Equation (b) is used to determine the expression ratio using the PCR efficiency and the crossing point (CP) values of the treated and control samples of the gene of interest (target) and a reference genes (Pfaffl 2001).

The average crossing point values (shown in Table 3.2) were used in equation (b) from Figure 3.3 to obtain the ratio of gene expression for each gene after UV damage. The lower the CP value, the higher the level of cDNA in the sample.

	Average CP Values of UV Damaged and Control Samples			
Gene	30 min	60 min	90 min	120 min
0771 Control	16.1	16.2	15.8	15.8
0771 UV Treated	14.9	12.2	12.0	12.7
0446 Control	14.5	14.7	15.3	14.6
0446 UV Treated	15.0	14.6	14.7	14.7
0946 Control	20.0	20.0	21.3	20.2
0946 UV Treated	20.6	21.1	21.7	21.4
0280 Control	16.7	17.3	17.2	16.3
0280 UV Treated	15.5	13.3	12.5	12.5
0959 Control	13.4	13.6	13.9	13.6
0959 UV Treated	13.9	13.5	13.8	13.8
1459 Control	13.6	13.6	13.8	13.7
1459 UV Treated	13.6	11.3	10.8	11.2
2079 Control	16.4	17.3	17.3	17.6
2079 UV Treated	16.8	16.6	16.5	16.6
2364 Control	9.5	9.2	9.4	9.8
2364 UV Treated	9.5	9.4	9.6	9.9
0121 Control	20.8	20.4	20.2	20.4
0121 UV Treated	20.7	20.3	19.3	18.1
0961 Control	11.5	12.4	12.0	11.9
0961 UV Treated	11.5	11.6	11.6	11.5
2506 Control	15.5	16.2	16.0	15.0
2506 UV Treated	14.6	15.1	15.0	15.0

Table 3.2 Crossing point values of genes in control and UV treated cultures.

Two ratios were obtained for each gene, one when the *sso0961* gene was used as a reference and one when the *sso2506* gene was used as a reference; the average of these two ratios was plotted on the graph shown in Figure 3.4.

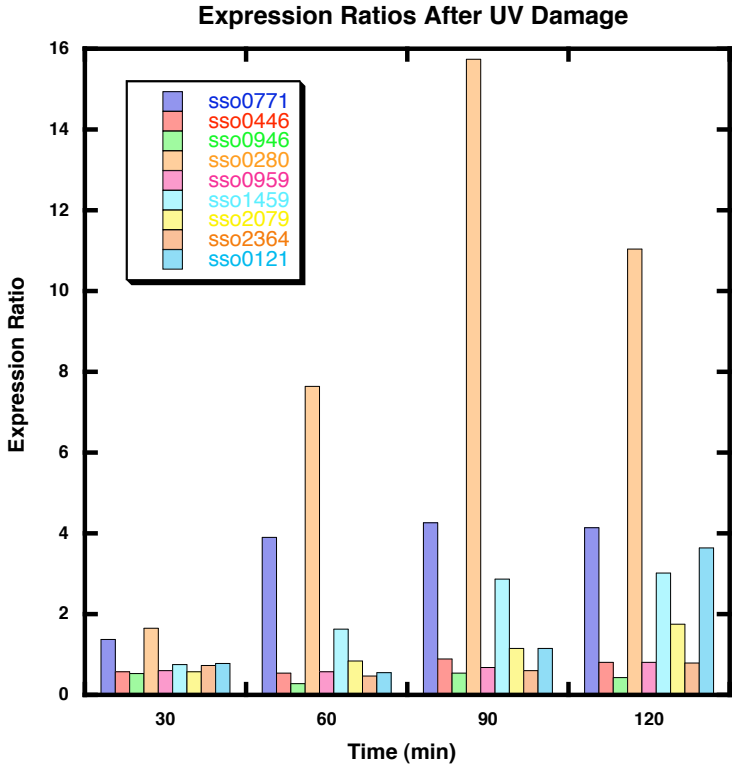


Figure 3.4 Graph of expression ratios after UV Damage
Ratios were obtained using the Pfaffl equation (Pfaffl 2001). Two ratios were obtained, one when *sso0961* was the reference gene and one when *sso2506* was the reference gene, the average of these two ratios was plotted.

3.4 Comparison of microarray and RT real time ratios

Once RT real time PCR ratios were obtained they were compared to the values from the microarray studies. The comparison and some discussion of the genes of interest follow.

3.4.1 Microarray and RT real time PCR ratios of *sso0771 cdc6-2*

Time After Damage (min)	30	60	90	120
sso0771 Microarray Ratio	1.99	7.66	10.1	6.90
sso0771 RT Real Time Ratio	1.37	3.90	4.26	4.14

Table 3.3 Ratios of expression of *sso0771* from microarray and RT real time PCR experiments.

Table 3.3 shows the ratios from the microarray and RT real time experiments for the *cdc6-2 gene, sso0771*. Generally replication and cell division proteins were down regulated in response to UV damage (Gotz, Paytubi *et al*. 2007), however gene *sso0771,* which encodes the cell division control protein (Cdc6-2) was up regulated. The Cdc6 proteins (Cdc6-1, Cdc6-2, Cdc6-3) are responsible for binding to replication origins and loading the replicative helicase MCM (Shin, Grabowski *et al*. 2003) and in normal conditions levels of all three proteins remains constant (Duggin, McCallum et al. 2008). However after UV damage down regulation of *cdc6-1* and *cdc6-3* was observed, while *cdc6-2* was up regulated. A role in control of replication provides one explanation for the differences in expression levels of the Cdc6 proteins. Cdc6-2 can bind at sites overlapping the binding sites for Cdc6-1 and Cdc6-3 at the origins of replication in *S. solfataricus,* suggesting a role as a negative regulator. While Cdc6-1 and Cdc6–3 act as promoters of replication, Cdc6-2 may act as a repressor by preventing binding of the other two Cdc6 proteins in conditions where replication needs to be inhibited (Robinson, Dionne *et al*. 2004). The up regulation of the *cdc6-2* gene was confirmed by real time PCR. The ratios from the microarray and RT real time data followed the same pattern however the ratios from the microarrays were higher, possible reasons for these differences are discussed in section 3.4.

3.4.2 Microarray and RT real time PCR ratios of *sso0446, sso0946, sso0280 tfb-1, 2* and *3*

Time After Damage (min)	30	60	90	120
***sso0446* Microarray Ratio**	0.97	1.29	1.75	1.70
***sso0446* RT Real Time Ratio**	0.57	0.54	0.89	0.81
***sso0946* Microarray Ratio**	0.61	0.33	0.42	0.46
***sso0946* RT Real Time Ratio**	0.53	0.28	0.54	0.43
***sso0280* Microarray Ratio**	1.11	2.82	8.26	8.03
***sso0280* RT Real Time Ratio**	1.65	7.64	15.7	11.0

Table 3.4 Ratios of expression of *sso0446, sso0946 and sso0280* from microarray and RT real time PCR experiments.

Table 3.4 shows the ratios from the microarray and RT real time experiments for the *tfb genes, sso0446, sso0946 and sso0280*. A transcriptional response after damage was expected as genes required to alleviate and repair damage are produced, and the transcription of other genes is decreased until the damage is repaired. Basal

transcription in archaea requires the TATA box binding protein (TBP), transcription factor II B homologue (TFB) and RNAP. Levels of TBP and RNAP transcript remained constant after UV damage, while there was a significant change in the expression of the TFB proteins. *S. solfataricus* possesses three TFB proteins named TFB-1,-2 and -3. TFB consists of an N-terminal domain containing a zinc ribbon and a B-finger, which interacts with the RNA polymerase; a flexible linker and a C-terminal domain containing two 90 amino acid repeats involved in binding to TBP and a helix-turn-helix motif used to bind the BRE (TFB Recognition Element) which orientates the transcription initiation complex (Bell, Kosa *et al*. 1999). Only TFB-1 and TFB-2 are full length, TFB-3 is missing the B-finger, the BRE binding helix-turn-helix motif and part of the core domain, see Figure 3.5.

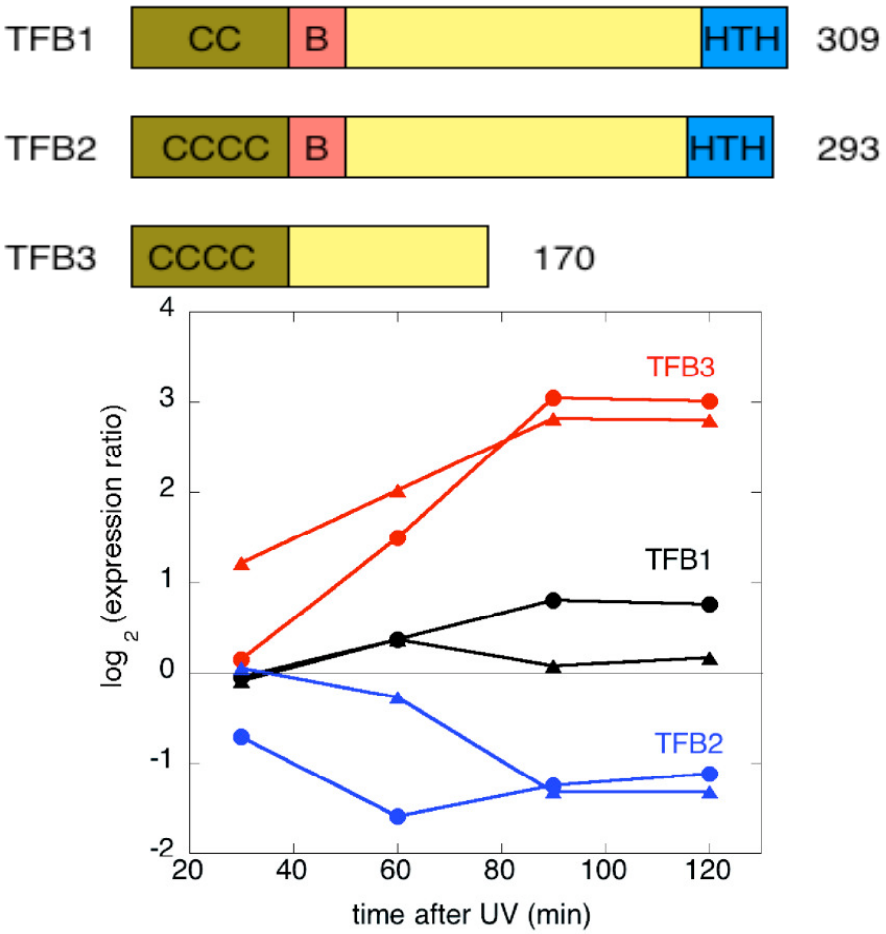


Figure 3.5 Three TFB proteins in *Sulfolobus solfataricus*
S. solfataricus encodes three TFB proteins. TFB-1 and TFB-2 are full length, but TFB-3 is truncated, lacking the B-finger and the helix-turn-helix domain. C's represent the cysteines implicated in zinc binding. The graph shows the log2 expression ratios of the three *tfb* genes after UV damage. The circles represent *S. solfataricus*, and triangles represent *S. acidocaldarius*. After UV damage the expression of TFB-1 remains relatively level, expression of TFB-2 decreases and expression of TFB-3 increases. Figure is taken from Götz, Paytubi *et al*. 2007.

Immunoprecipitation experiments using purified native RNA polymerase and recombinant TFB-1 and TFB-3, showed that TFB-3 competes with TFB-1 for binding of RNA polymerase. Addition of increasing concentrations of TFB-3 abolished the interaction of TFB-1 and RNA polymerase (Gotz, Paytubi *et al*. 2007) and based on these observations it was predicted that TFB-3 may act as a competitive inhibitor of transcription initiation. However, subsequent data (Paytubi unpublished) showed that TFB-3 had a stimulatory effect on transcription when added to transcription assay already containing the basal transcription initiation proteins (TBP, TFB-1 and RNAP).

The role of TFB-3 in this situation is not fully understood and further investigation is needed.

3.4.3 Microarray and RT real time PCR ratios of *sso0959 xpb-1*

Time After Damage (min)	30	60	90	120
***sso0959* Microarray Ratio**	1.00	0.99	0.98	0.95
***sso0959* RT Real Time Ratio**	0.60	0.57	0.68	0.81

Table 3.5 Ratios of expression of *sso0959* from microarray and RT real time PCR experiments.

Table 3.5 shows the ratios from the microarray and RT real time experiments for the *xpb-1 gene, sso0959*. The eukaryotic *xpb* gene codes for a helicase which, in complex with other proteins, forms the TFIIH complex that is essential for transcription and repair (Wouter L. de Laat 1999). *S. solfataricus* has two homologues of this protein, *xpb-1* and *xpb-2*. The expression of *xpb-1* has been shown to increase 10-fold two hours after UV damage (Salerno, Napoli *et al*. 2003). However expression of *xbp-1* observed in our microarray studies showed no increase in expression and this was mirrored in the RT real time PCR data. The reason for these differences is unknown, but the UV doses used were different which could have affected the results.

3.4.4 Microarray and RT real time PCR ratios of *sso1459 DNA polymerase II*

Time After Damage (min)	30	60	90	120
***sso1459* Microarray Ratio**	1.28	1.76	4.63	2.76
***sso1459* RT Real Time Ratio**	0.75	1.63	2.87	3.02

Table 3.6 Ratios of expression of *sso1459* from microarray and RT real time PCR experiments.

Table 3.6 shows the ratios from the microarray and RT real time experiments for the *DNA polymerase II gene, sso1459*. *S. solfataricus* contains 4 DNA polymerases (DpoI, II, III, and IV). DpoIV is a member of the Y-family *dinB* polymerases and is involved in lesion by-pass after DNA damage (Boudsocq, Iwai et al. 2001; Kokoska, McCulloch et al. 2003). Unlike its homologue in bacteria this gene was not up regulated after UV damage. The expression of DpoI, the main replicative DNA polymerase in *S. solfataricus,* was unaffected by UV damage. However Dpo II was up regulated, the function of this polymerase is not fully understood but the fact it is up

regulated after UV damage while other replication proteins are down regulated may indicate a role in repair of UV damage.

3.4.5 Microarray and RT real time PCR ratios of *sso2079 dps-like protection protein*

Time After Damage (min)	30	60	90	120
***sso2079* Microarray Ratio**	0.95	1.69	4.67	5.60
***sso2079* RT Real Time Ratio**	0.57	0.84	1.15	1.75

Table 3.7 Ratios of expression of *sso2079* from microarray and RT real time PCR experiments.

Table 3.7 shows the ratios from the microarray and RT real time experiments for the *dps-like gene, sso2079*. The Dps-like protein (DNA protection protein from starved cells) is a protection protein involved in guarding the cells against oxidative stress. The protein sequesters the components of the Fenton Reaction; hydrogen peroxide and ferrous iron, to prevent the production of extremely damaging hydroxyl radicals (Wiedenheft, Mosolf *et al*. 2005). The observed up regulation of this gene can be explained by the fact that although CPDs and 6-4 photoproducts are the most common forms of damage produced by UV irradiation, oxidative stress is also induced (Moriwaki and Takahashi 2008) leading to an up regulation of genes involved in protection against oxidative stress.

3.4.6 Microarray and RT real time PCR ratios of *sso2364 single stranded DNA binding protein*

Time After Damage (min)	30	60	90	120
***sso2364* Microarray Ratio**	2.28	1.24	1.07	0.95
***sso2364* RT Real Time Ratio**	0.73	0.47	0.60	0.79

Table 3.8 Ratios of expression of *sso2364* from microarray and RT real time PCR experiments.

Table 3.8 shows the ratios from the microarray and RT real time experiments for the *ssb gene, sso2364*. Repair mechanisms in archaea are not fully understood, in particular the way in which damage is detected and repair processes initiated, remains unknown. One candidate for this role is the single-stranded DNA-binding (SSB) protein, and there is early induction of this gene after UV damage. SSBs are present in all domains of life suggesting an important function, and the bacterial SSBs have been

shown to be involved in recognition of DNA damage and the SOS response (Meyer and Laine 1990). In *in vitro* experiments *S. solfataricus* SSB has been shown to detect a variety of DNA lesions, including cyclobutane pyrimidine dimers, common after UV irradiation (Cubeddu and White 2005).

The RT real time data for the SSB transcript showed no induction however the transcript levels were constitutively high with CP values of 9, the lowest values we witnessed for any of the genes, suggesting SSB is readily available and up regulation may not be needed.

3.4.7 Microarray and RT real time PCR ratios of *sso0121 pili system*

Time After Damage (min)	30	60	90	120
***sso0121* Microarray Ratio**	1.07	2.00	7.65	9.96
***sso0121* RT Real Time Ratio**	0.78	0.55	1.15	3.64

Table 3.9 Ratios of expression of *sso0121* from microarray and RT real time PCR experiments.

Table 3.9 shows the ratios from the microarray and RT real time experiments for the sso*0121 gene*. The up regulation of the operon containing *sso0121* was also observed by microarray experiments performed by Dr Schlepers lab (Frols, Gordon *et al.* 2007). *sso0121* is annotated as a hypothetical protein, however the other genes in its operon encode a pili assembly system. Schleper's lab showed that *S. solfataricus* aggregates after UV damage (Frols, Gordon *et al.* 2007), a phenomenon also observed in *S. acidocaldarius* (Schmidt, Beck *et al.* 1999). This aggregation and pili formation is predicted to allow DNA transfer by conjugation, facilitating repair of damage by homologous recombination (Ng, Zolghadr et al. 2008).

3.5 Transcription coupled repair in archaea

The fact that the *sso0121* operon was inducible by UV irradiation made it a useful tool to investigate the repair of transcribed and non-transcribed strands in *S. solfataricus* (Dorazi, Gotz *et al.* 2007). Archaea do not possess homologues of either the bacterial Mfd or human CSB proteins, responsible for detection of damage of transcribed strands (Li and Bockrath 1995), suggesting that either this organism does not have a transcription coupled repair system, or its Mfd/CSB equivalent has not yet

been discovered. To investigate whether transcribed and non-transcribed strands are repaired with the same efficiency in *S. solfataricus*, CPDs were induced in three operons by UV irradiation. Operons used for the experiment had to fit certain criteria; they had to be actively transcribed under the conditions of the experiment (in this case after UV irradiation), and the transcript produced had to be long enough to ensure at least one CPD would be present. The microarray data was used to select genes actively transcribed after UV irradiation. The three operons chosen were the pyr operon (*sso0610-0616*), *sso1397-1401* and the pili assembly system operon (*sso0117-0121*). RT real time PCR of the *sso0121* gene was used to confirm that the *sso0121* operon was being transcribed.

	Time After UV Irradiation (Hrs)			
Gene and Culture	**1**	**2**	**3**	**5**
Average CP Value *sso0121* control	18.1	17.7	18.6	17.2
Average CP Value *sso0121* UV treated	18.3	17.3	16.0	15.3
Ratio of Expression	0.90	1.26	5.35	3.42

Table 3.10 Crossing point values and ratio of expression of *sso0121* from RT real time PCR experiments.

Southern blot analysis was used to monitor the repair of CPDs from the transcribed and non-transcribed strands of the three operons. Any unrepaired CPDs were cleaved by T4 Endonuclease V (TEV), so only transcripts not containing CPDs were detected by the southern blot, this assay provided a way of tracking the repair of CPDs.

Results showed no preference for repair of the transcribed strand in any of the operons (Dorazi, Gotz *et al*. 2007). This data was backed up by work from the Ciaramella group who used three different genes in *S. solfataricus* to investigate TC-NER (Romano, Napoli *et al*. 2007). It is possible that *S. solfataricus* has no need for TC-NER. Observations of the rate of TC-NER in three other organisms (*E. coli, S. cerevisiae*, and human fibroblasts) showed that the repair of damage in non-transcribed strands in *S. solfataricus* is already as fast as the rate of TC-NER observed in these three organisms. The rate of global genome repair in *S. solfataricus* may be fast enough that there is no need for a faster transcription coupled repair pathway (Dorazi, Gotz *et al*. 2007).

3.6 Differences in microarray and RT real time PCR ratios

Although the RT real time and microarray data matched well (see Figure 3.6) there were instances where ratios produced by the two methods differed. There are a number of factors that could account for differences in the ratios produced by the two techniques. Due to the microarray experiments and RT real time experiments being performed months apart there was no possibility of using RNA from the same stocks for both experiments, as would have been the ideal situation. Different cells, differences in the extraction technique and quality of RNA extracted could all affect how well the data correlated.

Different reverse transcriptase enzymes were used for the two methods and differences in the efficiency of these enzymes could also be a factor. The way the data is normalized is also very different for the two techniques, in the microarrays normalization against the whole plate is performed while in the RT real time PCR the data was only normalized against two reference genes (Rajeevan, Vernon *et al.* 2001). The microarrays used contained *S. acidocaldarius* genes as well as *S. solfataricus* genes so there is a risk of cross hybridisation of very closely related genes. Previous studies looking at the validation of microarray data with RT real time PCR have shown that there is often a low correlation between the two techniques when the gene showed less than a 2-fold change, and also when the spot intensity was low. When the change being verified was a down regulation rather than an up regulation the correlation was generally lower, this was predicted to be due to the increased variability when data is taken from the later cycles of the RT real time PCR. Correlation is highest when the fold change in expression is above 2 and the spot intensity is good *ie* good signal strength, low background level, good spot morphology (Morey, Ryan *et al.* 2006).

When the log2 ratios for the microarray and RT real time data were compared for six of the genes (*cdc6-2, tfb-1, tfb-2, tfb-3, dpoII and ssb*) the data correlated well with an R-value of 0.91, see Figure 3.6.

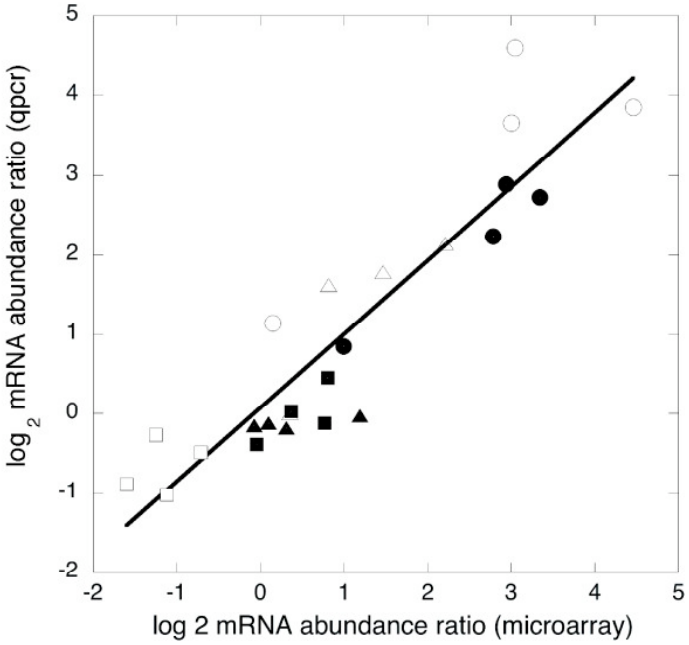


Figure 3.6 Comparison of expression ratios from microarray analysis and RT real time PCR

The changes in mRNA levels of six genes following UV irradiation were analysed by RT-PCR to provide confirmation of microarray data. The log2 expression values for *sso2364 - ssb* (closed triangles), *sso0771 - cdc6-2* (closed circles), *sso0446 - tfb-1* (closed squares), *sso0946 - tfb-2,* (open squares), *sso0280 - tfb-3* (open circles) and *sso1459 - dpoII* (open triangles) obtained from microarray and RT-PCR analysis were plotted on the x and y axis respectively. The data obtained from the two methods yields a linear fit with a slope of 0.93 and R-value of 0.91 (Gotz, Paytubi *et al*. 2007).

3.7 Discussion

Microarray analysis provides a vast amount of data and, in this case, allowed the response of *S. solfataricus* after UV irradiation to be investigated on a genome wide scale. Confirmation by RT real time PCR provided greater confidence in the data, and the correlations between the two techniques, for the genes we looked at, was good. An independent study by Schleper and colleagues (Frols, Gordon *et al*. 2007) produced similar results lending even more support to our data. Despite different UV doses being used in the two studies many of the same genes were identified as being up regulated and the differences in the expression profiles of the three *cdc6* genes was also observed by the Schleper group.

From the changes in gene expression observed we can conclude that *S. solfataricus* does have a transcriptional response to DNA damage. This transcriptional response

may be partly controlled by the TFB proteins; differences in the expression of these three proteins was observed, and further investigation into the role of TFB-3, possibly as a transcriptional repressor, is underway. It was interesting to note that the expression of repair proteins is not increased. This may be because, to deal with the high levels of damage the organism experiences, the repair proteins are constitutively expressed.

There also appears to be no transcription coupled repair in *S. solfataricus* (Dorazi, Gotz *et al*. 2007; Romano, Napoli *et al*. 2007). The rapid rate of global genome repair in this organism, which is equal to the rate of transcription coupled repair in *E. coli* and *S. cerevisiae,* may negate the need for a transcription coupled repair pathway.

The method of damage detection in *S. solfataricus* still remains unclear, it has been proposed that the single stranded DNA binding protein may play a role in damage detection and it has been shown to bind double strand breaks and CPD lesions (Cubeddu and White 2005). An early increase in the expression of this gene was observed after UV damage with the microarray, but this result was not mirrored in the RT real time data. However the extremely low CP values for the gene indicate that it is expressed at a constitutively high level in the cell, possibly negating the need to induce it after damage, as it is already readily available.

Chapter Four

Real Time PCR Detection of Changes In Gene Expression After Damage with Different Damaging Agents

4.1 Introduction

The process of detection and repair of DNA damage in archaea is still poorly understood. Homologues of the LexA protein required for the bacterial SOS response are absent from the majority of archaeal genomes (Robb, Antranikian *et al*. 2007). There is some evidence that the single-stranded DNA-binding protein could be an indicator of damage. It has been shown to melt DNA containing a mismatch or lesion *in vitro,* and interacts with proteins with predicted repair functions, such as XPB and the archaeal homologue of Rad50 (Cubeddu and White 2005) and it has also been shown to interact with reverse gyrase after DNA damage (Napoli, Valenti et al. 2005). There is evidence of early induction of the *ssb* gene after UV damage from microarray studies (Gotz, Paytubi *et al*. 2007). The induction is only seen at the 30 minute time point with very little change in expression at the later time points, however the crossing point values for both the treated and control samples were very low, suggesting a high abundance of the *ssb* gene transcript, even in non damage conditions (Gotz, Paytubi *et al*. 2007). This may explain why no up regulation was observed after UV damage, as the protein was already highly abundant.

S. solfataricus also possesses homologues for some of the proteins involved in the nucleotide excision repair pathways, however homologues of XPA and XPC (the NER damage detection proteins) do not appear to be present in the *S. solfataricus* genome. None of the NER homologues were up regulated in *S. solfataricus* after UV damage (see Chapter 3) and so it has been postulated that the repair proteins in archaea may be constitutively expressed in order to deal with the increased potential for damage because of the harsh conditions *S. solfataricus* inhabits (Gotz, Paytubi *et al*. 2007).

Previous experiments using microarray studies and RT real time PCR determined a group of nine genes whose expression was affected by UV irradiation (see Chapter 3). To determine whether these genes were part of a standard DNA damage response, or were specific to UV detection and repair, *S. solfataricus* was challenged with other

damaging agent that produced different types of damage, and the expression of these nine genes was investigated with RT real time PCR.

4.2 Determining non-lethal damage conditions

The first step in monitoring the transcriptional effect of each damaging agent on *S. solfataricus* was to determine a level of damage that would stress the cells into a damage response, but not prove fatal, allowing them to repair themselves.

Increasing concentrations of the damaging agents were added to exponentially growing cells and their subsequent growth monitored. The damaging agents chosen were Mitomycin C, Methyl methane sulfonate, Phleomycin and Hydrogen peroxide.

4.2.1 Mitomycin C (MMC)

Mitomycin C is a chemotherapy drug used as an anti-cancer antibiotic. It is a naturally occurring compound produced by a number of *Streptomyces* species. To exert a cytotoxic effect MMC requires reductive activation. MMC generates covalent inter-strand cross-links, and in doing so inhibits DNA synthesis which leads to cell death (Iyer and Szybalski 1963). Although the type of damage produced by MMC is different from that produced by UV irradiation both types of damage cause a stalling of replication and so may elicit a similar damage response.

MMC was added to growing cultures of *S. solfataricus*, optical density readings were taken at 600 nm and the growth monitored for the next 16 hours, see Figure 4.1. It was assumed that an increase in OD_{600} readings indicated that the cells were growing, however an increase in OD_{600} can also be observed if the cells are increasing in size, rather than in number (Samson, Obita et al. 2008). It was assumed in the following experiments that the increasing OD_{600} indicated a growing culture, however cells were not checked by microscopy.

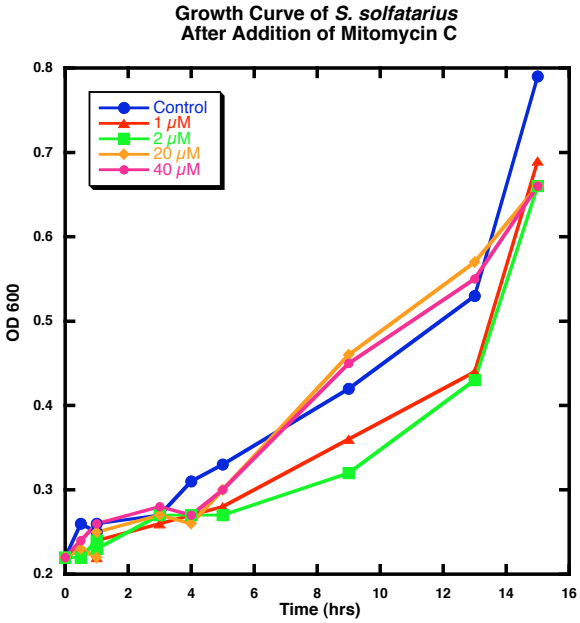


Figure 4.1 Growth curves of *S. solfataricus* after addition of increasing concentrations of Mitomycin C
MMC to final concentrations of 1, 2, 20 and 40 μM was added to growing cells. The two higher concentrations of MMC appeared to have less effect on the cells growth than the two lower concentrations.

Cells treated with higher concentrations of MMC grew more quickly than cells treated with lower concentrations. To determine if this trend continued, higher concentrations of MMC (100, 200 and 400 μM) were added to fresh cultures of cells and their growth measured by reading the optical density at 600 nm, see figure 4.2.

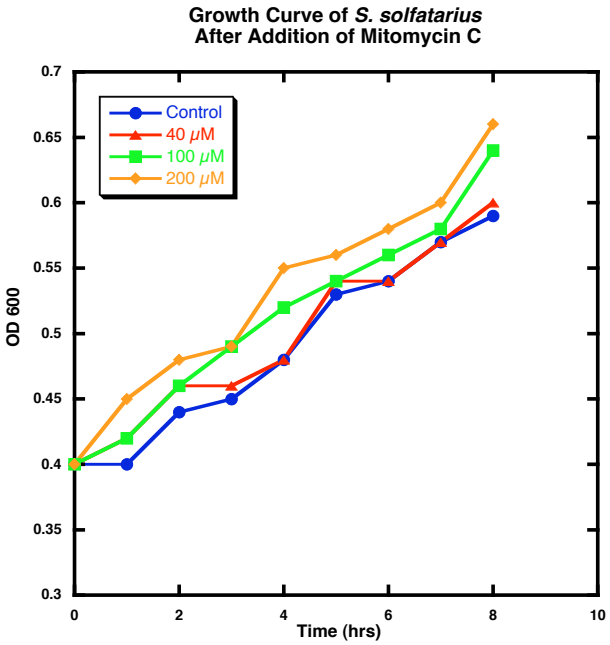


Figure 4.2 Growth curves of *S. solfataricus* after addition of increasing concentrations of Mitomycin C
Initial concentrations of MMC (1, 2, 20, 40 μM) had little effect on the cells; higher concentrations (100 and 200 μM) were added to determine if there was any effect.

The cells were treated with higher concentrations of MMC. The limited effect of MMC on the growth of *S. solfataricus* cells was surprising, as growth of *E. coli* is affected by as little as 2 μM MMC (Lee, Park *et al.* 2006). An experiment was set up

to determine whether the MMC was being destroyed by the low pH and/or high temperature conditions of the *Sulfolobus* growth media. Three *E. coli* cultures were grown to an OD_{600} of 0.4, one was kept as a control, one had 40 µM MMC added, the third culture also had 40 µM MMC added, however the MMC had been added to *Sulfolobus* media and incubated at 80 °C for 30 minutes prior to addition, see Figure 4.3.

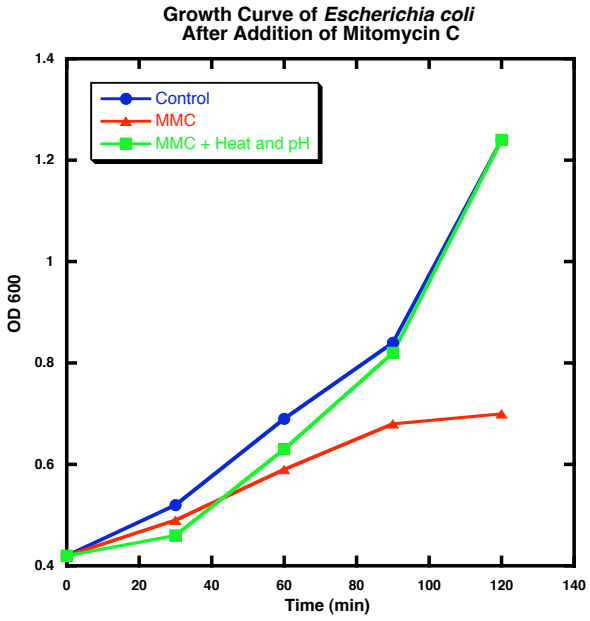


Figure 4.3 Graph showing growth of *E.coli* after addition of 40 µM Mitomycin C (normal and heat/pH treated) The high temperature, low pH conditions of the *Sulfolobus* media appear to be having an inhibitory or destructive effect on the MMC, as the growth of cells is not effect by the MMC that has been heat and pH treated before addition.

The growth of the control culture, and the culture containing the pH/heat treated MMC was very similar. The initial lag in growth of the culture containing the pH/heated treated MMC maybe due to lowering of the pH of the culture on addition of the MMC in *Sulfolobus* media, rather that any effect of the MMC. The growth of the culture containing the normal MMC was greatly reduced, showing that the MMC was effective. It appears that MMC is destroyed by either the low pH and/or the high temperature of the *Sulfolobus* media and so has little effect on the growth of the cells. It was decided that MMC to a final concentration of 2 µM would be used to damage the *S. solfataricus* cells and test the expression of two of the nine genes to determine if there was any effect on gene expression, this is discussed in section 4.3.1.

4.2.2 Methyl methane sulfonate (MMS)

Methyl methane sulfonate is an alkylating agent and a carcinogen. Alkylating agents attach small alkyl groups to DNA bases, which results in the DNA being targeted for degradation by repair enzymes, and fragmented as the repair enzymes attempt to replace the alkylated bases. Alkylated bases prevent DNA synthesis and transcription (Valenti, Napoli *et al*. 2006).

MMS was added to growing cultures of *S. solfataricus*, optical density readings were taken at 600 nm to monitor the growth for the next 10 hours, see Figure 4.4.

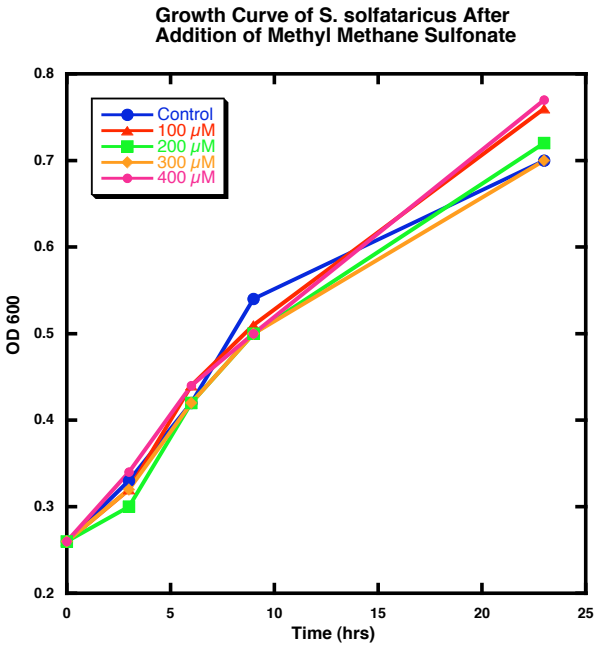


Figure 4.4 Growth curves of *S. solfataricus* after addition of Methyl methane sulfonate
MMS to final concentrations of 100, 200, 300, and 400 μM was added to growing cultures of *S. solfataricus*. The MMS appeared to have little effect on the growth of the cells.

MMS had little effect on the growth of the cells. It was decided that MMS to a final concentration of 300 μM would be used to damage the *S. solfataricus* cells in subsequent experiments, discussed in section 4.3.2.

4.2.3 Phleomycin

Phleomycin is a copper-containing antibiotic of the Bleomycin family. It is a naturally occurring compound and was isolated from a mutant strain of *Streptomyces verticillus*. It acts on membrane bound DNA affecting the integrity of the cell wall. It can also intercalate within the DNA strands causing first single, then double strand

breaks, leading to the arrest of DNA synthesis, and the fragmentation of the DNA (Reiter, Milewskiy *et al*. 1972).

Phleomycin was added to growing cultures of *S. solfataricus*, optical density readings were taken at 600 nm to monitor the growth for the next 10 hours, see Figure 4.5.

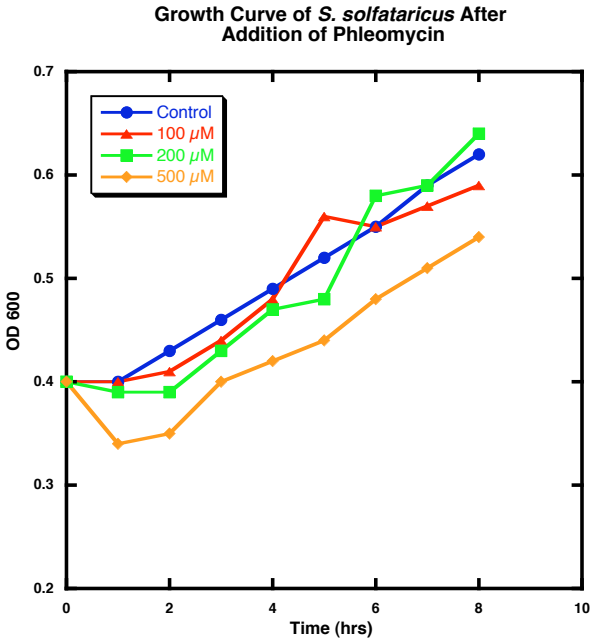


Figure 4.5 Growth curves of *S. solfataricus* after addition of phleomycin
Increasing concentrations of phleomycin (100, 200, 500 μM) caused increasing lag in the initial growth of the cultures.

Addition of phleomycin caused a lag in the initial cell growth that was more pronounced the higher the concentration. It was decided that phleomycin to a final concentration of 200 μM would be used in subsequent experiments to damage the *S. solfataricus* cells, discussed in section 4.3.3.

4.2.4 Hydrogen peroxide

Hydrogen peroxide is a natural by-product of oxygen metabolism. In the cell it is normally converted to water by catalase or peroxidases because it is a reactive oxygen species (ROS) and is toxic to the cell. It also has the potential, if it reacts with iron, to produce even more damaging hydroxyl radicals (Valko, Morris *et al*. 2005). Cells encounter hydrogen peroxide all the time, however its potentially damaging effects are kept in check by a plethora of enzymes (such as catalase and superoxide dismutase) and numerous antioxidants (such as vitamin C and glutathione) (Pedone, Bartolucci *et al*. 2004). When the balance of ROS and antioxidants is not maintained,

the cell must act to stop damage by ROS. Hydrogen peroxide was added to growing cultures of *S. solfataricus*, optical density readings were taken at 600 nm to monitor the growth for the next 10 hours, see Figure 4.6.

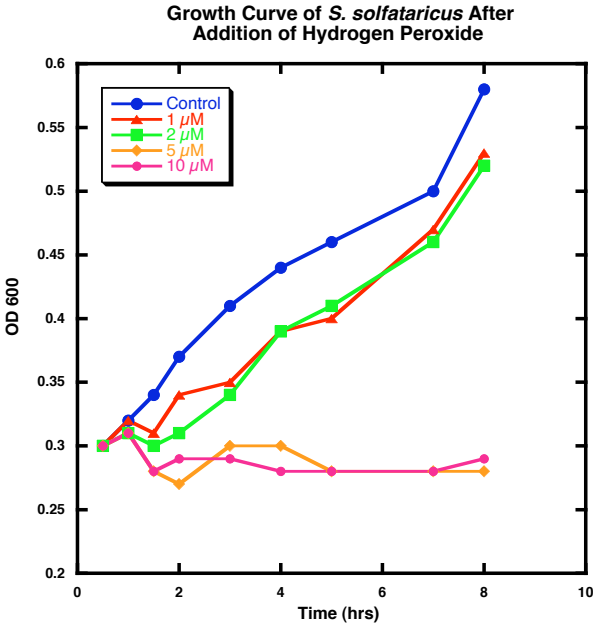


Figure 4.6 Growth curves of *S. solfataricus* after addition of hydrogen peroxide
Addition of hydrogen peroxide had a pronounced effect on the growth of the cells. Growth was retarded completely by addition of hydrogen peroxide over 5 μM.

Hydrogen peroxide elicited the strongest effect of any of the damaging agents, with as little as 5 μM having a dramatic effect on cell growth. It was decided that hydrogen peroxide to a final concentration of 1 μM would be used in subsequent experiments to damage the *S. solfataricus* cells, discussed in section 4.3.4.

4.3 The effects of different damaging agents on gene expression

The transcriptional response to the damaging agent was unknown so the use of reference genes was not possible. The calculation in Figure 4.7 was used to determine the expression ratio.

$$\text{Expression Ratio} = (E_{\text{target}})^{\Delta CP\text{target (control-sample)}}$$

Figure 4.7 Equation used to determine ratio of expression
Ratio of expression was determined using the efficiency of the PCR reaction (E), and the change in the crossing point (CP) values for the treated and control samples

Cultures were grown to early exponential phase, the culture was split into two pre-warmed flasks, one was treated with the damaging agent and the other kept as a control. Samples were taken from both cultures 30, 60, 90 and 120 minutes after the addition of the damaging agent. RNA was extracted from these samples and used in RT real time PCR reactions. The crossing point (CP values) for the control and treated samples and the PCR efficiency (E) were used to determine the expression ratio (see Figure 4.7 for calculation and Appendix B for CP values). This was performed for each damaging agent.

4.3.1 Gene expression changes after addition of Mitomycin C

An *S. solfataricus* culture was treated with 2 µM MMC and samples taken at 30, 60, 90 and 120 minutes after addition, from the treated and control cultures. After the limited effect MMC addition had on growth of the cells (due to it being destroyed by the acidic and/or high temperature of the *Sulfolobus* media (see Figures 4.1-4.3)), there was predicted to be little effect on gene expression. However two genes (*sso0771* and *sso0280*) were checked by RT real time PCR to see if there had been any effect on their expression.

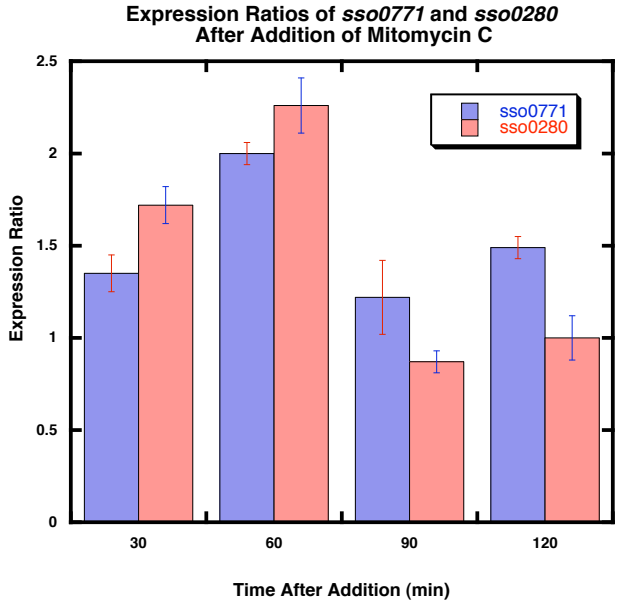


Figure 4.8 Graph of expression ratios of *sso0771* and *sso0280* after addition of Mitomycin C
There was some change in expression of the two genes at the earlier time points. A two-fold increase in expression of both genes was observed one hour after damage. The averages of triplicate measurements are plotted and standard errors are shown.

Although it was predicted that MMC would have little effect on gene expression, due to it being destroyed by the *Sulfolobus* media, there was a small increase in expression of the *sso0280* and *sso0771* genes, see Figure 4.8. The activity of MMC was obviously affected by the temperature or pH of the *S. solfataricus* culture (see figure 4.3). Changes in gene expression that were observed could be a reaction to the MMC before it was destroyed but because this could not be said with certainty the expression of the other genes after MMC treatment was not investigated.

4.3.2 Gene expression changes after addition of Methyl methane sulfonate

300 µM MMS was added to a growing culture of *S. solfataricus* and samples were taken for RNA extraction as described previously (see section 4.3). The addition of MMS had little effect on the expression of most of the genes looked at, however the expression of *sso0771* (*cdc6-2*), *sso0280 (tfb-3)* and *sso1459 (dpoII)* did show a significant increase, see Figure 4.9. Expression of *sso0771 (cdc6-2)* increased over three-fold, 90 minutes after damage. The increase of *sso1459 (dpoII)* was slower but an increase of over three-fold was also seen for this gene at the 120-minute time point. The increase in expression of *sso0280 (tfb-3)* was smaller though there was a two point five-fold increase in expression at the 120-minute time point. The increase in the expression of these genes is reminiscent of that seen after UV damage suggesting a similar response to the two types of damage.

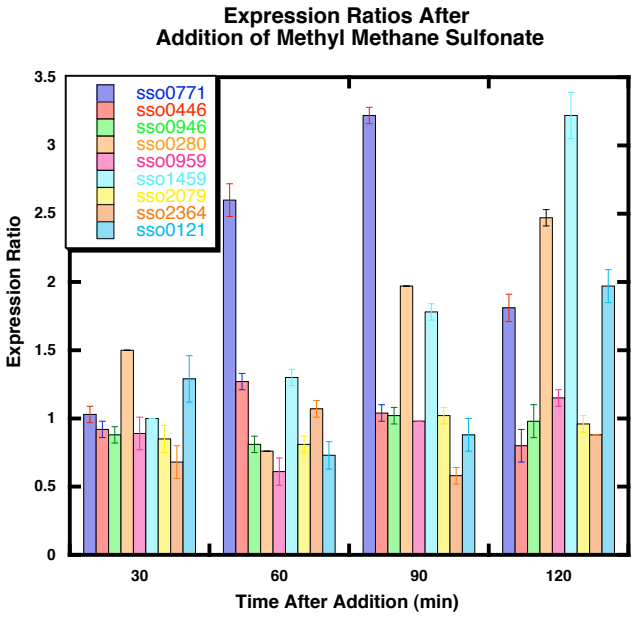


Figure 4.9 Graph of expression ratios after addition of 300 µM Methyl methane sulfonate
The addition of MMS appeared to have little effect on the expression of the majority of genes looked at, however an increase in the expression of *sso0771, sso0280* and *sso1459* was observed. The averages of triplicate measurements are plotted and standard errors are shown.

4.3.3 Gene expression changes after addition of Phleomycin

Phleomycin, to a final concentration of 200 μM, was added to growing *S. solfataricus* cells and samples taken for RNA extraction as described previously (see section 4.3).

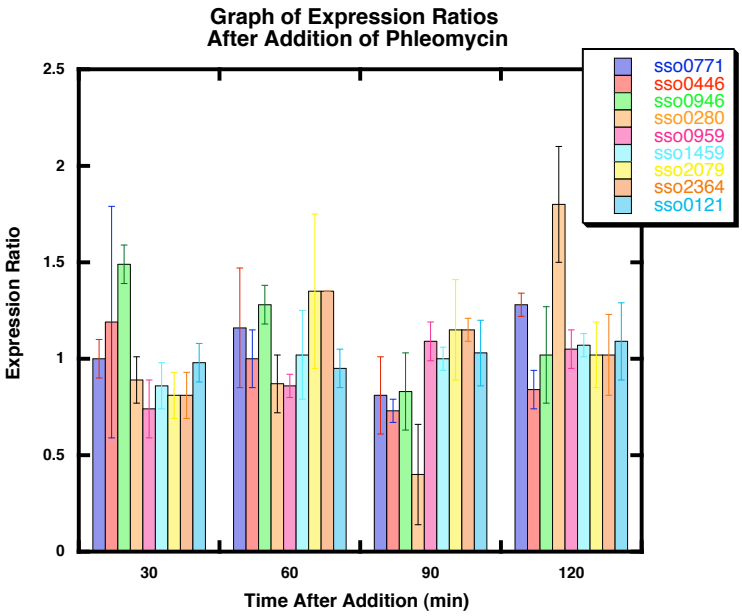


Figure 4.10 Graph of expression ratios after addition of 200 μM phleomycin
The addition of phleomycin appeared to have little effect on the expression of the genes. The averages of triplicate measurements are plotted and standard errors shown.

The addition of phleomycin appeared to have little effect on the expression of the genes looked at, see Figure 4.10. The only significant change was in the expression of *sso0280 (tfb-3)* whose expression increased by just under 2-fold by the 120-minute time point. It is possible that too low a concentration of the chemical was used for any damage response to be initiated, or that its effectiveness was inhibited by the low pH of the *Sulfolobus* media.

4.3.4 Gene expression changes after addition of Hydrogen peroxide

Hydrogen peroxide, to a final concentration of 1 μM, was added to growing *S. solfataricus* cells and samples taken for RNA extraction 30, 60, 90 and 120 minutes after addition. RT real time PCR was used to generate CP values that were then used to determine the changes in gene expression.

The addition of hydrogen peroxide to the *S. solfataricus* cells had a dramatic effect. Addition of as little as 5 μM results in the arrest of cell growth (see figure 4.6) and addition of 1 μM had a huge effect on gene expression, especially in the case of *sso2079* (*dps-like*) gene, see figure 4.11.

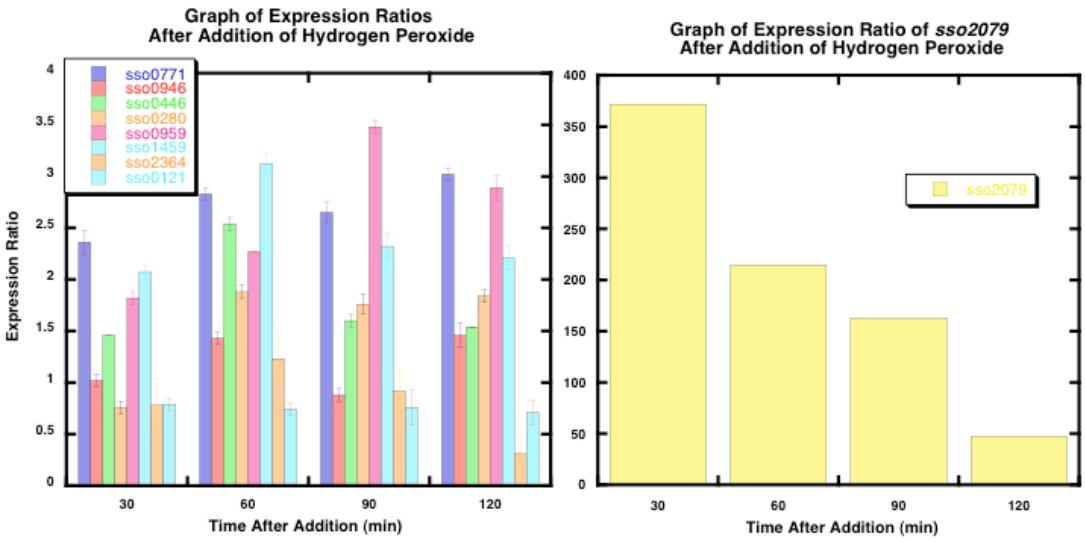


Figure 4.11 Graphs of expression ratios after addition of Hydrogen peroxide
The graph on the left shows the expression ratio of all the genes tested except *sso2079*, this gene had to be plotted on a separate graph because the increase in expression was so large. The averages of triplicate measurements are plotted and the standard errors shown.

There was a more significant response to hydrogen peroxide than there had been to the other damaging agents. A more than a 2-fold increase in expression was observed in *sso0771 (cdc6-2), sso0446 (tfb-1), sso0959 (xpb-1)* and *sso1459 (dpoII)*. However the increase in expression of gene *sso2079,* that codes for the *dps-like* (DNA protection protein from starved cells), was huge with over a 350-fold increase in expression at the 30-minute time point.

4.4 Response of *dps-like* (*sso2079*) to hydrogen peroxide treatment

The response of the *dps-like* gene to hydrogen peroxide is already well documented. The detoxifying effects of this gene product after oxidative stress are reported in many organisms, *E. coli* (Zhao, Ceci *et al*. 2002), bacteria (Andrews, Robinson *et al*. 2003), *P. furiosus* (Ramsay, Wiedenheft *et al*. 2006), *H. salinarum* (Reindel, Schmidt *et al*. 2006) and *S. solfataricus* (Wiedenheft, Mosolf *et al*. 2005).

The Dps protein was first discovered in 3-day old *E. coli* cells. It was shown to be abundant in starved cells, and formed extremely stable complexes with DNA

(Almiron, Link *et al*. 1992), hence the name DNA protection protein from starved cells (Dps).

S. solfataricus possesses a Dps-like protein encoded by *sso2079*, it is a 22 kDa protein, and twelve copies of the Dps protein assemble to form a hollow cage, with an inner diameter of ~5 nm and an outer diameter of ~10 nm, see Figure 4.12 (Wiedenheft, Mosolf *et al*. 2005).

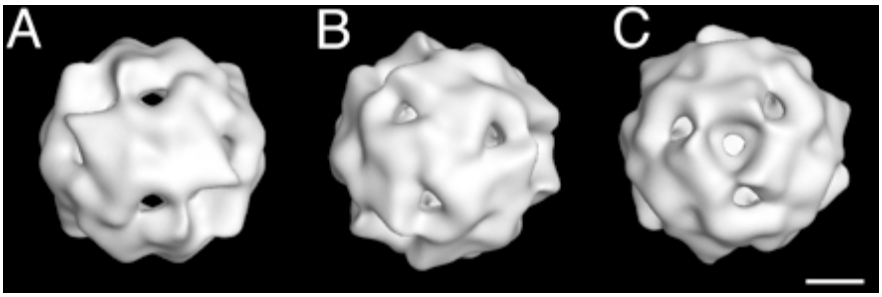


Figure 4.12 3D Image reconstruction of the *S. solfataricus* Dps-like protein
Image taken from Widenheft, Mosolf *et al*. 2005. Scale Bar 2.5 nm.

The structure of SsoDps is similar to the structure of ferritin proteins (iron storage proteins) and like the ferritin proteins SsoDps also sequesters iron within the interior of the protein cage. However, unlike the ferritins, which use oxygen as an oxidant to mineralise their stored iron (Arosio and Levi 2002), SsoDps uses hydrogen peroxide as an oxidant. In this way SsoDps is able to remove two of the components of the Fenton Reaction, see Figure 4.13, and so reduce the production of extremely damaging hydroxyl radicals (Gauss, Benas *et al*. 2006).

Fenton Reaction

$$H_2O_2 + Fe^{2+} \longrightarrow Fe^{3+} + OH^- + HO\cdot$$

Figure 4.13 Fenton reaction
The Fenton reaction occurs between hydrogen peroxide and ferrous iron and results in the formation of extremely damaging hydroxyl radicals.

The large increase in the expression of the *S. solfataricus dps-like* gene after hydrogen peroxide addition suggests that expression of the gene is controlled *in vivo* in normal (ie non-damage) condition. This prediction was confirmed by the results of

transcription assays performed by Dr Sonia Paytubi, which showed that expression from the *dps* promoter was strong *in vitro (*shown in Chapter 5, Figure 5.1). A closer look at the *ssodps-like* gene showed that it shares a divergent promoter with an *nramp* gene. Natural-resistance-associated-macrophage proteins (Nramp) are a family of divalent metal transporters involved in balancing the levels of metals within the cell (Forbes and Gros 2001). Metal ions are necessary cofactors for the activity of enzymes such as superoxide dismutase, which protect against oxidative stress. According to the *S. solfataricus* P2 complete genome sequencing project website (http://www-archbac.u-psud.fr/projects/sulfolobus/) there is a transposon interrupting the *nramp* gene. This is likely to adversely effect the expression of this gene and so the function of the protein, which may make the cells more sensitive to hydrogen peroxide damage. To determine whether our lab strain of *S. solfataricus* P2 contained the transposon, PCR primers for both halves of the gene, and the transposon were designed, see Figure 4.14.

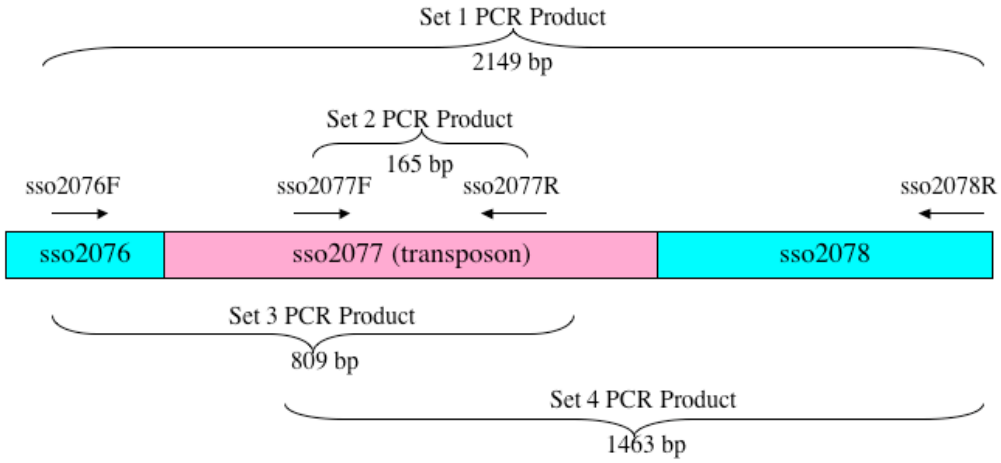


Figure 4.14 Diagram of the *nramp* gene and its transposon
The arrows represent primers. Primers were designed that would amplify the whole gene, including the transposon (sso2076F and Sso2078R). Primers for the transposon were also designed that could be used with the sso2076 and sso2078 primers to confirm the presence of the transposon.

PCR reactions were set up with the *nramp* and transposon specific primers; the results of the PCR can be seen in Figure 4.15.

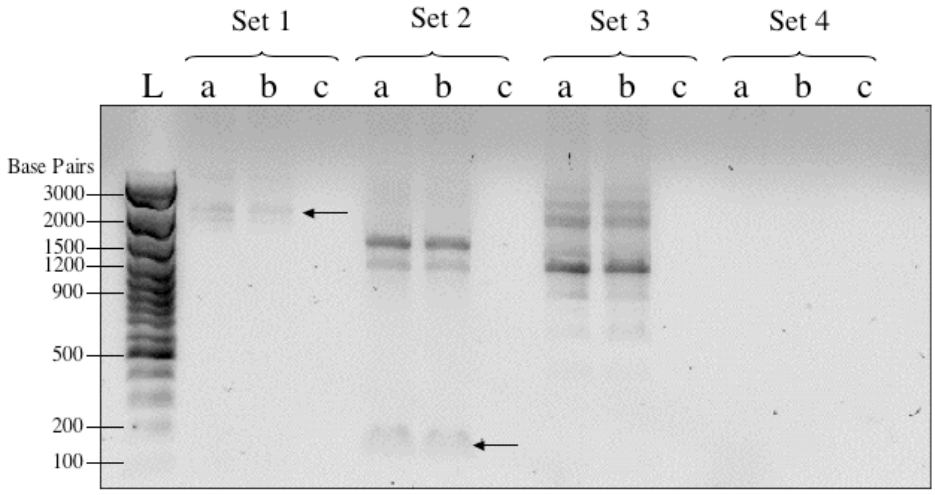


Figure 4.15 Agarose gel showing PCR products of *nramp*/transposon PCR
The expected product size with primers in Set 1 (sso2076F & sso2078R) if the transposon is present was 2149 base pairs; there was a very faint band at around this size. The expected product size of primers of Set 2 (sso2077F & sso2077R) was 165 base pairs; there is a faint band round that size however there are two larger bands of around 1200 and 1500 base pairs that are significantly brighter. Set 3 (sso2076F & sso2077R) produced multiple bands. The expected product size was 809 base pairs but the brightest of the multiple bands was bigger, around 1200 base pairs. There were no amplification products from Set 4 (sso2077F & sso2078R) the expected product size was 1463 base pairs. In each set a and b are replicates while c is the negative (no DNA) control.

The predicted size of the PCR product from Set 1 (sso2076F and sso2076R), if the transposon is present is 2149 base pairs, and if the transposon is absent is 1045 base pairs. There was a faint band running just above 2000 base pair band indicated by the ladder, suggesting the transposon is present. The primers within the transposon, Set 2 (sso2077F and sso2077F), produced multiple bands. This may be due to the fact that the *S. solfataricus* genome possesses more than one transposon (Redder, She et al. 2001). PCR reactions with primer set 3 (sso2076F sso2077R) which amplifies the first half of the *nramp* gene and half of the transposon, showed multiple products none of which appeared to be the predicted size (809 bases). There was no amplification with primer set 4, which amplifies half of the transposon and the second half of the *nramp* gene. The sso2077R primer appears to have more than one possible binding site in the genome, as PCR reactions with this primer resulted in multiple products. This could be because part of the transposon is inserted in other regions of the *S. solfataricus* genome (Martusewitsch, Sensen et al. 2000).

4.5 Discussion

After studying the genome wide response to UV damage using microarray technology and RT real time PCR a number of genes that maybe involved in repair were identified. To determine whether the induction of this set of genes was standard after all types of damage *S. solfataricus* was challenged with damaging agents that result in different types of damage, and the expression of the set of genes monitored by RT real time PCR.

The first damaging agent tested was Mitomycin C, which produces cross-links within the DNA, leading to stalling of replication. However, either the high temperature or acidic conditions of the *Sulfolobus* media appeared to inhibit its effect, or destroy the chemical altogether, as the effect on growth was minimal see Figure 4.1 - 4.3. The expression of *sso0280 (tfb-3)* and *sso0771 (cdc6-2)* did increase after the addition of MMC by 2-fold one hour after damage. Although this is a modest increase it still shows some effect of the MMC on these two genes, it is possible that there is a transient effect of the MMC before the *Sulfolobus* media destroys it.

Although there was little change in growth of the cells after Methyl methane sulfonate addition, there was a small induction of the *sso0771* (*cdc6-2*), *sso0280 (tfb-3)* and *sso1459 (dpoII*) genes. It has been suggested that *sso0771 (cdc6-2)* may be involved in control of replication and *sso0280 (tfb-3*) may be involved in controlling transcription (Gotz, Paytubi *et al*. 2007) if the prediction of these genes functions is correct, both would potentially be involved in the response to numerous kinds of damage.

The addition of Phleomycin had a noticeable effect on growth but strangely a minimal effect on gene expression. There was a small change in the expression of the *sso0280 (tfb-3)* gene, which increased around 1.8 fold by the 120 minute time point, but the ratio for the rest of the genes stayed around 1, indicating there was little change in expression in the control and treated samples.

Hydrogen peroxide had the most marked effect of any of the damaging agents and growth of the cells was severely affected by concentrations above 5 μM. Of the genes looked at, the expression of five increased more than 2-fold, *sso0771 (cdc6-2), sso0446 (tfb-1), sso0959 (xpb-1), sso1459 (dpoII)* and *sso2079 (dps-like)*. The expression ratio of the first four genes was between 2 - 3.5, however the increase in expression of *sso2079*, the *ssodps-like* gene was huge. At 30 minutes after damage there was a 350-fold increase in expression (see Figure 4.11). Hydrogen peroxide induces oxidative stress, as does UV irradiation (although to a lesser extent), which may explain why many of the same proteins were up regulated after both kinds of damage. Oxidative stress results from an imbalance between ROS and antioxidants in the cells, leading to numerous types of damage, such as DNA breaks, damage to lipids and proteins, as ROS scavenges electrons wherever they find them (Valko, Rhodes *et al*. 2006). Because of the indiscriminate damage caused by oxidative stress, genes involved in many different pathways could be induced after addition of hydrogen peroxide.

The set of genes as a whole does not appear to be induced as standard for all types of damage. However the results shown here indicate that *sso0280 (tfb-3)* and *sso0771 (cdc6-2)* are induced by all of the damaging agents tested. When faced with DNA damage the cells DNA replication and cell division processes are slowed down or stopped to allow repair or take place, and to prevent the transmission of damage to daughter cells. *S .solfataricus* possess three Cdc6 proteins and it has been predicted that they have a function in control of replication initiation. The Cdc6-2 protein is able to bind the Cdc6-1 and Cdc6-3 binding sites, possibly blocking access of these proteins in order to pause replication (Robinson, Dionne *et al*. 2004). The increased expression of *sso0771 (cdc6-2)* in all the damage condition tested lends further weight to this idea.

The difference in the levels of the three *S. solfataricus tfb* genes could provide some insight as to how this organisms controls transcription after being exposed to damage. After UV irradiation levels of *tfb1* remained constant, levels of *tfb2* decreased and

levels of *tfb3* increased. Increase in the levels of *tfb3* was observed with all damaging agent suggesting this protein is involved in the cells response to numerous types of damage. The TFB3 protein is a truncated version of TFB1 and 2, missing the B-finger and DNA binding helix-turn-helix domain, leaving it unable to bind to promoter DNA and presumably unable to stimulate RNAP as the B-finger is predicted to perform this function. Surprisingly TFB3 has been shown to have a stimulatory effect *in vitro* when added to assays containing the basal transcription proteins (Paytubi unpublished). The mechanism of this stimulation is not yet known.

sso1459 DpoII was also induced by methyl methane sulfonate and hydrogen peroxide, but the significance of this is unknown.

The response of the *dps-like* gene to hydrogen peroxide damage is well documented, but how the expression of this gene is controlled in *S. solfataricus* is not known. Transcription assays showed that transcription from the gene is strong *in vitro* (see Chapter 5, Figure 5.1) suggesting a repressor *in vivo*. Further investigation of the *dps* promoter and identification of a possible transcriptional repressor are presented in Chapter 5.

Chapter 5

The *Dps* Promoter and Purification and Characterization of its Predicted Transcriptional Repressor Sso2273 (and paralogue Sso0669)

5.1 Introduction

The huge increase in the transcription of the *Sulfolobus solfataricus dps-like* gene after hydrogen peroxide damage (see Chapter 4), suggests that the gene is being repressed under normal ie non-damaging conditions.

Closer inspection of the *dps* promoter revealed an inverted repeat sequence, which may produce an 8 base pair hairpin structure. These structures are known binding sites for other transcriptional regulators (Hill, Cockayne *et al*. 1998; Dussurget, Timm *et al*. 1999; Peeters, Thia-Toong *et al*. 2004). A PCR product 329 bases in length, which incorporated the inverted repeat, the TATA box and the predicted transcription start site of the *dps* gene was cloned into a 2.1 TOPO plasmid from Invitrogen. This plasmid was cut with *Xho*I to produce a linear piece of DNA for use in *in vitro* transcription assays.

In vitro transcription assays were performed by Dr Sonia Paytubi, the aim of these experiments was to determine whether the addition of the transcription factor TFE had a stimulatory effect on the transcription from a selection of *S. solfataricus* and *S. acidocaldarius* promoters, see Figure 5.1. The fact that strong transcription was observed from the *S. solfataricus* Dps promoter is of interest to this work as it supports the idea that its expression of the *dps* gene is being repressed under normal conditions in the cell.

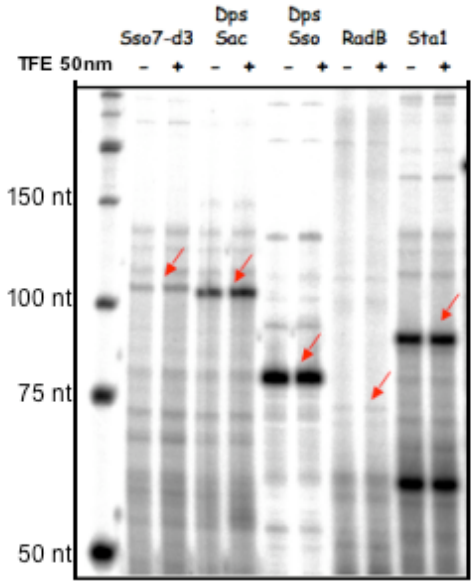


Figure 5.1 Gel showing transcription from a variety of *Sulfolobus* promoters, with and without TFE
In vitro transcription assays were performed by Dr Sonia Paytubi, and showed that, *in vitro*, there was strong transcription from the *dps* promoter.

5.2 Determining the *dps* promoter region

After looking at the promoter region of the *dps* gene an inverted repeat was identified, which was predicted to be a possible binding site for a transcriptional repressor. A 45 base pair region of the promoter containing the inverted repeat was biotinalyted (see Figure 5.2) and used as 'bait' in pull down assays in an attempt to isolate the transcriptional repressor.

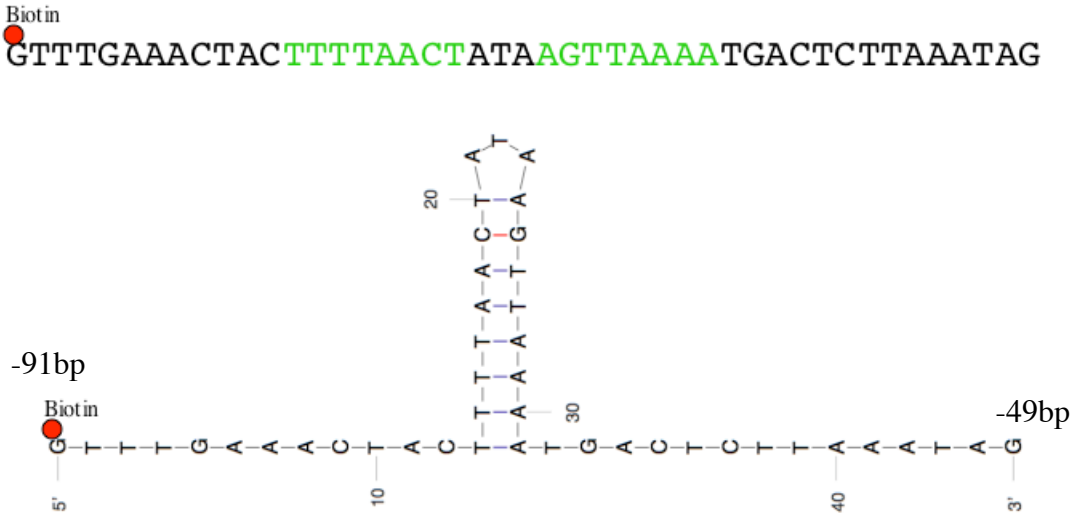


Figure 5.2 Diagram of the *dps* promoter
Diagram showing the sequence from the upstream region of the *dps* promoter used in the pull down assays. The sequence of DNA is 49 bases upstream of the genes predicted start site and contains an inverted repeat that is predicted to form a hairpin structure as shown. Double stranded DNA was used in the pull down assays.

5.3 Pull down assays with Biotin labelled *dps* promoter

The biotin labelled DNA was annealed with its complementary strand and incubated with cell lysate from either hydrogen peroxide treated or control (non-treated) *S. solfataricus* cultures. The biotin was then attached to streptavidin coated magnetic beads, which allowed the DNA to be isolated from the solution. The beads-DNA complex was washed with buffer containing increasing salt concentrations and the washes and beads analysed by SDS-PAGE, see Figure 5.3.

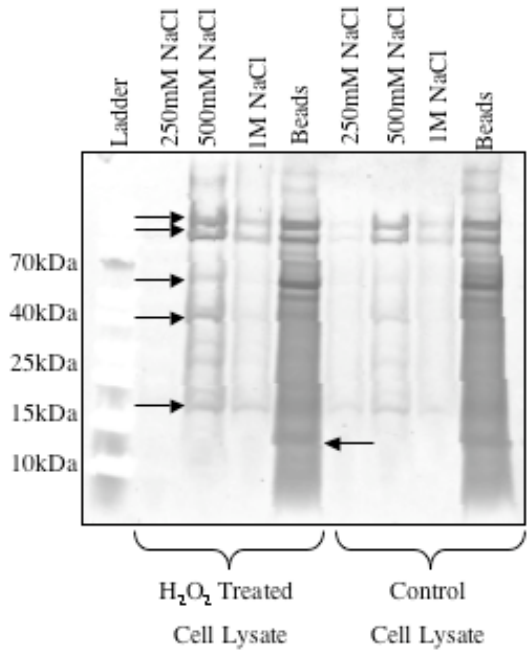


Figure 5.3 SDS-PAGE analyses of salt washes and beads from pull down assays
Once the biotin labelled *dps* promoter DNA was incubated with the cell lysate, buffer with increasing salt concentrations was used to wash the beads and remove any bound proteins. These washes, and the boiled beads, were analysed by SDS-PAGE and any bands of interest were cut out, and sent to the Mass spectrometry service at the University of St Andrews for identification. The arrows indicate bands that were cut out for identification; the corresponding bands were cut out of the control lanes.

5.4 Identification of a predicted transcriptional repressor by mass spectrometry

After the salt washes and beads had been analysed by SDS-PAGE and the gel stained with SYPRO ruby. Any bands of interest were cut out of the gel and sent to the Mass Spectrometry service at the University of St Andrews for identification. The majority of the bands from the 500 mM salt wash were identified as RNA polymerase subunits. The band from the beads lane was predicted to be Alba1 as it ran at the correct size and was expected to be abundant as it is a DNA binding protein. The most abundant protein in that band in the beads lane however, was identified as Sso2273. Bioinformatics showed that Sso2273 is closely related to a metal dependent repressor that controls transcription of the diphtheria toxin in *Corynebacterium diphtheria*, see

Figure 5.4. This repressor is called the diphtheria toxin repressor (DtxR) (Schmitt, Twiddy *et al*. 1992).

As well as regulating the transcription of the diphtheria toxin, DtxR controls proteins involved in iron metabolism (Brune, Werner *et al*. 2006). The iron paradox is well documented; iron is essential to cellular processes but in its reduced form is highly toxic. Therefore iron levels in the cell must be tightly controlled to prevent damage to the cell (Valko, Morris *et al*. 2005). Sso2273 is an attractive target as the repressor for the *dps* gene; it is a predicted transcriptional repressor and is homologous to proteins known to control Dps in other organisms. Its paralogue Sso0669 is also a predicted transcriptional repressor that share strong homology to the DtxR family of repressors

DtxR purifies as a monomer but readily forms a dimer through disulfide-bridges in its dimerization domain, each monomer contains two metal ion-binding sites (Spiering, Ringe *et al*. 2003; Love, vanderSpek *et al*. 2004) which must be occupied for the protein to adopt the correct conformation to bind DNA (Rangachari, Marin *et al*. 2005). *In vivo* the metal occupying the binding sites is ferrous iron, however in *in vitro* experiments most divalent metals were accommodated (Love, vanderSpek *et al*. 2004).

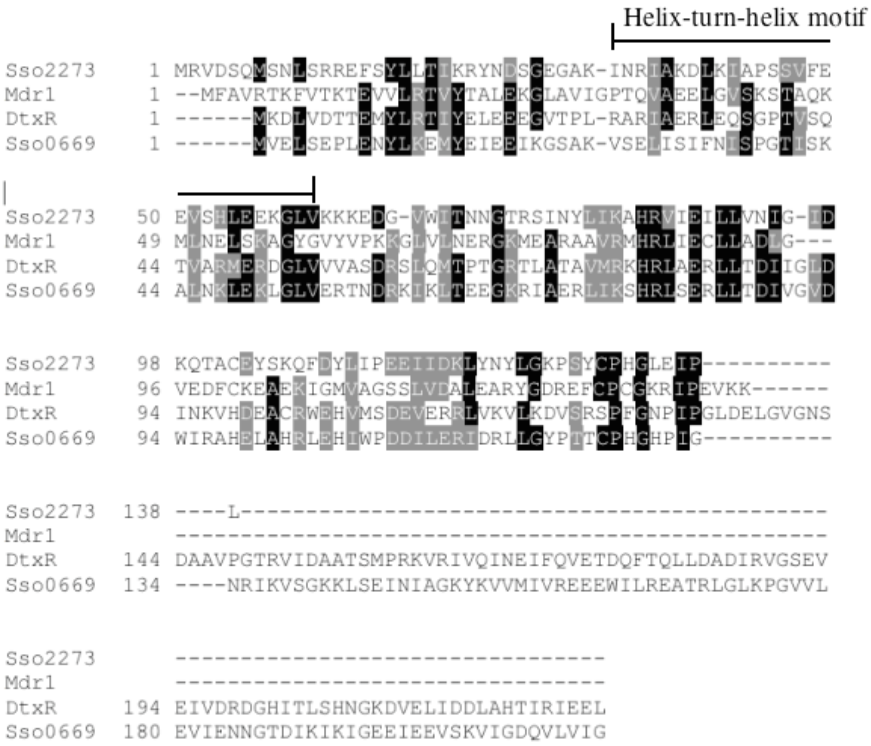


Figure 5.4 Protein sequence alignment of Sso2273 and other metal dependent repressors
Sequence alignment of Sso2273 (accession number D90397), Metal Dependent Repressor (MDR1) from *Archaeoglobus fulgidus* (accession number G69497), Diphtheria toxin repressor (DtxR) from *Corynebacterium diphtheria* (accession number P33120) and Sso0669 (accession number Q9UX58).

5.5 Expression and purification of Sso2273

Sulfolobus solfataricus genomic DNA was used to PCR amplify the gene *sso2273*. The PCR product was generated with primers that introduced restriction enzyme cut sites at the ends of the gene, to allow it to be cloned into a pET28c vector using restriction sites *Nco*I and *Sal*I. A sample of the plasmid, with insert, was sent to the University of Dundee's sequencing service to be checked for accuracy. The plasmid was transformed into *Escherichia coli* BL21 Rosetta cells for expression (see Chapter 2, section 2.8.2). The protein was heat stable, so the first step in purification was a heat treatment to remove the majority of *E. coli* proteins. Sso2273 was further purified through a Heparin column and by size exclusion chromatography (see Figure 5.5). Mass spectrometry confirmed the identity of the protein.

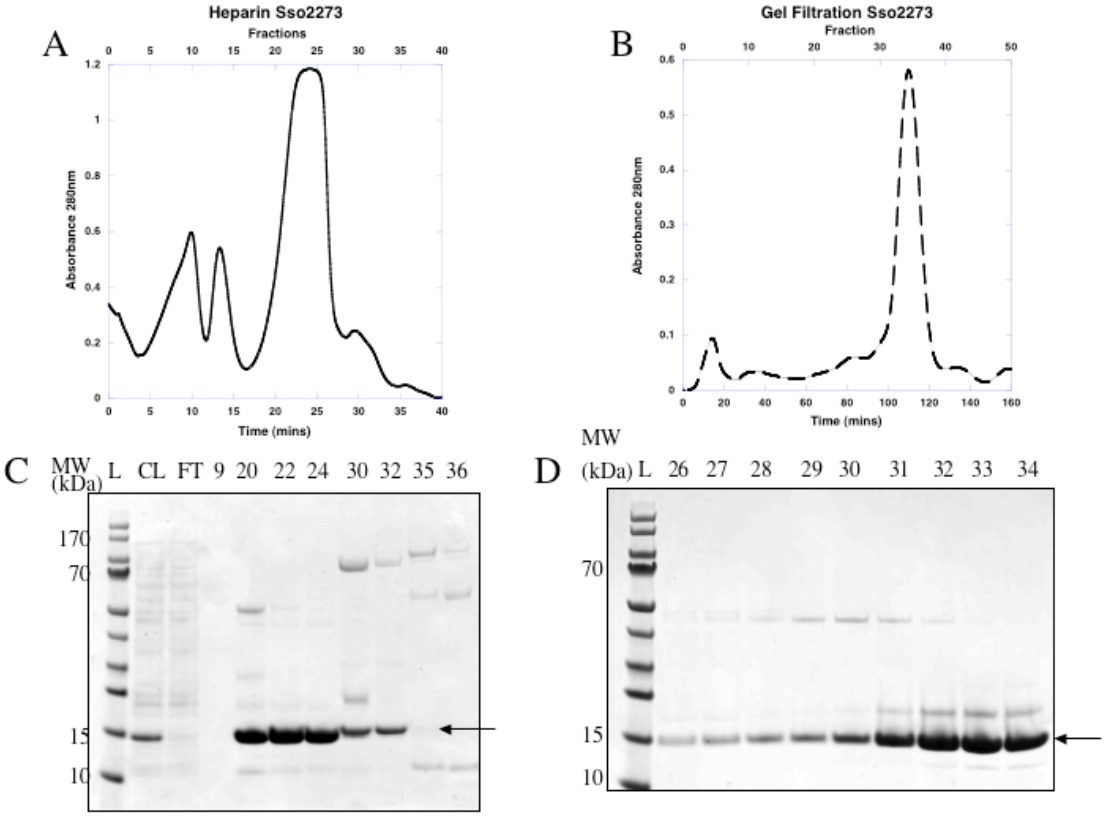


Figure 5.5 Purification of Sso2273
Elution chromatographs from (A) Heparin and (B) Gel Filtration columns. (C) and (D) show SDS-PAGE of fractions from the Heparin column and Gel Filtration columns, respectively. The Heparin column gel also shows the cell lysate (CL) and flow through (FT) from the column, which was collected in case the protein did not bind the column. Arrows indicate Sso2273.

A sample of the pure protein was sent to the Mass Spectrometry service at the University of St Andrews. The major peak in the mass spectrometry trace (see Figure 5.6) showed a protein of 15014.40 daltons, which is close to the theoretical mass of 15101.3 daltons expected for protein Sso2273. Purified Sso2273 will be named native Sso2273 through out this chapter.

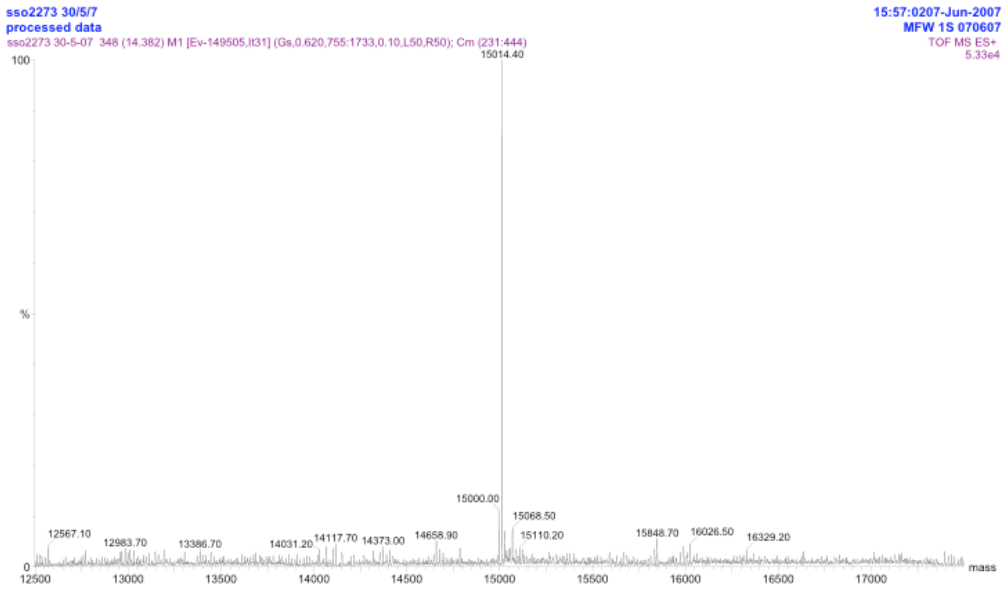


Figure 5.6 Whole mass determination of Sso2273 by Mass Spectrometry
The theoretical mass of Sso2273 according to ProtParam is 15101.3 Da. The major peak from the mass spectrometry trace matched well at 15014.40 Da

5.6 Expression and purification of Sso0669

Sulfolobus solfataricus genomic DNA was used to PCR amplify the gene *sso0669*. The PCR product was generated with primers that introduced restriction enzyme cut sites at the ends of the gene, to allow it to be cloned into a pET28c vector using restriction sites *Nco*I and *Sal*I. A sample of the plasmid, with insert, was sent to the University of Dundee's sequencing service to be checked for accuracy. The plasmid was transformed into *Escherichia coli* BL21 Rosetta cells for expression (see Chapter 2, section 2.8.2).

Sso0669 was also found to be heat stable, and so the first step in the purification of this protein was a heat treatment. The majority of contaminating proteins were removed by this heat treatment and further purification of Sso0669 did little to improve the purity of the protein. Sso0669 is annotated as an iron-dependent repressor in the *Sulfolobus* database; therefore it would be expected to bind DNA. The pI of this protein is relatively low and so at pH 8 the protein will be negatively charged. For these reasons it was expected that Sso0669 would bind to a strong anion exchange column like Mono Q column, however the majority of the protein did not bind the

column and came out in the flow through. Size exclusion chromatography was also used to further purify the protein (see Figure 5.7).

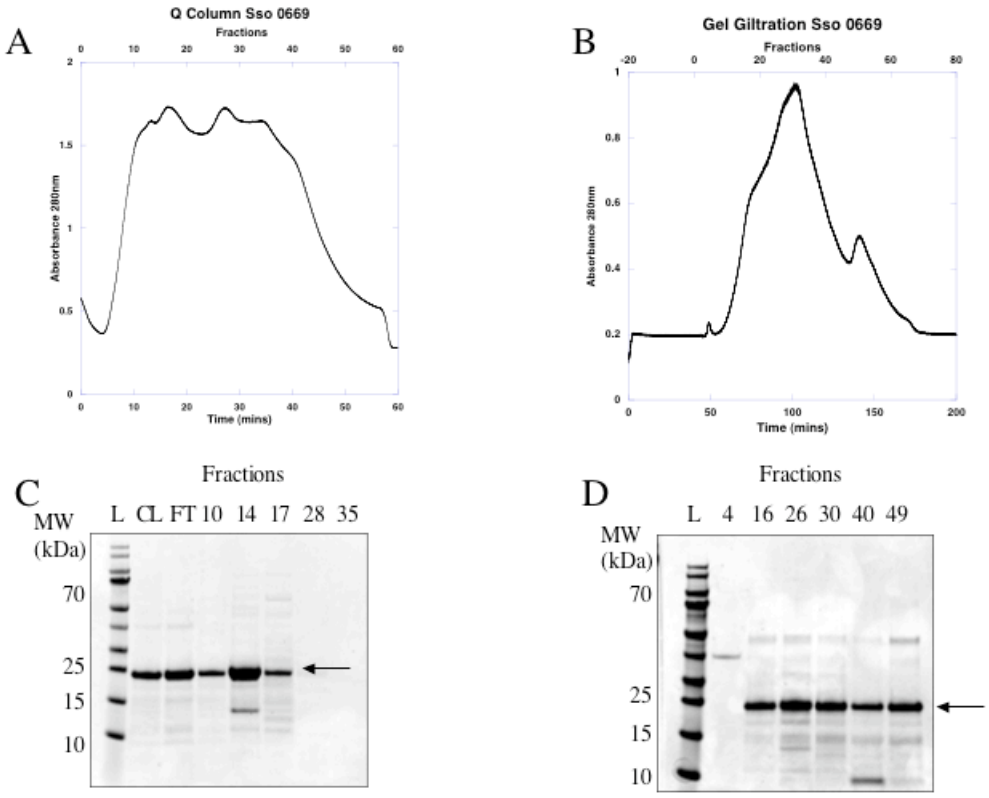


Figure 5.7 Purification of Sso0669
Elution chromatographs from (A) Mono Q Column and (B) Gel Filtration columns. (C) and (D) show SDS-PAGE of fractions from the Mono Q column and Gel Filtration columns, respectively. The Mono Q column gel also shows the cell lysate (CL) and flowthrough (FT) from the column, which shows that Sso0669 does not bind well to the Mono Q column, and much of the protein came out in the flow through. Arrows indicate Sso0669.

The trace of absorbance at 280nm from the Mono Q column shows something absorbing in the sample at elution volumes where the corresponding fractions show no protein is present. It is possible that there is DNA contamination in the sample leading to this false trace.

Sso0669 expressed well but purification of this protein was problematic, for this reason single band purity was never achieved for the protein, although it was the strongest band representing 80 % of the final purified sample. Purified Sso0669 will be named native Sso0669 through out this chapter.

5.7 Crystal structure of Sso2273

A sample of purified Sso2273 was given to the Scottish Structural Proteomics Facility (SSPF) at the University of St Andrews. Crystals were obtained in screen conditions containing zinc, see Figure 5.8, the structure was solved using anomalous signals present from the Zinc bound to the protein. The analysis presented here is the work of Dr Stephen McMahon's of the SSPF lab.

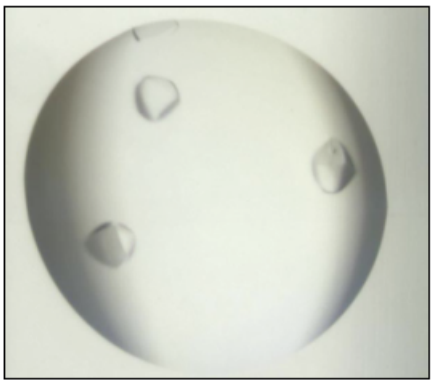


Figure 5.8 Sso2273 crystals
Work was performed in the SSPF lab at the University of St Andrews. Crystals were grown in the following screen conditions: 0.1 M Sodium cacodylate pH 6.5, 0.2 M zinc acetate, 10 % propan-2-ol.

The crystal structure is shown in Figure 5.9. The asymmetric unit contains four monomers of the protein forming two dimers. The monomers are essentially identical, superimposing on one another with an average rmsd of 0.81 Å. Each monomer contains two domains, an N-terminal DNA binding domain (residues 8 to 70) and a C-terminal dimerisation domain (residues 71 to 138). The N-terminal domain consists of three α helices (α 1, 2 and 3) and a two strand anti-parallel β sheet. Excluding α1, the remaining secondary structure elements of the N-terminal domain are arranged in a DNA binding winged helix-turn-helix motif. The C-terminal domain is connected to the N-terminal domain by one long α helix, α4. α4 is flanked at its termini by two short α helices, α5 and α6. Extending from α6 is a 14 residue loop which together with α helices 4, 5 and 6 constitutes the C-terminal domain. Two zinc atoms are bound by the C-terminal domain.

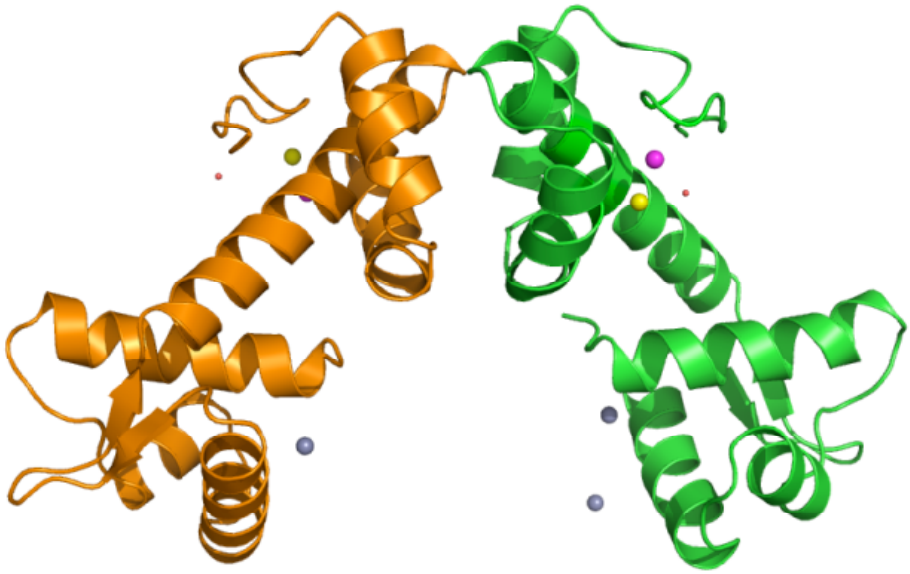


Figure 5.9 Crystal structure of Sso2273
Structure shows two monomers of Sso2273 in orange and green forming a dimer. The zinc atoms bound in the C-terminal domains are shown in yellow and pink, the smaller red sphere near the zinc atoms indicates a water molecule. The grey spheres near the N-terminal DNA-binding domains indicate zinc atom, however these are believed to be an artefact of the crystallisation process.

A structure similarity search using DALI showed that the overall dimer arrangement of Sso2273 resembles that of proteins from the Diphtheria toxin repressor family, this family include the manganese transport regulator of *Bacillus subtilis* MntR and the iron dependent repressor IdeR. As expected the general topology of Sso2273 compares well with that of MntR and IdeR, rmsd for corresponding C α is 2.4 (100 residues) and 2.2 (106 residues) respectively.

Analysis of the structure with the program PISA revealed a strong dimeric arrangement as the biological unit. The interface, organised at the C-terminus, buries approximately 800 Å^2 per monomer (approximately 10 % total monomer surface area), for comparison IdeR dimer buries 937 Å^2 and MntR 1254 Å^2. The dimer interactions are largely hydrophobic; L90, L91, I94, I96, F108, I112, I116 and L120 make most of the dimer contacts. In addition, there are up to four hydrogen bonds across the interface involving G95, Y104, E115 and K119. All dimer contacts are located in α helices 4, 5 and 6 and the connecting loops of the C-terminal domain, see Figure 5.10.

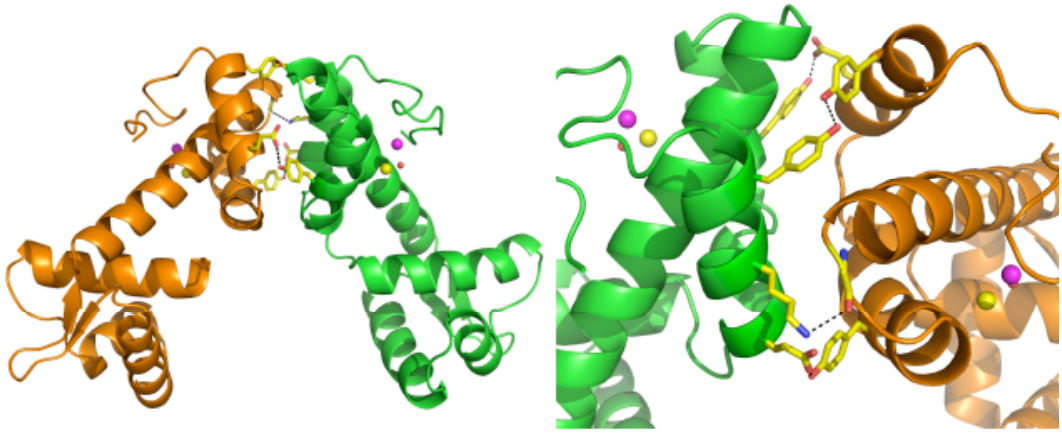


Figure 5.10 Dimer interface between Sso2273 monomers
The dimerisation interface of the two monomers is between their C-terminal domains. The dimer interface is largely hydrophobic involving leucines and isoleucines. The four hydrogen bonds, between residues glycine 95, tyrosine 104, glutamic acid 115 and lysine 119, are shown.

Each monomer tetrahedrally coordinates two zinc atoms in the C-terminal domain. The zinc atoms are separated on average by 3.28 Å. In total five residues participate in metal binding, H84, E88, C1020, C130, H132. Site one is coordinated by C130 and H132 and site 2 by H84 and C1020. In addition E88 and one solvent water molecule tethers both metal ions, see Figure 5.11. The average B factors for the metals (site 1 - 46.9, site 2 – 51.0) are comparable with that for all protein atoms in the structure.

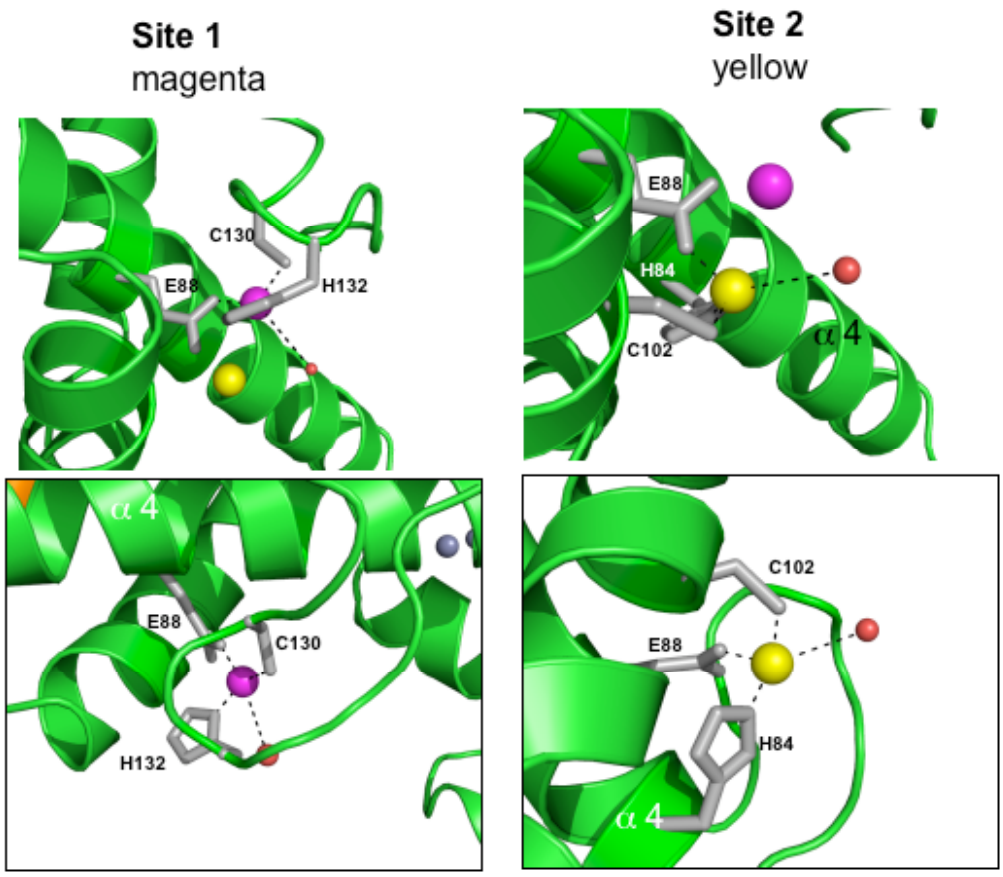


Figure 5.11 The two metal binding sites of Sso2273
In site 1 the zinc ion shown in magenta is coordinated by a cysteine 130 and histidine 132. In site 2 the zinc ion shown in yellow is coordinated by cysteine 102 and histidine 84, glutamic acid 88 and one solvent water molecule tethers both metal ions.

Sso2273 is a homologue of proteins in the Diphtheria toxin repressor family, the structures of two members of this family, the manganese transport regulator (MntR) from *Bacillus subtilis* and iron dependent repressor (IdeR) from *Mycobacterium tuberculosis* are seen in Figure 5.12.

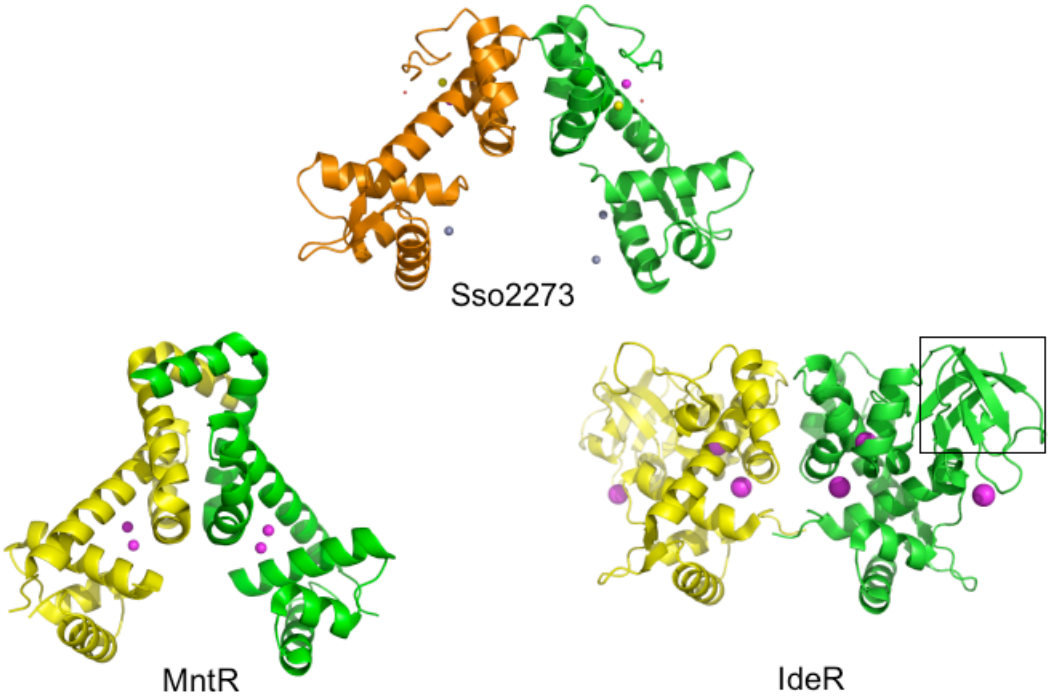


Figure 5.12 Structures of Sso2273 and its homologues MntR and IdeR
The overall structure of the three proteins is similar however there are distinct differences between them. The MntR has an extra C-terminal helices compared to Sso2273 and its metal binding sites are in a difference position. The IdeR contains a SH3 domain (in the black box) that is absent from Sso2273.

Analysis of the sequence alignment between Sso2273, IdeR and MntR suggested that for proteins so topologically similar the metal binding pattern of Sso2273 may be strikingly different from that observed in both IdeR and MntR. Firstly, the C-terminal SH3-like domain of IdeR that contributes two residues to binding site 1 is absent in Sso2273 (this domain is also missing from MntR) and the H98 residue involved in metal binding in IdeR (which is also conserved in MntR) is replaced by C1020 in Sso2273. None of the residue involved in coordinating Mn or Co at site 2 in IdeR, and MntR, are conserved in Sso2273.

Given this analysis it is not surprising to discover that neither binding sites in MntR resembles the binding sites in Sso2273. The distance between corresponding metal ions in Sso2273 and MntR is greater than 10 Å. The only residue in common to both proteins involved in metal binding is H84 (77), found in site 2 of Sso2273 and site 1 of MntR, see Figure 5.13.

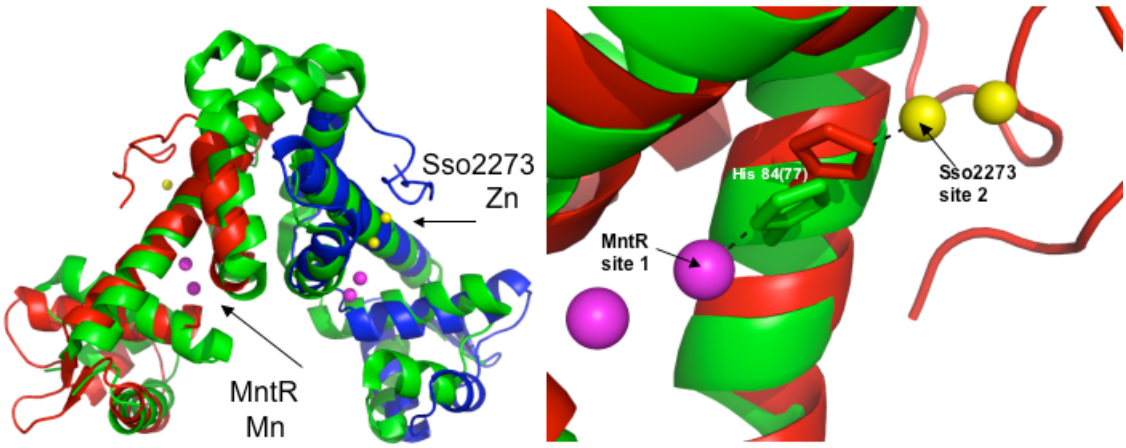


Figure 5.13 Comparison of metal ion binding sites of Sso2273 and MntR
In the left image one monomer of Sso2273 is shown in blue and one in red while MntR is shown in green. The zinc ions from Sso2273 are shown in yellow and the manganese ions from MntR are shown in pink. In the right hand image shows residue His84 (77) is the only residue involved in metal ion binding that both proteins share. His84 (77) coordinates metal ion binding in site 1 of MntR and site 2 of Sso2273

When comparing Sso2273 with IdeR (see Figure 5.14) it was surprising to discover some similarities between the metal binding sites. Overlaying the secondary structures placed metal binding site 1 of IdeR and metal binding site 2 of Sso2273 in a similar environment ~1.8 Å apart. H84 and E88 are conserved and are required in both proteins for metal coordination. Each residue has the same chemical binding pattern with their cognate metal (ion NE2 and OE2). Completing the binding site in Sso2273 is C102 which, although not conserved at the sequence level, structurally overlays with H98 (also involved in metal coordination) in IdeR. Additional residues from the SH3-like C-terminus complete site 1 in IdeR (Glu172 and Gln175) and do not overlay with any metal binding residues in Sso2273.

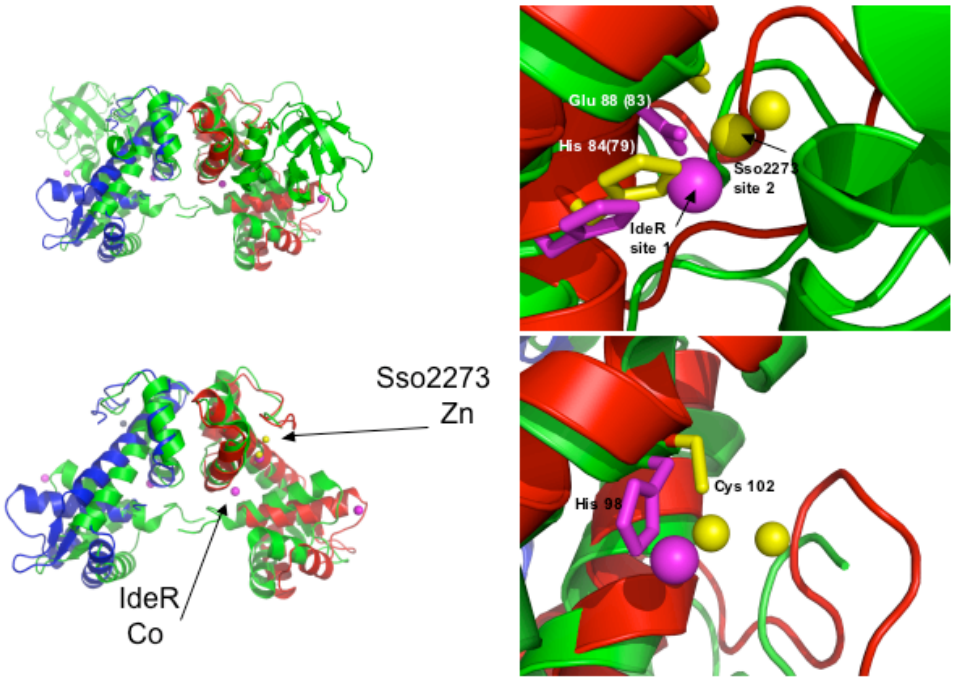


Figure 5.14 Comparison of the metal ion binding sites of Sso2273 and IdeR
In the left images one monomer of Sso2273 is shown in blue and one in red while IdeR is shown in green (with the SH3 domain in the top image, and missing the SH3 domain in the bottom image). The zinc ions from Sso2273 are shown in yellow and the cobalt ions from IdeR are shown in pink. Sso2273 does not have the SH3 domain found at the C-terminal of IdeR. The top right image shows Sso2273 metal binding site 2 and IdeR metal binding sites 1 overlaid. The bottom image shows the same image rotated slightly and with the SH3-like domain removed.

When the C-terminal region of Sso2273 is aligned relative to the C-terminal region of MntR or IdeR the DNA binding domains of the proteins do not aligned, see Figure 5.15. The poor alignment of the DNA binding domains may mean that helix-turn-helix of Sso2273 would not fit well in the major groove if the DNA, effecting its ability to bind DNA.

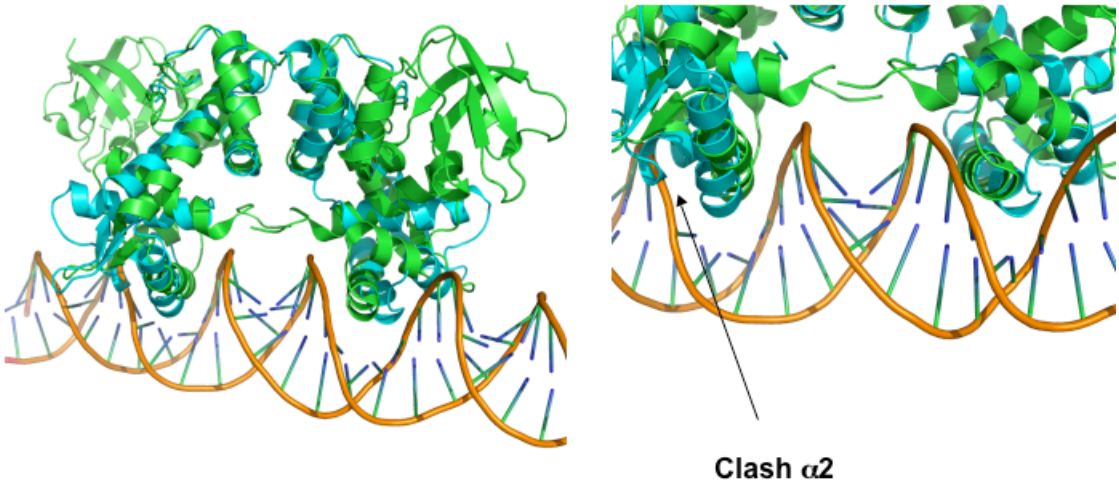


Figure 5.15 Alignment of the C-terminal domains of Sso2273 and IdeR
Sso2273 is shown in turquoise and IdeR is shown in green If the C-terminal domains of Sso2273 and IdeR are aligned the DNA binding domains of Sso2273 are not in the ideal position to bind DNA as the α2 helix clashes with the DNA.

There is also a clash between the DNA and the protein when either of the N-terminal recognition helices of Sso2273 and IdeR are aligned, as the α3 helix is completely misaligned, see Figure 5.16

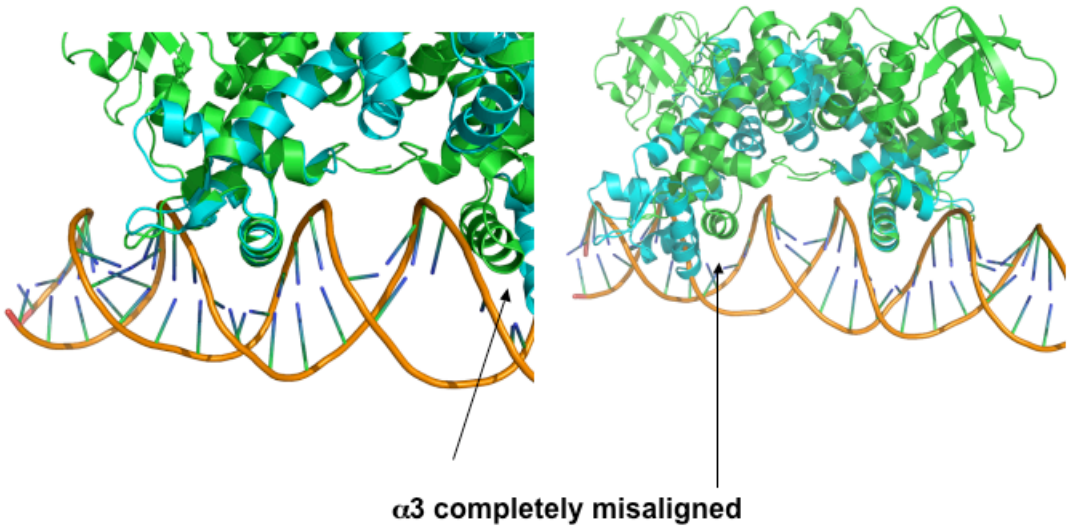


Figure 5.16 Alignment of the N-terminal recognition helix of Sso2273 and IdeR
When the N-terminal recognition helices of Sso2273 (in turquoise) and IdeR (in green) are aligned for one monomer, the α3 helix of the other monomer is in an unfavourable position for DNA binding and clashes with the DNA. The left image shows the alignment of the helices of the left monomer and the right image shows the alignment of the right monomer.

5.8 Metal reconstitution

It was unknown if any metal occupied the proteins metal binding sites after purification. Metal reconstitution was performed in an attempt to reintroduce divalent metals into the binding site, and so obtain the DNA binding form of the protein. Sso2273 was incubated with a number of different divalent metals, $ZnCl_2$, $NiCl_2$, $CoCl_2$ and $Fe(NH_4)_2SO_4$. After incubation, the protein-metal solution was applied to a PD-10 column to separate any unbound metal.

Sso2273 was also grown with $ZnCl_2$ (50 μM added at the point of IPTG induction) in the growth media. Zinc was the metal in the binding sites when the protein crystallised, by providing excess zinc it was hoped that the protein would bind the metal, and adopt the DNA binding form of the protein.

5.9 Analytical gel filtration (Superose 6)

Analytical gel filtration was performed to determine if metal reconstitution had had any effect on the elution volume (and so an indication of size change) of the protein. Native Sso2273, Sso2273 grown with $ZnCl_2$ or reconstituted with $NiCl_2$ or $CoCl_2$ was applied to the Superose 6 column; the same program was run in each case (see Figure 5.17). The volume at which the protein eluted was used to give an indication of the proteins size. Gel filtration can only give an estimate of the size of a protein, as the movement of a protein through the column is dependent on the shape of the protein as well as the size. The analytical column had previously been equilibrated with proteins of known size, and a standard curve was available.

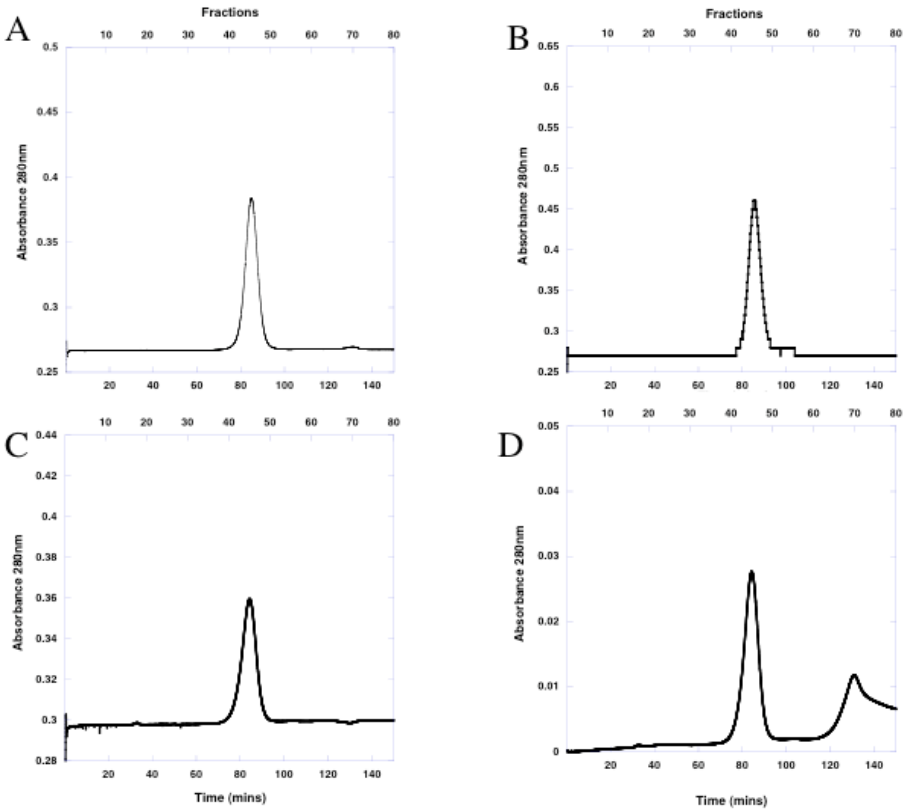


Figure 5.17 Chromatograph traces from the analytical gel filtration column (Superose 6) Chromatographs from the analytical gel filtration column for (A) Native Sso2273, (B) Sso2273 grown with $ZnCl_2$ (C) Sso2273 reconstituted with $MnCl_2$ and (D) Sso2273 reconstituted with $NiCl_2$

Using the equation for the standard curve (see figure 5.18) the size of the native and metal reconstituted Sso2273 proteins could be estimated. The elution volume for all the proteins was very similar, between 84.3 and 85.4 ml, giving molecular weights of between 16.5 and 15.1 kDa respectively, suggesting Sso2273 is a monomer in solution regardless of the metals added. DtxR is also a monomer in solution, forming a dimer once its metal ion binding sites are occupied (D'Aquino, Tetenbaum-Novatt et al. 2005).

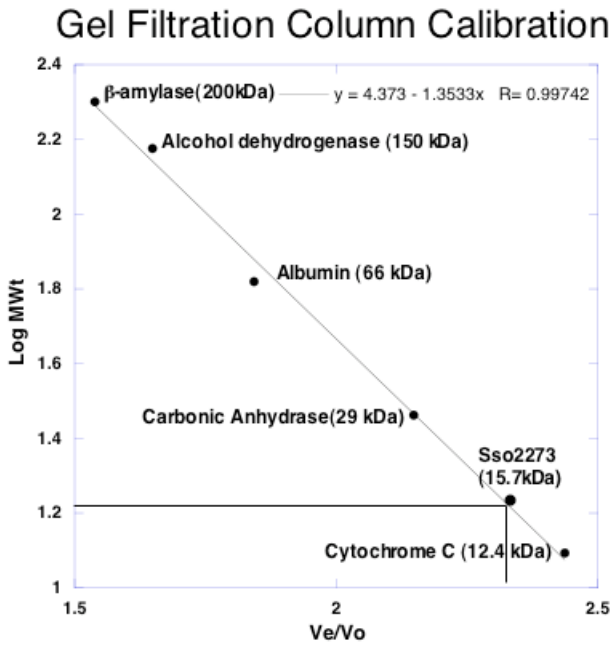


Figure 5.18 Gel filtration calibration
The analytical gel filtration column was calibrated with proteins of known size to produce a standard curve. The average elution volume of Sso2273 is indicated.

5.10 Determining if metal binding sites were occupied

The homologue of Sso2273 is a transcriptional repressor from *Corynebacterium diphtheria* (DtxR). DtxR is a dimer, and each monomer contains two metal binding sites that bind divalent metals. Ferrous iron is bound preferentially in the cell but *in vitro* the protein will accommodate most divalent metal ions in the binding site (Spiering, Ringe *et al*. 2003). A number of assays were performed to determine whether Sso2273 had purified with its metal ion binding sites occupied. In the crystal screens Sso2273 crystallised with zinc in the binding sites so it is possible that Sso2273 would bind other divalent metals.

5.10.1 Determination of Iron content using a Bathophenanthroline assay

Because of its ability to chelate iron producing a red complex, bathophenanthroline is used in an assay to indicate the presence of iron within proteins. First the native protein is treated with a strong reducing agent (e.g. ascorbic acid) that converts all iron ions present into their reduced form. Three molecules of Bathophenanthroline coordinate to form a complex with the iron. The number of iron molecules bound to the protein can be calculated by reading the absorbance of the solution at 535 nm and

using the extinction coefficient of bathophenanthroline, 22369 M^{-1} cm^{-1} (Pieroni *et al* (2001)).

Protein	Absorbance 535nm			Average Abs	Total Fe (pmol)	Fe/Sso2273
Sso2273	0.0037	0.0054	0.0044	0.0045	382	0.07

Table 5.1 Calculation of number of Iron molecules bound to Sso2273 using the bathophenanthroline detection method

However, when the assay was performed with Sso2273 the absorbance reading was 0.0045 (average of 3 readings), see Table 5.1. This low value suggests that no iron was bound to the protein. The total number of iron molecules (pmol) was worked out as follows; Average absorbance/0.022369 μM^{-1} x 1900 µl (final volume of the reaction). 5000 pmol of Sso2273 was used in the reaction, therefore the number of iron molecules per Sso2273 molecule is 382 pmol Fe/5000 pmol Sso2273, which gives 0.07, less than 1 iron per protein suggesting, as stated before, that no iron was bound to the native protein.

In the homologue, DtxR, the metal is released from the protein once oxidised (Spiering, Ringe *et al*. 2003). Sso2273 was not purified in anaerobic conditions so it was possible that any iron that may have been bound was oxidised and released during the purification process.

5.10.2 Inductively coupled plasma – optical emission spectroscopy (ICP-OES)

ICP-OES uses an inductively coupled plasma source (an ionized argon gas). As the sample is introduced into the plasma, through a nebuilzer system, it is completely atomised and the elements in the sample excited. As the elements relax back to their ground state, their emissions can be measured and the wavelength of emission is different for each element. By comparing the light emitted by a specific element, to measurements of standards for that element the concentration of the element in the sample can be calculated, see Figure 5.19.

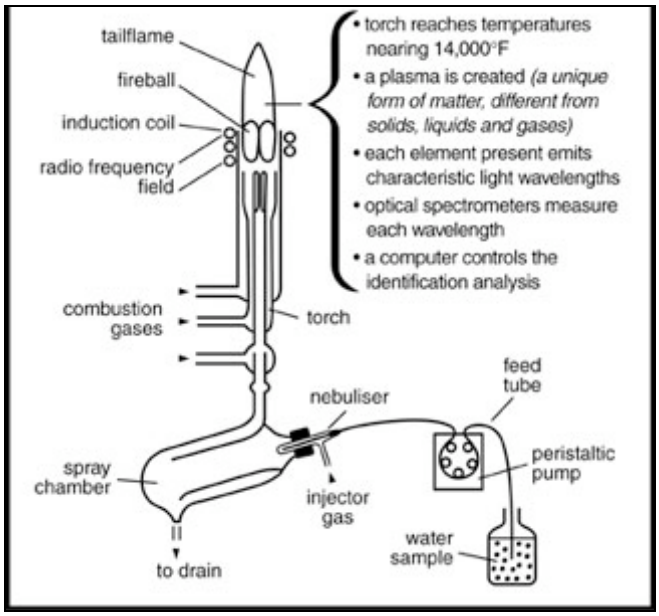


Figure 5.19 Diagram of Inductively coupled plasma – optical emission spectroscopy (ICP-OES)
Image from http://www.cleanwatertesting.com/news_NR149.htm

ICP-OES was performed on a Perkin-Elmer Optima 5330 ICP-OES machine at the Chemistry Department of the University of Edinburgh. Samples were purified and metal reconstituted as previously described. Standards for Iron and Manganese of 10 ppm, 5 ppm, 1 ppm, and 0.1 ppm were made to use for the standard curve. 60 μM native Sso2273 and Sso2273 reconstituted with Fe and Mn were analysed on the machine and compared to standard curves for Fe or Mn. Buffer that the protein was reconstituted in was used as a blank.

However, when the metal levels in the samples were compared to the metal levels in the blank, the difference was negligible, see Figure 5.20. Even for the iron reconstituted protein only 0.4 μM iron was detected, 60 μM Sso2273 was used in each sample, therefore if both metal binding sites in each monomer were occupied 120 μM of iron would be expected. These results were interpreted to mean that no iron metal had bound the protein. In the case of iron this could be because it is difficult to keep the iron in its ferrous state (state which it is predicted to bind the protein). Another possibility is that the binding of the metal to the protein is weak and is disrupted during the preparation of the sample for ICP-OES. These results do not

rule out the possibility that other metals are bound to the protein, as only wavelengths for iron and manganese were analysed.

Wavelength	Metal	Abs (minus blank)	Sso2273 Fe (Mg/L)		Wavelength	Metal	Abs (minus blank)	Sso2273 Mn (Mg/L)
238.204	Fe	1748	0.022		259.372	Mn	32	0.000
259.939	Fe	2105.5	0.022		260.568	Mn	23	0.000

Wavelength	Metal	Abs (minus blank)	Sso2273 native (Mg/L)
238.204	Fe	1380.7	0.017
259.939	Fe	1668.3	0.017
259.372	Mn	45.2	0.000
260.568	Mn	40.8	0.000

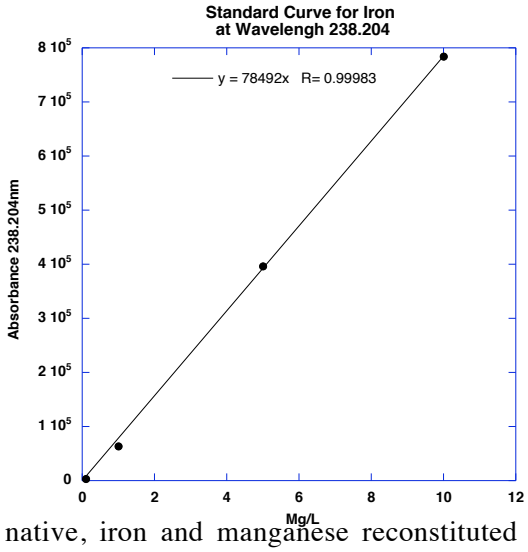


Figure 5.20 Values obtained from ICP-OES
The tables show the absorbance values for the native, iron and manganese reconstituted samples of Sso2273. The samples were read at two different wavelengths for each metal. The absorbance reading shown is minus the absorbance recorded for the blank, in this case the buffer the proteins were reconstituted in. The graph shows an example of the standard curve produced from solutions containing metals of known concentration. A standard curve was produced for each wavelength and the equation used to determine the Mg/L of metal for each sample. No manganese was detected in the samples and very low levels of iron suggesting the metal binding is extremely weak, if occurring at all. 0.022 Mg/L of iron in the iron-reconstituted sample equates to 0.4 μM iron in the sample. 60 μM of Sso2273 was used per sample so if two metals bound per monomer 120 μM of iron would be expected.

5.11 Sso2273 does not bind *dps* promoter DNA

5.11.1 Electrophoretic mobility shift assay, with and without divalent metals

The promoter region of the *dps* gene contains an 8 bases pair inverted repeat (see figure 5.2) that could be a binding site for Sso2273. A 45 base pair segment of this promoter was [γ^{32}P] ATP-labelled and used in band shift assays to determine whether Sso2273, either native or metal reconstituted, would bind to the promoter DNA. A scrambled version of the *dps* promoter (see appendix A for sequence), that contained the same bases but in an different order, abolishing the inverted repeat, was also [γ^{32}P]

ATP labelled and used as a control to establish whether any binding observed was specific to the *dps* promoter.

Band shift assays were first performed with the native form of Sso2273, no binding was observed (see Figure 5.21).

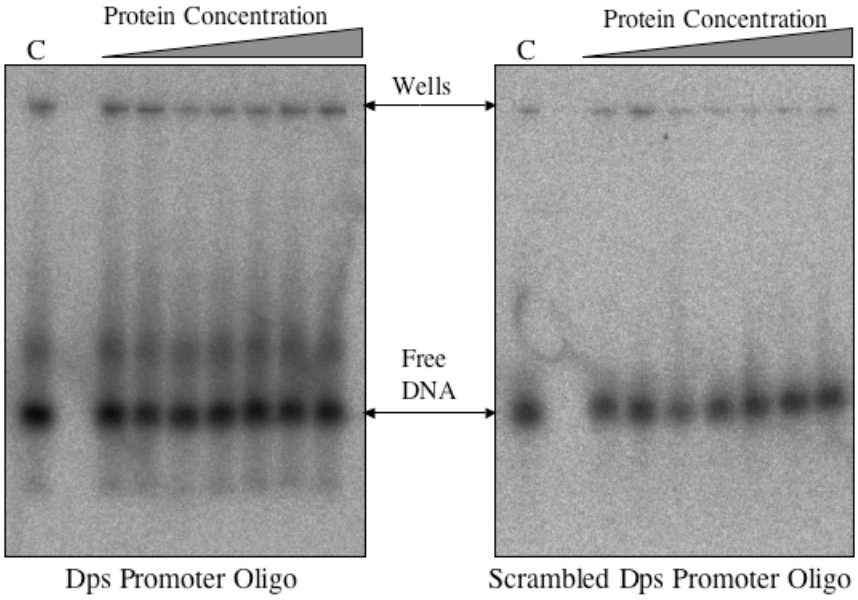


Figure 5.21 EMSA with *dps* promoter and scrambled *dps* promoter oligonucleotides and native Sso2273
EMSA run on a native 12% acrylamide gel. 10 nM [γ^{32}P] ATP labelled oligonucleotide (either *dps* promoter or Scrambled *dps* promoter) was incubated with increasing concentrations (1 nM, 10 nM, 50 nM, 250 nM, 1 μM, 5 μM, 20 μM) of native Sso2273. A control (C) containing no protein was also run.

Further band shift assays were performed with Sso2273 that had been reconstituted with a variety of divalent metals (see figure 5.22). A longer piece of DNA was used in an attempt to give the protein more opportunity to bind. A PCR product 329 base pairs long, that incorporated the 45 base pair fragment and the predicted transcription start site of Sso2273, was produced (see Appendix A for primers) and [γ^{32}P] ATP labelled. DNA binding was observed where Sso2273 had been reconstituted with iron sulphate and, to a lesser extent, with cobalt.

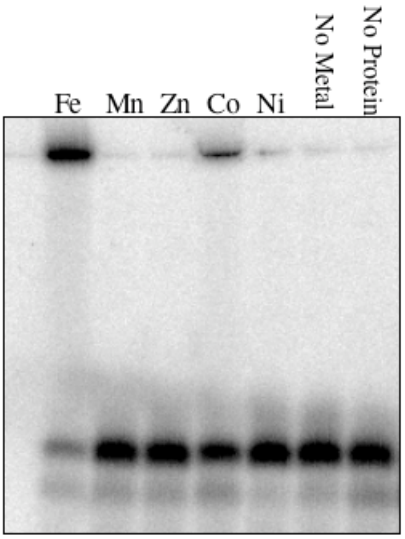


Figure 5.22 EMSA with *dps* promoter PCR product and metal reconstituted Sso2273
EMSA was run on a native 6 % acrylamide gel. 60 ng of [γ^{32}P] labelled *dp*s promoter PCR product was incubated with 1 μg Sso2273. The protein was reconstituted with $Fe(NH_4)2SO_4$, $MnCl_2$, $ZnCl_2$, $CoCl_2$, and $NiCl_2$

To confirm the band shift that was observed, assays were run with $Fe(NH_4)_2SO_4$, $FeCl_3$ and $CoCl_2$. Two different concentrations of Sso2273 were used to determine whether an increase in protein would produce an increase in binding to the *dps* promoter, see Figure 5.23.

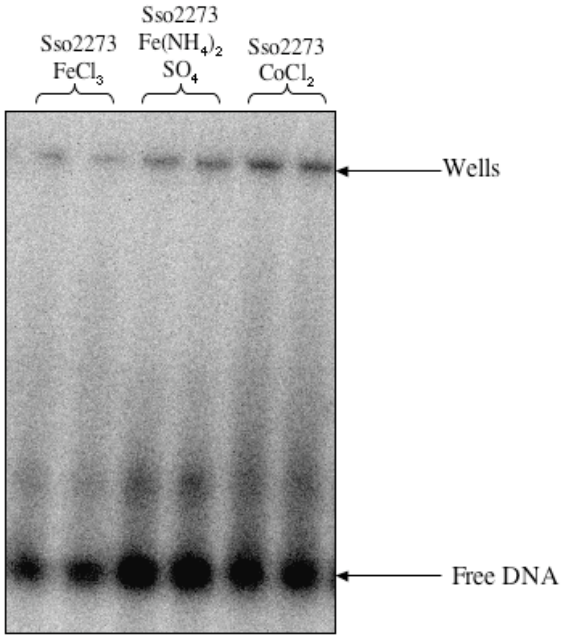


Figure 5.23 EMSA with *dps* promoter, Sso2273 and metal salts
EMSA run on an 8% native acrylamide gel. 30 nM of [γ^{32}P] labelled *dps* promoter oligonucleotide was incubated with 50 nM or 5 μM Sso2273 and 1 mM metal salt. For each set the concentration of Sso2273 is 50 nM in the first lane and 5 μM in the second.

A shift was observed in the lanes containing $Fe(NH_4)_2SO_4$ and $CoCl_2$. The shift was most prominent in the lanes containing $CoCl_2$. Further assays were performed with increasing concentration of Sso2273 reconstituted with $CoCl_2$. The *dps* promoter oligo

and the *dps* scrambled oligo were used as substrates to determine whether the binding was promoter specific (see Figure 5.24).

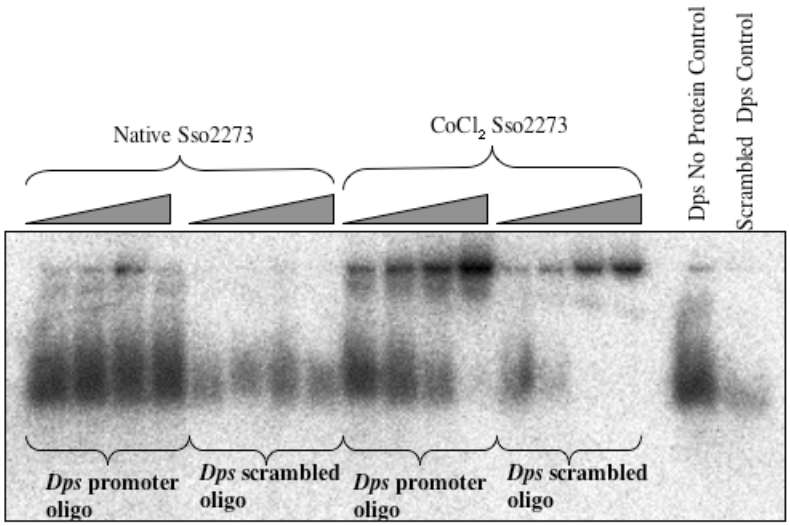


Figure 5.24 EMSA with *dps* promoter and scrambled *dps* promoter with native and cobalt reconstituted Sso2273
EMSA run on native 8 % acrylamide gel. 30 nM [γ^{32}P] labelled (either *dps* promoter or scrambled *dps* promoter) was incubated with increasing concentrations of Sso2273 (2.5 μM, 5 μM, 10 μM, 20 μM). Sso2273 was either native, or metal reconstituted with $CoCl_2$. Controls containing no protein were also run.

It is possibly that some binding was observed with the $CoCl_2$ reconstituted protein, however the DNA did not run out of the wells, making it inconclusive. The possible binding was observed with the *dps* and scrambled *dps* oligonucleotides so even if we accept that there was binding; it is of a non-specific nature.

A control containing DNA and iron only was run in the subsequent band shift experiments, which showed that the retardation of the DNA was due to the presence of the metal, not the presence of Sso2273 (see figure 5.25).

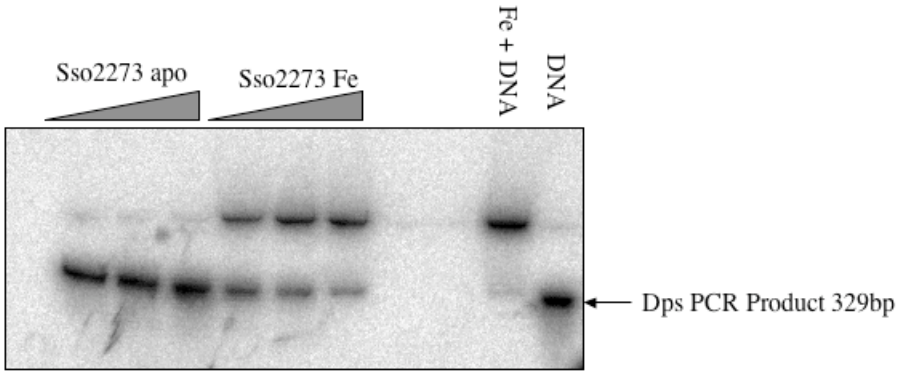


Figure 5.25 EMSA with native and iron reconstituted Sso2273 with *dps* promoter PCR product as substrate
EMSA run on native 6 % acrylamide gel. 60 ng [γ^{32}P] labelled *dps* promoter PCR product was incubated with increasing concentrations of Sso2273 (10 μM, 20 μM, 40 μM). Sso2273 was either native, or metal reconstituted with $Fe(NH_4)_2SO_4$. Controls containing 0.5 mM $Fe(NH_4)_2SO_4$+DNA and DNA alone were also run. The iron+DNA control shows that the band shift observed was due to the presence of iron and not to the presence of Sso2273.

Figure 5.25 shows that although the iron reconstituted Sso2273 appears to be causing a shift in the DNA the lane containing iron and DNA shows shifting to an equal degree.

Sso2273 expressed well and was purified by heat treatment, heparin column and gel filtration. The native form of the protein was assumed to have no metals bound in the metal binding site and Bathophenanthroline assays were used to confirm there was no iron present. In an attempt to obtain the metal bound conformation needed for DNA binding, the protein was reconstituted with a variety of divalent metals and ICP-OES was performed in an attempt to determine if the metal reconstitution was successful. The results indicated that the protein had not bound metal, in the case of iron this maybe due to the redox sensitivity of the metal and the difficulty in keeping it in its ferrous form.

Electrophoretic mobility shift assays were performed with the native and metal reconstituted forms of the protein, initially these assays suggested that Sso2273 could bind the *dps* promoter if reconstituted with cobalt or iron, however after further

controls were run it became apparent that the shift observed was due to the presence of the metal rather than the protein.

5.12 Sso2273 and Sso0669 primer extension transcription assays

It was known from previous experiments that under normal conditions (ie with the basal transcription apparatus TBP, TFB-1 and RNAP) transcription from the *dps* promoter was strong (work by Dr S.Paytubi see figure 5.1). If Sso2273 were in fact the repressor of the *dps* gene it would be expected that inhibition of transcription would be observed if Sso2273 was added to transcription assays with the *dps* promoter.

Sso0669 is a paralogue of the Sso2273 protein, and aligns well with other metal dependent repressor (see figure 5.4). It may have a similar function to Sso2273, and for that reason was added to transcription assays with a number of promoters to investigate its possible role as a transcriptional repressor.

5.12.1 Sso2273 does not inhibit transcription

Primer extension transcription assays were used to determine whether Sso2273 could inhibit transcription from the *dps* promoter. The *dps* promoter was pre-incubated with Sso2273 before the transcription proteins (TBP, TFB-1 and RNAP) were added; this was to give Sso2273 time to bind the promoter. The binding site of Sso2273 is not yet known, DtxR binds a 30 base pair region of the tox promoter which covers the –10 element of the *tox* gene (Schmitt, Twiddy *et al*. 1992) and the archaeal repressor MDR1 binds downstream of the TATA box, blocking recruitment of RNA polymerase (Bell, Cairns et al. 1999).

Transcription assays using a number of different promoters were performed (see figure 5.26). There appeared to be no inhibition with either the native, or metal reconstituted forms of Sso2273 from the *dps* promoter. The iron reconstituted Sso2273 does appear to inhibit the transcription from the T6 promoter, however this result was never reproduced and maybe due to errors during loading. Transcription

from the SSB promoter was poor so any effect of the addition of Sso2273 was hard to determine.

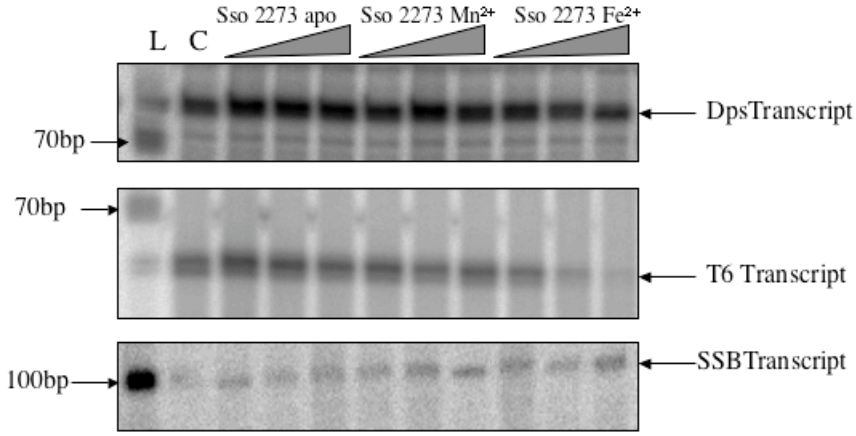


Figure 5.26 Transcription from *dps, T6* and *ssb* promoters after addition of Sso2273
Primer extension assays were performed using the *dps*, *T6* and *SSB* promoters to see if Sso2273 specifically inhibited transcription from the *dps* promoter. The transcription from the *dps* promoter appeared consistent in all conditions. Transcription from the *T6* promoter appears to be slightly inhibited by increasing concentrations of iron reconstituted Sso2273, however this result was never repeated and so the repression seen on this gel is put down to errors during loading. Transcription from the *SSB* promoter was very weak, so it was difficult to determine any differences.

5.12.2 Sso0669 inhibits transcription non-specifically

Transcription assays were also performed with Sso0669; the protein was first used in its native form (see figure 5.27), which showed inhibition of transcription with all the promoters tried.

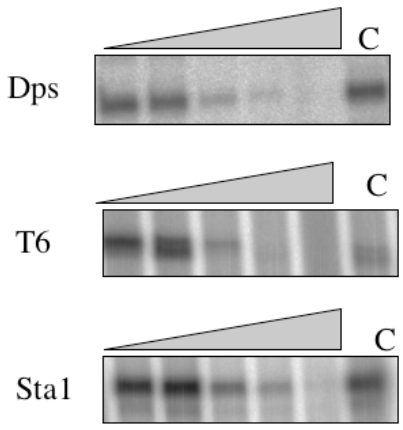


Figure 5.27 Transcription from *dps*, *T6* and *sta1* promoter is inhibited by increasing concentrations of Sso0669
Transcription from three different *S. solfataricus* promoters appears to be inhibited by increasing concentration of native Sso0669. Sso0669 at 0.5, 1, 5, 10 and 20μM was incubated with the promoter DNA for 10 mins at 65°C before the transcription proteins were added. C denotes the control where no Sso0669 was added.

Sso0669 was reconstituted with manganese (as it is a more redox stable metal than iron) and used in transcription assays to see if the reconstituted form of the protein increased the level of inhibition. No difference between the inhibition produced by the native and metal reconstituted forms was observed (see figure 5.28).

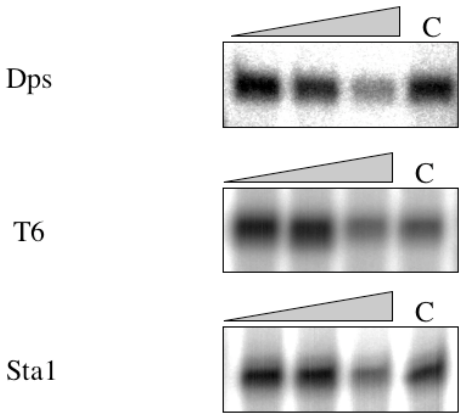


Figure 5.28 Inhibition of transcription from *dps*, *T6* and *sta1* promoters is not increased after metal reconstitution of Sso0669
Transcription from three different promoters appears to be inhibited by Sso0669 metal reconstituted with Mn. 0.5, 1 and 5μM of Sso0669 Mn were used and a control sample containing no Sso0669 was also included. There does not appear to be any increase in the degree of inhibition with the metal reconstituted protein and again the inhibition is non-specific.

Although the apparent inhibition of transcription by Sso0669 was an interesting result, due to the difficulty in obtaining pure protein, Sso0669 was not investigated further.

5.13 Chromatin immunoprecipition with Sso2273 antibodies

Chromatin immunoprecipitation is a technique used to determine the binding site of DNA binding proteins *in vivo* (Haring, Offermann *et al*. 2007). Formaldehyde is used to crosslink the proteins to the DNA *in vivo*, and the cells are then lysed by sonication. The *in vivo* nature of this technique negates the problem of the Sso2273 protein's redox sensitivity. Once the cells are lysed the DNA is fragmented by sonication, and immunoprecipitation is used to isolate the protein of interest, and the DNA it is bound to. The cross links are reversed, releasing the DNA from the protein. The DNA can then be used in PCR reactions with primers specific to genes, or regions, predicted to be bound by the protein of interest.

Sso2273 was predicted to bind to the promoter of genes *sso2078* and *sso2079* (*nramp* and *dps-like* respectively) so primers for the promoter region of these genes were designed. Work in *C. diphtheria* (Brune, Werner *et al*. 2006) showed that DtxR (the protein Sso2273 is homologous to) also binds proteins involved in iron metabolism,

so primers were designed for the promoter regions of the *Sulfolobus* equivalents of these genes, *sso2121* (a putative peroxiredoxin), *sso2244* (a putative ferric uptake regulator), *sso2642* (Rubrerythrin), and *sso0927* (an ABC type transporter). Primers were also designed for the promoter region of the *sso2273* gene to investigate whether the protein was self-regulated.

Antibodies for the Alba1 protein were used as a positive control; Alba1 acts as a chromatin protein in Archaea and so binds readily throughout the whole genome (Wardleworth, Russell *et al*. 2002; White and Bell 2002), and would provide a strong signal. Pre-immune antibodies were used for the negative controls.

ChIP was performed with hydrogen peroxide treated and control cultures of *Sulfolobus solfataricus*. The hypothesis being that Sso2273 would be bound to the *dps* promoter in the control culture, inhibiting transcription, and so a band would be visible after a PCR with the ChiP DNA and *dps* promoter primers. No band would be visible after amplification with the same primers with DNA from the hydrogen peroxide treated culture as, Sso2273 would have been oxidised and released from the promoter. However this did not appear to be the case (see Figure 5.29)

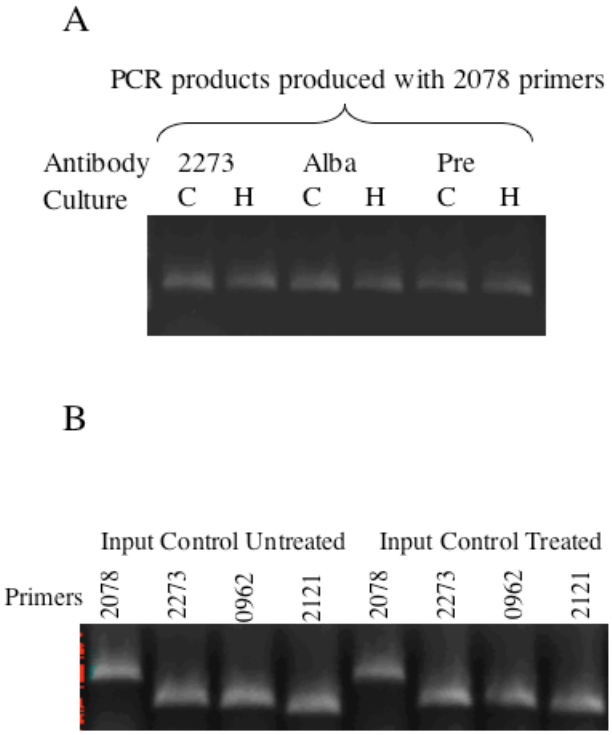


Figure 5.29 Chromatin immunoprecipitation with Sso2273 antibodies
A shows an agarose gel run with the PCR products produced with primers for the gene *sso2078* (*nramp*). The DNA had been immunoprecipitated with antibodies for Sso2273, Alba, or the pre-immune antibodies. Control (C) and Hydrogen peroxide treated cultures were used.
B shows an agarose gel containing PCR products produced from the input control DNA, ie the DNA before the immunoprecipitation.

The ChIP PCR produced bands of similar intensity for all samples, only products amplified with primers for the promoter region of the 2078 gene are show, but the gel is representative of those produced with the other primers.

The DNA isolated during ChIP was also used in real time PCR. Because of the sensitivity of real time PCR it was hoped that differences in the levels of DNA in the samples would be more apparent with this method. PCR products run on an agarose gel give an indication of the DNA in the sample once the PCR reaction has performed 30 cycles and, in many cases, reached an end point. With real time PCR the crossing point value indicate the point at which the DNA in the sample has crossed a threshold set by the machine using a calculation involving readings from the first 10 cycles of the PCR. This method is much more sensitive to the starting levels of DNA in the sample.

	2273 Antibodies Control Culture	2273 Antibodies H_2O_2 Treated Culture	Alba Antibodies Control Culture	Alba Antibodies H_2O_2 Treated Culture	Pre immune Antibodies Control Culture	Pre immune antibodies H_2O_2 Treated Culture
Ave CP 2078 primers	26.4	26	25	26.4	25.6	25.5
Ave CP 2273 primers	24.1	22.9	23.5	24.7	24.3	23
Ave CP 0121 primers	24.3	23.2	23.7	24.5	24.6	23.3

Table 5.2 Crossing point values from real time PCR of ChIP DNA
The DNA isolated from the ChIP was used in real time PCR. The table shows the average crossing point (CP) values for DNA from control and hydrogen peroxide treated cultures immunoprecipitated with Sso2273, Alba1 or pre immune antibodies. Real time PCR was performed with primers for the *nramp* gene (*sso2078*), *sso2273* and a *peroxiredoxin* gene (*sso2121*)

CP values were compared for the control and hydrogen peroxide treated samples, for each set of primers with the different antibodies (see Table 5.2). Very little difference was noted, the largest change in CP values was between the control and treated samples, immunoprecipitated with Alba when the DNA was amplified with *sso2273* primers. A difference of 1.2 cycles is observed and if we assume that for each cycle the DNA in the sample doubles, this indicates there was 2-fold more DNA for the *sso2273* gene in the treated sample compared to the control. This does not agree with what is expected if the hypothesis for Sso2273 is correct. It was hypothesised that Sso2273 would be bound to the promoter region in normal conditions and be released after oxidative stress. If this is true less DNA would be expected in the hydrogen peroxide treated samples as no protein would have been present on the DNA to allow it to be immunoprecipitated by protein specific antibodies.

The fact that the intensity of all of the bands and the CP values for all of the samples was the same suggested that the DNA being amplified was carried over during the process of the experiment, rather than DNA specifically isolated by immunoprecipitation.

5.14 Discussion

The promoter region of the *dps* gene contains an inverted repeat, these structures in DNA have been shown to be binding sites for regulatory proteins. MDR1 binds to a region of dyad symmetry in its own promoter (Bell, Cairns *et al*. 1999) and DtxR binds a palindromic sequence in the tox promoter (Pennella and Giedroc 2005). There were no sequences consistent with MDR-1 or DtxR binding sites found on the promoter region of the dps gene, however it was predicted that the inverted repeat present would be a likely binding site for the genes repressor. The sequence from the *dps* promoter including the inverted repeat was used in pull down assays with magnetic beads, to 'fish out' the *dps* gene repressor. Salt washes of the beads, and the beads themselves were analysed by SDS-PAGE, and a likely candidate, Sso2273, was identified by mass spectrometry. Sso2273 was determined to be a likely candidate because it was the top hit after mass spectrometry analysis of the beads and because of its homology to another transcriptional repressor, DtxR, from *C. diphtheria*.

Sso2273 was expressed and purified. The purified protein has a molecular weight of 15.01 kDa, suggesting it purifies as a monomer as the predicted molecular weight obtained from ProtPram was 15.10 kDa. In the homologue metal reconstitution is required before the correct DNA binding, dimeric form of the protein can be produced (Rangachari, Marin et al. 2005). Sso2273 was metal reconstituted with a variety of divalent metals, but as observed by analytical gel filtration metal reconstitution had little effect on the molecular weight suggesting that it was unsuccessful in producing the dimer form of the protein. The metal reconstituted form of the protein was also analysed by ICP-OES at the University of Edinburgh. The analysis showed no difference between the blank and the samples, suggesting that the metal reconstitution was not successful and metal was not present in the binding sites. The metal that binds preferentially in the homologue (DtxR) is ferrous iron (Spiering, Ringe *et al*. 2003). The redox sensitivity of ferrous iron and its propensity to convert to its oxidised, ferric form in normal (ie non-anaerobic) conditions, led to problems with the reconstitution and could explain the lack of metal being observed in the binding sites. It is also possible that, due to the weak interaction of the protein and the metal, the

metal binding was disrupted during sample preparation. However DtxR was successfully reconstituted with other more redox stable divalent metals such as manganese, it is unclear why this was not successful for Sso2273. The dimer form of the protein was crystallised with zinc bound in the metal binding sites, zinc reconstitution was tried but with no success.

Analysis of the crystal structure of Sso2273 (performed by Dr Stephen McMahon of the SSPF lab) shows that it is similar to proteins of the Diphtheria toxin repressor family. Sso2273 possessed strong structural similarity to two proteins of this family, MntR and IdeR, however there were significant differences in the Sso2273 metal ion binding sites, compared to those in MntR and IdeR. Neither of the metal ion binding sites of MntR resembles either of those found in Sso2273. Site 1 of IdeR and site 2 of Sso2273 did overlay well, and were only 1.8 Å apart, however Sso2273 does not possess the SH3 domain present in IdeR and two of the residues involved in metal ion binding are in this domain.

When the C-terminal domains of Sso2273 and IdeR are aligned the N-terminal DNA binding domains do not match up, this could have implications for Sso2273 ability to bind DNA, as the α2 helix of its DNA binding domain clashes with the DNA when the C-terminal domains are aligned, see Figure 5.15. There is also a clash between the DNA and the protein N-terminal helices when the N-terminal recognition helix of one of the Sso2273 monomers is aligned with the N-terminal recognition helix of IdeR, see Figure 5.16. It is possibly that the DNA binding form of the protein is not what has been crystallised and differences in the metal ion binding sites of Sso2273 compared to other proteins in the DtxR family could indicate that Sso2273 does not function in the same way as the other DtxR proteins. It is possible that instead of forming a homodimer containing two Sso2273 monomers the functional repressor actually contains a heterodimer containing Sso2273 and it paralogue, Sso0669. The possibility of dimerization between these two proteins should be investigated.

Electrophoretic mobility shift assays were performed with the *dps* promoter, Sso2273 and metals to see whether the protein would bind to the inverted repeat as predicted.

Initial assays appeared to show binding to the DNA when the Sso2273 had been reconstituted with either iron or cobalt. However after more controls were run, it became evident that the metal, and not the protein, was causing the retardation of the DNA. This may be due to the formation of DNA metal ion complexes called M-DNA which can affect the mobility of the DNA (Hartzell and McCord 2005).

No inhibition of transcription in *in vitro* transcription assays was observed with Sso2273. In contrast, Sso0669 showed inhibition of transcription from a number of promoters. This inhibition did not seem to be dependent on metal reconstitution, and the concentrations of protein needed to cause inhibition did not decrease after the protein was metal reconstituted.

Chromatin immunoprecipitation was performed in the hope that the redox sensitivity of the protein would not be an issue as any protein bound to the DNA is crosslinked *in vivo*, and oxidation of the protein would not result in its release from the DNA in this case. The hypothesis was that Sso2273 would be bound to the *dps* promoter in the control cultures, and a band would be visible after PCR with the immunoprecipitated DNA. In the hydrogen peroxide treated cultures, Sso2273 would be released from the promoter and no band would be visible. However, all bands appeared the same intensity, when the PCR products were run on an agarose gel. The DNA was also analysed by real time PCR, it was hoped that this more sensitive method would indicate changes that would not be evident by running the end point PCR on a gel. However the CP values of the treated and control samples was very close, which did not match with the hypothesis. Alternatively the presence of bands in all the samples could suggest the DNA being amplified was carried over through the assays rather than DNA that was specifically isolated by immunoprecipitation.

Due to the difficulties in binding the appropriate ferrous iron metal in the metal binding sites of Sso2273, the DNA binding form of the protein was not obtained. Carrying the experiments out in anaerobic conditions would remove the redox sensitivity problems and allow more effective investigation of the role of Sso2273. It

is also possible that Sso2273 is not in fact the repressor of the *dps* gene, and performs another function entirely in the cell. A strain of *S.solfataricus* missing the *sso2273* gene was produced and used in microarray studies in an attempt to elucidate Sso2273 role in the cell. This work is discussed in chapter 6.

Chapter 6

Construction and Characterization of the *sso2273* Knockout

6.1 Introduction

Techniques for knocking out genes in prokaryotes and eukaryotes have been around for decades, but only in the last few years has a reliable system for producing gene knockouts in hyperthermophilic archaea been made available (Schelert, Dixit et al. 2004; Albers and Driessen 2007; Berkner and Lipps 2008). One method uses a strain derived from *Sulfolobus solfataricus 98/2* called PBL2025 to produce strains with gene knockouts. This strain has a deletion from *SSO3004* to *SSO3050*, the region of the genome that contains the *lacS* gene, which means the strain cannot grow with lactose as its only sugar source; this fact is exploited during the selection process. The first step in creating a knockout is to insert the up and downstream flanking regions of the gene to be knocked out, into the plasmid see figure 6.1.

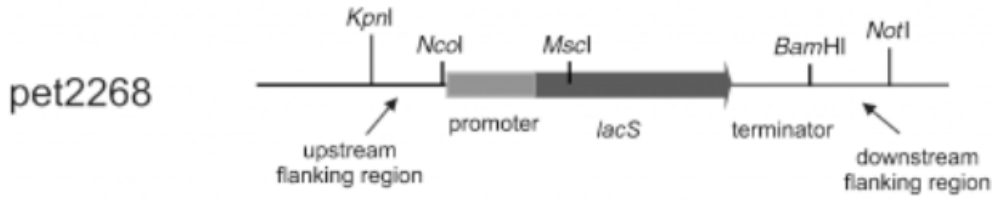


Figure 6.1 Diagram showing a region of the pET2268 plasmid
The pET2268 plasmid contains restriction digestion sites that allow insertion of the flanking regions (of the gene to be knocked out) into the plasmid. The flanking regions are inserted either side of the *lacS* gene on the plasmid (Albers and Driessen 2007).

The plasmid is then transformed into the PBL2025 cells where the homologous flanking regions line up and the gene is 'popped out' by homologous recombination.

It was hoped that this method could be used to make an *S. solfataricus* strain missing the *sso2273* gene which would allow the investigation of this genes role in the cell. Microarray analysis could then be utilised to gain a more genome wide view of the effect of the *sso2273* knockout. A diagram of the process of microarray analysis is shown in Figure 6.2.

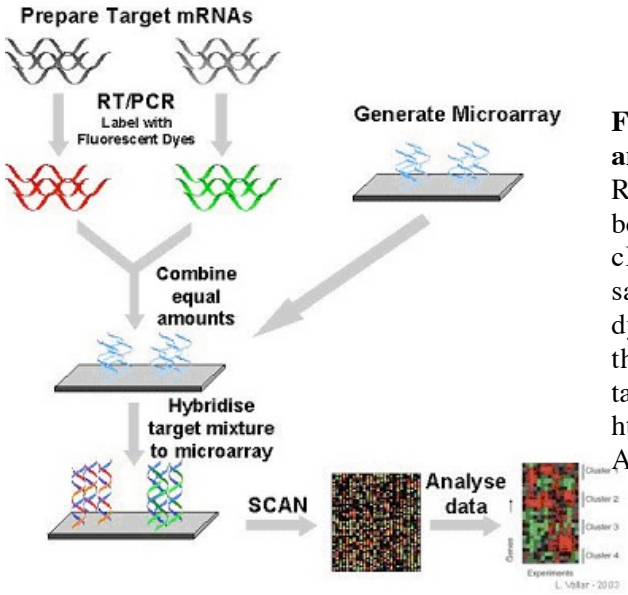


Figure 6.2 Process of microarray analysis
RNA is extracted from cultures of cells to be compared (ie damaged and control cells) cDNA is produced and the different samples labelled with different fluorescent dyes. The labelled cDNA is hybridised to the array and the plate scanned. Image taken from http://www.microarray.lu/en/MICROARRAY_Overview.html.

First RNA is extracted from growing culture, converted to cDNA and labelled with Cy dye. By labelling RNA from different conditions with different Cy dyes (ie RNA from the wild type strain with Cy3 and RNA from the knockout strain with Cy5) the differences in gene expression can be determined by comparing the signals for the two different dyes. The Cy dyes are read with a scanner and because different laser emission wavelengths are used for each dye (532 nm for Cy3 and 635 nm for Cy5) the samples can be easily distinguished, and background contamination is minimised. By comparing the samples a ratio is obtained and the level of up or down regulation determined for each gene.

6.2 Cloning and selection of *sso2273* knockouts

The system described above was utilised to make a *Sulfolobus solfataricus* PBL2025 strain missing the *S. solfataricus* P2 *sso2273* gene. By removing this gene it was hoped that its function could be better investigated and its role in the response to hydrogen peroxide damage and oxidative stress determined. It was predicted from previous results (see Chapter 5) that the Sso2273 protein was a transcriptional repressor, inhibiting the production of the Dps protein in normal conditions. Therefore, it was thought that in the knockout strain *dps* would be constitutively expressed at a high level. It was also predicted that the levels of the *nramp* gene would be affected in a similar way, as these two genes are on a divergent promoter.

Production of a knockout strain missing the *sso0669* gene was also attempted, but was unsuccessful.

The PBL2025 strain is a derivative of the *S. solfataricus* 98/2 strain, so the strain used for production of the knockout, and the strain the genes are from are different. The two strains were isolated from different geographical locations (P2 from Pisciarelli Solfatara in Italy and 98/2 from Yellowstone Park in America) but are predicted to have evolved from a common ancestor (Chae, Kim et al. 2007). There are obvious disadvantages in using the PBL2025 strain to knockout genes originating from the P2 strain, as differences in their genome may mean the flanking regions of some genes are not similar enough in sequence for homologous recombination to occur. This may explain why a knockout strain for *sso0669* could not be obtained.

The vector used to produce the knockout was pET2268 (see Figure 6.3), this plasmid contains the *lacS* gene, so successfully transformed cells were able to grow with lactose as a sugar source. A PCR product of the 1000 bases upstream of *sso2273* was inserted between restriction sites *Nco*I and *Kpn*I on the plasmid, and a PCR product of the 1000 bases downstream of *sso2273* was inserted between restriction sites *BamH*I and *Not*I. The resulting plasmid contains the *sso2273* up and down stream flanking regions either side of the *lacS* gene (see Figure 6.1 and 6.3)

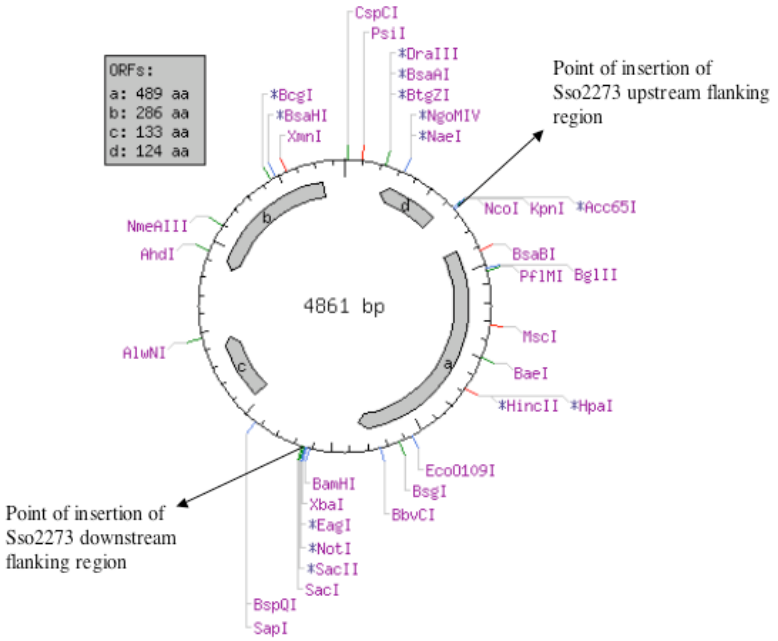


Figure 6.3 pET2268 vector
Illustration of the pET2268 vector, showing the restriction digestion sites and points where the up and down stream flanking regions of the *sso2273* gene were inserted. The grey section of the plasmid marked 'a' represents the *lacS* gene. Figure was produced using NEBcutter. http://tools.neb.com/NEBcutter2/index.php

The upstream flanking region was inserted first, and the plasmid was transformed into *E. coli*. The cells were grown up and the plasmid extracted with a Qiagen Mini Prep kit. The purified plasmid was then digested with *Kpn*I and *Nco*I (the restriction enzymes used to insert the plamid) and an agarose gel was run to check for the presence of the insert (see Figure 6.4).

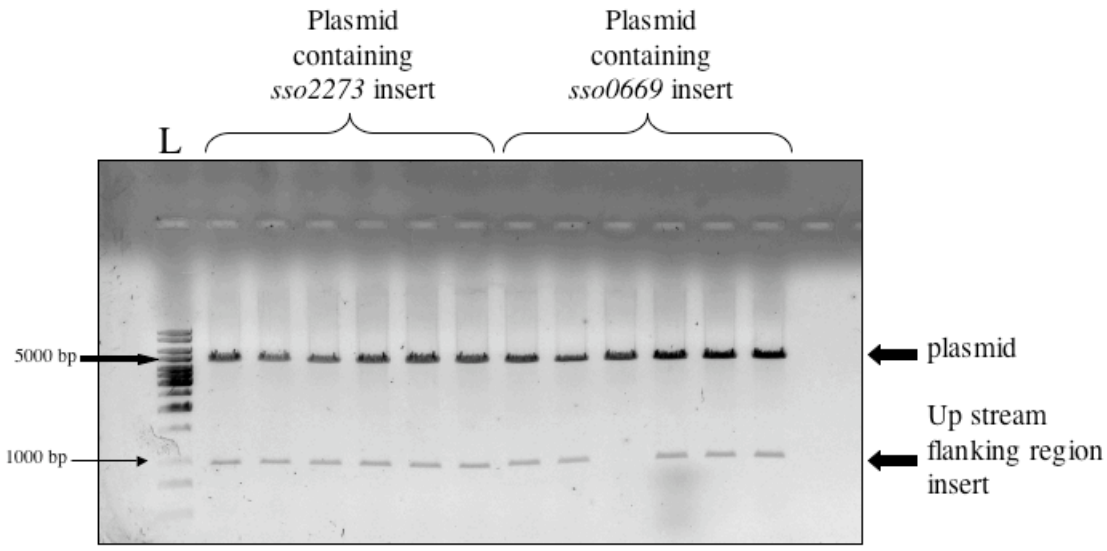


Figure 6.4 Agarose gel of restriction digestion of plamids from transformants
Transformants containing the pET2268 plamid were grown up and the plamid extracted using a Qiagen mini prep kit. The purified plasmid was then digested with *Kpn*I and *Nco*I and if the flanking region had been inserted correctly it would be 'popped out'

A sample of the plasmid was sent to the Dundee Sequencing Service to check the accuracy of the insert. The downstream flanking region was then inserted, and the process repeated. Once both flanking regions had been successfully inserted into the plasmid it was sent to Dr Sonja Albers at the University of Groningen where the rest of the knockout procedure was performed.

The plasmid containing the flanking regions was transformed, by electroporation, into the PBL2025 strain and the flanking regions line up with the complementary region in the genome, and the gene is 'popped out' by homologous recombination, and replaced with the *lacS* gene (Sonia-Verena Albers 2007).

After electroporation the cells were left to grow for 8-14 days in lactose media. Once growth was observed an aliquot of the culture was inoculated into new lactose media. Growth would be expected within a few days if the cells had been transformed successfully and were able to grow using lactose as a sugar source, cells from this culture were then plated and allowed to grow. Once colonies were visible, the plate was sprayed with X-gal and any single blue colonies picked.

6.3 Confirmation of knockouts

Once blue colonies were obtained, genomic DNA was extracted and PCR reactions performed to confirm the knockout had been successful (see Figure 6.5).

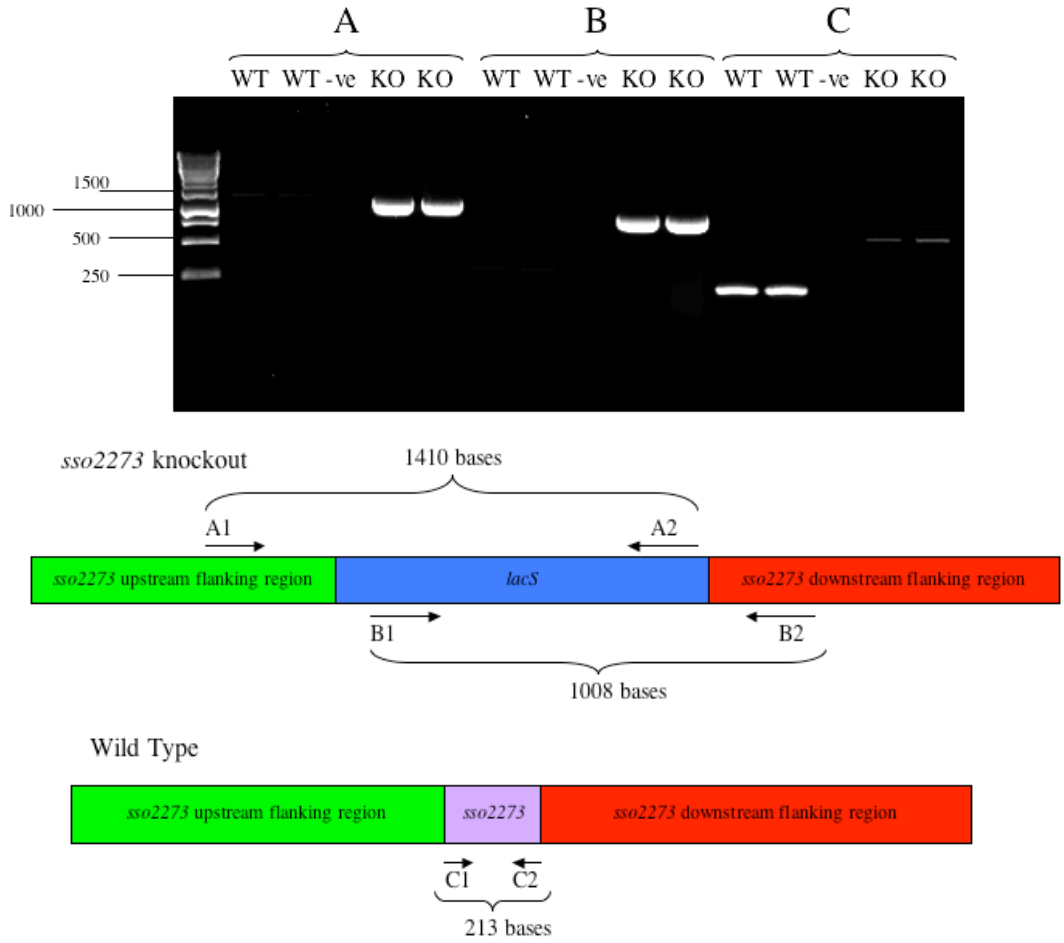


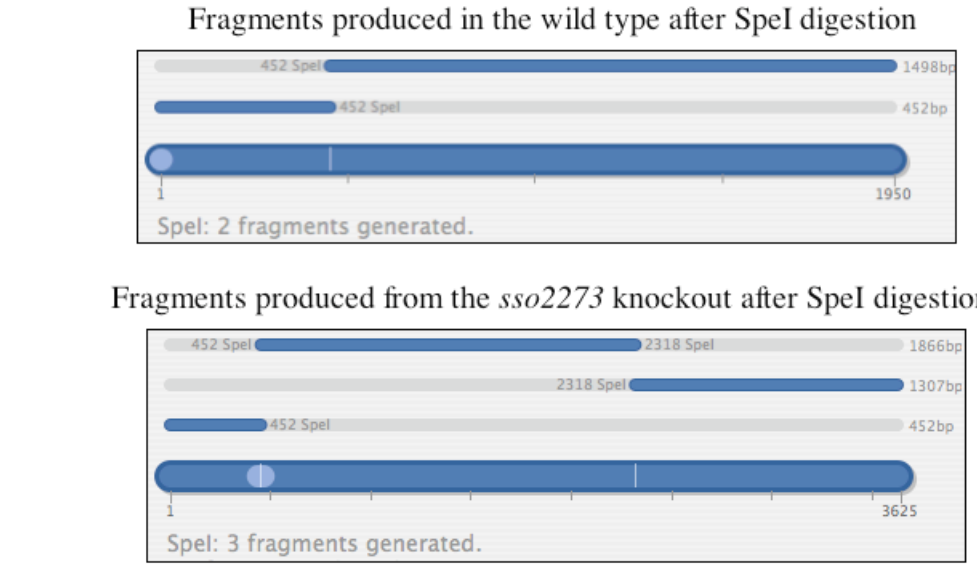
Figure 6.5 Agarose gel of PCR products confirming *sso2273* knockout
The gel shows the PCR products produced after amplification reactions were performed with genomic DNA from PBL2025 wild type and *sso2273* knockout strains. Primer set A used a forward primer specific to a sequence in the upstream flanking region of the *sso2273* gene (primer A1 in the diagram) and a reverse primer specific to a region of the *lacS* gene (primer A2 in the diagram). Primer set B used a forward primer specific to a region of the *lacS* gene (primer B1 in the diagram) and a reverse primer specific for the downstream flanking region of the *sso2273* gene (primer B2 in the diagram). Primer set C used forward and reverse primers specific to the *sso2273* gene (primers C1 and C2 in the diagram). Predicted sizes of the PCR products for each primer set are shown.

If the knockout was successful the *sso2273* gene would be 'popped out' by homologous recombination and replaced by the *lacS* gene. Primers specific for the *sso2273* flanking regions, *lacS* and *sso2273* genes were used to determine whether the *lacS* gene had been inserted. In PCR reactions using genomic DNA from the wild type strain no products were obtained when using the primer sets A and B that

contained primers specfic for the *lacS* gene suggesting the *lacS* gene was not present. Strong bands were obtained for the *sso2273* gene in the wildtype strain, showing this gene was still present.

In PCR reactions using genomic DNA from the *sso2273* knockout strain strong bands were obtained using primer sets A and B showing that the *lacS* gene was present in this strain, and the sizes of the PCR products also confirmed that the gene had inserted in the correct orientation. There was a faint band of around 500 bases produced when using the primers for *sso2273*, this maybe due to non-specific binding of the primers with another region of the genome.

As further confimation of the success of the knockout, PCR products were produced from both the wild type and knockout genomic DNA using primers specific for the upstream and downstream flanking regions of the *sso2273* gene (primers A1 and B2 in the digram in figure 6.5). These PCR products were run on a gel to check they were the predicted size and then purified out of the gel using a Qiagen gel purification kit. The cleaned PCR products were then digested with *Spe*I. The predicted fragment sizes had been determined using the EnzymesX program. The digests were run on an agarose gel, the fragement sizes for the wild type and knockout were different, and matched with the fragment sizes predicted by the EnzymesX program, see Figure 6.6.



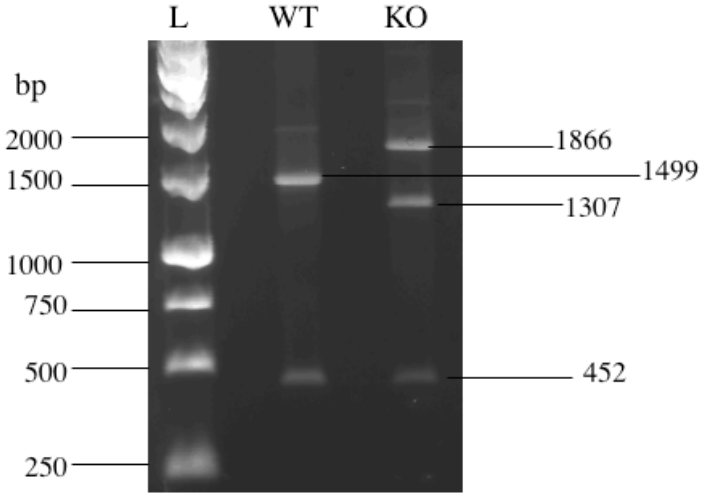


Figure 6.6 Restriction digestions of wild type and knockout PBL2025
Diagram showing the predicted fragment sizes produced after *Spe*I digestion of PCR fragments from wild type and knockout strains. An agarose gel showing the *Spe*I digested fragments; fragment sizes for the wild type and knockout were different and matched well with the sizes predicted by the EnzymeX program.

Ideally a southern blot showing the presence of the *sso2273* gene in the wild type strain and its absence in the knockout would have been produced as proof of the knockout. Attempts to obtain a southern blot showing this were unsuccessful, due to issues with detection, and so, despite the evidence above, the knockout has not been proven to be missing the *sso2273* gene.

6.4 Determining *sso2273* knockout phenotype

Once the *sso2273* knockout was obtained, its growth was compared to that of the wild type. The effect of iron availability and treatment with hydrogen peroxide, on the growth of the cell and gene expression of the *dps* and *nramp* genes was investigated.

6.4.1 Growth of PBL2025 wild type and *sso2273* knockout in normal conditions

Growth curves of the wild type and *sso2273* knockout strain in normal conditions shows the two strains grew the same initially, however in the later stages of growth the knockout grew more quickly overtaking the wild type, see Figure 6.7.

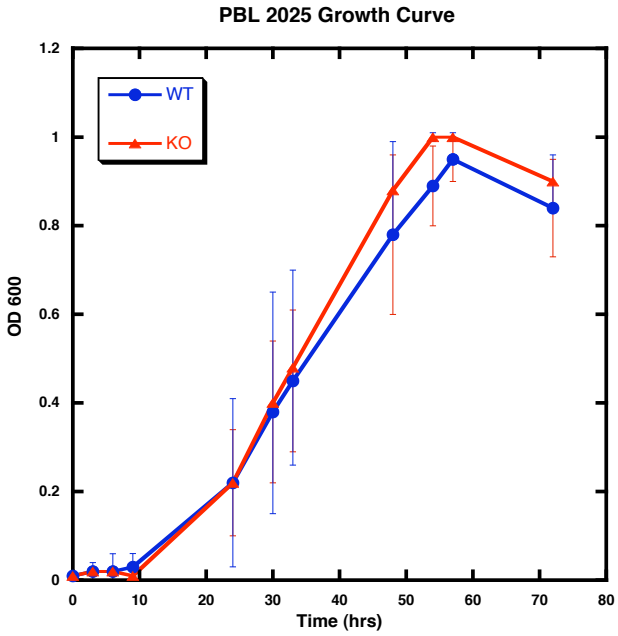


Figure 6.7 Growth curves of PBL2025 wild type and knockout strains in normal conditions
The strains grow the same in the initial stages. The knockout eventually reaches a higher OD_{600} than the wild type. The averages of triplicate experiments are plotted and standard errors shown.

6.4.2 Growth of PBL2025 wild type and *sso2273* knockout after hydrogen peroxide damage

Hydrogen peroxide damage causes a large increase in the transcription of the *dps* gene (see Chapter 4). This gene protects the cell from damage by sequestering hydrogen peroxide and free ferrous iron that can react together to produce extremely damaging hydroxyl radicals (Valko, Rhodes et al. 2006). If Sso2273 were the repressor of the *dps* gene, transcription of the gene in the *sso2273* knockout would be expected to be constitutively high. Due to the constant availability of the Dps protein, the knockout may be more proficient at surviving damage by hydrogen peroxide. The wild type and *sso2273* knockout were treated with different concentrations of hydrogen peroxide to determine what affect it had on the growth of the two strains (see Figure 6.8).

Treatment with 100 μM hydrogen peroxide had a very damaging effect on both strains and no growth was observed. The growth of the control and treated cultures, after addition of 10 μM hydrogen peroxide, was similar initially but the knockout

strain grew to a higher final OD_{600}. In the non-treated cultures, again the two strains grew the same initially however this time the wildtype strain reached the higher final OD_{600}. This could mean that the knockout copes better after hydrogen peroxide damage.

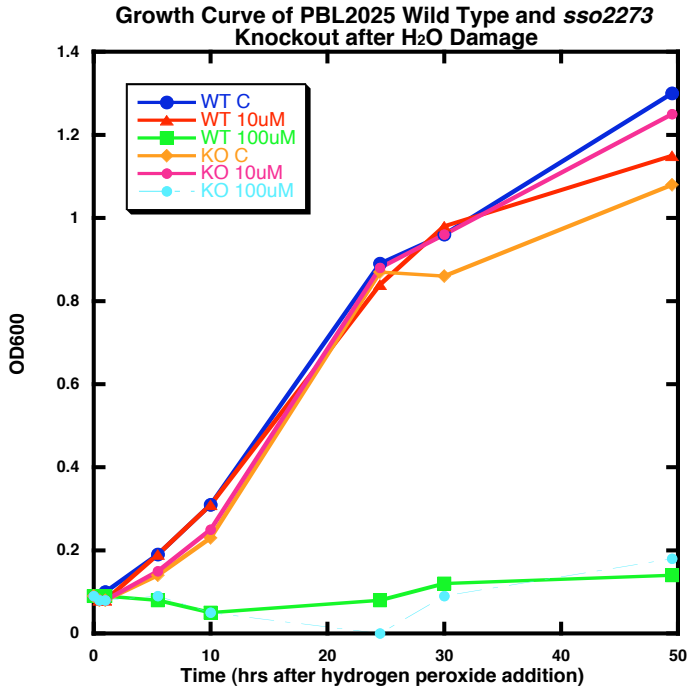


Figure 6.8 Growth curves of PBL2025 wild type and knockout strains after addition of different concentrations of Hydrogen peroxide
No growth is observed in either strain after addition of 100 µM hydrogen peroxide. In the control and 10 µM hydrogen peroxide treated cultures both strains grew the same initially, with the wild type reaching the highest final OD_{600} in the control conditions and the knockout reaching the highest final OD_{600} in the treated condition.

However in subsequent experiments using 5 µM hydrogen peroxide the knockout strain grew to the higher final OD_{600} in both control and treated condition, although the difference in the final OD_{600} of all strains was minimal, see Figure 6.9.

In all subsequent experiment 5 µM hydrogen peroxide was added to the treated cultures as this level of damage showed some difference in the growth of the cells, indicating an effect of the hydrogen peroxide but no retardation of growth. The addition of 1 µM hydrogen peroxide (the concentration used previously see Chapter 4) had no effect on the growth of the cells. The experiments described in Chapter 4 were performed with *S. solfataricus* P2, the PBL2025 strain used in the experiments described here appears to be less sensitive to hydrogen peroxide.

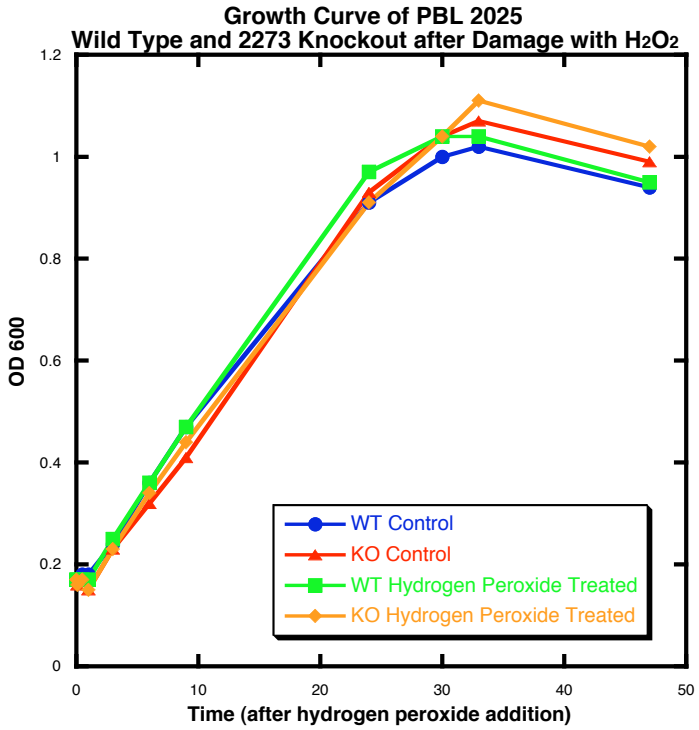


Figure 6.9 Growth curves of PBL2025 wild type and knockout strains after addition of 5 μM Hydrogen peroxide
Treated and control cultures for both strains grew the same initially, with the treated cultures growing to a higher final OD_{600} than the control in both the wild type and knockout strains

6.4.3 Real time PCR to study changes in gene expression after hydrogen peroxide damage

Sso2273 was predicted to be the repressor for the Dps protein, and consequently the repressor for the NRAMP protein as these genes are on a divergent promoter. By knocking out the repressor it was predicted that the transcription of the *dps* and *nramp* genes would remain at a constitutively high level.

Cells were grown to an OD_{600} of 0.1 before addition of 5 μM hydrogen peroxide. Samples were taken for RNA extraction 10 and 30 minutes after the addition of the hydrogen peroxide. The RNA extracted was then used in real time PCR to determine any changes in expression of the *dps* and *nramp* genes, in the wild type and knockout strains, see Figure 6.10.

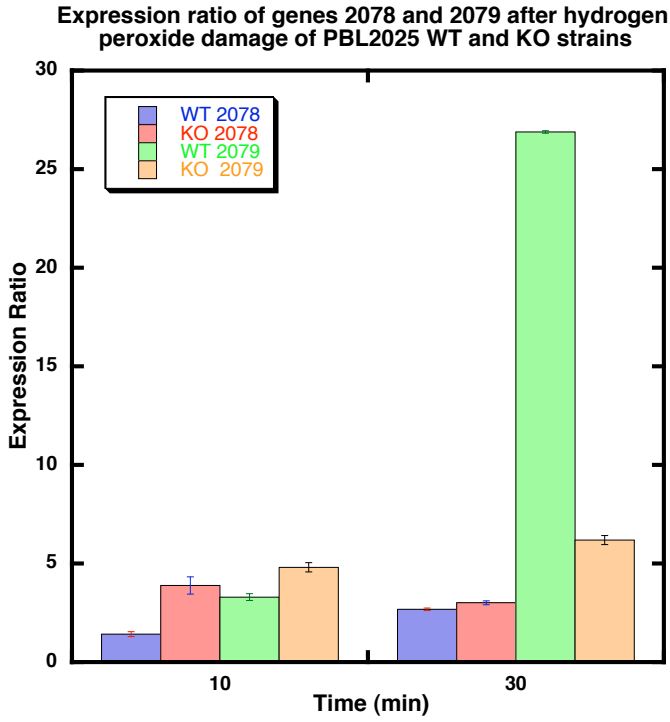


Table of Real Time PCR Crossing Point Values for *sso2078* and *sso2079*

Strain/ Gene	Average CP 10 min	Average CP 30 min
Wild type control *sso2078*	17.3	18.0
Wild type treated *sso2078*	16.8	16.0
Knockout control *sso2078*	18.8	19.4
Knockout treated *sso2078*	16.8	17.8
Wild Type control *sso2079*	19.8	20.4
Wild type treated *sso2079*	18.0	15.7
Knockout control *sso2079*	21.0	21.5
Knockout treated *sso2079*	18.7	18.8

Figure 6.10 Graph showing expression changes in *sso2078* and *sso2079* after 5 µM hydrogen peroxide damage and table showing crossing point values used to determine expression ratios
In the wild type expression of *sso2078* went up slightly and *sso2079* went up significantly after hydrogen peroxide damage. In the knockout there was a slight decrease in the expression of *sso2078* and a slight increase in the expression of *sso2079*. Average CPs are from triplicate experiments and standard errors are shown.

As has been shown previously (see Chapter 4) expression of the *dps* gene increases after hydrogen peroxide damage. The increase was larger in the wild type that in the *sso2273* knockout. The expression of these genes in the knockout was expected to be unchanged as they would be constitutively expressed when the repressor was absent, however the CP values were not significantly lower in the knockout compared to the wild type as would be expected if the genes were being continuously expressed.

6.4.4 Growth of PBL 2025 wild type and *sso2273* knockout in different Iron concentrations

Sso2273 is predicted to be a metal dependent repressor, similar to the DtxR protein from *C. diphtheria*. In *C. diphtheria*, transcription of the *tox* gene, that the DtxR protein represses, is regulated by iron levels in the growth media and its expression is highest in low iron conditions (Schmitt, Twiddy et al. 1992). The wild type and knockout strains were grown in culture with different iron concentrations to see what effect this had on the growth of the organisms, see Figure 6.11. RNA was also extracted from these cultures and used in real time PCR reactions with primers for the *dps* and *nramp* genes to determine whether expression of these genes was affected by the iron availability in the growth media.

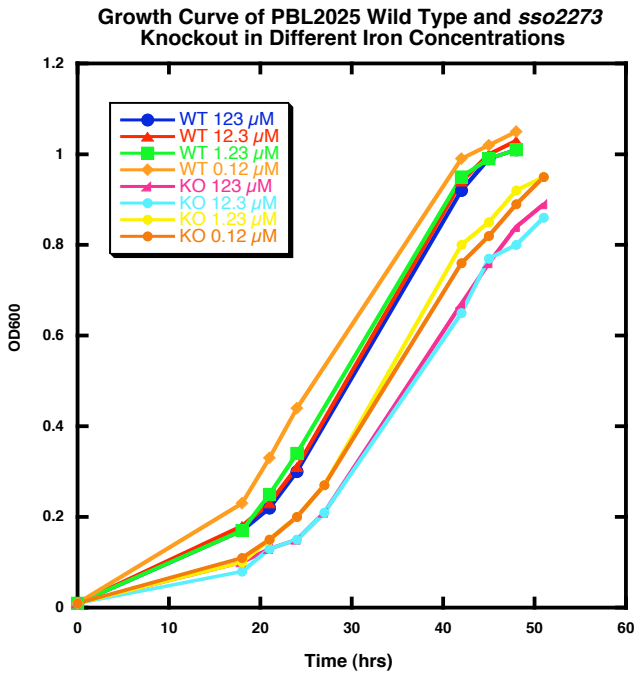
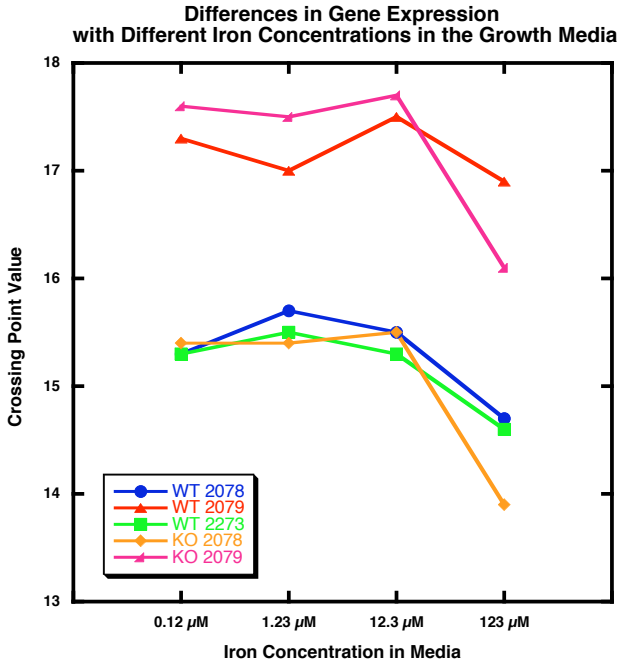


Figure 6.11 Graph showing growth curves of PBL2025 wild type and *sso2273* knockout in media with different Iron concentrations
The wild type strain grew more quickly than the knockout but over all the concentration of iron in the media had little effect on the growth of either strain.

Iron is essential for growth, but if levels are too high it is toxic, and if any more than double the normal amount of iron was added to the media the iron began to precipitate out, making OD_{600} readings inaccurate. Growth in the lower concentration iron media appeared to be slightly quicker for both the wild type and the knockout strains.

6.4.5 Real time PCR to study the effect of Iron availability on gene expression

RNA was also extracted from these culture and Real time PCR used to investigate the effect different iron conditions had on the expression of the *dps* and *nramp* genes, see Figure 6.12.



Gene / Strain	CP 0.12 μM (-30x Fe)	CP 1.23 μM (-20xFe)	CP 12.3 μM (-10x Fe)	CP 123 μM (Normal Fe)
2078 WT	15.3	15.7	15.5	14.7
2078 KO	15.4	15.3	15.5	13.9
2079 WT	17.2	17.0	17.5	16.9
2079 KO	17.6	17.4	17.6	16.0
2273 WT	15.3	15.5	15.2	14.5

Figure 6.12 Graph showing changes in gene expression in the PBL2025 wild type and *sso2273* knockout strains with different Iron concentrations in the media and table of the CP values.
Graph showing the changes in gene expression, of genes *sso2078, sso2079* and *sso2273* in the wild type strain and *sso2078* and *sso2079* in the *sso2273* knockout strains when cells are grown in media containing different iron concentrations. The x-axis shows the CP values, the lower the CP value the higher the gene expression.

There is an increase in expression of the three genes in the wild type as iron concentrations increase. This is in contrast with data from Mark Young's lab which showed, using western blot analysis, that Dps expression increases under iron-limiting growth conditions (Wiedenheft, Mosolf et al. 2005). In the homologue the expression of the gene that DtxR represses is highest in low iron conditions, Sso2273 is predicted

to repress the *dps* and *nramp* genes, if Sso2273 acts like DtxR we would predict that the expression of the *dps* and *nramp* genes would increase in low iron media, however the opposite is true. The transcription of another metal dependent repressor (MDR1) in the Archaeal species *Archaeoglobus fulgidus* has been shown to increase when metal ions are chelated out of the growth media by EDTA (Bell, Cairns et al. 1999). The results shown here do not follow what has been previously observed. Expression of the *sso2078* and *sso2079* genes appears to be induced by higher iron concentrations. This maybe in response to the damaging effects higher iron levels have on the cell, rather than the effect the iron levels are having on the repressor.

6.5 Microarray analysis of PBL2025 wild type and *sso2273* knockout

Microarrays containing the whole genome of *S. solfataricus* P2 are used routinely in the laboratory of Professor Rolf Bernander. Working in collaboration with Prof Bernander, microarray analysis was used to investigate the effect deletion of *sso2273* had on the PBL2025 strain. The microarrays contain genes from the *Sulfolobus solfataricus* P2 strain, and the knockouts were produced in the PBL2025 strain. As mentioned previously this could cause problems with complementarity if gene sequences from the two strains are very different. The spots on the array contain nucleotide sequences between 100-1000 bases (Andersson, Bernander et al. 2005), because of the length of the sequences some differences in gene sequence between the two strains will be tolerated.

The technique was used to compare the cDNA from hydrogen peroxide treated cells to cDNA from control cells for the wild type and knockout strains. This allowed analysis of the changes in gene expression after hydrogen peroxide damage. cDNA from the wild type and knockout strains were also compared, in control and hydrogen peroxide treated conditions. Due to time constraints only two sets of biological replicates was performed for the hydrogen peroxide versus control experiments, and only technical replicates were performed for the knockout versus wildtype experiments.

The RNA extraction, cDNA synthesis and Cy dye labelling was performed at the University of St Andrews. The labelled cDNA was then taken to the University of Uppsala where the hybridization and scanning of the plates was performed.

For each experiment dye swaps were performed to account for any gene specific dye biased. The chances of bias in this case are reduced as indirect labelling was used to attach the dye, however dye swaps were performed as a control. Once the arrays were scanned they were checked for any inferior spots (spots that had not been printed properly) and any dust or marks on the plate. These discrepancies were then removed from the analysis.

Further analysis of the arrays involved inputting all data into The Linnaeus Centre for Bioinformatics (LCB) Laboratory Information Management System (LIMS). The data was then transferred to the LCB Data Warehouse where analysis of the microarray data can be performed. Normalization within the array and merging of replicates and dye swaps was performed.

6.5.1 Microarray results for Hydrogen peroxide treated versus control cultures from wild type and knockout strains

By comparing the hydrogen peroxide treated cultures to the control cultures, genes induced by oxidative stress could be determined. Cultures of the wild type and *sso2273* knockout strains were grown to an OD_{600} of 0.1, each of the cultures were split into 2 pre-warmed flasks, 5 μM of hydrogen peroxide was added to one and the other was kept as a control. 20 ml samples were taken at 10 and 30 minutes after damage. RNA was extracted, converted to cDNA and labelled with either Cy3 or Cy5 dye. Arrays were hybridised and scanned and once the analysis had been performed a list of the 20 most up regulated and 20 most down regulated genes was compiled see Tables 6.1 – 6.4, colour coding of expression ratios is as follows; Red >3, Orange 2.5 – 3, Light orange 2 - 2.5, Yellow 1 – 2, Green <1. In the left hand column genes in the same operon are shown in the same colour.

Sso	Description	Function	Ratio 10 min	Ratio 30 min
2079	Conserved Hypothetical	DPSL-type antioxidant enzyme	21.78	21.02
2078	Putative NRAMP family protein	Mn2+ and Fe2+ Transporter	7.02	2.79
2080	Archaeal Rieske-type ferredoxin	Energy metabolism	4.22	1.96
2121	Bcp-2	Peroxiredoxin, bacterioferritin	4.04	14.96
2643	Glycerol-3-phosphate dehydrogenase chain c	Energy metabolism	3.09	3.51
2644	Hypothetical Protein	?	2.98	3.38
2568	Membrane conserved hypothetical protein	Cell Envelope	2.82	1.76
2847	Sugar-binding periplasmic protein	Transport	2.57	0.79
0761	mRNA 3'end processing factor	Transcription	2.48	1.02
0405	Proliferating cell nuclear antigen (PCNA)	Replication and repair	2.34	0.88
1399	Conserved Hypothetical	Uncharacterized predicted to be involve in DNA repair	2.24	0.91
2642	Rubrerythrin	Energy Metabolism	2.15	1.87
2110	Hypothetical Protein	?	2.13	1.49
2087	Carboxymethylenebutenolidase putative	Amino acid biosynthesis	2.10	0.94
0527	Phosphoglycerate kinase putative	Energy metabolism	2.02	0.89
2815	Conserved Hypothetical	Zn dependent hydrolase	1.97	0.78
2515	NAD-dependent malic enzyme (malate oxidoreductase)	Energy metabolism	1.97	1.05
2869	2-oxoacid-ferredoxin oxidoreducatase alpha chain	Energy metabolism	1.97	0.98
2539	Hypothetical	?	1.94	0.92
1668	HuyA-like	Amino acid Biosynthesis	1.94	1.02

Table 6.1 Microarray data showing genes up regulated after Hydrogen peroxide damage in PBL2025 wild type
Table shows the top 20 up regulated genes 10 minutes after hydrogen peroxide was added, the ratios for these genes 30 minutes after addition is also shown. Ratios were obtained by comparing values from hydrogen peroxide treated and control samples.

The top 20 up regulated genes in the wild type after hydrogen peroxide damage are shown in table 6.1. The three most up regulated genes in the wild type after hydrogen peroxide damage were the *dps* and *nramp* genes and a Rieske-type ferredoxin, these genes are within the same operon, and are involved in protecting the DNA from ROS after oxidative stress, metal iron transport and electron transfer respectively. This

corroborates the up regulation of the *dps* gene already observed using real time PCR (see Chapter 4).

Another operon that was up regulated (*sso2642 – sso2644*) contained genes involved in energy metabolism and a hypothetical protein. One of the genes in the operon (*sso2644*) codes for a rubrerythrin protein, these protein are known to be involved in the oxidative stress response in anaerobic bacteria and some Archaea (Sztukowska, Bugno et al. 2002; Weinberg, Jenney et al. 2004). A peroxiredoxin gene was also up regulated, peroxiredoxins protect the DNA by removing hydrogen peroxide (Limauro, Pedone et al. 2008). Microarrays looking at the effects of hydrogen peroxide damage in *S. solfataricus* P2 have been performed in Professor Mark Young's laboratory and the genes up regulated in their arrays matched with this data (unpublished observation), suggesting the issue of genetic similarity between the P2 and PBL2025 strains was not barrier to the microarray analysis.

Sso	Description	Function	Ratio 10 min	Ratio 30 min
9180	7kD DNA binding protein (Sso7D)	DNA binding	0.23	2.41
9535	7kD DNA binding protein (Sso7D)	DNA binding	0.29	2.02
0637	Hypothetical Protein	?	0.30	1.25
2827	Conserved Hypothetical Protein	Predicted transcriptional regulator	0.31	1.88
0368	Thioredoxin	Energy metabolism	0.32	1.34
0079	Bacterial-like DNA primase	Replication and repair	0.32	1.05
3246	Hypothetical Protein	?	0.35	2.09
0454	Hypothetical Protein	?	0.35	1.30
0961	Conserved hypothetical	?	0.36	1.41
0425	SSU ribosomal protein	Translation	0.36	1.91
0284	Hypothetical Protein	?	0.39	1.31
6418	LSU ribosomal protein L37AE	Translation	0.40	1.37
6397	LSU ribosomal protein L29AB	Transation	0.40	1.22
5559	Hypothetical Protein	Ubiquitin-like protein	0.41	1.16
5098	Terminal oxidase, small hydrophobic unit	Energy metabolism	0.41	1.13
2712	Hypothetical Protein	ABC-type sugar transport system	0.42	1.05
0567	ATP synthase subunit K	Energy metabolism	0.42	1.27
0451	Conserved hypothetical	Implicated in secretion	0.42	1.02
6817	LSU ribosomal protein S30E	Translation	0.42	1.48
1100	Hypothetical Protein	?	0.43	1.25

Table 6.2 Microarrray data showing genes down regulated after Hydrogen peroxide damage in PBL2025 wild type
Table shows the genes that were most down regulated 10 minutes after hydrogen peroxide treatment, the ratios for these genes 30 minutes after treatment is also shown. Ratios were obtained by comparing values from hydrogen peroxide treated and control samples.

The top 20 down regulated genes in the wild type after hydrogen peroxide damage are shown in Table 6.2. The two most down regulated genes after hydrogen peroxide damage were *Sso7D* genes. These genes code for small (7kDa), highly abundant, monomeric, DNA binding proteins. They have a small hydrophobic core and many solvent exposed hydrophobic residues and bind non-specifically to DNA (Guagliardi, Cerchia et al. 2002). It has been suggested that these proteins function as chromatin proteins in *Sulfolobus*, in the place of histone proteins that are not found in the Crenarchaea. They promote annealing of complementary DNA, induce negative supercoiling and act as chaperones facilitating the disassembly and renaturation of protein aggregates (Napoli, Zivanovic et al. 2002). These genes may have been down regulated in response to damage due to their role in DNA packaging; transcription of

genes to protect against and repair the oxidative stress would be needed and negative supercoiling would have to be repressed to allow their transcription. They have also been shown to be down regulated after UV damage (Gotz, Paytubi et al. 2007).

Transcription of a thioredoxin gene was also down regulated, this seems strange as this protein is an antioxidant and would be expected to be up regulated to help redress the imbalance between ROS and antioxidants that causes oxidative stress. Transcription of this gene did increase at the 30-minute time point however, so it is possible that it is induced later in the damage response. A predicted transcriptional repressor (Sso2827) was also down regulated. This protein contains a winged helix-turn-helix motifs, proteins containing this motif often function in the response to antibiotics, oxidative stress, virulence factors and regulation of aromatic catabolic pathways (Wilkinson and Grove 2006). A number of larger subunit (LSU) ribosomal proteins were also down regulated suggesting a general down regulation in translation.

Sso	Description	Function	Ratio 10 min	Ratio 30 min
2079	Conserved Hypothetical	DPSL-type antioxidant enzyme	25.86	29.3
2121	Bcp-2	Peroxiredoxin, bacterioferritin	19.28	14.5
2078	Putative NRAMP family protein	Mn2+ and Fe2+ Transporter	14.40	5.46
2644	Hypothetical Protein	?	5.86	3.97
2080	Archaeal Rieske-type ferredoxin	Energy metabolism	4.36	2.54
2643	Glycerol-3-phosphate dehydrogenase chain c	Energy metabolism	4.23	3.99
2244	Ferric uptake regulation protein	Transport	3.28	2.61
2568	Membrane conserved hypothetical protein	Cell Envelope	2.97	2.72
0757	Spermidine Synthase	Central intermediary metabolism	2.74	0.96
0515	Transposase	-	2.43	1.03
0390	Hypothetical protein	?	2.26	0.90
2645	Hypothetical protein	?	2.16	2.07
3246	Hypothetical protein	?	2.13	0.95
0368	Thioredoxin	Energy metabolism	2.10	0.85
2115	Hypothetical protein	?	2.05	1.12
0961	Conserved Hypothetical Protein	?	2.00	0.94
2261	Sulfide-quinone reductase related protein	Energy metabolism	1.89	2.22
1740	Hypothetical protein	?	1.87	1.11
2665	Muconate cycloisomerase related protein	Central intermediary metabolism	1.85	1.06
2243	Purine nucleosidase putative	Nucleotide transport and metabolism	1.83	1.44

Table 6.3 Microarray data showing genes up regulated after Hydrogen peroxide damage in PBL2025 *sso2273* knockout
Table shows the 20 most up regulated genes 10 minutes after hydrogen peroxide treatment, the ratios for these genes 30 minutes after treatment is also shown. Ratios were obtained by comparing values from hydrogen peroxide treated and control samples.

The top 20 up regulated genes in the knockout after hydrogen peroxide damage are shown in Table 6.3. The list of genes up regulated in the *sso2273* knockout was very similar to the list for the wild type. A ferric uptake regulator (*sso2244*) was up regulated in the knockout to a greater degree than in the wild type. There was early up regulation of a thioredoxin (antioxidant) gene, however this gene as then down regulated at the 30 minute time point. This gene was in the top 20 down regulated list for the wild type, which is unexpected as antioxidant proteins help rectify the

imbalance cause by oxidative stress and would be expected to be up regulated, the expression of this gene did increase in the wild type at the 30 minute time point.

Sso	Description	Function	Ratio 10 min	Ratio 30 min
10224	Hypothetical protein	?	0.42	0.99
0302	Chorismate mutase/ prehenate dehydratase	Amino acid biosynthesis	0.49	0.95
0397	Proliferating cell nuclear antigen (PCNA)	Replication and repair	0.54	0.78
2199	Hypothetical protein	Predicted DNA replication, recombination and repair	0.54	1.05
0564	ATP synthase subunit	Energy metabolism	0.56	0.92
0179	DNA repair endo/exonuclease FEN-1 (RAD2)	Replication and repair	0.59	1.00
0078	Tyrosyl-tRNA synthetase	Translation	0.60	0.96
0202	D-arabino 3-hexulose 6-phosphate formaldehyde lyase	Energy metabolism	0.60	0.98
0767	Conserved Hypothetical	Phospholipid-binding protein	0.61	0.96
0352	Apoptosis-related Tfar10 related protein	Cellular processes	0.62	1.03
2250	DNA repair exonuclease (rad32/mre11)	Exonuclease	0.62	1.01
2315	Flagella-related protein	Cellular processes	0.62	1.13
2803	Rieske iron-sulfur protein-1	Energy metabolism	0.63	0.81
2390	Inorganic pyrophosphatase	Energy metabolism	0.63	0.88
1353	Hypothetical Protein	Predicted bile acid beta-glucosidase	0.63	0.97
0353	SSU ribosomal protein S19E	Translation	0.63	0.94
0400	Riboflavin synthase beta chain	Cofactor biosynthesis	0.63	1.34
0947	Hypothetical Protein	Uncharacterized coiled-coil protein	0.64	1.11
5763	LSU ribosomal protein L14E	Translation	0.65	0.97
0632	Amidophosphoribosyltransferase	Nucleotide transport and metabolism	0.66	0.90

Table 6.4 Microarray data showing genes down regulated after hydrogen peroxide damage in PBL2025 *sso2273* knockout
Table shows the 20 most down regulated genes 10 minutes after hydrogen peroxide treatment, the ratios for these genes 30 minutes after treatment is also shown. Ratios were obtained by comparing values from hydrogen peroxide treated and control samples.

The top 20 down regulated genes in the knockout after hydrogen peroxide damage are shown in Table 6.4. The list of down regulated genes in the knockout is very different from the list of down regulated genes in the wild type, and generally the genes that were down regulated were not as tightly repressed as in the wild type.

The most down regulated gene *sso10224*, codes for a hypothetical protein, the protein is only 60 amino acids long and a BLAST search produced only one hit, an uncharacterized protein from *Metallosphaera sedula*.

A proliferating cell nuclear antigen (PCNA) gene (*sso0397*) was down regulated in the knockout. PCNA has been described as a molecular tool belt (Williams, Johnson et al. 2006) it acts as a processivity factor for a number of other proteins, FEN-1 (*sso0179*) is one of these proteins and this protein is also down regulated in the knockout. These proteins are involved in replication so the down regulation witnessed could indicate a general down regulation of replication in the cells.

The *sso7D* genes that were so strongly down regulated in the wild type strain (0.23 and 0.29), which are also down regulated after UV damage (Frols, Gordon et al. 2007; Gotz, Paytubi et al. 2007) showed very little change in expression in the knockout with ratios of 1.67 and 1.05 for 10 and 30 minutes respectively.

6.5.2 Microarray results for *sso2273* knockout versus wild type strains in normal and Hydrogen peroxide treated conditions.

Changes in gene expression between the wild type and knockout were looked at in normal conditions and at two time points after hydrogen peroxide treatment, to determine what affect the absence of *sso2273* had on gene expression. If Sso2273 was the repressor for the *dps* gene, an increase in the expression of this gene was predicted in the knockout in normal, ie non-damage, conditions. An increase in the expression of the *nramp* gene was also predicted as it shares a divergent promoter with the *dps* gene. However the results obtained did not support the hypothesis (see Tables 6.5 – 6.8).

Sso	Description	Function	Ratio
2241	BPS2	DNA replication recombination and repair	3.19
0352	Apoptosis-related Tfar19 related protein	Cellular processes	2.92
0712	SSU ribosomal protein S32AB	Translation	2.84
0233	Conserved Hypothetical protein	Adenine/guanine phosphoribosyltransferases	2.70
0633	Amidophosphoribosyltranferase	Metabolism	2.68
0441	Hypothetical Protein	?	2.54
0632	Amidophosphoribosyltranferase	Metabolism	2.44
2319	Hypothetical Protein	?	2.40
0697	LSU ribosomal protein L30AB	Translation	2.39
0100	Phenylalanyl-tRNA synthetase alpha subunit	Translation	2.35
0726	2-haloalkanoic acid dehalogenase	?	2.31
0899	Valy;-tRNA synthetase	Translation	2.30
0351	Translation initiation factor 6	Translation	2.26
1131	Heterodisulfide reductase subunit A	Energy metabolism	2.22
0274	Queuine/archaeosine-tRNA ribosyltransferase	Translation	2.22
0641	Carbamoyl-phosphate synthase large subunit	Amino acid biosynthesis	2.20
0629	FGAM synthase II	?	2.19
0626	SAICAR synthetase	?	2.18
2199	Conserved Hypothetical protein	?	2.17
0400	Riboflavin synthase beta chain	Cofactor biosynthesis	2.16

Table 6.5 of Microarray data showing the top 20 genes up regulated in the knockout compared to the wild type in normal conditions
Table shows the top 20 up regulated genes in the *sso2273* knockout at OD_{600} 0.1. Ratios were obtained by comparing values from knockout and wild type samples.

The top 20 up regulated genes in the knockout compared to the wildtype in normal conditions are shown in Table 6.5. The most up regulated gene in the knockout when compared to the wild type in normal (ie non damage) conditions was a *bps2* gene. This gene is predicted to be an ATPase involved in DNA replication, recombination and repair. The protein contains an ATP-binding domain, a metal-binding domain, a coiled coil domain and a zinc-hook domain. The protein shows similarity to rad50, a BLAST search shows a rad50 protein from another Archaeal species, *Natronomonas pharaonis* as the fifth hit. Rad50 is involved, along with Mre11, in recognising and repairing double strand breaks (Lavin 2007). Interestingly this gene is less induced after hydrogen peroxide damage in either the knockout or the wild type and is down regulated in response to UV damage with a ratio of 0.25 30 minutes after damage

(data from Dr Götz arrays). Proteins involved in glutamine metabolism and various translation proteins are also up regulated.

Sso	Description	Function	Ratio
2909	Sulfite reductase hemoprotein beta	Central intermediary metabolism	0.08
2911	PAPS reductase	Central intermediary metabolism	0.09
2912	Sulfate adenylyltransferase	Central intermediary metabolism	0.10
1817	Thiosulfate sulfurtransferase	Inorganic ion transport and metabolism	0.14
3089	Hypothetical Protein	?	0.26
2908	Uroporphyrin-III C-methyltransferase	Cofactor biosynthesis	0.31
0242	IMP asparate ligase	Ligase	0.31
0366	Glutamine synthetase	Amino acid biosynthesis	0.32
0244	Conserved hypothetical Protein	Uncharacterized FAD-dependent dehydrogenase	0.32
3189	Amino acid transporter	Transport	0.33
1469	Hypothetical Protein	Methyl-accepting chemotaxis protein	0.38
2527	Prolidase	Protease	0.43
3246	Hypothetical Protein	?	0.43
0503	Conserved hypothetical Protein	?	0.46
0515	Transposase	-	0.48
2612	HtrA like serine protease (periplasmic)	Protease	0.48
3085	Conserved hypothetical Protein	Predicted transport component	0.50
2915	Hypothetical Protein	?	0.50
2972	Quinol oxidase-2-sulfocyanin	Energy metabolism	0.51
1288	Hypothetical Protein	ABC-type dipeptide transport system	0.52

Table 6.6 Microarray data showing the genes most down regulated in the knockout compared to the wild type in normal conditions
Table shows the 20 most down regulated genes in the *sso2273* knockout at OD_{600} 0.1. Ratios were obtained by comparing values from knockout and wild type samples. The ratios of Sso2273 were around 1 showing no change in expression; the spot intensities from the Sso2273 gene were also low for the knockout cultures.

The top 20 down regulated genes in the knockout compared to the wild type in normal conditions are shown in Table 6.6. The up regulated genes in the knockout compared to the wild type in normal condition showed a relatively low level of up regulation, in

contrast there was much stronger control of the genes down regulated in the knockout compared to the wild type in normal condition.

The top three down regulated gene belong to an operon (*sso2908-2912*), these genes are involved in metabolism of sulphur. For organisms, such as *Sulfolobus solfataricus,* which inhabit volcanic hot springs (Brock TD 1972) the oxidation and reduction of sulphur are important energy-yielding reactions (Kletzin, Urich et al. 2004).

Sso	Description	Function	Ratio 10 min	Ratio 30 min
1127	Heterodisulfide reductase subunit C	Energy Metabolism	1.88	0.90
1135	Heterodisulfide reductase subunit B	Energy Metabolism	1.86	0.94
1251	Conserved hypothetical protein	Predicted nucleotidyltransferases	1.84	1.11
1131	Heterodisulfide reductase subunit A	Energy Metabolism	1.79	1.08
12018	Conserved hypothetical protein	Regulator of stationary/sporulation gene expression	1.67	0.89
1125	Conserved hypothetical protein	Predicted peroxiredoxins	1.57	0.84
0353	SSU Ribosomal protein S19E	Translation	1.56	1.64
2078	Putative NRAMP family protein	Metal transport	1.56	1.36
5561	SSU Ribosomal protein S17E	Translation	1.53	0.99
2079	DPSL-type antioxidant enzymes	DNA protection	1.49	1.54
0632	Amidophosphoribosyltranaferase	Nucleotide transport and metabolism	1.48	1.43
6223	Phosphoribosyl-ATP cyclohydrolase	Amino acid biosynthesis	1.46	1.21
0515	Transposase		1.45	0.98
0352	Apoptosis-related Tfar10 related protein	Cellular processes	1.44	2.22
1092	Hypothetical Protein	?	1.43	1.57
5577	DNA-directed RNA polymerase subunit L	Transcription	1.42	1.06
0629	FGAM synthase II	Nucleotide transport and metabolism	1.41	1.74
5576	DNA-directed RNA polymerase subunit M	Transcription	1.41	1.32
1129	Heterodisulfide reductase subunit B	Energy Metabolism	1.40	1.00
1134	Heterodisulfide reductase subunit C	Energy Metabolism	1.39	0.85
0348	Signal recognition particle (docking protein)	Cellular Processes	1.37	2.25
0633	Amidophosphoribosyltranaferase	Nucleotide transport and metabolism	1.37	1.47

Table 6.7 Microarray data showing genes up regulated in the knockout compared to the wild type after Hydrogen peroxide damage
Table shows the top 20 up regulated genes in the *sso2273* knockout 10 minutes after hydrogen peroxide treatment. Ratios for these genes 30 minute after damage are also shown. Ratios were obtained by comparing values from knockout and wild type samples.

The top 20 up regulated genes in the knockout compared to the wild type after hydrogen peroxide treatment are shown in Table 6.7. The ratios of the top 20 up

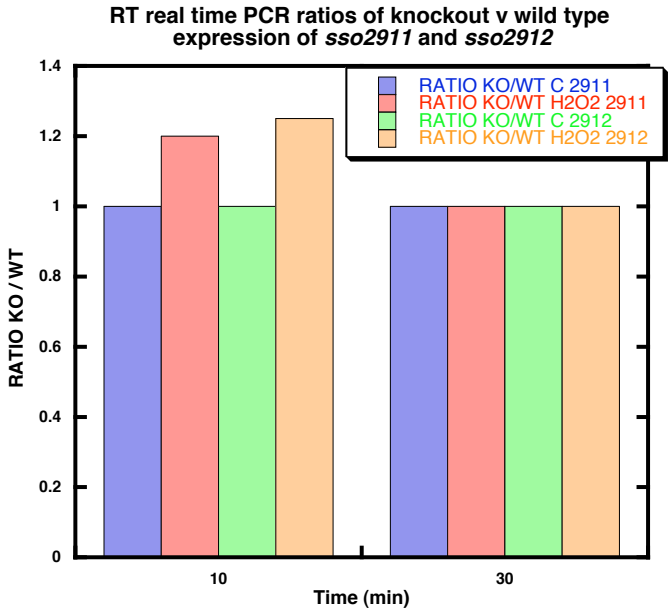
regulated genes in the knockout compared to the wild type after hydrogen peroxide damage were relatively small, suggesting the up regulated genes in the knockout and wild type after damage were similar. Many of the genes that appear on the list are involved in energy metabolism and the electron transport chain. The *dps* and *nramp* genes are slightly more up regulated in the knockout compared to the wild type.

Sso	Description	Function	Ratio 10 min	Ratio 30 min
2912	Sulfate adenyltransferase	Central intermediary metabolism	0.29	0.24
2909	Sulfite reductase hemoprotein beta	Central intermediary metabolism	0.29	0.23
2911	3'phosphoadenosine 5'-phosphosulfate sulfotransferase	Central intermediary metabolism	0.44	0.26
1183	Inorganic phosphate transporter	Transport	0.44	0.59
2971	Quinol oxidase-2-rieske iron-sulphur protein	Energy metabolism	0.48	0.89
2970	Quinol oxidase-2-cytochrome b	Energy metabolism	0.48	1.07
0366	Glutamine synthetase	Amino acid biosynthesis	0.50	0.36
2973	Quinol oxidase-2-subunitI/II cytochrome	Energy metabolism	0.50	0.61
1817	Thiosulfate sulfurtransferase	Central intermediary metabolism	0.52	0.32
2972	Quinol oxidase-2 sulfocyanin	Energy metabolism	0.54	0.81
0638	Argininosuccinate synthetase	Amino acid biosynthesis	0.56	0.77
0503	Conserved hypothetical protein	?	0.57	0.46
0260	Conserved hypothetical protein	Uncharacterized Zn ribbon-containing protein	0.58	0.85
3019	Beta-glycosidase	Central intermediary metabolism	0.59	0.50
2087	Carboxymethylenebutenolidase	Amino acid biosynthesis	0.59	0.74
2632	Conserve hypothetical protein	?	0.59	0.78
2908	Uroporphyrin III C-methyltransferase	Cofactor biosynthesis	0.63	0.36
3051	Alpha-glucosidase	Central Intermediary metabolism	0.63	0.88
2704	Permease multidrug efflux	Transport	0.63	0.64
3085	Hypothetical Protein	Predicted transported component	0.64	0.54

Table 6.8 Microarray data showing genes down regulated in the knockout compared to the wild type after Hydrogen peroxide damage
Table shows the 20 most down regulated genes in the *sso2273* knockout 10 minutes after hydrogen peroxide treatment. Ratios for these genes at 30 minutes after damage are also included. Ratios were obtained by comparing values from knockout and wild type samples.

The top 20 down regulated genes in the knockout compared to the wild type after hydrogen peroxide treatment are shown in Table 6.8. The *sso2908-2912* operon that was down regulated in the non-damage condition in the knockout strain was also down regulated after hydrogen peroxide damage, although to a lesser extent.

RT real time PCR was used to check the down regulation of genes *sso2911* and *sso2912*, see Figure 6.13



Strain/ Gene	Average CP 10 min	Average CP 30 min
Wild type control *sso2911*	20.8	24.4
Wild type treated *sso2211*	16.8	20.7
Knockout control *sso2911*	21.6	24.8
Knockout treated *sso2911*	20.8	21.6
Wild Type control *sso2912*	19.9	25.2
Wild type treated *sso2912*	16.2	20.6
Knockout control *sso2912*	21.6	26.0
Knockout treated *sso2912*	20.1	20.9

Figure 6.13 Graph showing ratio of KO v WT for genes *sso2911* and *sso2912* in control and 5 µM Hydrogen peroxide treated cultures, the table shows crossing point values used to determine the ratios.
RT real time was used to confirm the down regulation of he 2909-2912 operon witnessed in the microarray. Crossing point values were used to obtain a ratio (KO v WT) for each gene in control and hydrogen peroxide treated conditions. There was very little change in ratio in any of the samples.

The ratios obtained by RT real time PCR were around one for all samples. This suggests there is very little change in expression between the two strains and in either

condition. It is more difficult to confirm the down regulation of a gene due to the increased variability when data is taken from the later cycles of the RT real time PCR**.** Due to the strong repression of the *sso2911* operon in the knockout compared to the wild type it is possible that Sso2273 is having some effect on its expression. It could be a direct effect, Sso2273 could be an activator of the *sso2911* operon, and so when it is knocked out expression of the *sso2911* operon is markedly reduced. To test this hypothesis the promoter region of *sso2911* (see Figure 6.13) (which also has an inverted repeat like the one found in the *dps* promoter, although its sequence is different (see Chapter 5)) was used in EMSA assays with the Sso2273 protein (see Figure 6.14).

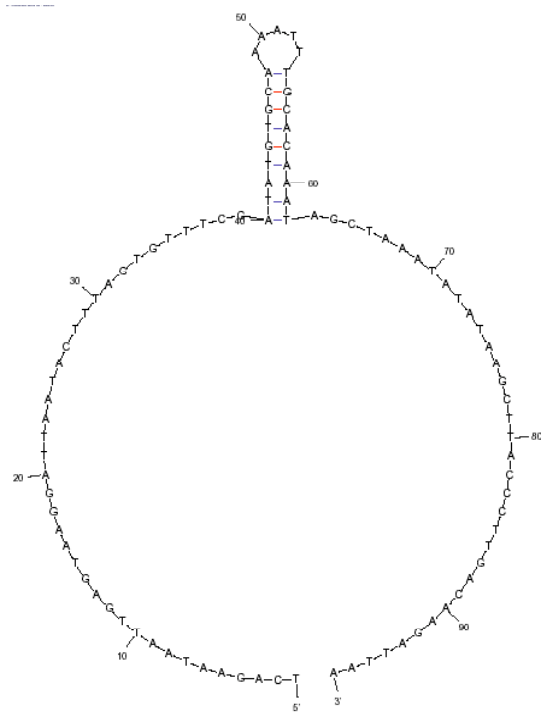


Figure 6.14 A graphical representation of the inverted repeat in the promoter region between the *sso2911* and *sso2912* genes
The region between the *sso2911* and *sso 2912* genes contains an inverted repeat, which maybe a possible binding site for any proteins that control transcription of these genes. The graphic was produced using mfold http://frontend.bioinfo.rpi.edu/zukerm/cgi-bin/rna-index.cgi.

A 44 base pair region of the *sso2911/2912* promoter was used in band shift assays to determine whether Sso2273 may be effecting transcription from this promoter, as the microarray data from the knockout compared to the wild type suggested the absence of Sso2273 had an inhibiting effect on the *sso2911* operon.

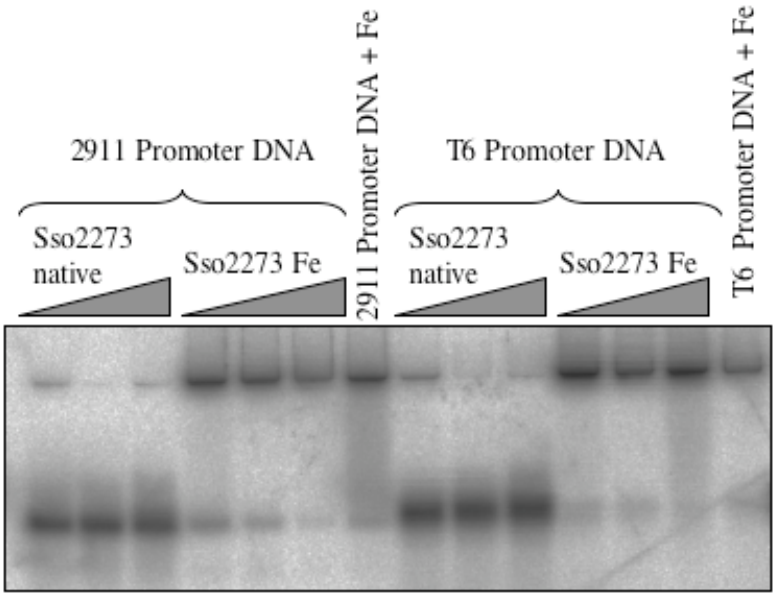


Figure 6.15 EMSA with *sso2911* and *T6* promoter regions and Sso2273 with and without iron

EMSA run on native 12 % acrylamide gel. 60 ng [γ^{32}P] labelled *dps* promoter or 60 ng T6 promoter were incubated with increasing concentrations of Sso2273 (10 μM, 20 μM, 40 μM). Sso2273 was either with or without addition of 0.5mM $Fe(NH_4)_2SO_4$. Controls containing 0.5 mM $Fe(NH_4)_2SO_4$+DNA. The iron+DNA control shows that the band shift observed was due to the presence of iron and not to the presence of Sso2273.

The T6 promoter region was included in the assay to determine whether Sso2273 was specific for the 2911 promoter. DNA + iron controls were run, as it was previously observed that the DNA could be retarded by the presence of the metal alone. The gel shows that Sso2273 is not specific to the 2911 promoter and the shifting is due to the presence of metal, not the presence of protein. These assays are not an effective way of assessing the binding of Sso2273, it is possible that Sso2273 is involved in the control of the *sso2911* operon but further investigation is needed. It is also possible that Sso2273 does not directly control the *sso2911* operon, it may repress a repressor of this operon and so when *sso2273* is knocked out, the repressor is no longer controlled and reduces transcription of the *sso2911* operon. Further investigation into this operon is needed but due to time constrains it was not possible within this PhD.

6.6 Discussion

The genetic techniques that have recently become available for thermophilic archaea open a number of avenues for further study of these organisms. The gene knockout technique developed by Paul Blum was utilised, with generous help from Dr Sonja Albers, to produce a strain of *Sulfolobus solfataricus* PBL2025 missing the *sso2273* gene.

It was hoped that with this knockout the role of Sso2273 in the cell could be elucidated. Growth curves comparing the knockout and wild type strains showed that the growth of the two strains was very similar, the culture which grew to the higher final optical densities was different when different size flasks were used to grow the cultures. The growth of the strains was also similar after the cultures were treated with hydrogen peroxide, and real time PCR of the *dps* and *nramp* genes showed an increase in their expression in the wild type strain, although the increase in the expression of the *dps* gene was greater. There was little change in the expression of these genes in the knockout, which fitted with the hypothesis that the genes would be constitutively expressed because no repressor was present. However the CP values for the gene were similar in the wild type and knockout, and if the genes were being constitutively expressed you would expect the CP values in the knockout to be lower, the lower the CP value the higher the level of transcript. It is possible that once *sso2273* was knocked out its paralogue *sso0669* takes over its role and so, without knocking out both *sso2273* and *sso0669* the effect of these repressors cannot be properly studied.

The level of the repression exacted by the Sso2273 homologue, DtxR, on the tox gene that it represses is dependent on the iron levels in the growth media. To determine whether this was also true for Sso2273, growth curves with cultures containing different iron concentrations were performed and RT real time PCR was used to look at the expression levels of the *dps* and *nramp* genes. In the wild type the expression of these genes, and the expression of *sso2273* itself increased with increasing iron concentrations. This is the opposite of what is seen in the homologue, as the highest expression of the *tox* gene occurs in low iron conditions.

Due to the number of results that contradict the hypothesis that the Sso2273 protein is the transcriptional repressor of the *dps* and *nramp* genes, it is likely that we were wrong in our prediction of the function of this protein. In an attempt to determine the role of Sso2273, microarray experiments were performed. One set of arrays was used to look at the changes in gene expression in the wild type and knockout after hydrogen peroxide damage, and another set were used to look at the differences in expression between the knockout and wild type stains.

Results from the hydrogen peroxide treated versus control cultures arrays showed that the genes up regulated after damage were similar in the wild type and knockout out strains, with *dps* and *nramp* genes being most highly up regulated. A number of genes involved in protection and repair were also up regulated and the results matched well with microarray results from Mark Young who has performed similar experiments in *Sulfolobus solfataricus* P2 (personal communication). The genes down regulated however were very different with transcription and translation proteins being down regulated in the wild type that did not show the same down regulation in the knockout. Conversely genes for proteins involved in replication, PCNA and FEN-1, were down regulated in the knockout after damage but were not down regulated to the same extent in the wild type.

In the arrays that looked at the knockout strain versus the wild type strain, expression in normal and hydrogen peroxide damage conditions were looked at. In normal conditions the most up regulated gene in the knockout compared to the wild type was a *bps2* gene. This gene is predicted to be involved in repair of double strand breaks, as the protein shows strong similarity to Rad50, a protein known to repair double strand breaks in conjunction with Mre11, however the Mre11 homologue is not up regulated in the knockout. The majority of other genes up regulated were involved in metabolism. The ratios of up regulation were relatively close to 1 suggesting that the same proteins were up regulated in the knockout and control. However there was a marked down regulation of one operon in both the control and treated conditions. This operon (*sso2909 – sso2911*) contained genes involved in sulphur assimilation and repression of the operon, especially in the control condition was strong. Whether

Sso2273 acts directly on this operon, or indirectly by affecting a protein involved in its control is unclear. Band shift assays were performed with the promoter region of the operon and Sso2273 (see Figure 6.11). The protein was used native and with metal added as the protein needs metal ions bound to adopt the DNA binding conformation (see Chapter 5). However the results could only confirm that the retardation of the DNA was caused by the iron present in the assays, rather than the protein. Further investigation is needed to determine what role if any Sso2273 plays in the expression of the *sso2911* operon.

Chapter Seven

Conclusions and Future Work

7.1 Conclusions and future work

The initial project of this thesis was to provide confirmation of microarray data showing the expression change in a number of genes after UV irradiation. After this confirmation was obtained the expression of the same set of genes was investigated after exposure to different damaging agents. RT real time PCR experiments showed that genes *sso0771 (cdc6-2)* and *sso0280 (tfb-3),* were induced by all the damaging agents tested. It is predicted that Cdc6-2 has a role in controlling the cell cycle, inhibiting binding of the other two Cdc6 proteins to the replication start sites after damage, to stop the cell cycle and allow repair to occur (Robinson, Dionne et al. 2004). A strong induction of *cdc6-2* was seen after UV, while cdc6-1 and cdc6-3 were down regulated (cdc6-1 to a greater extent). Western blot analysis confirmed that protein levels of these three genes followed the transcript levels witnessed (Gotz, Paytubi et al. 2007). There is some contention as to the levels of the Cdc6 proteins during the normal cell cycle. Work from the lab of Professor Rolf Bernander has shown that Cdc6-1 and Cdc6-3 are induced in early G1 phase (Lundgren and Bernander 2007) while recent work from Professor Stephen Bell's lab has shown that the levels of all three Cdc6 proteins remains constant across the cell cycle. These labs use different techniques to produces synchronised cell cultures, in Prof Bernander's lab cells are arrested in the G2 phase with sodium acetate, then released from this arrest by dilution in fresh media, while in Prof Bell's lab the 'baby machine' is used. Cells are grown on a membrane that is continually perfused with media. Newly divided cells, in early G1 phase, drop off the membrane and are held in liquid media on ice, holding them in G1 phase. It maybe possible that changes in the levels of the Cdc6 proteins witnessed by Bernander 's lab are a consequence of the cells response to the acetate treatment.

TFB-3 was one of the most highly up regulated genes after UV damage, and was up regulated in response to all the damaging agents tested. TFB-3 is one of three TFB proteins in *S. solfataricus*. It is a truncated form of the other two proteins and is missing the DNA binding helix-turn-helix domain and B-finger, used to stimulate RNAP. It was initially suggested that this protein may have an inhibitor role, as it was shown to compete with TFB-1 for binding of RNAP (Gotz, Paytubi et al. 2007).

However work from Prof Malcolm White's lab has shown that TFB-3 actually has a stimulatory effect on transcription (Paytubi unpublished). TFB-1 lacks the zinc ribbon seen in most other archaea and eukaryotes, however TFB-1 can still initiate transcription. It is suggested that in response to damage the zinc ribbon domain is provided in *trans* by TFB-3, which stimulates transcription. Thus, transcription is functional in the absence of TFB-3 but is stimulated in its presence, allowing transcriptional activation to be fine-tuned by altering the levels of the two TFBs.

When hydrogen peroxide, which induces oxidative stress, was added to *S. solfataricus* there was the huge induction of the *dps* gene. This gene codes for a protein involved in protecting the cell from oxidative stress. Because transcription from the *dps* gene had been shown to be strong in *in vitro* transcription assays, it was predicted a repressor was inhibiting transcription of the gene *in vivo*.

A potential repressor, Sso2273, was isolated from pull down assays using the promoter of the *dps* gene as bait. Sso2273 is a homologue of the diphtheria toxin repressor from *Corynebacterium diphtheria* (Schmitt, Twiddy et al. 1992), this repressor binds to the *tox* promoter in low iron conditions inhibiting expression of the gene. When the metal ions contained within the protein are released a change in the protein's conformation causes it to release the DNA (Rangachari, Marin et al. 2005). Sso2273 was reconstituted with a variety of divalent metals in an attempt to form the DNA binding conformation of the protein. Sso2273 did not bind DNA and the metal reconstitution was shown to have been unsuccessful by ICP-OES analysis. The EMSA assays performed with the Sso2273 protein were unsuccessful, perhaps because the DNA binding form of the protein was never reconstituted in stable form due to the redox sensitivity of ferrous iron. Investigation of the protein in anaerobic conditions might solve the redox sensitivity problems and possibly produce more conclusive results on the function of this protein.

In vitro transcription assays using a variety of promoters showed Sso2273 had no effect on transcription, perhaps due to the difficulty obtaining the DNA binding form of the protein. However its paralogue Sso0669 did appear to inhibit transcription from

all promoters tried. Unfortunately a knockout strain missing the *sso0669* gene could not be obtained.

A knockout strain of PBL2025 missing the *sso2273* gene was produced in an attempt to determine the function of this gene within the cell. Microarray analysis comparing the *sso2273* knockout and wild type strains in control and hydrogen peroxide damage conditions was performed. Results were not as expected, the absence of *sso2273* did not lead to a constitutively high level of the *dps* gene, as was predicted if the repressor of this gene was absent. There was however, an effect on the expression of an operon containing genes involved in sulphur assimilation. This operon *sso2909-2912* was strongly inhibited in control and, to a lesser extent, damage conditions in the *sso2273* knockout compared to the wild type. There was also a difference in the expression of the *sso7D* genes between the wild type and knockout. In the wild type these genes were the most down regulation after hydrogen peroxide damage at the 10-minute time point, followed by a 2-fold increase in expression at the 30-minute time point. However in the *sso2273* knockout these genes show hardly any change in expression, with ratios of around 1 at both the 10 and 30-minute time points.

The microarray data has produced some interesting results that warrant further investigation. The expression of the *sso2909-2912* operon is clearly affected by the absence of the *sso2273* gene, however EMSA assays with Sso2273 indicate that it does not bind the *sso2909-2912* promoter directly, although until the correct DNA binding form of the protein can be obtained this result is not conclusive. It is possible that Sso2273 is acting indirectly, for example by repressing a repressor of the *sso2909-2912* operon. The lack of inhibition of the *sso7D* genes in the knockout was another interesting result from the microarray data; these proteins do the job of histones, which are absent from the crenarchaea, and have been shown to be down regulated after damage by UV (Gotz, Paytubi et al. 2007) and oxidative stress (Chapter 6). These histone like proteins have many roles within the cell, performing chromatin functions (White and Bell 2002), inducing negative supercoiling (Agback, Baumann et al. 1998; Napoli, Zivanovic et al. 2002) and increasing the thermal

stability of dsDNA (Agback, Baumann et al. 1998) and their down regulation in the knockout should be investigated further.

The crystal structure of Sso2273 showed the high degree of structural similarity the protein shares with proteins of the Diphtheria toxin repressor family. The protein crystallised as a dimer with zinc bound in its metal ion binding sites. Comparison of Sso2273 with two proteins of the DtxR family (MntR and IdeR) showed that despite strong overall structural similarity the metal ion binding sites of Sso2273 differed from those of the other two proteins. Interestingly when the C-terminal portion of Sso2273 and IdeR were aligned the DNA binding helix turn helix domains did not align well, which may have implications for Sso2273 ability to bind DNA. Further analysis of Sso2273 metal binding site is in progress.

The paralogue of Sso2273, Sso0669, appeared to inhibit transcription from all the promoters tried. Despite expressing well this protein was difficult to purify, and was only obtained at 80 % purity. Attempts to produce knockouts of this gene were also unsuccessful, possibly because of differences in flanking regions sequences between the P2 strain (used to amplify the flanking regions) and the PBL2025 strain (used to produce the knockout). Genetic techniques for *Sulfolobus* are progressing rapidly; one option is to product a knockout of the gene homologous to *sso0669* in *S. acidocaldarius*. If successful a knockout missing both the *sso2273* and *sso0669* homologues could be produced in *S. acidocaldarius*. It is possible that these genes perform overlapping functions within the cell and their action in the cell cannot be fully investigated until both are knocked out.

The work presented in this thesis sheds some light on the transcriptional response of a number of genes in *S. solfataricus* to different kind of damage. The potential transcriptional repressor Sso2273 was investigated and its crystal structure examined. A knockout strain missing the *sso2273* gene was used in microarray studies to assess the genome wide response to hydrogen peroxide, of the wild type and *sso2273* knockout in an attempt to determine the role of Sso2273 within the cell.

References

Agback, P., H. Baumann, et al. (1998). "Architecture of nonspecific protein-DNA interactions in the Sso7d-DNA complex." Nat Struct Biol **5**(7): 579-84.

Albers, S. J. and A. J. M. Driessen (2007). "Conditions for gene disruption by homologous recombination into the *Sulfolobus solfataricus* genome." Archaea **2**: 145-149.

Allers, T. and M. Mevarech (2005). "Archaeal genetics - the third way." Nat Rev Genet **6**(1): 58-73.

Almiron, M., A. J. Link, et al. (1992). "A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli." Genes Dev **6**(12B): 2646-54.

Andersson, A., R. Bernander, et al. (2005). "Dual-genome primer design for construction of DNA microarrays." Bioinformatics **21**(3): 325-32.

Andrews, S. C., A. K. Robinson, et al. (2003). "Bacterial iron homeostasis." FEMS Microbiol Rev **27**(2-3): 215-37.

Arosio, P. and S. Levi (2002). "Ferritin, iron homeostasis, and oxidative damage." Free Radic Biol Med **33**(4): 457-63.

Ashwell, S. and S. Zabludoff (2008). "DNA damage detection and repair pathways--recent advances with inhibitors of checkpoint kinases in cancer therapy." Clin Cancer Res **14**(13): 4032-7.

Baliga, N. S., Y. A. Goo, et al. (2000). "Is gene expression in Halobacterium NRC-1 regulated by multiple TBP and TFB transcription factors?" Mol Microbiol **36**(5): 1184-5.

Barns, S. M., C. F. Delwiche, et al. (1996). "Perspectives on archaeal diversity, thermophily and monophyly, from environmental rRNA sequences." Proc Natl Acad Sci U S A **93**: 9188-9193.

Bartlett, M. S. (2005). "Determinants of transcription initiation by archaeal RNA polymerase." Curr Opin Microbiol **8**(6): 677-84.

Bell, S. D. (2005). "Archaeal transcriptional regulation--variation on a bacterial theme?" Trends Microbiol **13**(6): 262-5.

Bell, S. D., A. B. Brinkman, et al. (2001). "The archaeal TFIIEalpha homologue facilitates transcription initiation by enhancing TATA-box recognition." EMBO Rep **2**(2): 133-8.

Bell, S. D., S. S. Cairns, et al. (1999). "Transcriptional regulation of an archaeal operon in vivo and in vitro." Mol Cell **4**(6): 971-82.

Bell, S. D. and S. P. Jackson (2001). "Mechanism and regulation of transcription in archaea." Curr Opin Microbiol **4**(2): 208-13.

Bell, S. D., P. L. Kosa, et al. (1999). "Orientation of the transcription preinitiation complex in archaea." Proc Natl Acad Sci U S A **96**(24): 13662-7.

Berkner, S. and G. Lipps (2008). "Genetic tools for Sulfolobus spp.: vectors and first applications." Arch Microbiol.

Bertram, C. and R. Hass (2008). "Cellular responses to reactive oxygen species-induced DNA damage and aging." Biol Chem **389**(3): 211-20.

Binz, S. K., A. M. Sheehan, et al. (2004). "Replication protein A phosphorylation and the cellular response to DNA damage." DNA Repair **3**: 1015-1024.

Binz, S. K., A. M. Sheehan, et al. (2004). "Replication protein A phosphorylation and the cellular response to DNA damage." DNA Repair (Amst) **3**(8-9): 1015-24.

Boudsocq, F., S. Iwai, et al. (2001). "Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4): an archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic poleta." Nucleic Acids Res **29**(22): 4607-16.

Brinkman, A. B., S. D. Bell, et al. (2002). "The Sulfolobus solfataricus Lrp-like protein LysM regulates lysine biosynthesis in response to lysine availability." J Biol Chem **277**(33): 29537-49.

Brinkman, A. B., I. Dahlke, et al. (2000). "An Lrp-like transcriptional regulator from the archaeon Pyrococcus furiosus is negatively autoregulated." J Biol Chem **275**(49): 38160-9.

Brinkman, A. B., T. J. Ettema, et al. (2003). "The Lrp family of transcriptional regulators." Mol Microbiol **48**(2): 287-94.

Brochier, C., B. Boussau, et al. (2008). "Mesophilic crenarchaeota: proposal for a third archaeal phylum the Thaumarchaeota." Nat Rev Microbiol **6**: 245-252.

Brochier, C. and P. Forterre (2006). "Widespread distribution of archaeal reverse gyrase in thermophilic bacteria suggests a complex history of vertical inheritance and lateral gene transfers." Archaea **2**: 83-93.

Brochier, C., S. Gribaldo, et al. (2005). "Nanoarchaea:representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales?" Genome Biol **6**.

Brock TD, B. K., Belly RT, Weiss RL (1972). "Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature." Arch Microbiol **84**(1): 54-68.

Browning, D. F. and S. J. Busby (2004). "The regulation of bacterial transcription initiation." Nat Rev Microbiol **2**(1): 57-65.

Brune, I., H. Werner, et al. (2006). "The DtxR protein acting as dual transcriptional regulator directs a global regulatory network involved in iron metabolism of Corynebacterium glutamicum." BMC Genomics **7**(1): 21.

Bustin, S. A. (2000). "Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays." J Mol Endocrinol **25**(2): 169-93.

Butala, M., D. Zgur-Bertok, et al. (2008). "The bacterial LexA transcriptional repressor." Cell. Mol. Life Sci.

Cadet, J., E. Sage, et al. (2005). "Ultraviolet radiation-mediated damage to cellular DNA." Mutat Res **571**(1-2): 3-17.

Chae, J. C., E. Kim, et al. (2007). "Comparative analysis of the catechol 2,3-dioxygenase gene locus in thermoacidophilic archaeon Sulfolobus solfataricus strain 98/2." Biochem Biophys Res Commun **357**(3): 815-9.

Chen, L., K. Brugger, et al. (2005). "The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota." J Bacteriol **187**(14): 4992-9.

Christmann, M., M. T. Tomicic, et al. (2003). "Mechanisms of human DNA repair: an update." Toxicology **193**: 3-34.

Cuadrado, M., B. Martinez-Pastor, et al. (2006). "ATM regulates ATR chromatin loading in response to DNA double-strand breaks." J Exp Med **203**(2): 297-303.

Cubeddu, L. and M. F. White (2005). "DNA damage detection by an archaeal single-stranded DNA-binding protein." J Mol Biol **353**(3): 507-16.

D'Amours, D. and S. P. Jackson (2002). "The Mre11 complex: at the crossroads of dna repair and checkpoint signalling." Nat Rev Mol Cell Biol **3**(5): 317-27.

D'Aquino, J. A., J. Tetenbaum-Novatt, et al. (2005). "Mechanism of metal ion activation of the diphtheria toxin repressor DtxR." Proc Natl Acad Sci U S A **102**(51): 18408-13.

Dajkovic, A., A. Mukherjee, et al. (2008). "Investigation of regulation of FtsZ assembly by SulA and development of a model for FtsZ polymerization." J Bacteriol **190**(7): 2513-2526.

DeDecker, B. S., R. O'Brien, et al. (1996). "The crystal structure of a hyperthermophilic archaeal TATA-box binding protein." J Mol Biol **264**(5): 1072-84.

Delong, E. F. (1992). "Archaea in coastal marine environments." Proc Natl Acad Sci U S A **89**: 5685-5689.

Demple, B. and L. Harrison (1994). "Repair of oxidative damage to DNA: enzymology and biology." Annu. Rev Biochem **63**: 915-948.

Dismukes, G. C., V. V. Klimov, et al. (2001). "The origin of atmospheric oxygen on Earth: the innovation of oxygenic photosynthesis." Proc Natl Acad Sci U S A **98**(5): 2170-5.

Dorazi, R., D. Gotz, et al. (2007). "Equal rates of repair of DNA photoproducts in transcribed and non-transcribed strands in Sulfolobus solfataricus." Mol Microbiol **63**(2): 521-9.

Duggin, I. G., S. A. McCallum, et al. (2008). "Chromosome replication dynamics in the archaeon Sulfolobus acidocaldarius." Proc Natl Acad Sci U S A **105**(43): 16737-42.

Dussault, A. A. and M. Pouliot (2006). "Rapid and simple comparison of messenger RNA levels using real-time PCR." Biol Proced Online **8**: 1-10.

Dussurget, O., J. Timm, et al. (1999). "Transcriptional control of the iron-responsive fxbA gene by the mycobacterial regulator IdeR." J Bacteriol **181**(11): 3402-8.

Early, A., L. S. Drury, et al. (2004). "Mechanisms involved in regulating DNA replication origins during the cell cycle and in response to DNA damage." Philos Trans R Soc Lond B Biol Sci **359**(1441): 31-8.

Elkins, J. G., M. Podar, et al. (2008). "A korarchaeal genome reveals insights into the evolution of the archaea." Proc Natl Acad Sci U S A **105**(23): 8102-8107.

Fanning, E., V. Klimovich, et al. (2006). "A dynamic model for replication protein A (RPA) function in DNA processing pathways." Nucleic Acids Res **34**(15): 4126-37.

Fleck, O. and O. Nielsen (2004). "DNA repair." J Cell Sci **117**(Pt 4): 515-7.

Forbes, J. R. and P. Gros (2001). "Divalent-metal transport by NRAMP proteins at the interface of host-pathogen interactions." Trends Microbiol **9**(8): 397-403.

Fortini, P., B. Pascucci, et al. (2003). "The base excision repair: mechanisms and its relevance for cancer susceptibility." Biochimie **85**(11): 1053-71.

Frols, S., P. M. Gordon, et al. (2007). "Response of the hyperthermophilic archaeon Sulfolobus solfataricus to UV damage." J Bacteriol **189**(23): 8708-18.

Fuhrman, J. A., K. McCallum, et al. (1992). "Novel major archaebacterial group from marine plankton." Nature **356**: 148-149.

Fujihashi, M., N. Numoto, et al. (2007). "Crystal structure of archaeal photolyase from *Sulfolobus tokodaii* with tow FAD molecules: implication of a novel light-harvesting cofactor." J. Mol. Biol **365**: 903-910.

Gauss, G. H., P. Benas, et al. (2006). "Structure of the DPS-like protein from Sulfolobus solfataricus reveals a bacterioferritin-like dimetal binding site within a DPS-like dodecameric assembly." Biochemistry **45**(36): 10815-27.

Gotz, D., S. Paytubi, et al. (2007). "Responses of hyperthermophilic crenarchaea to UV irradiation." Genome Biol **8**(10): R220.

Gourse, R. L., W. Ross, et al. (2000). "UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition." Mol Microbiol **37**(4): 687-695.

Grogan, D. W. (1997). "Photoreactivation in an archaeon from geothermal environments." Microbiology **143**: 1071-1076.

Grogan, D. W. (2004). "Stability and repair of DNA in hyperthermophilic archaea." Curr. Issues Mol Biol **6**: 137-144.

Grogan, D. W., G. T. Carver, et al. (2001). "Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon Sulfolobus acidocaldarius." Proc Natl Acad Sci U S A **98**(14): 7928-33.

Guagliardi, A., L. Cerchia, et al. (2002). "The Sso7d protein of Sulfolobus solfataricus: in vitro relationship among different activities." Archaea **1**(2): 87-93.

Hahn, S. (2004). "Structure and mechanism of the RNA polymerase II transcription machinery." Nat Struct Mol Biol **11**(5): 394-403.

Haring, M., S. Offermann, et al. (2007). "Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization." Plant Methods **3**: 11.

Hartzell, B. and B. McCord (2005). "Effect of divalent metal ions on DNA studied by capillary electrophoresis." Electrophoresis **26**(6): 1046-56.

Hazra, T. K., T. Izumi, et al. (2003). "The discovery of a new family of mammalian enzymes for repair of oxidatively damaged DNA, and its physiological implications." Carcinogenesis **24**(2): 155-157.

Hickey, A. J., E. Conway de Macario, et al. (2002). "Transcription in the archaea: Basal factors, regulation, and stress gene expression." Crit Rev Biochem Mol Biol **37**(4): 199-258.

Hill, P. J., A. Cockayne, et al. (1998). "SirR, a novel iron-dependent repressor in Staphylococcus epidermidis." Infect Immun **66**(9): 4123-9.

Hirata, A., B. J. Klein, et al. (2008). "The X-ray crystal structure of RNA polymerase from Archaea." Nature **451**(7180): 851-4.

Hochheimer, A., R. Hedderich, et al. (1999). "The DNA binding protien Tfx from *Methanobacterium thermoautotrophicum*: structure, DNA binding properties and transcriptional regulation." Mol Microbiol **31**(2): 641-650.

Hoeijmakers, J. H. (2001). "Genome maintenance mechanisms for preventing cancer." Nature **411**(6835): 366-74.

Hopkin, B. B. and T. T. Paull (2008). "The P. furiosus Mre11/Rad50 complex promotes 5' strand resection at a DNA double-strand break." Cell **135**: 250-260.

Huber, H., M. J. Hohn, et al. (2002). "A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont." Nature **417**: 63-67.

Huffman, J. L. (2006). DNA damage recognition, Taylor and Francis Group, Boca Raton, Fl.

Iyer, V. N. and W. Szybalski (1963). "A Molecular Mechanism of Mitomycin Action: Linking of Complementary DNA Strands." Proc Natl Acad Sci U S A **50**: 355-62.

Jackson, S. P. (2002). "Sensing and repairing DNA double-strand breaks." Carcinogenesis **23**(5): 687-696.

Jackson, S. P. (2002). "Sensing and repairing DNA double-strand breaks." Carcinogenesis **23**(5): 687-96.

Janion, C. (2008). "Inducible SOS response system of DNA repair and mutagenesis in Escherichia coli." Int J Biol Sci **4**(6): 338-44.

Jolivet, E., F. Matsunaga, et al. (2003). "Physiological response of the hyperthermophilic archaeon Pyrococcus abyssi to DNA damage caused by ionizing radiation." J Bacteriol **185**(13): 3958-3961.

Justice, S. S., D. A. Hunstad, et al. (2008). "Morphological plasticity as a bacterial survival strategy." Nat Rev Microbiol **6**(2): 162-8.

Kadonaga, J. T. (2004). "Transcription in Eukaryotes."

Kang, J. and M. J. Blaster (2006). "Bacterial populations as perfect gases: genomic integrity and diversification tensions in Helicobacter pylori." Nature **4**: 826-836.

Kelman, Z. and M. F. White (2005). "Archaeal DNA replication and repair." Curr Opin Microbiol.

Kletzin, A., T. Urich, et al. (2004). "Dissimilatory oxidation and reduction of elemental sulfur in thermophilic archaea." J Bioenerg Biomembr **36**(1): 77-91.

Kokoska, R. J., S. D. McCulloch, et al. (2003). "The efficiency and specificity of apurinic/apyrimidinic site bypass by human DNA polymerase eta and Sulfolobus solfataricus Dpo4." J Biol Chem **278**(50): 50537-45.

Lavin, M. F. (2007). "ATM and the Mre11 complex combine to recognize and signal DNA double-strand breaks." Oncogene **26**(56): 7749-58.

Lee, Y. J., S. J. Park, et al. (2006). "An in vivo analysis of MMC-induced DNA damage and its repair." Carcinogenesis **27**(3): 446-53.

Leonard, P. M., S. H. Smits, et al. (2001). "Crystal structure of the Lrp-like transcriptional regulator from the archaeon Pyrococcus furiosus." Embo J **20**(5): 990-7.

Li, B. H. and R. Bockrath (1995). "Benefit of transcription-coupled nucleotide excision repair for gene expression in u.v.-damaged Escherichia coli." Mol Microbiol **18**(4): 615-22.

Limauro, D., E. Pedone, et al. (2008). "Peroxiredoxins as cellular guardians in Sulfolobus solfataricus: characterization of Bcp1, Bcp3 and Bcp4." Febs J **275**(9): 2067-77.

Limauro, D., E. Pedone, et al. (2006). "Identification and characterization of 1-Cys peroxiredoxin from Sulfolobus solfataricus and its involvement in the response to oxidative stress." Febs J **273**(4): 721-31.

Lodish, H., A. Berk, et al. (1999). Molecular Cell Biology, Freeman.

Love, J. F., J. C. vanderSpek, et al. (2004). "Genetic and biophysical studies of diphtheria toxin repressor (DtxR) and the hyperactive mutant DtxR(E175K) support a multistep model of activation." Proc Natl Acad Sci U S A **101**(8): 2506-11.

Lundgren, M. and R. Bernander (2007). "Genome-wide transcription map of an archaeal cell cycle." Proc Natl Acad Sci U S A **104**(8): 2939-44.

Macris, M. A. and P. Sung (2005). "Multifaceted role of the Saccharomyces cerevisiae Srs2 helicase in homologous recombination regulation." Biochem Soc Trans **33**(Pt 6): 1447-50.

Magill, C. P., S. P. Jackson, et al. (2001). "Identification of a conserved archaeal RNA polymerase subunit contacted by the basal transcription factor TFB." J Biol Chem **276**(50): 46693-6.

Malik, P. S. and L. S. Symington (2008). "Rad51 gain-of-function mutants that exhibit high affinity DNA binding cause DNA damage sensitivity in the absence of Srs2." Nucleic Acids Res **36**(20): 6504-10.

Martusewitsch, E., C. W. Sensen, et al. (2000). "High spontaneous mutation rate in the hyperthermophilic archaeon Sulfolobus solfataricus is mediated by transposable elements." J Bacteriol **182**(9): 2574-81.

McCready, S., J. A. Muller, et al. (2005). "UV irradiation induces homologous recombination genes in the model archaeon, Halobacterium sp. NRC-1." Saline Systems **1**: 3.

Menck, C. F. (2002). "Shining a light on photolyases." Nat Genet **32**(3): 338-9.

Meyer, R. R. and P. S. Laine (1990). "The single-stranded DNA-binding protein of Escherichia coli." Microbiol Rev **54**(4): 342-80.

Missiakas, D. and S. Raina (1998). "The extracytoplasmic function sigma factors: role and regulation." Mol Microbiol **28**(6): 1059-66.

Morey, J. S., J. C. Ryan, et al. (2006). "Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR." Biol Proced Online **8**: 175-93.

Moriwaki, S. and Y. Takahashi (2008). "Photoaging and DNA repair." J Dermatol Sci **50**(3): 169-76.

Napoli, A., A. Valenti, et al. (2005). "Functional interaction of reverse gyrase with single-strand binding protein of the archaeon Sulfolobus." Nucleic Acids Res **33**(2): 564-76.

Napoli, A., J. van der Oost, et al. (1999). "An Lrp-like protein of the hyperthermophilic archaeon Sulfolobus solfataricus which binds to its own promoter." J Bacteriol **181**(5): 1474-80.

Napoli, A., Y. Zivanovic, et al. (2002). "DNA bending, compaction and negative supercoiling by the architectural protein Sso7d of Sulfolobus solfataricus." Nucleic Acids Res **30**(12): 2656-62.

Ng, S. Y., B. Zolghadr, et al. (2008). "Cell surface structures of archaea." J Bacteriol **190**(18): 6039-47.

Ouhammouch, M., R. E. Dewhurst, et al. (2003). "Activation of archaeal transcription by recruitment of the TATA-binding protein." Proc Natl Acad Sci U S A **100**(9): 5097-102.

Ouhammouch, M., G. E. Langham, et al. (2005). "Promoter architecture and response to a positive regulator of archaeal transcription." Mol Microbiol **56**(3): 625-37.

Pedone, E., S. Bartolucci, et al. (2004). "Sensing and adapting to environmental stress: the archaeal tactic." Front Biosci **9**: 2909-26.

Peeters, E., T. L. Thia-Toong, et al. (2004). "Ss-LrpB, a novel Lrp-like regulator of Sulfolobus solfataricus P2, binds cooperatively to three conserved targets in its own control region." Mol Microbiol **54**(2): 321-36.

Pennella, M. A. and D. P. Giedroc (2005). "Structural determinants of metal selectivity in prokaryotic metal-responsive transcriptional regulators." Biometals **18**(4): 413-28.

Pfaffl, M. W. (2001). "A new mathematical model for relative quantification in real-time RT-PCR." Nucleic Acids Res **29**(9): e45.

Pfeifer, G. P., Y. H. You, et al. (2005). "Mutations induced by ultraviolet light." Mutat Res **571**(1-2): 19-31.

Pitcher, R. S., N. C. Brissett, et al. (2007). "Nonhomologous end-joining in bacteria: A microbial perspective." Annu. Rev. Microbiol **61**: 259-82.

Podar, M. and A. L. Reysenbach (2006). "New opportunities revealed by biotechnological explorations of extremophiles." Curr Opin Biotechnol **17**(3): 250-5.

Quaiser, A., C. Florence, et al. (2008). "The Mre11 protein interacts with both Rad50 and the HerA bipolar helicase and is recruited to DNA following gamma irradiation in the archaeon Sulfolobus acidocaldarius." BMC Molecular Biology **9**(25).

Qureshi, S. A. and S. P. Jackson (1998). "Sequence-specific DNA binding by the S. shibatae TFIIB homolog, TFB, and its effect on promoter strength." Mol Cell **1**(3): 389-400.

Radonic, A., S. Thulke, et al. (2004). "Guideline to reference gene selection for quantitative real-time PCR." Biochem Biophys Res Commun **313**(4): 856-62.

Rajeevan, M. S., S. D. Vernon, et al. (2001). "Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR." J Mol Diagn **3**(1): 26-31.

Ramsay, B., B. Wiedenheft, et al. (2006). "Dps-like protein from the hyperthermophilic archaeon Pyrococcus furiosus." J Inorg Biochem **100**(5-6): 1061-8.

Rangachari, V., V. Marin, et al. (2005). "Sequence of ligand binding and structure change in the diphtheria toxin repressor upon activation by divalent transition metals." Biochemistry **44**(15): 5672-82.

Redder, P., Q. She, et al. (2001). "Non-autonomous mobile elements in the crenarchaeon Sulfolobus solfataricus." J Mol Biol **306**(1): 1-6.

Reindel, S., C. L. Schmidt, et al. (2006). "Expression and regulation pattern of ferritin-like DpsA in the archaeon Halobacterium salinarum." Biometals **19**(1): 19-29.

Reiter, H., M. Milewskiy, et al. (1972). "Mode of action of phleomycin on Bacillus subtilis." J Bacteriol **111**(2): 586-92.

Richard, D. J., S. D. Bell, et al. (2004). "Physical and functional interaction of the archaeal single-stranded DNA-binding protein SSB with RNA polymerase." Nucleic Acids Res **32**(3): 1065-74.

Robb, F., G. Antranikian, et al. (2007). Thermophiles. Biology and Technology at High Temperatures, CRC Press.

Robinson, N. P., I. Dionne, et al. (2004). "Identification of two origins of replication in the single chromosome of the archaeon Sulfolobus solfataricus." Cell **116**(1): 25-38.

Romano, V., A. Napoli, et al. (2007). "Lack of strand-specific repair of UV-induced DNA lesions in three genes of the archaeon Sulfolobus solfataricus." J Mol Biol **365**(4): 921-9.

Salerno, V., A. Napoli, et al. (2003). "Transcriptional response to DNA damage in the archaeon Sulfolobus solfataricus." Nucleic Acids Res **31**(21): 6127-38.
Samson, R. Y., T. Obita, et al. (2008). "A Role for the ESCRT System in Cell Division in Archaea." Science.
Scandalios, J. G. (2005). "Oxidative stress: molecular perception and transduction of signals triggering antioxidant gene defenses." Braz J Med Biol Res **38**(7): 995-1014.
Schelert, J., V. Dixit, et al. (2004). "Occurrence and characterization of mercury resistance in the hyperthermophilic archaeon Sulfolobus solfataricus by use of gene disruption." J Bacteriol **186**(2): 427-37.
Scheuch, S., L. Marschaus, et al. (2008). "Regulation of gvp genes encoding gas vesicle proteins in halophilic Archaea." Arch Microbiol **190**(3): 333-9.
Schmidt, K. J., K. E. Beck, et al. (1999). "UV stimulation of chromosomal marker exchange in Sulfolobus acidocaldarius: implications for DNA repair, conjugation and homologous recombination at extremely high temperatures." Genetics **152**(4): 1407-15.
Schmitt, M. P., E. M. Twiddy, et al. (1992). "Purification and characterization of the diphtheria toxin repressor." Proc Natl Acad Sci U S A **89**(16): 7576-80.
She, Q., R. K. Singh, et al. (2001). "The complete genome of the crenarchaeon Sulfolobus solfataricus P2." Proc Natl Acad Sci U S A **98**(14): 7835-40.
Shin, J. H., B. Grabowski, et al. (2003). "Regulation of minichromosome maintenance helicase activity by Cdc6." J Biol Chem **278**(39): 38059-67.
Sonia-Verena Albers, A. J. M. D. (2007). "Conditions for gene disruption by homologous recombination into the *Sulfolobus solfataricus* genome." Archaea **2**: 145-149.
Soppa, J. (1999). "Transcription initiation in Archaea: facts, factors and future aspects." Mol Microbiol **31**(5): 1295-305.
Spiering, M. M., D. Ringe, et al. (2003). "Metal stoichiometry and functional studies of the diphtheria toxin repressor." Proc Natl Acad Sci U S A **100**(7): 3808-13.
Sztukowska, M., M. Bugno, et al. (2002). "Role of rubrerythrin in the oxidative stress response of Porphyromonas gingivalis." Mol Microbiol **44**(2): 479-88.
Szutorisz, H., N. Dillon, et al. (2005). "The role of enhancers as centres for general transcription factor recruitment." TRENDS in Biochemical Sciences **30**(11).
Thomas, M. C. and C. M. Chiang (2006). "The general transcription machinery and general cofactors." Crit Rev Biochem Mol Biol **41**(3): 105-78.
Thompson, C. L. and A. Sancar (2002). "Photolyase/cryptochrome blue-light photoreceptors use photon energy to repair DNA and reset the circadian clock." Oncogene **21**: 9034-9056.
Truglio, J. J., D. L. Croteau, et al. (2006). "Prokaryotic nucleotide excision repair: the UvrABC system." Chem Rev **106**(2): 233-52.
Valenti, A., A. Napoli, et al. (2006). "Selective degradation of reverse gyrase and DNA fragmentation induced by alkylating agent in the archaeon Sulfolobus solfataricus." Nucleic Acids Res **34**(7): 2098-108.
Valko, M., H. Morris, et al. (2005). "Metals, toxicity and oxidative stress." Curr Med Chem **12**(10): 1161-208.
Valko, M., C. J. Rhodes, et al. (2006). "Free radicals, metals and antioxidants in oxidative stress-induced cancer." Chem Biol Interact **160**(1): 1-40.

Van der Vossenberg, J., A. Driessen, et al. (1998). "The essence of being extremophilic: the role of unique archaeal membrane lipids." Extremophiles **2**: 163-170.

Van Gent, D. C. and M. Van der Burg (2007). "Non-homologous end-joining, a sticky affair." Oncogene **26**: 7731-7740.

Vierke, G., A. Engelmann, et al. (2003). "A novel archaeal transcriptional regulator of heat shock response." J Biol Chem **278**(1): 18-26.

Vitzthum, F., G. Geiger, et al. (1999). "A quantitative fluorescence-based microplate assay for the determination of double-stranded DNA using SYBR Green I and a standard ultraviolet transilluminator gel imaging system." Anal Biochem **276**(1): 59-64.

Wardleworth, B. N., R. J. Russell, et al. (2002). "Structure of Alba: an archaeal chromatin protein modulated by acetylation." Embo J **21**(17): 4654-62.

Weinberg, M. V., F. E. Jenney, Jr., et al. (2004). "Rubrerythrin from the hyperthermophilic archaeon Pyrococcus furiosus is a rubredoxin-dependent, iron-containing peroxidase." J Bacteriol **186**(23): 7888-95.

Werner, F. (2007). "Structure and function of archaeal RNA polymerases." Mol Microbiol **65**(6): 1395-404.

Werner, F. and R. O. Weinzierl (2005). "Direct modulation of RNA polymerase core functions by basal transcription factors." Mol Cell Biol **25**(18): 8344-55.

White, A., X. Ding, et al. (1998). "Structure of the metal-ion-activated diphtheria toxin repressor/tox operator complex." Nature **394**(6692): 502-6.

White, M. F. (2003). "Archaeal DNA repair: paradigms and puzzles." Biochem Soc Trans **31**(Pt 3): 690-3.

White, M. F. and S. D. Bell (2002). "Holding it together: chromatin in the Archaea." Trends Genet **18**(12): 621-6.

Wiedenheft, B., J. Mosolf, et al. (2005). "An archaeal antioxidant: characterization of a Dps-like protein from Sulfolobus solfataricus." Proc Natl Acad Sci U S A **102**(30): 10551-6.

Wilkinson, S. P. and A. Grove (2006). "Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins." Curr Issues Mol Biol **8**(1): 51-62.

Williams, E., T. M. Lowe, et al. (2007). "Microarray analysis of the hyperthermophilic archaeon Pyrococcus furiosus exposed to gamma irradiation." Extremophiles **11**(1): 19-29.

Williams, G. J., K. Johnson, et al. (2006). "Structure of the heterotrimeric PCNA from Sulfolobus solfataricus." Acta Crystallogr Sect F Struct Biol Cryst Commun **62**(Pt 10): 944-8.

Wiseman, H. and B. Halliwell (1996). "Damage to DNA by reactive oxygen and nitrogen species: role in inflammatory disease and progresion of cancer." Biochem. J. **313**: 17-29.

Woese, C. R. and G. E. Fox (1977). "Phylogenetic structure of the prokaryotic domain: the primary kingdoms." Proc Natl Acad Sci U S A **74**(11): 5088-90.

Woese, C. R., O. Kandler, et al. (1990). "Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya." Proc Natl Acad Sci U S A **87**: 4576-4579.

Wouter L. de Laat, N. G. J. J., Jan H.J Hoeijmakers (1999). "Molecular mechanisms of nucleotide excision repair." Genes & Dev **13**: 768-785.

Yellaboina, S., S. Ranjan, et al. (2004). "Prediction of DtxR regulon: identification of binding sites and operons controlled by Diphtheria toxin repressor in Corynebacterium diphtheriae." BMC Microbiol **4**: 38.

Zhao, G., P. Ceci, et al. (2002). "Iron and Hydrogen Peroxide Deoxification Properties of DNA-binding Protein from Starved Cells." Journal of Biolgical Chemistry.

Zimmermann, P. and F. Pfeifer (2003). "Regulation of the expression of gas vesicle genes in Haloferax mediterranei: interaction of the two regulatory proteins GvpD and GvpE." Mol Microbiol **49**(3): 783-94.

Zou, Y., Y. Liu, et al. (2006). "Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses." J Cell Physiol **208**(2): 267-73.

Appendix A – Oligonucleotide Sequences

Name	Primer Sequence 5' to 3'
Real time PCR Primers	
0280 Forward	TTA GAT TCG CGT TAA ATA ATG G
0280 Reverse	CAA ATA CGA TCG CTT TCT TCG
0771 Forward	CCA CAT AGA GAA GAG AAG ATT AAG G
0771 Reverse	GTA GCT GTT TTC CCA GTA CC
2506 Forward	ACC AAT AAG GCT GGT GTT GC
2506 Reverse	GGC CTA TGC ACT TGA GCT TC
0446 Forward	GGC CAG AAC TTT GGA TGA GA
0446 Reverse	CCA GCA GTT AAC CCA GAA CC
0946 Forward	CGC GTT GAA AAG AGT CCA AT
0946 Reverse	GGA AGC TGC GCT CAA AGA TA
2364 Forward	AGT TTT GGA AGC AAG CGA AG
2364 Reverse	GTG GTC CAC GCG TTT TCT AT
0121 Forward	TAC TGG TTG GAG GGA GAT CG
0121 Reverse	GCT AAA CCA ACA GTT CCG TCA
0961 Forward	GGA AAG CGA AAA TAG GGG AAA TG
0961 Reverse	CAG CAA CAT AGT GGA GTT CTT G
1459 Forward	CCG CCT AGG GAT AAA ACC AT
1459 Reverse	CCT CAA CTT CAG GCT TTT CG
2079 Forward	GAG AAA CCC CAA GAA CCA AAG G
2079 Reverse	CCA GTT AGG TGC ATT CTC AAT ATG G
2395 Forward	CAA GTC AAC AAA ACG GGA GTT
2395 Reverse	TGG GAC TAG CCG TTT AGG AA
2506 Forward	GGC TGG TGT TGC TTA CGG TTT G
2506 Reverse	CTT ATT CTG GCT CAT CCA TTC C
0959 Forward	TGA ATG CAG GGG TTC TTG TT
0959 Reverse	AGT TTT GCT TGC TTG CCA TT
2273 Real Time Forward	ATG GCG TGT GGA TTA CCA AT
2273 Real Time Reverse	TCT CTA ACC CGT GTG GAC AG
0927 Forward	GTA CAG CGC CAC AAC TCA AA
0927 Reverse	ATT TCC ATC AGC CCA TGC TC
2121 Forward	AGC CAT CCT GGA GAT TTC AC
2121 Reverse	GTA GAG GAC TGC GCA TGG AT
2244 Forward	GCA AAC GGC ATA ACT TGG TT
2244 Reverse	AAC CAA GTT ATG CCG TTT GC
2642 Forward	CCC TGA AAT TGC TGG ATT GT
2642 Reverse	GGG TAC ATT TGC GTC CAT TC
Cloning Primers	
2273 Cloning Primer Forward	GCG CGG ATC CCC CAT GGC TAA CTT ATC ACG AAG AGA ATT CTC ATA TCT GC
2273 Cloning Primer Reverse	GGC CGT CGA C TT AGA GCG GTA TCT CTA ACC CGT GTG G

0669 Cloning Primer Forward	GCG CGG ATC CCC C AT GGT TGA ACT TTC GGA ACC TTT AG
0669 Cloning Primer Reverse	GGC CGT CGA CTC ATC CAA TTA CTA GCA CCT GGT CGC C
Transcription Primers and EMSA	
Dps Promoter Forward	GAG TAA GTG CAA ATG TTG TA
Dps Promoter Reverse	GAT ATC TAA CCC AGA TTT TTC
Primers and Oligos for Pull Downs	
DpsPromter forward (biotin labelled for pull downs)	GTT TGA AAC TAC TTT TAA CTA TAA GTT AAA ATG CCT CTT AAA TAG
Dps Promoter Reverse (for pull downs)	CTA TTT AAG AGT CAT TTT AAC TTA TAG TTA AAA GTA GTT TCA AAC
Primers for the *nramp* gene	
2067 Forward	GCT CGG CAA ATA CTC TTC CA
2077 (transposon) Forward	TTG CCT ACA TAC CCC GAA AG
2077 (transposon) Reverse	GGA GAG TCC TCG AGT TCA CG
2078 Forward	CAT CGC AGT AAA CGT TGG TG
2078 Reverse	GGT TAA GAA CCC AGC GCT T
Chromatin Immunoprecipitation Primers	
2078 Forward	ATC TTA AAA CAC CTC GCT TA
2078 Reverse	TTC ATA CAC CCT AAT AGT C
0927 Forward	GGA AGA TCA TAG CAG AAG G
0927 Reverse	CTT GGC ATC TAT GGT AGC G
0962 Forward	CCA CCA CAG TTA GGG CAG G
0962 Reverse	GTT CAT TAC TGG TTT CTT TCC
2121 Forward	CGG TCT AAC GTT ATC TCT C
2121 Reverse	CCA TTT CTG GGA ATC TTT CTC C
2244 Forward	GCG AAA GAA CAA TAT CAA TTG
2244 Reverse	GCG ATT CTT TGT GGT GTC AC
2273 Forward	CCA CTA GTG CCA AGA ACC C
2273 Reverse	GGT AAG CAG ATA TGA GAA TTC
2642 Forward	CTC CAG AGC CTC GCA TT
2642 Reverse	GCT TCT CCG CAG AAT GC
Primers for Confirmation of Microarray	
2911 Forward	CGT GCT GTG AAG TGA GGA AA
2911 Reverse	CCA TAC CTG TTC CCA TGT CC
2912 Forward	ACG TTG TCA TTG GGG AGA AG
2912 Reverse	GAG CCT CTC TAG GTC CAG CA
2911 Promoter	CTTTAGTGTTTCGATATGTGCAAAATTTG CACAAATAGCTAAAT
2911 Promoter (complement)	ATTTAGCTATTTGTGCAAATTTTGCACAT ATCGAAACACTAAAG

Appendix B - Crossing Point Values for RT Real Time PCR

Sample	Time (hrs)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change	Ratio/ 2506	Ratio/ 0961	Average of 2 ratios
sso0121 control	1	18	18.3	18.1	18.1	1.89	-0.17	0.90	0.67	0.40	0.53
	2	17.7	17.5	17.8	17.7	1.89	0.37	1.26	1.00	1.21	1.10
	3	18.5	18.6	18.8	18.6	1.89	2.63	5.35	4.70	5.35	5.02
	5	17	16.9	17.7	17.2	1.89	1.93	3.42	3.98	3.76	3.87
sso0121 UV	1	18.3	18.6	18	18.3	1.89					
	2	17.2	17.8	16.9	17.3	1.89					
	3	15.8	15.9	16.3	16.0	1.89					
	5	15.2	15.5	15.1	15.3	1.89					
sso0961 control	1	14.2	14.4	14.1	14.2	2					
	2	13.3	13.6	13.3	13.4	2					
	3	14.5	14.6	14.7	14.6	2					
	5	13.7	13.7	13.6	13.7	2					
sso0961 UV	1	13.1	13.1	13	13.1	2					
	2	13.3	13.4	13.3	13.3	2					
	3	14.2	14.7	14.9	14.6	2					
	5	13.7	13.7	14	13.8	2					
ssso2506 control	1	15	15.1	14.9	15.0	1.9					
	2	14.8	14.9	14.7	14.8	1.9					
	3	14.6	15.2	15.3	15.0	1.9					
	5	14.4	14.3	14.6	1.4	1.9					
sso2506 UV	1	14.7	14.8	14.1	14.5	1.9					
	2	14	14.6	14.7	14.4	1.9					
	3	14.7	14.6	15.2	14.8	1.9					
	5	14.5	14.8	14.7	14.7	1.9					

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change	Ratio/ 2506	Ratio/ 0961	Average of 2 ratios
sso0280 control	30	17.0	16.5	16.6	16.7	1.97	1.17	2.21	1.19	2.11	1.65
	60	17.1	17.1	17.6	17.3	1.97	3.93	14.40	7.42	7.86	7.64
	90	17.1	17.1	17.3	17.2	1.97	4.70	24.21	12.74	18.74	15.74
	120	16.3	16.3	16.2	16.3	1.97	3.73	12.57	12.57	9.51	11.04
sso0280 UV	30	15.5	15.5	15.6	15.5	1.97					
	60	13.4	13.5	13.1	13.3	1.97					
	90	12.6	12.7	12.1	12.5	1.97					
	120	12.5	12.6	12.5	12.5	1.97					
sso0771 control	30	16.0	16.1	16.2	16.1	1.64	1.23	1.84	0.99	1.76	1.37
	60	16.3	16.1	16.2	16.2	1.64	4.03	7.35	3.79	4.02	3.90
	90	16.0	15.8	15.6	15.8	1.64	3.80	6.55	3.45	5.07	4.26
	120	15.8	15.8	15.8	15.8	1.64	3.13	4.71	4.71	3.56	4.14
sso0771 UV	30	15.0	14.8	14.8	14.9	1.64					
	60	12.2	12.2	12.1	12.2	1.64					
	90	12.1	12.0	11.9	12.0	1.64					
	120	12.7	12.6	12.7	12.7	1.64					
ssso0961 control	30	11.5	11.6	11.5	11.5	2.01	0.07	1.05			
	60	12.3	12.5	12.5	12.4	2.01	0.87	1.83			
	90	12.0	11.9	12.1	12.0	2.01	0.37	1.29			
	120	11.9	11.8	12.0	11.9	2.01	0.40	1.32			
sso0961 UV	30	11.5	11.5	11.4	11.5	2.01					
	60	11.4	11.6	11.7	11.6	2.01					
	90	11.6	11.6	11.7	11.6	2.01					
	120	11.4	11.5	11.6	11.5	2.01					

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change	Ratio/ 2506	Ratio/ 0961	Average of 2 ratios
sso2079 conrol	30	16.5	16.3	16.4	16.4	1.99	-0.40	0.76	0.41	0.72	0.57
	60	17.4	17.2	17.3	17.3	1.99	0.67	1.58	0.82	0.86	0.84
	90	17.5	17.2	17.3	17.3	1.99	0.83	1.77	0.93	1.37	1.15
	120	17.7	17.6	17.5	17.6	199	1.00	1.99	1.99	1.51	1.75
sso2079 UV	30	16.8	16.8	16.8	16.8	1.99					
	60	16.6	16.7	16.6	16.6	1.99					
	90	16.6	16.6	16.3	16.5	199					
	120	16.5	16.6	16.7	16.6	1.99					
sso2506 control	30	15.2	15.7	15.7	15.5	1.90	0.97	1.86			
	60	15.9	16.5	16.1	16.2	1.90	1.03	1.94			
	90	15.7	16.3	16.0	16.0	1.90	1.00	1.90			
	120	15.6	14.8	14.7	15.0	.190	0.00	1.00			
sso2506 UV	30	14.6	14.4	14.8	14.6	1.90					
	60	15.1	15.4	14.9	15.1	1.90					
	90	15.5	14.9	14.6	15.0	1.90					
	120	15.1	14.7	15.3	15.0	1.90					
sso0959 control	30	13.3	13.3	13.7	13.4	1.67	-0.43	0.80	0.43	0.76	0.60
	60	13.5	13.6	13.8	13.6	1.67	0.13	1.07	0.55	0.58	0.57
	90	13.8	13.9	14.0	13.9	1.67	0.10	1.05	0.55	0.81	0.68
	120	13.7	13.5	13.7	13.6	1.67	-0.17	0.92	0.92	0.69	0.81
sso0959 UV	30	13.8	13.9	13.9	13.9	1.67					
	60	13.6	13.6	13.3	13.5	1.67					
	90	13.7	13.8	13.9	13.8	1.67					
	120	13.9	13.7	13.8	13.8	1.67					

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change	Ratio/ 2506	Ratio/ 0961	Average of 2 ratios
sso1459 control	30	13.6	13.6	13.6	13.6	1.64	0.00	1.00	0.54	0.95	0.75
	60	13.6	13.5	13.7	13.6	1.64	2.27	3.07	1.58	1.68	1.63
	90	13.8	13.8	13.9	13.8	1.64	3.00	4.41	2.32	3.41	2.87
	120	13.7	13.7	13.6	13.7	1.64	2.50	3.44	3.44	2.61	3.02
sso1450 UV	30	13.6	13.6	13.6	13.6	1.64					
	60	11.3	11.4	11.3	11.3	1.64					
	90	10.8	10.8	10.9	10.8	1.64					
	120	11.1	11.1	11.3	11.2	1.64					
sso2364 control	30	9.6	9.4	9.5	9.5	1.88	-0.03	0.98	0.53	0.93	0.73
	60	9.4	9.2	9.1	9.2	1.88	-0.20	0.88	0.45	0.48	0.47
	90	9.3	9.6	9.4	9.4	1.88	-0.13	0.92	0.48	0.71	0.60
	120	9.9	9.5	9.9	9.8	1.88	-017	0.90	0.90	0.68	0.79
sso2364 UV	30	9.7	9.4	9.5	9.5	1.88					
	60	9.6	9.5	9.2	9.4	1.88					
	90	9.8	9.7	9.2	9.6	1.88					
	120	10.0	9.9	10.0	9.9	1.88					
Ssso0121 control	30	20.6	20.8	20.9	20.8	1.89	0.07	1.04	0.56	1.00	0.78
	60	20.4	20.2	20.5	20.4	1.89	0.07	1.04	0.54	0.57	0.55
	90	20.2	20.8	20.2	20.2	1.89	0.90	1.77	0.93	1.37	1.15
	120	20.3	20.5	20.3	20.4	1.89	2.23	4.14	4.14	3.13	3.64
Sso0121 UV	30	20.8	20.8	20.5	20.7	1.89					
	60	20.2	20.4	20.3	20.3	1.89					
	90	19.0	19.0	198	19.3	1.89					
	120	18.2	18.0	18.2	18.1	1.89					

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change	Ratio/ 2506	Ratio/ 0961	Average of 2 ratios
sso0446 control	30	14.3	14.7	14.4	14.5	1.68	-0.53	0.76	0.41	0.72	0.57
	60	14.8	14.8	14.4	14.7	1.68	0.03	1.02	0.52	0.56	0.54
	90	15.2	15.2	15.4	15.3	1.68	0.60	1.37	0.72	1.06	0.89
	120	14.5	14.6	14.6	14.6	1.68	-0.17	0.92	0.92	0.69	0.81
sso0446 UV	30	14.5	14.7	15.8	15.0	1.68					
	60	14.5	14.6	14.8	14.6	1.68					
	90	14.6	14.7	14.7	14.7	1.68					
	120	14.5	14.8	14.9	14.7	1.68					
sso0946 control	30	20.1	20.1	19.8	20.0	1.76	-0.60	0.71	0.38	0.68	0.53
	60	20.0	20.1	19.9	20.0	1.76	-1.10	0.54	0.28	0.29	0.28
	90	21.3	21.2	21.5	21.3	1.76	-0.33	0.83	0.44	0.64	0.54
	120	20	20.1	20.3	20.2	1.76	-1.27	0.49	0.49	0.37	0.43
sso0946 UV	30	20.6	20.8	20.4	20.6	1.76					
	60	20.9	20.9	21.5	21.1	1.76					
	90	21.8	21.6	21.6	21.7	1.76					
	120	21.2	21.5	21.6	21.4	1.76					

Sample	Time (hrs)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change
sso0121 control	30	15.5	15.4	15.3	15.4	1.97	0.70	1.61
	60	16.0	16.1	16.3	16.1	1.97	1.33	2.47
	90	14.7	14.8	14.7	14.7	1.97	-0.17	0.89
	120	15	14.8	14.8	14.9	1.97	0.07	1.05
sso0121 MMC	30	14.7	14.7	14.7	14.7	1.97		
	60	14.8	14.8	14.8	14.8	1.97		
	90	14.9	14.9	14.9	14.8	1.97		
	120	14.8	14.8	14.8	14.8	1.97		
sso0961 control	30	16.0	16.1	15.9	16.0	1.64	0.70	1.41
	60	15.7	15.8	15.7	15.7	1.64	1.43	2.03
	90	14.5	14.7	14.3	14.5	1.64	0.20	1.10
	120	14.5	14.4	14.5	14.5	1.64	0.77	1.46
sso0961 MMC	30	15.3	15.3	15.3	15.3	1.64		
	60	14.3	14.3	14.3	14.3	1.64		
	90	14.3	14.3	14.3	14.3	1.64		
	120	13.7	13.7	13.7	13.7	1.64		

Sample	Time (hrs)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change
sso0280 control	30	15.6	15.6	15.6	15.6	1.97	0.60	1.50
	60	14.6	14.6	14.6	14.6	1.97	-0.40	0.76
	90	14.8	14.8	14.8	14.8	1.97	1.00	1.97
	120	14.0	14.0	14.1	14.0	1.97	1.33	2.47
sso0280 MMS	30	15.0	15.0	15.0	15.0	1.97		
	60	15.0	15.0	15.0	15.0	1.97		
	90	13.8	13.8	13.8	13.8	1.97		
	120	12.7	12.7	12.7	12.7	1.97		
sso0771 control	30	15.1	15.1	15.2	15.1	1.64	0.07	1.03
	60	14.7	14.7	14.6	14.7	1.64	1.93	2.60
	90	15.0	15.3	15.1	15.1	1.64	2.37	3.22
	120	13.8	13.8	13.8	13.8	1.64	1.20	1.81
sso0771 MMS	30	15.0	15.1	15.1	15.1	1.64		
	60	12.8	12.6	12.8	12.7	1.64		
	90	12.8	12.8	12.7	12.8	1.64		
	120	12.7	12.6	12.5	12.6	1.64		
ssso2079 control	30	16.1	16.2	16.3	16.2	1.99	-0.23	0.85
	60	16.0	16.0	16.1	16.0	1.99	-0.30	0.81
	90	16.5	16.5	16.4	16.5	1.99	0.03	1.02
	120	16.6	16.7	16.6	16.6	1.99	-0.07	0.96
sso2079 MMS	30	16.4	16.4	16.5	16.4	1.99		
	60	16.4	16.2	16.4	16.3	1.99		
	90	16.5	16.5	16.5	16.4	1.99		
	120	16.7	16.7	16.7	17.7	1.99		

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change
sso0959 control	30	13.7	13.5	13.7	13.6	1.67	-0.23	0.89
	60	13.6	13.5	13.4	13.5	1.67	-0.97	0.61
	90	14.4	14.4	14.4	14.4	1.67	-0.03	0.98
	120	14.1	14.0	14.1	14.1	1.67	0.27	1.15
sso0959 MMS	30	13.9	13.9	13.8	13.9	1.67		
	60	14.4	14.4	14.6	14.5	1.67		
	90	14.5	14.3	14.5	14.4	1.67		
	120	13.9	13.6	13.9	13.8	1.67		
sso1459 control	30	13.5	13.8	13.8	13.7	1.64	0.00	1.00
	60	13.6	13.5	13.6	13.6	1.64	0.53	1.30
	90	13.7	13.5	13.6	13.6	1.64	1.17	1.78
	120	13.9	13.8	13.9	13.9	1.64	2.37	3.22
sso1459 MMS	30	13.7	13.7	13.7	13.7	1.64		
	60	13.0	13.0	13.1	13.0	1.64		
	90	12.4	12.4	12.5	12.4	1.64		
	120	11.6	11.3	11.6	11.5	1.64		
ssso2364 control	30	11.2	11.0	11.2	11.1	1.88	-0.60	0.68
	60	11.2	11.3	11.3	11.3	1.88	0.10	1.07
	90	11.8	11.9	11.8	11.8	1.88	-0.87	0.58
	120	8.1	8.1	8.1	8.1	1.88	-0.20	0.88
sso2364 MMS	30	11.9	11.7	11.6	11.7	1.88		
	60	11.2	11.3	11.0	11.2	1.88		
	90	12.7	12.7	12.7	12.7	1.88		
	120	8.3	8.4	8.2	8.3	1.88		

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change
sso0121 control	30	18.6	18.6	18.5	18.6	1.89	0.40	1.29
	60	17.8	17.7	17.7	17.7	1.89	-0.50	0.73
	90	18.6	18.5	18.9	18.7	1.89	-0.20	0.88
	120	18.7	18.7	18.4	18.6	1.89	1.07	1.97
sso0121 MMS	30	18.2	18.2	18.1	18.2	1.89		
	60	18.2	18.3	18.2	18.2	1.89		
	90	18.8	18.9	18.9	18.9	1.89		
	120	17.5	17.6	17.5	17.5	1.89		
sso0446 control	30	15.7	15.4	5.7	15.6	1.68	-0.17	0.92
	60	15.0	14.9	14.9	14.9	1.68	0.47	1.27
	90	15.2	14.9	15.2	15.1	1.68	0.07	1.04
	120	15.3	15.3	15.0	15.2	1.68	-0.43	0.80
sso0446 MMS	30	15.8	15.7	15.8	15.8	1.68		
	60	14.5	14.4	14.5	14.5	1.68		
	90	15.0	15.1	15.0	15.0	1.68		
	120	15.5	15.7	15.7	15.6	1.68		
ssso0946 control	30	18.6	18.7	18.6	18.6	1.76	-0.23	0.88
	60	18.2	18.0	18.1	18.1	1.76	-0.37	0.81
	90	19.0	18.8	18.8	18.9	1.76	0.03	1.02
	120	17.6	17.9	17.9	17.8	1.76	-0.03	0.98
sso0946 MMS	30	18.9	18.8	18.9	18.9	1.76		
	60	18.5	18.4	18.5	18.5	1.76		
	90	188	18.8	18.9	18.8	1.76		
	120	17.9	17.9	17.7	17.8	1.76		

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change
sso0280 control	30	15.6	15.5	15.6	15.6	1.97	-0.17	0.89
	60	15.8	15.5	15.9	15.7	1.97	-0.20	0.87
	90	14.0	14.6	13.9	14.2	1.97	-1.33	0.40
	120	16.4	16.5	16.5	16.5	1.97	0.87	1.80
sso0280 Phleomycin	30	15.8	15.8	15.6	15.7	1.97		
	60	16.1	15.9	15.8	15.9	1.97		
	90	15.2	15.7	15.6	15.5	1.97		
	120	15.3	15.6	15.9	15.6	1.97		
sso0771 control	30	15.4	15.5	15.6	15.5	1.64	0.00	1.00
	60	15.4	15.0	15.6	15.3	1.64	0.30	1.16
	90	14.6	14.4	14.2	14.4	1.64	-0.43	0.81
	120	15.4	15.3	15.3	15.3	1.64	0.50	1.28
sso0771 Phleomycin	30	15.6	15.4	15.5	15.5	1.64		
	60	15.0	15.1	15.0	15.0	1.64		
	90	15.0	14.6	14.9	14.8	1.64		
	120	14.7	14.9	14.9	14.8	1.64		
ssso2079 control	30	16.3	16.2	16.3	16.3	1.99	-0.30	0.81
	60	16.4	16.4	16.6	16.5	1.99	0.43	1.35
	90	16.8	17.0	16.0	16.6	1.99	0.20	1.15
	120	16.8	17.0	16.7	16.8	1.99	0.03	1.02
sso2079 Phleomycin	30	16.7	16.5	16.5	16.6	1.99		
	60	16.5	15.8	15.8	16.0	1.99		
	90	16.6	16.1	16.5	16.4	1.99		
	120	16.6	16.9	16.9	16.8	1.99		

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change
sso0959 control	30	14.4	14.3	14.1	14.3	1.67	-0.60	0.74
	60	14.0	13.9	13.9	13.9	1.67	-0.30	0.86
	90	14.2	14.3	14.1	14.2	1.67	0.17	1.09
	120	14.4	14.3	14.5	14.4	1.67	0.10	1.05
sso0959 Phleomycin	30	14.9	15.0	14.7	14.9	1.67		
	60	14.2	14.2	14.3	14.2	1.67		
	90	14.1	14.0	14.0	14.0	1.67		
	120	14.1	14.6	14.2	14.3	1.67		
sso1459 control	30	13.9	13.9	13.7	13.8	1.64	-0.30	0.86
	60	14.5	14.1	14.1	14.2	1.64	0.03	1.02
	90	14.6	14.5	14.5	14.5	1.64	0.00	1.00
	120	14.6	14.6	14.7	14.6	1.64	0.13	1.07
sso1459 Phleomycin	30	14.0	14.2	14.2	14.1	1.64		
	60	14.2	14.4	14.0	14.2	1.64		
	90	14.5	14.5	14.6	14.5	1.64		
	120	14.8	14.6	14.1	14.5	1.64		
ssso2364 control	30	10.5	10.3	10.3	10.4	1.88	-0.50	0.73
	60	9.7	9.7	9.7	9.7	1.88	-0.20	0.88
	90	10.1	10.2	10.2	10.2	1.88	0.07	1.04
	120	10.6	10.5	10.2	10.4	1.88	0.10	1.07
sso2364 Phleomycin	30	11.0	10.9	10.7	10.9	1.88		
	60	10.0	9.8	9.9	9.9	1.88		
	90	10.1	10.2	10.0	10.1	1.88		
	120	10.2	10.2	10.6	10.3	1.88		

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change
sso0121 control	30	17.5	17.4	17.6	17.5	0.89	0.13	0.98
	60	17.7	17.6	17.8	17.7	0.89	0.43	0.95
	90	17.2	17.2	17.5	17.3	0.89	-0.23	1.03
	120	16.6	16.4	16.8	16.6	0.89	-.073	1.09
sso0121 Phleomycin	30	17.6	17.5	17.0	17.4	0.89		
	60	17.2	17.0	17.6	17.3	0.89		
	90	17.4	17.5	17.7	17.5	0.89		
	120	17.0	17.4	17.6	17.3	0.89		
sso0446 control	30	15.2	14.0	14.6	14.6	1.68	0.33	1.19
	60	14.0	14.2	14.3	14.2	1.68	0.00	1.00
	90	13.5	13.6	13.6	13.6	1.68	-0.60	0.73
	120	13.7	13.6	13.5	13.6	1.68	-0.33	0.84
sso0446 Phleomycin	30	14.3	14.3	14.2	14.3	1.68		
	60	14.2	14.2	14.1	14.2	1.68		
	90	14.2	14.3	14.0	14.2	1.68		
	120	13.8	13.9	14.1	13.9	1.68		
ssso0946 control	30	19.9	19.8	19.4	19.7	1.76	0.70	1.49
	60	19.3	19.2	19.2	19.2	1.76	0.43	1.28
	90	18.4	18.5	18.8	18.6	1.76	-0.33	0.83
	120	18.3	18.0	18.3	18.2	1.76	0.03	1.02
sso0946 Phloemycin	30	19.0	19.1	18.9	19.0	1.76		
	60	18.9	18.8	18.7	18.8	1.76		
	90	18.9	19.1	18.7	18.9	1.76		
	120	17.9	18.4	18.2	18.2	1.76		

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change
sso0280 control	30	15.2	15.2	15.1	15.2	1.97	-0.40	0.76
	60	16.3	16.4	16.3	16.3	1.97	0.93	1.88
	90	15.9	16.0	15.8	15.9	1.97	0.83	1.76
	120	15.7	15.7	15.6	15.7	1.97	0.90	1.84
sso0280 H_2O_2	30	15.5	15.6	15.6	15.6	1.97		
	60	15.4	15.4	15.4	15.4	1.97		
	90	15.0	15.2	15.0	15.1	1.97		
	120	14.7	14.7	14.9	14.8	1.97		
sso0771 control	30	17.6	17.6	17.4	17.5	1.64	1.73	2.36
	60	16.7	16.6	16.7	16.7	1.64	2.10	2.83
	90	16.2	16.3	16.1	16.2	1.64	1.97	2.65
	120	16.6	16.6	16.5	16.6	1.64	2.23	3.02
sso0771 H_2O_2	30	15.8	15.8	15.8	15.8	1.64		
	60	14.6	14.6	14.5	14.6	1.64		
	90	14.3	14.1	14.3	14.2	1.64		
	120	14.2	14.4	14.4	14.3	1.64		
ssso0121 control	30	19.4	19.4	19.5	19.4	0.89	2.03	0.79
	60	19.6	19.5	19.6	19.6	0.89	2.57	0.74
	90	19.0	19.3	19.3	19.2	0.89	2.37	0.76
	120	18.7	18.9	18.9	18.8	0.89	2.90	0.71
sso0121 H_2O_2	30	17.4	17.4	17.4	17.4	0.89		
	60	17.1	17.1	17.1	17.0	0.89		
	90	16.8	16.8	16.9	16.8	0.89		
	120	16.0	15.9	15.9	15.9	0.89		

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change
sso0446 control	30	15.3	15.3	15.2	15.3	1.68	0.73	1.46
	60	15.9	16.0	15.9	15.9	1.68	1.80	2.54
	90	15.6	15.5	15.6	15.6	1.68	0.90	1.60
	120	15.3	15.5	15.5	15.4	1.68	0.83	1.54
sso0446 H_2O_2	30	14.6	14.6	14.4	14.5	1.68		
	60	14.2	14.2	14.0	14.1	1.68		
	90	14.7	14.6	14.7	14.7	1.68		
	120	14.6	14.6	14.6	14.6	1.68		
sso0946 control	30	19.5	19.5	19.5	19.5	1.76	0.03	1.02
	60	20.1	20.0	20.1	20.1	1.76	0.63	1.43
	90	19.4	19.3	19.4	19.4	1.76	-0.23	0.88
	120	19.2	19.2	19.2	19.2	1.76	0.67	1.46
sso0946 H_2O_2	30	19.3	19.5	19.6	19.5	176		
	60	19.5	19.3	19.5	19.4	1.76		
	90	19.6	19.6	19.6	19.6	1.76		
	120	18.6	18.5	18.5	18.5	1.76		
ssso0959 control	30	14.9	15.0	14.9	14.9	1.67	1.17	1.82
	60	15.7	15.8	15.6	15.7	1.67	1.60	2.27
	90	15.5	15.8	15.8	15.7	1.67	2.43	3.48
	120	15.8	15.8	15.8	15.8	1.67	2.07	2.89
sso0959 H_2O_2	30	13.8	13.7	13.8	13.8	1.67		
	60	14.1	14.1	14.1	14.1	1.67		
	90	13.3	13.2	13.3	13.3	1.67		
	120	13.8	13.8	13.6	13.7	1.67		

Sample	Time (mins)	CP1	CP2	CP3	Mean CP	E	dCP	Expr change
sso1459 control	30	15.2	15.4	15.4	15.3	1.64	1.47	2.07
	60	15.8	15.8	15.8	15.8	1.64	2.30	3.12
	90	15.5	15.6	15.6	15.6	1.64	1.70	2.32
	120	15.4	15.4	15.2	15.3	1.64	1.60	2.21
sso1459 H_2O_2	30	13.8	13.9	13.9	13.9	1.64		
	60	13.4	13.6	13.5	13.5	1.64		
	90	13.8	13.8	14.0	13.9	1.64		
	120	13.8	13.6	13.8	13.7	1.64		
sso2079 control	30	17.6	17.5	17.6	17.6	1.99	8.50	346.94
	60	17.6	17.6	17.6	17.6	1.99	7.87	224.38
	90	18.1	18.3	18.3	18.2	1.99	7.30	151.92
	120	18.3	18.0	18.3	18.2	1.99	5.50	44.02
sso2079 H_2O_2	30	9.0	9.2	9.0	9.1	1.99		
	60	9.6	9.8	9.8	9.7	1.99		
	90	11.0	10.9	10.9	10.9	1.99		
	120	12.7	12.7	12.7	12.7	1.99		
ssso2364 control	30	20.4	20.8	21.2	20.8	1.88	-0.40	0.78
	60	20.9	20.1	20.7	20.6	1.88	0.33	1.23
	90	20.5	20.7	20.7	20.6	1.88	-0.13	0.92
	120	19.0	19.2	19.6	19.3	1.88	-1.87	0.31
sso2364 H_2O_2	30	21.4	20.9	21.3	21.2	1.88		
	60	20.3	20.0	20.4	20.2	1.88		
	90	21.0	20.6	20.7	20.8	1.88		
	120	20.9	21.2	21.3	21.1	1.88		