The role of McmD and the C-terminal of McmD in *Methanococcus maripaludis*

Juliet M. Borgia

MSc. by research

University of York
Biology

September 2014

Abstract

The minichromosome maintenance (MCM) proteins play an essential role in the replication of DNA, functioning as the genomic helicase in eukaryotes and archaea. In eukaryotes the MCMs comprise a family of six highly related proteins; MCM 2-7, which interact together to form heterohexameric ring complexes. The hexameric structure of MCM contributes to the highly complex and tightly regulated assembly and processivity of the replisome. Significant similarities can be observed between the structure of eukaryotic MCM 2-7 and homologous MCM complexes found in archaea, presenting an intriguing, simplified and biochemically tractable model. The majority of archaeal species have been shown to possess a single functional MCM that forms homohexameric complexes. In previous studies we have discovered that species in the order Methanococcales possess between two and eight MCM genes. Our model organism *Methanococcus maripaludis* S2 possesses four MCM genes (McmA, B, C & D), which may offer insight into additional motifs required for MCM function. We have generated deletions of individual and combinations of MCM genes, and determined that McmA appears to be essential. Previous studies have shown that McmD may be implicated in DNA damage response and the deletion of McmD causes a DNA damage repair defect. A series of experiments employing UV exposure and ionising radiation as agents of DNA damage were planned, designed to investigate possible roles the protein may play in DNA damage repair. This study was focused around McmD and the insert in the C-terminal of the protein; investigations found that, extrachromosomal overexpression of McmD caused colonies to become sick and possibly caused lethality to cells. Therefore 'rescue' of the McmD mutant under a strong constitutive promoter located on plasmid clones (pJB002/pJB003) could not be investigated. McmD gene (MMP1024) and MMP1025 is present in 16 Methanococcales species. MMP1024 and MMP1025 were included in operonic arrangements in 8 of the Methanococcales species analysed. It was possible to produce double and triple *M. maripaludis* MCM knockout strains (BDΔ, CDΔ, BCΔ, and BCDΔ). For purposes of the domain swap / expression of fusion protein investigation; a clone comprising the N-terminal of McmB and the C-terminal of McmD in commercial expression vector pPROEXHTa was produced. Future work would continue with cloning, co-expression, co-purification of MMP1024 and MMP1025 and biochemical characterisation of McmD.

Contents

List of Tables

List of Figures

Acknowledgments

I would like to thank James Chong for providing guidance, encouragement, support and enthusiasm throughout the project. I would also like to thank all members of Chong lab who contributed, Ally Walters, who's pioneering work with *M. maripaludis* in Chong Lab, knowledge and hard-work made this project possible, support and advice from Chitvan Amin (*nee* Bochiwal), Richard Parker, Emma Hesketh, Henry Nicholls, and Alex Leach was always gratefully received. Thanks also go to C. Hirst and S. Noble for their contributing work with the double knockouts. I would also like to thank Members of the Coverley Lab for their support and advice, Gillian Higgins, Rose Wilson and Dawn Coverley.

I would like to thank Prof. Ken Jarrell, who visited our lab during the beginning of the project, for offering a wealth of experience and knowledge and the invaluable opportunity of contributing to a review on the subject of archaea, cited in this project.

Thanks go to my friends and family for all the support they have offered me and to my husband Tom; for his patience, support, excellent IT skills, and taking care of our wonderful little daughter Aphra during the write-up. Thanks also go to Jon Borgia and Sue Shepherd for their support and contribution of funds towards this degree.

Authors Declaration

The work in this thesis is entirely my own, with the exception of referenced constructs, *M. maripaludis* MCM double knockouts, and probes used for Southern blots. Sequencing was carried out by Genomics, Technology Facility, University of York. None of the work described herein has been presented previously.

Chapter 1

Introduction

Archaea

The archaea are a diverse group of single-celled organisms phylogenetically distinct from bacteria and eukaryotes comprising the third domain of life. These organisms were first identified as being distinct from bacteria by differences in their 16S rRNA sequences (Woese and Fox, 1977). Formally known as 'archaebacteria' the archaea as they are now known, are divided into five phyla the korarchaeota, nanoarchaeota, thaumarchaeota, euryachaeota, crenarchaeota, of which the latter two have been most extensively studied (Luo et al., 2009).

Methanogenic archaea

Methanogens (Methane-producing) of the phylum euryachaeota, were the first group of archaea to be classified (Woese and Fox, 1977). Five orders of methanogen have been identified (Boone, 2001) including the Methanococcales, key to this study. The order Methanococcales encompasses two families, Methanocaldococcaceae and Methanococcaceae (Liu, 2010). The former are all hyperthermophilic, while the latter are extremely thermophilic or mesophilic. Members of this order are all capable of forming methane by CO_2 reduction with H_2 in anoxic conditions (Thauer et al., 2008); (Bapteste et al., 2005); (Weiss and Thauer, 1993); (Jarrell et al., 2011). Formate may also act as an electron donor in these organisms for review see (Weiss, 1993). Methanogens are common in wetlands, marshlands and are found in the digestive tract of animals and humans (Miller, 1986). some species are found at elevated temperatures in hot springs and deep sea hydrothermal vents (Chaban, 2006). Phylogenetic evidence proposed an ancient divergence between bacteria and archaea occurring prior to that of archaea from the later emerging eukaryotic domain, explaining the closer lineal link between the archaea and eukaryotes in comparison to bacteria. Methanogens are of an ancient lineage thought to have evolved ~3 billion years ago when the earth's atmosphere was still rich in hydrogen and carbon dioxide (Garrett, 2007)

Archaea as model organisms

Initial analyses in *Methanocaldococcus jannashii*, the first archaeon to have its complete genome sequenced (Bult et al., 1996), revealed that archaea possess many biochemical characteristics (e.g. chromosome repair and recombination) in common with bacteria however, archaeal genes involved with information processing are significantly more comparable to the equivalents in eukaryotic systems (Edgell and Doolittle, 1997). Although many similarities may be observed between archaeal and eukaryotic systems, the archaea usually employ simplified versions, leading to an increase in the

use of archaeal species as models for a catalogue of eukaryotic processes. (Costa et al., 2006); (Chong et al., 2000). DNA replication is a good example of how archaeal information processing systems are comparable but simplified versions of their eukaryotic counterparts.

DNA replication in the three domains

The process of faithful chromosome replication, cell division and proliferation is vital and common to all life. Cellular DNA replication occurs in a precisely controlled defined period, the 'S phase', ensuring the entirety of the genome is copied only once during each cell cycle. Cell cycle progression and S phase duration varies greatly between species, cell type, and developmental stages of an organism. In eubacteria rapid growth is fundamentally dependent upon micro-environmental conditions and genotoxic stress. During the cell cycle, DNA replication is carefully coordinated by mechanisms of positive and negative regulation. Positive regulation is necessary for initiation signalling and negative regulation ensures that origins are fired only once per cycle (Kelman and Kelman, 2003). DNA replication can be divided into three stages: initiation, the binding of initiation proteins to origins, duplex unwinding and the recruitment of replication apparatus; elongation, where DNA synthesis occurs; and termination, where replication folks collide and concatenated DNA molecules are separated into two daughter molecules.

Initiation of DNA synthesis

In all organisms, the unit of DNA replication, the replicon consists of a cis-acting element known as the replicator (origin of replication) and a trans-acting element known as the initiator (initiation protein complex) for review see (Bell and Dutta, 2002). Origins of replication have been identified to contain long AT-rich stretches with inverted repeat (IR) elements and purine / pyrimidine stretches (Boulikas, 1996); (Pearson et al., 1996).

In *E.coli*, a single origin oriC, contains 260bp five DnaA boxes downstream of an AT-rich region consisting of three 13-mer repeats (Messer, 2002). DnaA proteins bind to the five DnaA boxes in oriC, causing localising bending or 'melting' of the DNA due to torsional stresses (Gille and Messer, 1991). Two double hexamers of the bacterial helicase DnaB is loaded onto to DNA facilitated by DnaA and helicase loader DnaC. Once DnaB is loaded onto the DNA, DnaC disassociates from the complex facilitated by ATP hydrolysis, causing a conformational change which activates the helicase activity of DnaB. Both DnaB hexamers slide past each other in a 5' to 3' direction expanding the replication initiation bubble, for review see (Messer, 2002). DnaA disassociates from the origin as replication begins.

In budding yeast *Saccharomyces cerevisiae* multiple origins of replication have been identified as AT rich sequences between 150-200 bp containing a consensus of 11 bp known as the A-element or autonomously replicating sequences (ARS) (Fangman and Brewer, 1991); (Bell, 2002). ARS sequences are bound by the origin recognition complex (ORC) a complex of six related polypeptides Orc1-6 (Bell and Dutta, 2002). Minichromosome maintenance proteins MCM2-7 are recruited by ORC along with replication factors Cdc6 and Cdt1 (Barry and Bell, 2006). MCM associates with GINS complex and Cdc45, essential for the initiation and elongation phase of DNA replication (Gambus et al., 2006). DNA synthesis is initiated by the activation of pre-recognition complex by Cdc7 and Cdc28 protein kinases (Lei and Tye, 2001). In higher eukaryotes, replication initiation sites are not defined as repeat consensus sequences and are located broadly within particular chromosomal regions, known as 'initiation zones' for review see (Hamlin et al., 2008) however assembly of the preRC appears to follow a similar scheme to that repeated for yeast.

In eukaryotic systems the MCM2-7 complex is localised to the nucleus and binds to chromatin in an 'inactive' form during early G1 of the cell cycle. During S-phase on assembly of the replisome, MCM is converted to the 'active' form and DNA synthesis occurs. MCM remains bound to chromatin until late S-phase (Kuipers et al., 2011).

Minichromosome maintenance proteins (MCM)

Minichromosome maintenance proteins were first discovered in yeast mutant screens, where genes supporting the maintenance of small chromosome-like structure were defective (Maine et al., 1984). The heterohexameric minichromosome maintenance (MCM) complex is widely believed to be the replicative helicase in eukaryotes and archaea (Bell and Dutta, 2002), which unwinds genomic DNA prior to chromosome replication (Labib et al., 2000) (Prokhorova and Blow, 2000). In bacteria the replicative helicase is DnaB (an AAA+ protein) functioning as a homohexameric complex. The MCM complex comprises six highly related protein subunits MCM2-7 interacting together to form ring-like oligomeric heterohexameric complexes (Bochman and Schwacha, 2009); (Jenkinson and Chong, 2006). MCM proteins are classed as AAA+ N-loop NTPases, possessing the classic Walker A and Walker B motifs facilitating the binding and cleavage of NTPs. Functional MCM proteins possess two key hairpins in the secondary structure of the monomer; the first is located in the N-terminal domain and the second is known as the pre-sensor-1 β hairpin (PS1BH) located in the ATPase catalytic domain. MCMs are classified as members of the superclade PS1BH helicases, further categorised by the presence of a β-α-β insert in helix 2 between the Walker A and Walker B motifs (Chong, 2005). The primary sequence of MCM proteins can be divided into three major domains (See Figure 2), the N-terminal, AAA+ catalytic domain and the C-terminal. MCMs possess highly conserved functional

motifs, a zinc finger in the N-terminal; two hairpin structures; an arginine finger which have all been shown to be involved in the binding of DNA and the regulation of ATP hydrolysis (Chong, 2005). The function of the C-terminal domain of MCM is not fully understood. However, biochemical studies in *Methanothermobacter thermautotrophicus* MCM (*Mth*MCM) showed mutations in the C-terminal domain resulted in increased unwinding processivity, suggesting that this region of the protein is involved in regulation of the complex (Jenkinson and Chong, 2006). Winged helix-turn-helix motifs, an example of which is found in the C-terminal of MCM monomers, have been implicated in both DNA and origin binding in *Sulfolobus solfataricus* Cdc6 proteins (Robinson et al., 2004). There is evidence that all six proteins are required for function *in vivo* (Bochman and Schwacha, 2008). A range of different MCM complexes have been isolated from eukaryotic systems (Forsburg, 2004). Furthermore, in *vitro* biochemical evidence suggests containing MCM4, MCM6 and MCM7 have also been shown to possess helicase activity in *vitro*, suggesting that the complex is involved in the initiation of DNA replication (You et al., 1999). However, the complex nature of these helicases has proved challenging during comprehensive biochemical characterisation. In archaea (not including members of the order Methanococcales), a single MCM gene is present on the genome encoding for MCM proteins that interact together to form homohexameric complexes (Barry and Bell, 2006), the crystal structure of a near-full length archaeal MCM complex is illustrated in Figure 1, adapted from (Brewster et al., 2008). MCM2-7 and archaeal MCM complexes possess ATP dependant DNA helicase activity (Bochman and Schwacha, 2007); (Davey and O'Donnell, 2003); (Ishimi, 1997); (Jenkinson and Chong, 2006). Model species *Sulfolobus* and *Methanothermobacter* have been extensively studied in archaeal MCM characterisation. Motifs required for MCM function have been identified using ATP hydrolysis, DNA binding and DNA helicase activity assays (Barry et al., 2007); (Carpentieri et al., 2002); (Chong et al., 2000); (Jenkinson and Chong, 2006); (Kasiviswanathan et al., 2004); (Kelman et al., 1999); (Poplawski et al., 2001); (Shechter et al., 2000). Previous studies have shown that, of the eight species analysed in the order Methanococcales, all possessed between 2 and 8 MCM homologues (Walters and Chong, 2010). Furthermore, all possessed motifs required for function except for two truncated proteins which do not contain the arginine finger required for ATP hydrolysis. One of these proteins is found on an extra chromosomal element in *Methanocaldococcus jannaschii*, the other is on the main chromosome of *Methanococcus vannielii*. A third protein also shows significant changes, ORF 0961 of *Methanocaldococcus jannaschii* contains a large insertion between the first two cysteines of the zinc finger and lacks the second pair of cysteines.

DNA replication initiation in archaea

The archaeal DNA replication machinery has been shown to be strikingly similar to those present in eukaryotic systems (Barry and Bell, 2006). In most cases, the archaeal replicative apparatus

comprises simplified versions of eukaryotic equivalents e.g. origin recognition complex and MCM protein complexes. Nonetheless archaea appear to be far more closely related to eukaryotes than bacteria. Origins in archaea contain repeat sequences known as origin recognition boxes (ORBs), (Gaudier et al., 2007). Examples of multiple origins of replication in archaea have been identified e.g. in *Sulfolobus solfataricus* (Robinson et al., 2004). In most cases, origin-binding proteins in archaea are homologues of the related AAA+ proteins Orc1/Cdc6 (Barry and Bell, 2006). Once bound to the origin, Orc1/Cdc6 causes localised melting of the DNA. ATP bound Cdc6/Orc1 associates with the Cdc6/Orc1-origin complex and the MCM protein complex is recruited. ATP is hydrolysed and the Cdc6/Orc1 complex releases the helicase (Kelman and Kelman, 2003).

In most cases archaeal genomes include at least one homologue of Orc1/Cdc6 with the exception of the methanogenic species; *Methanocaldococcus jannaschii*, *Methanococcus maripaludis* and *Methanopyrus kandleri*, initiator proteins in these species remain unknown (Barry and Bell, 2006). In eukaryotes, MCM is loaded onto DNA in a process that requires ORC, Cdc6 and Cdt1. Origin binding proteins bind to DNA without the requirement of ATP hydrolysis, loading of MCM however does require ATP hydrolysis. As previously mentioned, in bacteria DnaB loading is facilitated by DnaA and DnaC. For archaea, little is known about MCM loading. Although, studies suggest a preliminary version of the complex may form around the DNA, as is the case with DnaB (Davey and O'Donnell, 2003); (Gomez-Llorente et al., 2005). In another study, the in *vitro* MCM2-7 complex in yeast appeared to show an ‘opening’ or ‘gate’ between MCM2 and MCM5 in the presence of duplex DNA, presenting another possible mechanism for MCM loading (Bochman and Schwacha, 2010); (Bochman and Schwacha, 2007). The understanding of the mechanism by which the MCM complex translocates along DNA remains incomplete, however binding of the MCM complex has been shown to be oriented so that the motor domain of the protein faces duplex DNA (McGeoch et al., 2005). The central channel of the MCM complex is large enough to accommodate dsDNA and ssDNA (Bochman and Schwacha, 2009), all functional complexes possess motifs involved with the binding of DNA e.g. the winged helix-turn-helix motif in the C-terminal domain, and the zinc-finger motif in the N-terminal domain.

Why study MCMs?

MCM proteins are fundamentally and intrinsically linked with core cellular processes such as DNA replication and cell cycle progression, attracting broad scientific interest in recent years. An increased understanding of these proteins could also lead to medical innovation; various studies have suggested that several subunits of the MCM2-7 complex may be useful diagnostic and prognostic biomarkers in a wide range of human cancers (Gonzalez et al., 2005); (Lei, 2005). MCMs

are good potential biomarkers because they are present in both actively proliferating cells and cells which are licensed for replication. This means they can be used to identify precancerous cells prior to malignancy. MCMs are absent in quiescent , differentiated and senescent cells (Musahl et al., 1998); (Stoeber et al., 2001), but are highly abundant, stable and immunogenic in cycling cells. MCMs 2, 3, 5, and 7 have all been shown to be good biomarkers in a wide range of cancers (Freeman et al., 1999).

***M. maripaludis* S2 possesses four MCM homologues; ORFs MMP0030-McmA, MMP0470-McmB, MMP0748-McmC, MMP1024-McmD**

The mesophilic species *Methanococcus maripaludis* S2 served as the model species in this study. This species presents a particularly interesting subject as it has been found to possess multiple MCM genes (McmA, -B, -C and –D) (Walters and Chong, 2010). An alignment of all four *M. maripaludis* sequences of the motifs required for function is detailed in Figure 4. This organism was relatively easily manipulated in the laboratory, was cultured at 37^{o}C, a fully sequenced genome (Hendrickson et al., 2004) and a range of genetic and biochemical tools were available in support of analysis. Phylogenetic analyses showed there are at least two MCM homologues that play important functional roles in the Methanococcales, McmD and McmA (Walters and Chong, 2010). These two homologues are present in the eight species analysed and appear to have been created by an ancient duplication pre-dating the emergence of modern Methanococcales species. This ancient divergence appears to have given rise to two major groups of MCMs. One of these groups contains McmD from *M. maripaludis* S2 and was given the name McmD group. All eight species analysed in this group contained an identified conserved 20 amino acid insert which was both of similar size and position of an insert contained within MCM3. Interestingly this insert has been implicated in the regulation of the eukaryotic MCM2-7 complex (Lin et al., 2008); (Shi et al., 2007); (Takei et al., 2002) so it is possible that this insertion plays a role in regulating MCM activity in the Methanococcales. This highly conserved insert is not found in any other archaeal MCM implicating the Methanococcales as key species for improved understanding of heterohexameric MCM complexes going forward. The second group has undergone further duplications creating several subgroups of MCMs, this group contained McmA from *M. maripaludis* S2 and was therefore named the McmA group. Full-length sequences of all Methanococcales MCMs were compared. The comparison showed that the MCMs falling outside the McmA and McmD groups were found close to an integrase gene. An 83 amino acid ORF was present upstream of the integrase gene in several MCM regions and was also found in the cryptic plasmid pURB500 which was first isolated from *M. maripaludis* C5 (Tumbula et al., 1997). In most cases the integrase homologue was annotated as a phage integrase (Walters and Chong, 2010), and the emergence of *M. maripaludis* McmB –C was

attributed to phage mediated events and mobile elements. It is unclear why MCM sequences have been conserved following duplication in the Methanococcales. It could be that they are conserved but that only a single MCM is functional as is the case in other archaea. However shotgun proteomics analysis carried out in *M. maripaludis* S2 detected peptides for McmA –B and –D indicating that multiple MCMs are expressed (Xia et al., 2006). Peptides were not detected for McmC, but this is not to say that McmC is not expressed. It could be that expression occurs under specialised conditions. All four *M. maripaludis* MCMs have been conserved following ancient duplications and all retain motifs required for function suggesting either, that these duplications occurred recently in evolutionary terms and have not had sufficient time to accumulate mutations, or the presence of multiple functional MCMs is somehow advantageous to the cells. In the case of the integrase associated MCM, functional motifs may be required for plasmid maintenance, requiring a functional MCM encoded on either an extra chromosomal element or integrated on the genome. Another explanation for the presence of conserved multiple MCM is that they provide functions which are beneficial to chromosome replication as may be the case for non-integrase related MCMs, McmA and –D. It is possible that multiple MCM are conserved in the Methanococcales because they have diversified to perform specialised roles similarly to those observed in eukaryotes. Eukaryotic MCM3, most closely related to McmD, appears to have subtly different roles in the regulation of helicase loading and unwinding (Sheu and Stillman, 2006); (Tsuji et al., 2006) and potentially could be implicated in the triggering of damage checkpoints (Cortez et al., 2004); (Tsao et al., 2004). In order for cells to proliferate, the genome must be faithfully replicated and segregated in order to support viable daughter cells. Chromosomally damaging agents for example; UV rays and ionising radiation can interfere with these processes resulting in the introduction of lethal levels of mutation. Unpublished in-house data suggests the deletion of *M. maripaludis* McmD (McmDΔ mutant) results in greater sensitivity to UV exposure. This 'damage phenotype' showed approximately two orders of magnitude higher occurrence in cell death when exposed to 30 Joules m^2 in comparison with the wildtype. There did not appear to be a significant difference between the wildtype and the other mutants in similar experiments (McmBΔ, –CΔ). It is possible that McmD provides a regulatory role in this species interacting with repair proteins in complex damage response checkpoint pathways. An alignment of all 4 *M. maripaludis* MCM sequences reveals eight inserts are present in McmD which are absent from the rest, the longest of which is a 20 amino acid insert in the C-terminal (see Figure 20. & Figure 21). As previously mentioned this insert is of a similar length and position to that of an insert conserved in MCM3 which is implicated in the regulation of MCM2-7 complex. A genetic, biochemical and bioinformatic approach has been employed in order to investigate possible specialised roles of *M. maripaludis* S2

McmD in DNA damage checkpoint pathways. This study also aimed to biochemically characterise purified *M. maripaludis* S2 recombinant proteins, namely McmD with the scope of investigating helicase, ATPase in *vitro* activity and complex stoichiometry within the limits of gel filtration assays (SEC-MALLS).

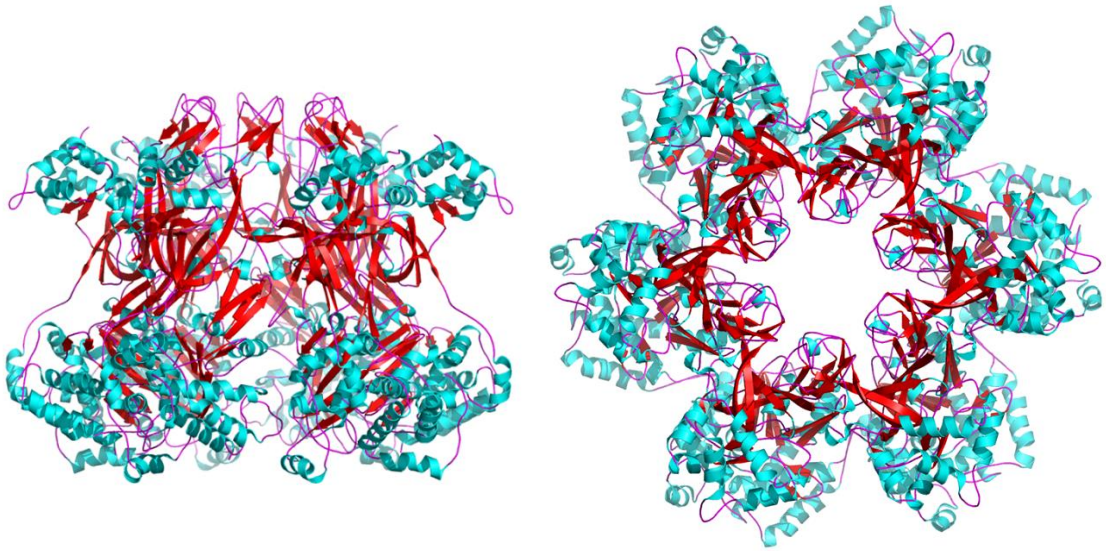

Figure 1. Crystal structure of a near –full- length archaeal MCM

Adapted from (Brewster et al., 2008), redrawn in PyMOL. Individual MCM subunits can be seen interacting together forming the characteristic 'ring-shaped' complexes. Side view (left) top view (right).

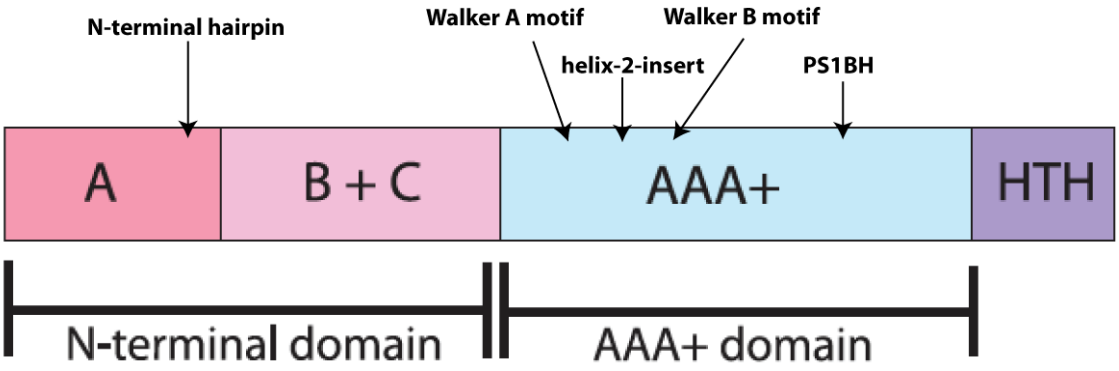

Figure 2. Domain organisation of MCM monomer.

The N-terminal consists of three domains, A, B and C. The N-terminal hairpin located in domain A is labelled. The AAA+ catalytic domain is highlighted in light blue. Walker A, Walker B motifs, the helix-2-insert and the PS1BH are labelled. The helix-turn-helix motif in the C-terminal is highlighted in violet, adapted from (Barry and Bell, 2006).

Previous work with *M. maripaludis* McmDΔ strain shows acute sensitivity in the presence of UV rays (Walters, 2010 unpublished). Figure 3. (top left), shows the effects of exposing *M. maripaludis* MCM single knockout strains to varying doses of ultra violet rays (Joules per m^2). Figure 3. (top right), shows the results of a similar experiment with ionising radiation as the source of genotoxic stress. Strains McmBΔ and McmCΔ show similar responses to wildtype in the presence of DNA damaging agents, However the McmDΔ strain shows greater and more pronounced occurrence of cell death, particularly when it is considered the results are shown on a log scale. This could indicate that the deletion of McmD has a fundamental effect on DNA damage repair, revealing a damage sensitive phenotype.

Figure 3., flow cytometry figure (bottom) shows a comparison of McmBΔ (first two columns to the left), CΔ (middle two columns), DΔ (last two columns, right) *M. maripaludis* knockout strains in terms of cell size and DNA content. The black trace shows the mutant curve, with the grey area representing the wildtype curve. The McmDΔ mutant curve shows the most deviation from wildtype. Cell size and DNA content are significantly increased in the mutant indicating continual replication cycles possibly unregulated by cell cycle checkpoints. Eventually resulting in cell death as unsuccessful chromosome segregation and cell division has occurred. These data suggest the McmDΔ strain possesses poor genome stability in comparison with the other mutant strains (McmB -CΔ) which may present an interesting target for investigations into cell cycle regulation in this species.

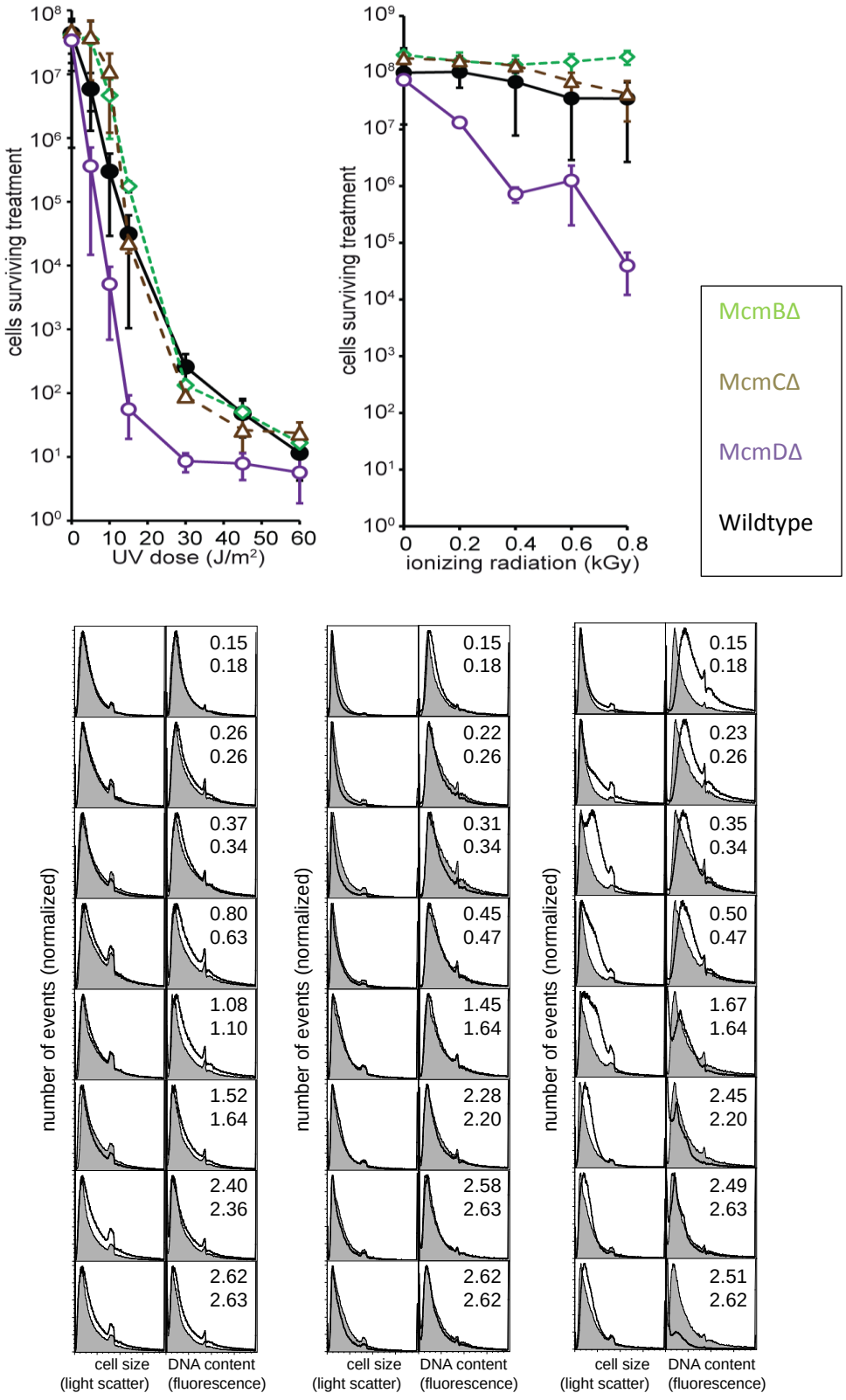

Figure 3. UV / ionising radiation exposure (Top) and flow cytometry data

flow cytometry data: Left to right, McmBΔ, McmCΔ, McmDΔ, single M. maripaludis knockout mutants. (unpublished, Walters, A.D., Chong, J.P.J).

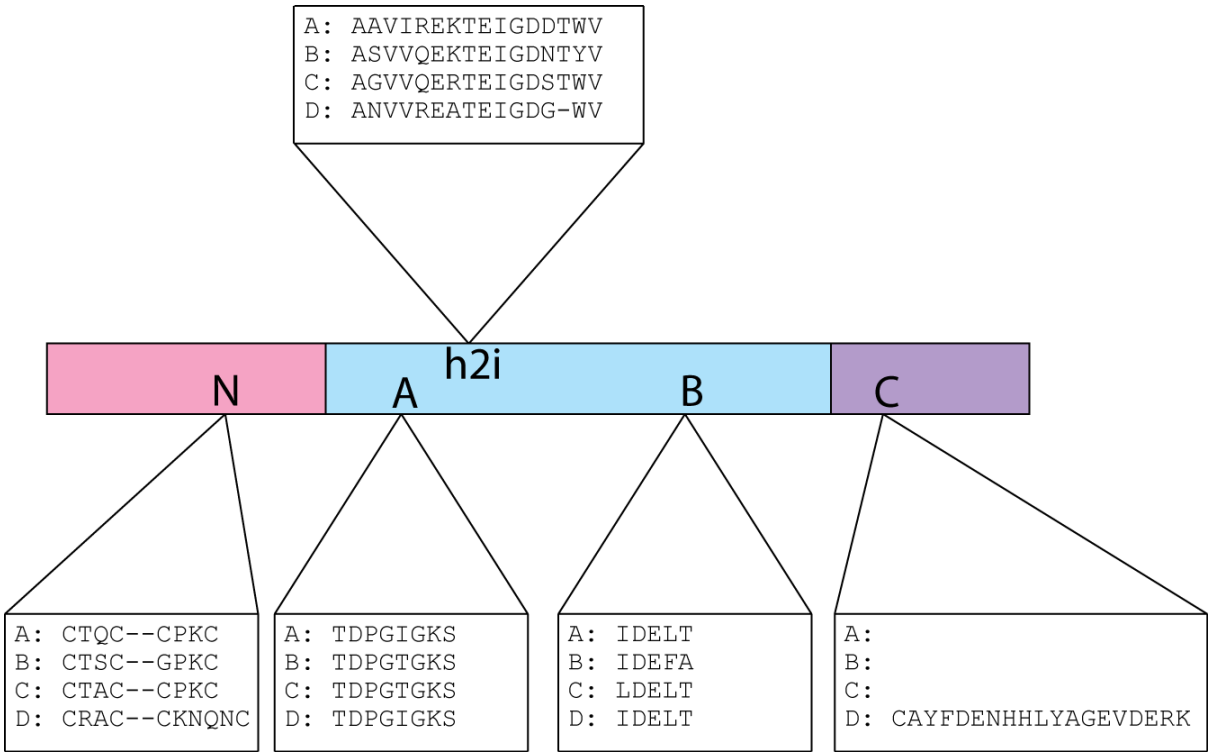

Figure 4. Domain organisation of MCM monomer (*M. maripaludis* S2)

Highlighted sequences alignments show a comparison of functional motifs (McmA-D) from left to right; N-terminal hairpin, Walker A Motif, helix-2 insert, Walker B motif, 20-amino acid insert present in McmD sequence only (C-terminal).

Chapter 2

Methods

Production of constructs

For the purposes of this study, *M. maripaludis* McmD (full-length gene) and $McmD_{20}$ (McmD ORF including a 20 amino acid in-frame deletion at the C-terminal) coding regions were amplified from holding plasmids pAW22 and pAW23 (Walters, unpublished) for continued manipulation in cloning strategy.

pJB001:

$McmD_{20}$ was amplified with Platinum ® Pfx DNA Polymerase (Life Technologies) from holding vector pAW23 (Walters, unpublished) with primers (Eurofins) including engineered restriction sites for restriction enzymes PspOMI (NEB) at the N-terminal and AvrII (NEB) at the C-terminal (primer 1. see Table 3). Amplicon was gel-purified on 1% agarose DNA gel, band excised and PCR components were removed with PCR clean-up kit 'QIAquick' (Qiagen). Blunt-ended PCR product was treated with Taq DNA Polymerase (NEB) and dATP to create 'A-tails'. Insert was cloned in pGEMT (Promega) and transformed in competent Nova Blue *E.coli* cells. 100µl of transformation mixture was spread-plated on LB agar *amp*, and growth overnight at 37^{o}C. Individual colonies were isolated, and cultured overnight at 37^{o}C with shaking (~200 r.p.m) in culture tubes (Thermo Scientific). Plasmids were purified using (Nucleospin ®Plasmid miniprep kit – Macherey Nagel)* prior to diagnostic restriction digests and sequence confirmation (primers for sequencing pGEMT-insert junctions were provided by Genomics, Technology Facility, University of York).

McmD was amplified with Platinum ® Pfx DNA Polymerase (Life Technologies) from holding vector pAW22 (Wallters, 2009) with primers (Eurofins) including engineered restriction sites for restriction enzymes PspOMI (NEB) at the N-terminal and AscI (NEB) at the C-terminal (see Table 3). Several attempts at cloning McmD in pGEMT were made but were unsuccessful. pAW22, from a previous project was eventually used in the cloning of pJB003.

pJB002

Following correct sequence confirmation of $McmD_{20}$, insert was amplified with Phusion® High fidelity DNA polymerase (NEB) from pJB001 with primer 2. (see Table 3.). Insert was gel purified (1% agarose), PCR components removed with PCR clean-up kit 'QIAquick' (Qiagen). A sample of purified pAW42 (vector) was prepared (streaked from glycerol, approx. 5x 5ml overnight cultures prepared, plasmid purified using Nucleospin ®Plasmid miniprep kit – Macherey Nagel), measured for DNA

concentration and cleaved at multiple cloning site with restriction enzymes PspOMI and AvrII (NEB) complementary to those of the insert. A ratio of 1:1 insert to vector ligation reaction was prepared. Ligation mixture was transformed in competent Nova Blue *E. coli* cells. 100µl of transformation mixture was spread-plated on LB agar *amp*, and growth overnight at 37^{o}C. Individual colonies were isolated, and cultured overnight at 37^{o}C with shaking (~200 r.p.m) in culture tubes (Thermo Scientific). Plasmids were purified using (Nucleospin ®Plasmid miniprep kit – Macherey Nagel)* prior to diagnostic restriction digests. Adjoining junctions between vector and insert were analysed and confirmed with primers 6. & 7. (see Table 3.).

pJB003

McmD coding region was amplified with primer 1. (see Table 3.) with Phusion® High fidelity DNA polymerase (NEB) from holding vector pAW22. Insert was gel purified (1% agarose), PCR components removed with PCR clean-up kit 'QIAquick' (Qiagen). A sample of purified pAW42 (vector) was prepared, (streaked from glycerol, approx. 5x 5ml overnight cultures prepared, plasmid purified using Nucleospin ®Plasmid miniprep kit – Macherey Nagel), measured for DNA concentration and cleaved at multiple cloning site with restriction enzymes PspOMI and AscI (NEB) complementary to those of the insert. A ratio of 1:1 insert to vector ligation reaction was prepared. Ligation mixture was transformed in competent Nova Blue *E. coli* cells. 100µl of transformation mixture was spread-plated on LB agar *amp*, and growth overnight at 37^{o}C. Individual colonies were isolated, and cultured overnight at 37^{o}C with shaking (~200 r.p.m) in culture tubes (Thermo Scientific). Plasmids were purified using (Nucleospin ®Plasmid miniprep kit – Macherey Nagel)* prior to diagnostic restriction digests. Adjoining junctions between vector and insert were analysed and confirmed with primers 6. & 7. (see Table 3.).

pJB004

A fusion construct was produced by amplifying the N-terminal (and catalytic) domain of McmB and the C-terminal only of McmD (see Figure 17., Figure 18.). Oligos spanning the junctions between vector and McmB at the N-terminal and vector and McmD at the C-terminal along with the junctions between McmB and McmD adjoining regions were used in order to amplify inserts of interest. The PCR reaction produced both inserts which were purified and ligated in commercial protein expression vector pProEXHTa (Invitrogen) using In-fusion HD cloning kit (Clontech®). In-Fusion HD Cloning Kits were employed for rapid, directional cloning of multiple fragments of DNA into appropriate linearised vector (pProEXHTa). Target fragments (including 15bp overhangs, complementary with correct directional positioning of inserts in vector) were cloned in pProEXHTa

facilitated by homologous recombination. Nicks and breaks in recombinant strands were repaired in *vivo*.

Vector pProEXHTa (Invitrogen) was digested with restriction enzymes NcoI and KpnI, and gel-purified (Nucleospin gel clean-up kit, Macherey Nagel). His-tagged McmB N-terminal and His-tagged McmD C-terminal inserts were amplified by PCR (Phusion® High fidelity DNA polymerase, NEB) with primers 8., 9., 10. & 11., (see Table 3.), including 15bp extensions complementary to the ends of linearised vector and adjacent insert. Inserts were gel purified, PCR components removed with PCR clean-up kit 'QIAquick' (Qiagen). Inserts and linearised vector were combined in cloning reaction (In-Fusion® HD cloning kit, Clontech), incubated for 15 min at 50^{o}C and placed on ice. Reaction mixture was transformed in XL 10 gold (Stratagene) competent *E. coli* cells. 100µl of transformation mixture was spread-plated on LB agar *amp*, and growth overnight at 37^{o}C. Individual colonies were isolated, and cultured overnight at 37^{o}C with shaking (~200 r.p.m) in culture tubes (Thermo Scientific). Plasmids were purified using (Nucleospin ®Plasmid miniprep kit – Macherey Nagel)* prior to diagnostic restriction digests. Adjoining junctions between vector and inserts were analysed and confirmed with primers 12., 13., & 14. (see Table 3.).

*Prepared plasmid samples were eluted in distilled, autoclaved water, stored at 4^{o}C.

Protein expression of *M. maripaludis* S2 MCMs

M. maripaludis S2 MCM coding regions were amplified by PCR from genomic DNA and cloned into protein expression vector pProEXTHa (Invitrogen). Clones were transformed into *E. coli* Arctic Express RIL competent cells (Stratagene). Cells were grown with shaking (200 r.p.m) at 37°C until the OD_{600nm} reached 0.8 and then chilled on ice for 20 min before induction with 0.1mM IPTG. Cultures were incubated for a further 24h at 12°C with 200 r.p.m shaking before cells were harvested. Cells were resuspended in 2 ml g^{-1} lysis buffer (25 mM Tris, pH 8.0, 150 mM NaCl, 10% v/v, glycerol, 1 mM PMSF, ^{1} µg leupeptin ml^{-1}, 1 µg pepstatin ml^{-1}, 1 µg aprotonin ml^{-1}), supplemented with 0.75 mg lysozyme ml^{-1} and 5 µg DNase ml^{-1}. Cells were lysed by sonication and the lysate was clarified by centrifugation at 50,000 g for 30 min.

Protein purification of *M. maripaludis* S2 MCMs

1ml Co^{2+} resin was washed with 5 ml lysis buffer, centrifuged at 3220 g for 3 min. The buffer was discarded, process repeated twice more. Cleared lysate was added and mixed continually for 30 min at 4°C. The mixture was centrifuged at 3220 g at 4°C for 3 min. The supernatant was retained as 'flow through' sample and frozen at -80°C. All performed at 4°C; the resin was resuspended in 2 ml lysis buffer and transferred to a column (Pierce). The resin was washed with 10ml wash buffer (50

mM Tris, pH 8.0, 300 mM NaCl, 5% v/v, glycerol, 10 mM imidazole, 0.1 mM PMSF, 1 µg pepstatin ml^{-1}, 1 µg leupeptin ml^{-1}, 1 µg aprotinin ml^{-1}).

Flow-through sample was retained and stored at -80°C. Protein was eluted in 10x 1 ml fractions with elution buffer (50 mM, Tris pH 8.0, 300 mM NaCl, 5% glycerol, 150 mM imidazole, 0.1 mM PMSF, 1 µg pepstatin ml^{-1}, 1 µg leupeptin ml^{-1}, 1 µg aprotinin ml^{-1}), fractions were snap frozen in liquid N_2 and stored at -80°C prior to SDS page DGGE analysis.

Transformation of recombinant constructs in *M. maripaludis* S2

M.maripaludis cultures (to be transformed) were grown anaerobically to an optimal OD_{600nm} of 0.7 – 1.0 and pressurised to 30 psi with $4H_2$ CO_2. 5µg of plasmid DNA in TE buffer (and equivalent volume of blank TE buffer) were stored in anaerobic conditions ≥ 2 hours prior to transformation. Cultures were centrifuged at 1500g at room temperature for 15 min. Supernatant was removed through stopper with surgical needle. Pellet was resuspended in 5 ml transformation buffer and re-pressurised to 30 psi with $4H_2$ CO_2. Cultures were centrifuged at 1500g at room temperature for 15 min. Supernatant was removed. Pellet was resuspended in 0.375 mL transformation buffer. Cultures were transferred to anaerobic chamber (AC), stopper was removed, plasmid and TE blanks samples were added, cultures were mixed, stopper replaced and removed from AC. 0.225 ml 40% PEG was added. Cultures were flushed with 100% N_2 for 1 min and pressurised with 100% N_2. Cultures were incubated at 37^{o}C for 1 hour. Cultures was transferred to 5 ml fresh media containing sulphide and pressurised to 30 psi with $4H_2$ CO_2. Cultures were centrifuged at 1500g at room temperature for 25 min. Supernatant was removed, 5 ml fresh media containing sulphide was added. Cultures were flushed with $4H_2$ CO_2 and pressurised to 30 psi with $4H_2$ CO_2. Cultures were incubated at 37^{o}C with gentle shaking overnight. After overnight recovery, cultures were transferred to media containing puromycin selection (2.5 µg/ml). After 2-3 days cultures were plated on solid McCas media containing puromycin selection.

Under anaerobic conditions, individual colonies were picked into McCas media containing puromycin selection, grown to an optimal OD_{600nm} of 0.7-1.0, sub cultured with no selection and retained as glycerol stock.

Previous data showed varying responses to genotoxic stress in the form UV exposure between the single knockouts *M. maripaludis* S2 McmBΔ, McmCΔ and McmDΔ. The procedure was performed thus;

UV irradiation procedure of *M. maripaludis* S2 test strains

UV irradiation procedure of McmDΔ strains containing rescue constructs pJB002, pJB003 was not used due to unsuccessful transformation of test constructs pJB002, pJB003, added for completeness.

Constructs pJB002, pJB003 were transformed in *M. maripaludis* S2 McmDΔ. Transformed cells were subjected to 10-fold serial dilutions (10^{-0} -10^{-6}) and then 4-µl aliquots were spotted onto McCas solid media with puromycin selection (2.5 µg/ml). Under dark conditions, test colonies were treated with UV doses varying in intensity from 0 J/m^2– 35 J/m^2 prior to incubation in pressure vessels containing H_2/CO_2 (4:1) at 20 lb/in^2 for 3 or 4 days at 37°C.

Standard UV irradiation procedure of *M. maripaludis* S2 test strains

Under anaerobic conditions, mutant strains were revived from glycerol in McCas media + 25µg/ml Na_2S and grown to an OD_{600nm} of 0.8-1.00 in a shaking water bath set to 37°C. Test strains were subjected to 10-fold serial dilutions (10^{-0}– 10^{-6}), 4-µl aliquots were spotted onto McCas solid media. Test colonies were treated with UV exposure varying in intensity from 0 J/m^2– 35 J/m^2 before being incubated in a pressure vessel containing H_2/CO_2 (4:1) at 20 lb/in^2 for 3 or 4 days at 37°C.

Deletion of McmC in knockout mutant McmBΔ, deletion McmB in double knockout mutant McmCDΔ

Preparation of knock-in constructs (constructs prepared by A. Walters, used in study). *M. maripaludis* S2 MCM (McmC, -B) was amplified including sufficient flanking regions ~500bp on each side to mediate homologous recombination. Insert was cloned in pGEMT (Invitrogen). A PCR reaction was set up which amplified the edge of the target gene, the flanking DNA and vector, without the internal region of the gene. The ends of the PCR were ligated forming a circular plasmid containing the MCM read through sequence, with 500bp flanking regions up and downstream. Construct was sub-cloned from pGEMT into *M. maripaludis* transformation plasmid pCRPrtNeo (Moore and Leigh, 2005) using restriction enzyme NotI (NEB).

Preparation of MCM knockout Mutants in *M. maripaludis* S2 (Mm900)

Markerless mutagenesis in M. maripaludis

M. maripaludis transformation plasmid pCRPrtNeo (Moore and Leigh, 2005) contained two selectable markers for neomycin (*neo*) and 8-azahypoxanthine (*hpt*). Clones were selected with *neo* (500µg/ml on plates). This resulted in a single recombination event and a merodiploid that contained both the wildtype and the mutant copies of the gene with the vector in between. Colonies were streak purified on *neo*, colonies were picked into McCas + no selection, to allow for looping out. Colonies were grown overnight on solid McCas + *hpt* (250µg/ml). This selected for strains in

which a second recombination event had eliminated the vector, several putative mutants colonies were inoculated into McCas broth and were screened. Appropriate colonies were streak purified onto fresh McCas agar. Southern blots were performed to verify no vector backbone was left behind and only the mutant copy of the gene was present. A result of 50:50 ratio of mutant to wildtype was expected.

LIC Duet Cloning of ORFs MMP1024 2133bp and MMP1025 504bp

ORFs MMP1024 and MMP1025 were amplified by PCR from genomic *M. maripaludis* DNA (primers 15 and 16 see Table of primers), with PrimeSTAR GXL DNA Polymerase, Takara. Inserts were gel purified, PCR components removed with PCR clean-up kit 'QIAquick' (Qiagen). Purified PCR product was stored in buffer 10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). Target sequences were treated with T4 DNA polymerase; components were combined on ice, the reaction was started by incorporating the enzyme, the reaction was incubated at 22^{o}C for 30 min. The enzyme was inactivated by incubating the reaction at 75^{o}C for 20 min. Treated inserts were annealed with LIC Duet minimal adaptor kit and pRSF LIC vector (LIC Duet™Adaptor kits, Novagen). LIC Duet adaptors were supplied possessing LIC overhangs therefore treatment with T4 DNA Polymerase was not required. Annealing reaction was prepared, incubated at 70^{o}C for 30 seconds, placed at room temperature for 2 min to cool, 1µl 25mM EDTA was added for a final volume of 7µl. The reaction was incubated for 5 min at 22^{o}C. Annealing mixture was transformed in XL 10 gold (Stratagene) competent *E. coli* cells recombinant plasmids were isolated and purified from individual colonies (Nucleospin ®Plasmid miniprep kit – Macherey Nagel) prior to diagnostic restriction digests and sequence confirmation.

Methods for co-expression, detection, purification and quantification of target proteins made possible by the LIC construct produced in this study has not been included due to the project being limited to the cloning stages. Recommended protocols are obtainable from Novagen technical bulletins (see Technical Bulletin 163 for a listing).

Standard transformation procedure for recombinant constructs in *E. coli*

30 – 50 ng of target plasmid was added to 50 µl competent *E. coli* culture (e.g. Nova Blue, XL 10 gold-Stratagene, Rosetta). Reaction was incubated on ice for 15 min, transferred to a hot block and incubated for 30 seconds at 42^{o}C. Reaction was transferred on ice for a further 2 min. 250 µl SOC media was added to reaction tube, incubated in shaking water bath for 1-2 hours at 37^{o}C. 100 µl reaction mixture was plated on solid media with appropriate selection (e.g. ampicillin), incubated overnight at 37^{o}C.

Preparing anaerobic solutions

Anaerobic Na_2S solutions were prepared by rinsing Na_2S crystals with distilled H_2O, then weighing them before taking them into the anaerobic chamber where they were dissolved in anaerobic water in 100 ml glass vials to make 2.5% and 25% solutions. Vials were sealed using butyl stoppers (Bellco Glass Co.) then crimped with aluminium caps (Wheaton).

All other solutions were made anaerobic by placing them in glass vial that could be sealed using 20 mm butyl stoppers. A needle attached to the gassing manifold was placed through the butyl stopper and a short length of thin-walled PTFE tubing (Cole Palmer) was placed onto the needle. The stopper was placed on the vial, allowing the rubber tubing to reach the bottom of the solution. N_2 was passed through the solution at low pressure (10 psi) and a second open-ended needle was placed through the stopper to act as a gas outlet. Solutions of 100 ml or less was gassed for up to 3 hours. Solutions of greater than 100 ml were gassed overnight. After gassing the outlet needle was removed before removing the gassing needle and the vial was crimped with an aluminium cap.

Media preparation

E. coli

Luria Broth (LB) Added to 1 litre distilled, autoclaved water 10g Tryptone, 10g NaCl, 5g yeast extract, autoclaved and cooled.

LB agar Added to 1 litre distilled, autoclaved water 10g Tryptone, 10g NaCl, 5g yeast extract, 15g agar autoclaved and cooled.

SOC media Added to 1 litre distilled, autoclaved water 20g Tryptone, 5g yeast extract, 2ml 5M NaCl, 2.5ml 1M KCl, 10ml 1M MgCl2, 10ml MgSO4, 20ml 1M glucose autoclaved and cooled.

M. maripaludis

McCas 250ml water, 250ml general salts solution, 2.5g $NaHCO_3$, 11g NaCl, 5ml $K2HPO_4$ solution,2.5ml $FeSO_4$ solution,0.5ml trace vitamins (1000X), 5ml Vitamin solution (100X), 0.5ml Resazurin solution, 0.7g NaAcetate 3H2O, 1g casamino acids, *0.1g DTT, 0.25g cysteine H_2O HCl (100 µl 2.5% Na_2S added prior to use).

*Added after other components

All tubes and stoppers were in anaerobic were placed in anaerobic chamber for a minimum of 1 hour prior to medium dispensing.

Components were combined in large stoppered boiling vessel, under a stream of 20:20 N_2CO_2. DTT was added just before boiling, 0.25g cysteine H2O HCl added just after boiling. Solution was swirled until colour change from pink to clear occurred. Solution was continually streamed with 20:20 N_2CO_2 whilst allowed to cool (5-10 min). Gassing canular was removed and stopper taped down, media was placed in anaerobic chamber. 5ml of medium was dispensed per tube, swirled occasionally to keep precipitate evenly suspended. Rubber stopper were inserted into the tubes, all tubes were removed from the anaerobic chamber. All tubes were sealed with metal crimps, gas in tubes was exchanged from 20:20 N_2CO_2 to 80:20 N_2CO_2, removed by vacuum pump and replaced three times. Tubes were filled with 80:20 N_2CO_2, pressurised to 12 psi and autoclaved.

Solid McCas

Components were combined in large stoppered conical. 1.5 g Difco Noble agar was added per 100ml of media required. 0.1g DTT and 0.25g cysteine H_2O HCl was added. Media was autoclaved, transferred to a 40^oC water bath under a stream of 100% N_2 and gassed for a minimum of 3 hours. Media (stoppered) was transferred to anaerobic chamber. Na_2S and appropriate selection was added. Plates poured in anaerobic chamber and allowed to set.

Standard Methanococcus reagents (all solutions prepared anaerobically)

General salts solution Added to 1 litre autoclaved, distilled water KCl 0.67 g $MgCl_2$·6H2O 5.5 g $MgSO_4$·7H_2O 6.9 g $CaCl_2$·2H_2O 0.28 g NH_4Cl 1g (omitted from N-free stock).

K2HPO4 solution Added to 1 litre autoclaved, distilled water K_2HPO_4 14g

FeSO4 solution Added to 100 ml of 10 mM HCl; $FeSO_4$·7H_2O 0.19g

Trace vitamins (1000X) Added to 100 ml autoclaved, distilled water Na_3Citrate·2H_2O 2.1g, adjust pH to 6.5 $MnSO_4$·2H_2O 0.5g $CoSO_4$ or $CoCl_2$ (·6H_2O) 0.1g $ZnSO_4$ (·7H_2O) 0.1g $CuSO_4$·5H_2O 0.01g AlK (SO_4)2 0.01g H_3BO_4 0.01g Na_2MoO_4·2H_2O 0.1g $NiCl_2$·6H_2O 0.025g Na_2SeO_3 0.2g V(III)Cl 0.01g Na_2WO_4·2H_2O 0.0033g

Vitamin solution (100X) Added to 1 litre autoclaved, distilled water biotin 2 mg folic acid 2 mg pyridoxine HCl 10 mg thiamine HCl 5 mg riboflavin 5 mg nicotinic acid 5 mg DL-calcium pantothenate 5 mg vitamin B12 0.1 mg p-aminobenzoic acid 5 mg lipoic acid 5 mg

Rezasurin solution Added to 1 litre autoclaved, distilled water, Rezasurin 1g

Sulphide solution 2.5% (w/v) Na_2S·9H_2O crystals were rinsed with dH_2O blot dried and weighed, Dissolved in anaerobic dH_2O.

Transformation buffer 50 mM Tris, 0.35 M sucrose, 0.38 M NaCl, 1 mM MgCl2, 0.00001 % rezasurin, adjusted to pH 7.5

PEG 40% solution Transformation buffer with 40% (wt/vol) PEG8000.

Inoculation method and growth conditions for *M. maripaludis*

Small scale growth in liquid media was carried out in standard volume of 5 ml. Additions to medium tubes were made using syringes that were pre-gassed with N_2 and small gauge needles to in order to maintain anaerobic conditions. Immediately prior to inoculation, 0.1ml of 2.5% Na_2S (w/v) was added to each tube, appropriate selection or base analogues were added. 0.1-0.5 ml of cell culture was used to inoculate each tube. After inoculation, culture tubes were pressurised to 40 psi using H_2/CO_2 (80:20) ratio and then incubated at 37^{o}C overnight in a water bath with lateral shaking at 110 r.p.m.

Antibiotics

E coli

All antibiotic stock solutions were filter sterilised before being added to growth media. Antibiotics were added to liquid and solid media to give the following final concentrations: ampicillin 100 µg/ml, kanamycin 30 µg/ml.

M. maripaludis

All antibiotic and base analogue solutions were filter sterilised and made anaerobic before being added to growth media. Neomycin was added to a final concentration of 1 mg/ml in solid medium and 0.5 mg/ml in liquid medium. 8-azahypoxanthine was dissolved in 100 mM NaOH and heated to 60^{o}C before each use to ensure it was in solution. 8-azahypoxanthine was added to a final concentration of 0.25 mg/ml in both liquid and solid media.

Isolation of *M. maripaludis* genomic DNA

2 ml of later log-phase *M. maripaludis* culture was centrifuged at 4000xg for 10 min at room temperature, supernatant was removed. The pellet was resuspended on 500 µl of T.E pH 8.0 and frozen at -20^{o}C. The cell suspension was thawed at 37^{o}C then 25 µl of 10% SDS and 2.5µl of 20 mg/ml proteinase K were added. The cells were mixed gently by inversion, incubated at 37^{o}C for 1 hour and then 90 µl of 5M NaCl and 75 µl of CTAB/NaCl solution (10% cetyl trimethylammonium bromide, 0.7M NaCl) were added. The cells were mixed gently then incubated at 65^{o}C for 20 min. 500 µl chloroform:1AA (24:1) was added, the contents of the tube were mixed then centrifuged at 600xg for 10 min at room temperature. The aqueous layer was removed to a fresh tube using a wide

bore pipette to avoid shearing the DNA, 300 µl isopropanol was added. The tube was mixed gently until the DNA precipitated. The DNA was pelleted by centrifugation at 10000xg for 5 min. the supernatant was removed and the pellet was allowed to dry before being resuspended in 200 µl of TE pH 8.0. The concentration of genomic DNA was measure using A_{260} on the NanoDrop ND-1000 spectrophotometer (Thermo Scientific) and DNA quality was checked by gel electrophoresis.

Gel electrophoresis of DNA

DNA was electrophoresed using 1% agarose gels at 100V for 45 min. the Q-step IV ladder (York Bioscience) was used as a molecular weight marker. Samples were prepared for loading by the addition of 6x DNA loading dye (30% glycerol, 0.1% bromophenol blue) to a final concentration of 1x.

SDS-PAGE

SDS PAGE gels were run using a minigel system (CBS Scientific) all gels contained 12.5% polyacrylamide. For a 12.5% resolving gel, Protogel pre-mix (30% solution of 37.5:1 acrylamide:bis-acrylamide, National Diagnostics) was added to a final concentration of 12.5% acrylamide along with 375 mM Tris pH 8.7 and 0.1% SDS. For a 5ml gel, 5µl TEMED and 16µl 10% APS were added to set the gel. For a stacking gel, Protogel pre-mix (National Diagnostics) was added to final concentration of 5% acrylamide, along with 50mM Tris pH 6.9 and 0.1% SDS. For 2ml of stacking gel, 5µl of TEMED and 16µl of 10% APS were added. Samples were prepared by adding 6x Laemmli buffer (30% 2-mercaptoethanol, 12% SDS, 10% glycerol, 0.1% bromophenol blue, 440 mM Tris pH 6.8) to a final concentration of 1x running buffer (25 mM Tris base, 192 mM glycine and 0.1% SDS).

Gels were stained with Coomassie blue R250 (40% methanol, 10% acetic acid, 0.1% Coomassie Blue, Fisher) and destained in 40% methanol, 10% acetic acid. Gels were photographed using the geldoc system (Fuji) and stored in gel preservative (18% ethanol, 5% glycerol).

Southern blotting

Probe labelling Relevant probes (complementing flanking region for *M. maripaludis* McmB and McmC flanking regions) were used in this study, previously prepared (Walters, 2009). Probes were stored long-term at -20^{o}C.

Gel electrophoresis and blotting 1 µg of each genomic DNA sample was digested using the restriction enzymes SacI and PvuII (NEB) at 37^{o}C overnight. DNA loading dye was added and the DNA was separated by electrophoresis on a 0.8% TBE gel (0.8g agarose in 90 mM Tris-borate, 2mM EDTA, pH 8.0) in a 1x TBE running buffer. DIG-labelled markers (Roche) were also loaded onto the gel. After running, the gel was stained using 0.5 µg/ml ethidium bromide in 1x TBE, rinsed briefly was H_2O then photographed to check that the DNA had been digested. The gel was denatured by incubated in 0.5

M NaOH, 1.5 M NaCl for 2x 15 min, washed with H_2O then neutralised by incubation in 0.5 M Tris pH 7.5, 1.5 M NaCl for 2x 15 min. The gel equilibrated in 20x SSC buffer (2 M NaCl, 0.3 M sodium citrate, pH 7.0) for 10 min and the DNA was transferred to a positively charged nylon membrane (BioTrans+) using capillary transfer in a buffer of 20x SSC overnight. After transfer the membrane was placed on 3 mm filter paper (Whatman) pre-soaked in 2x SSC and the DNA was fixed using cross linking at 254nm for 1 min. the membrane was rinsed in H_2O and allowed to dry completely before being placed in hybridisation bottle and incubated in hybridisation buffer (Roche) for 30 min at 39^{o}C with slow rotation of the bottle. 25 ng of labelled probe per ml of hybridisation buffer was added to 50µl of H2O then boiled at 100^{o}C for 5 min, immediately placed on ice and then added to pre-warm hybridisation buffer was poured out of the hybridisation bottle and the fresh buffer containing the probe was added. The membrane was incubated with the probe overnight at 39^{o}C with slow rotation of the bottle then washed 2x 5 min in 2x SSC/0.1% SDS at room temperature and 2x 15 min in 0.5x SSC/0.1% SDS at 65^{o}C. The membrane was rinsed for 2 min in wash buffer (0.1 M maleic acid, 0.15 M NaCl, 0.3% v/v Tween 20, pH 7.5), incubated in 1x blocking solution (Roche) for 1 hour then incubated in anti-DIG antibody (Roche) diluted 1:10000 in 1x blocking solution for 30 min at room temperature. The membrane was incubated in wash buffer for 2x at min then equilibrated in detection buffer (0.1 M Tris pH 9.5, 0.1 M NaCl) for 3 min at room temperature. The membrane surface was covered with CPSD detection reagent (Roche), incubated for 5 min at room temperature then the excess buffer was removed and the membrane was washed in a plastic envelope. After a 10 min incubation at 37^{o}C, the membrane was exposed to photographic film (GE Healthcare).

Western blotting

Gels for Western blot analysis were loaded with a pre-stained protein molecular weight marker so that efficient transfer could be easily confirmed. Western transfers were carried out using the Transblot-SD semi-dry transfer cell (BioRad). Two pieces of 3 mm filter paper (Whatman) soaked in anode buffer 1 (0.3 M Tris pH 10.4, 10% methanol) were placed on the anode plate. One piece of filter paper soaked in anode buffer 2 (25 mM Tris pH 10.4, 10% methanol) was placed on top of the first two sheets and then a piece of PVDF membrane (GE Healthcare) pre-soaked in methanol was added to the stack. The minigel was placed on top of the membrane and 3 pieces of filter paper soaked in cathode buffer (25 mM Tris pH 9.4, 40 mM glycerol, 10% methanol) were placed on top of the gel. The cathode electrode was placed on top of the stack and stained with Ponceau S (Biorad) and lane positions were marked. The membrane was washed for 2x 10 minutes in TBS wash buffer (10 mM Tris pH 8.0, 150 mM NaCl) then blocked for 1 hour in blocking buffer (10% milk in TBS). After blocking, the membrane was washed for 3x 10 min in wash buffer, and then incubated with primary antibody overnight at room temperature (1:4000 Anti- TetraHis, Qiagen, 5% milk in TBS). The

primary antibody solution was removed and the membrane was washed for 6x 5 min in wash buffer, and then incubated with the secondary antibody for 1 hour (1:20000 Anti-mouse-HRP, Zymed, in TBS containing 0.1% Tween 20). Finally, the membrane was washed for 5x 10 min using wash buffer before being developed using the Supersignal ECL kit (Pierce) and imaged using photographic film (Thermo Scientific).

Bioinformatics tools

Sequence Alignment M. maripaludis McmA-D

A multiple alignment of the four *M. maripaludis* S2 MCMs was carried out using ClustalX. The alignment highlights additional inserts present in the McmD sequence (highlighted in blue) the longest (highlighted in pink) of which, is located in the C-terminal (See Figure 20. & Figure 21.).

Phylogenetic trees

Phylogenetic trees were created using NBCI BLAST analysis for homologous genes for MMP1024, showing relatedness of in the Methanococcales (Figure 11.) and MMP1025 (see Figure 12. -species highlighted in red denote ORFs MMP1024 / MMP1025 included in an operonic arrangement), sequences were then aligned using ClustalX with default parameters and tree visualised.

Gene map - MMP1024 (MMP1025)

A gene map for ORFs MMP1025 and MMP1024 on the *M. maripaludis* genome was prepared in Lasergene (from DNA Star).

Chapter 3

Results

Characterisation of McmD (MMP1024)

Production of constructs

Work completed during this study was centred on the investigation of the *M. maripaludis* protein McmD with further focus on the C-terminal. In order to facilitate McmD protein characterisation, a series of molecular constructs and tools were required. These included, pAW22 (Walters, unpublished); clone of full-length McmD coding region in commercial holding vector pGEMT (Promega), pJB001; clone of McmD coding region with 20 amino acid in-frame deletion of insert at the C-terminal* (pGEMT), pJB002; insert from pJB001 cloned in shuttle vector pAW42 (Walters et al., 2011), pJB003; insert from pAW22 cloned in shuttle vector pAW42, pJB004; contained two inserts, N-terminal (His-tagged) of *M. maripaludis* McmB and C-terminal (His tagged) of McmD cloned in commercial expression vector pProEXHTa (Invitrogen) (see Table 1.).

*Coding region hereby denoted as $McmD_{20}$

Production of constructs required for 'rescue' experiment

The 'rescue' of the sensitive phenotype (*M. maripaludis* McmDΔ) by constitutive overexpression of McmD under UV exposure required prior production of several constructs. McmD and $McmD_{20}$ coding regions were cloned into *M. maripaludis* shuttle vector pAW42 (Walters et al., 2011) (see Figure 5.). Shuttle vector pAW42 was developed to enable transfer of recombinant constructs between *E. coli* and *M. maripaludis*, model species used in this study. McmD coding region was amplified from pAW22, purified and ligated into pAW42, confirmed by DNA gel (agarose, 1%) and sequencing analysis resulting in the production of pJB003 see (Figure 6.). $McmD_{20}$ coding region was amplified, purified and ligated into pGEMT (Invitrogen), the sequence including the in-frame deletion was analysed for accuracy (pAW23) (see Table 1.). Following sequence confirmation, $McmD_{20}$ was cloned into pAW42 resulting in the production of pJB002 (see Figure 7.). pJB002 and pJB003 were sequenced across insert junctions confirming plasmids were cloned correctly. However, these data are included for pJB002 only (see Figure 8.), as data sequence data for pJB003 was unfortunately misplaced. The coding region for McmD was confirmed by genomics for pJB001 and pJB003 (data not shown). In order to further add to these data. Figure 9. shows a restriction digest of pJB002 and pJB003 indicating the presence of both inserts; McmD coding region 2123bp, $McmD_{20}$ coding region 2062bp and vector 4956bp.

Constructs used from previous studies	Description	Sequenced?
pAW22*	MMP McmD coding sequence in pGEMT *amp* (Invitrogen)	Yes
pAW23*	MMP $McmD_{20}$ coding sequence in pGEMT amp (Invitrogen)	Yes
pAW42**	MMP shuttle vector; containing *E.coli* origin, puromycin and ampicillin resistance markers	Yes
pAW18***	MMP McmB (His-tag) coding sequence in pProExHTa *amp* (Invitrogen)	Yes
pAW28***	MMP McmD (His-tag) coding sequence in p-ET28a *kan* (Novagen)	Yes

Constructs prepared and used during study	Description	Sequenced?	Used in strain(s)
pJB001	MMP McmD coding sequence with in frame deletion of 20 amino acid insert from C-terminal in pGEM-T *amp* (Invitrogen) (holding /sequencing vector only used for sub-cloning)	Yes	Nova Blue *E.coli*
pJB002	MMP McmD coding sequence with in frame deletion of 20 amino acid insert from C-terminal in pAW42	Yes	MMP McmDΔ (& MMP wildtype as over expression control)
pJB003	MMP McmD coding sequence in pAW42	Yes	MMP McmDΔ (& MMP wildtype as over expression control)
pJB004	N-terminal of *M. maripaludis* McmB and C-terminal of McmD in pProEXHTa *amp* (Invitrogen)	Yes	*E. coli* Rosetta, protein expression strain

Table 1. Table of constructs

*Walters, A.D. unpublished, **(Walters et al., 2011), ***(Walters, 2009)

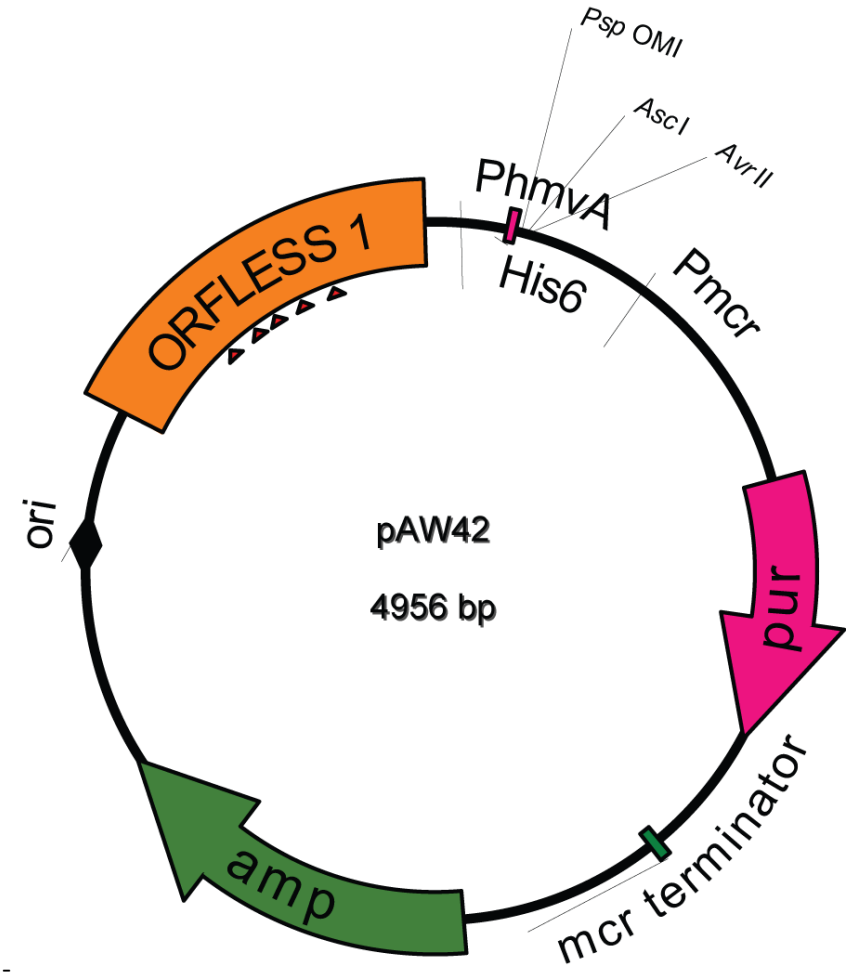

Figure 5. Plasmid map of pAW42

Shown for completeness, plasmid map of pAW42, adapted from (Walters et al., 2011) containing selectable markers for *E.coli* (amp) and *M. maripaludis* (pur), ORFLESS1 region containing putative origin for *M. maripaludis*, Ori, *E. coli* origin, multiple cloning site, His-tag, PhmvA, histone gene promoter from *M. voltae*, mcr gene terminator from *M. voltae* (Moore and Leigh, 2005).

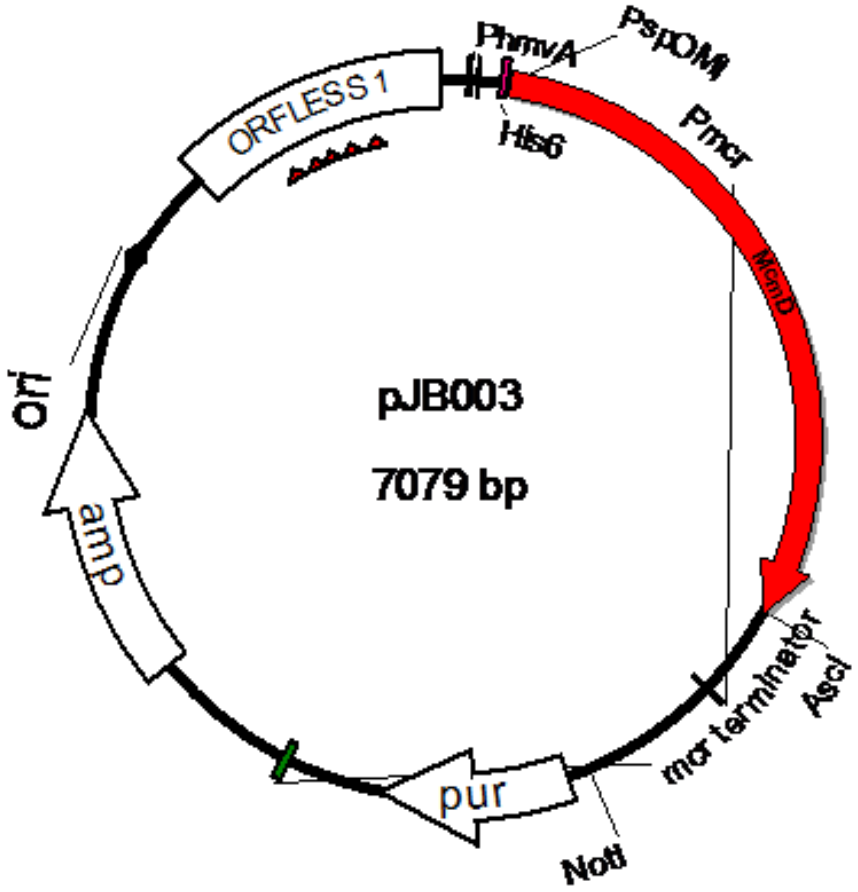

Figure 6. Plasmid map of construct pJB003

For use in rescue experiments, McmD coding region was sub-cloned into shuttle vector pAW42 (Walters et al., 2011) (Figure 5.). McmD coding region was amplified and purified from parent plasmid pAW22. pJB003 contained selectable markers for *E.coli* (*amp*) and *M. maripaludis* (*pur*), ORFLESS1 region containing putative origin for *M. maripaludis*, Ori, *E. coli* origin, multiple cloning site, His-tag, PhmvA, histone gene promoter from *M. voltae*, mcr gene terminator from *M. voltae* (Moore and Leigh, 2005).

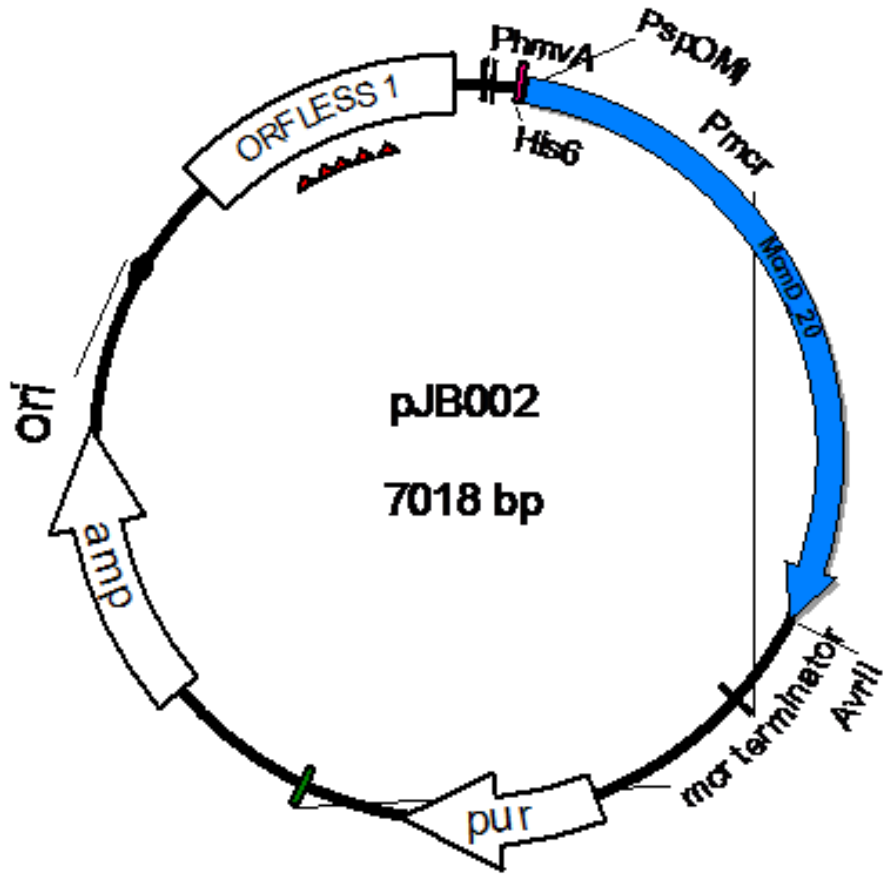

Figure 7. Plasmid map of construct pJB002

For use in rescue experiments, $McmD_{20}$ coding region was sub-cloned into shuttle vector pAW42 (Walters et al., 2011) (Figure 5.) from holding /sequencing vector pJB001. pJB002 contained selectable markers for *E.coli* (*amp*) and *M. maripaludis* (*pur*), ORFLESS1 region containing putative origin for *M. maripaludis*, Ori, *E. coli* origin, multiple cloning site, His-tag, PhmvA, histone gene promoter from *M. voltae*, mcr gene terminator from *M. voltae* (Moore and Leigh, 2005).

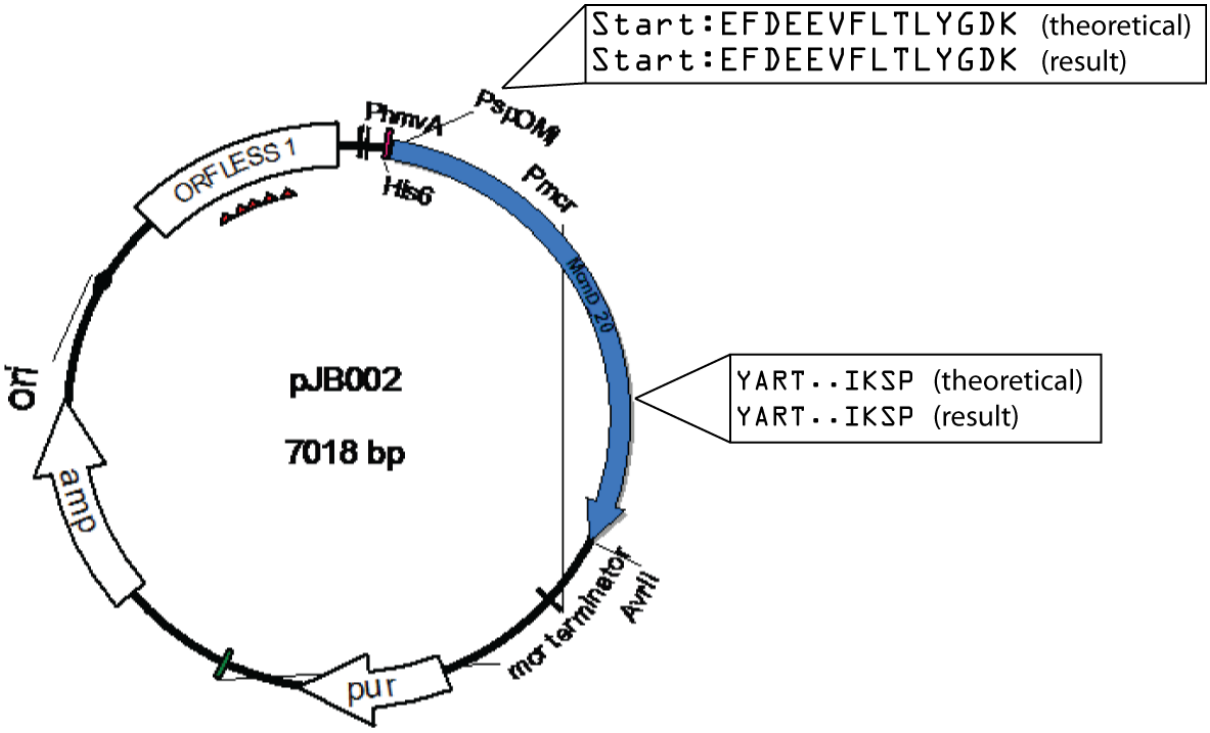

Figure 8. Sequencing results for pJB002

Plasmid clone pJB002 was analysed by DNA sequencing, the figure above is a representation of the raw data generated by sequencing ORF $McmD_{20}$ (shown in blue). Sequencing primers for vector pAW42 (primers 6. & 7., see Table 3.), were used to sequence complete cloned $McmD_{20}$ ORF including the N-terminal (Top box) and C-terminal junctions. The insert deletion join at the C-terminal was also checked (Bottom box). Primers were compatible with pAW42 vector located 23bp upstream / downstream of the multiple cloning site. However, there were problems with the C-terminal read which contained mis-reads causing the result to be inconclusive. Sequence confirmation for the C-terminal junction would therefore require repeating.

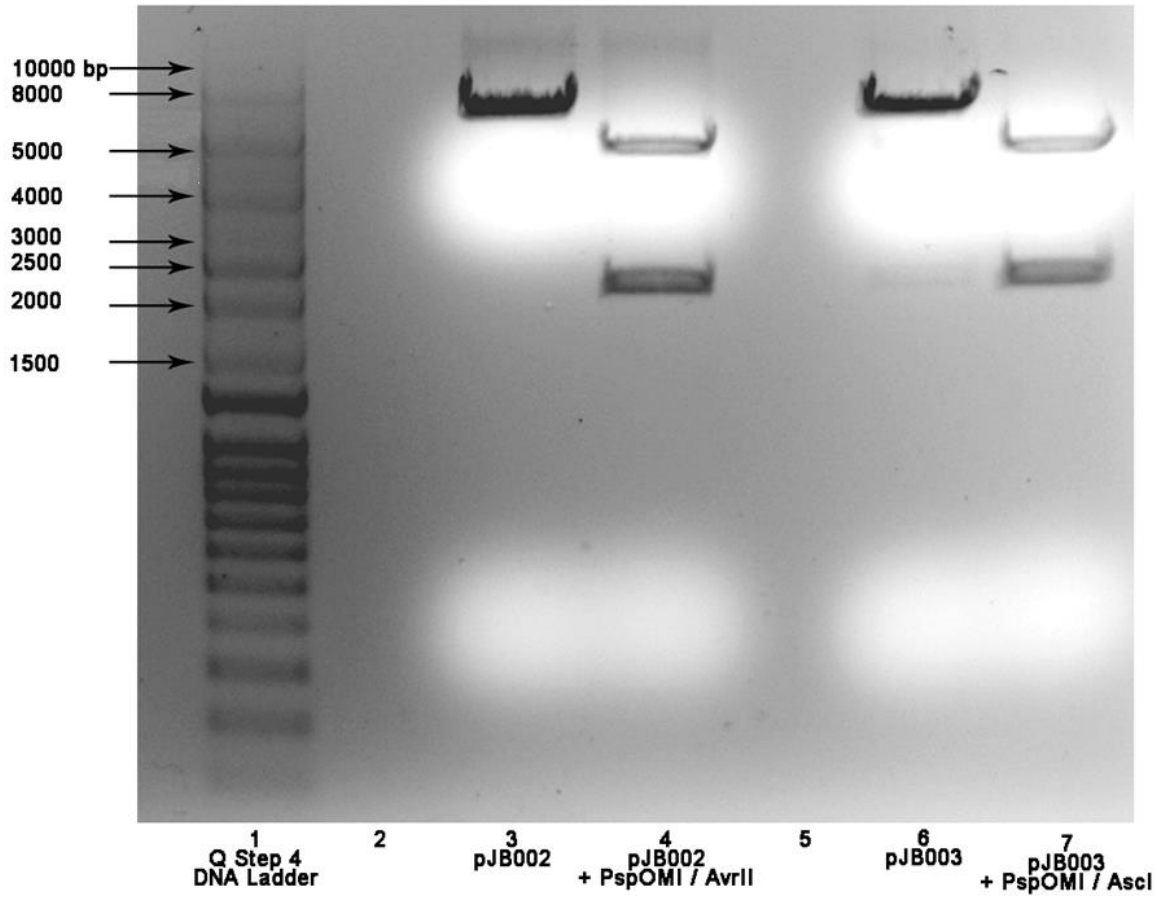

Figure 9. Restriction digest of pJB002 and pJB003

Restriction digest of pJB002 and pJB003 with PspOMI (at the N-terminal) and AscI/AvrII (at the C-terminal). Lane 1 Q-step 4 Ladder (Yorkshire Biosciences Ltd), Lane 2 (no sample), Lane 3 undigested pJB002 expected band pattern was 7018bp, Lane 4 double digest of pJB002 expected band patterns were 2062bp and 4956bp for $McmD_{20}$ and pAW42 respectively, Lane 5 (no sample), Lane 6 undigested pJB003 expected band pattern was 7079bp, Lane 7 double digest of pJB003 expected band patterns were 2123bp and 4956bp for McmD and pAW42 respectively. Correct band patterns were observed, correct sequences confirmed (University of York Genomics).

Rescue of sensitive phenotype (McmDΔ) by McmD overexpression

In order to investigate previous findings with strain McmDΔ, A 'rescue' experiment was designed where McmD was overexpressed via plasmids pJB002 and pJB003 in *M. maripaludis* McmDΔ strain with the aim of rescuing the phenotype under UV exposure.

Table 2. shows the results from the attempted rescue experiment, pJB003 (vector+ McmD coding region), pJB002 (vector + $McmD_{20}$ coding region) and vector (pAW42), were transformed in *M. maripaludis* strain McmDΔ (S0001DΔ) (Walters et al., 2011). All plasmid constructs were also transformed with wildtype strain (S0001) (Walters et al., 2011) in order to assess issues with overexpression. The results show that is was possible to transform S0001DΔ and S0001 successfully with vector (pAW42). Transformation of S0001 with plasmids pJB002 and pJB003 was not successful with final selection with puromycin yielding no growth. Initially, transformation of S0001ΔD with plasmids appeared successful with growth occurring at final selection stage so glycerol stocks were prepared. However, in the case of both pJB002 and pJB003 transformations, colonies could not be subsequently revived from glycerol. Colonies appeared sick; overexpression of McmD (& $McmD_{20}$) resulted in lethality. An explanation for this is that the expression of McmD under cellular levels is low and therefore overexpression results in cell death. The next step following on from these findings involved the re-cloning of pJB002 and pJB003 with an upstream (from McmD and $McmD_{20}$ ORFs) endogenous *M. maripaludis* MCM promoter located on the genome upstream of (wildtype) McmD (see Figure 10.). Further investigations revealed McmD (MMP1024) is found throughout the order Methanococcales (Figure 11.) usually downstream of gene MMP1025 and interestingly clustered within an operon in 8 of the 16 species analysed (see Figure 12.). Figure 13. shows the results of the analysis of *M. maripaludis* genome for a putative McmD promoter. McmD (MMP1024) in located 12bp downstream of MMP1025. A 12bp sequence is unlikely to contain a promoter sequence suggesting the promoter for MMP1024 and MMP1025 is located upstream of MMP1025 within the 143bp stretch, revealing MMP1024 and MMP1025 are most likely expressed in an operon.

Construct / Strain	Growth?
pAW42 + S0001ΔD	Growth possible from glycerol
pJB002 + S0001ΔD	Glycerol taken, cannot be revived
pJB003 + S0001ΔD	Glycerol taken, cannot be revived
pAW42 + S0001	Growth possible from glycerol
pJB002 + S0001	Transformation not successful, No growth in puromycin selection
pJB003 + S0001	Transformation not successful, No growth in puromycin selection
TE buffer (control) + S0001 +puromycin	No growth
TE buffer (control) + S0001ΔD + puromycin	No growth

Table 2. Results from rescue experiment

Results from rescue experiment demonstrating problems with growth of recombinant clones.

Construct key:

pAW42 – shuttle vector for *M. maripaludis* (Walters et al., 2011) (infers puromycin resistance)

pJB002 – pAW42 + $McmD_{20}$

pJB003 – pAW42 + McmD (full-length gene)

***M. maripaludis* strains key:**

S0001: ORF1 located on genome (Walters et al., 2011) (*M. maripaludis* strain; maintains pAW42)

S0001ΔD: As above, McmD deleted

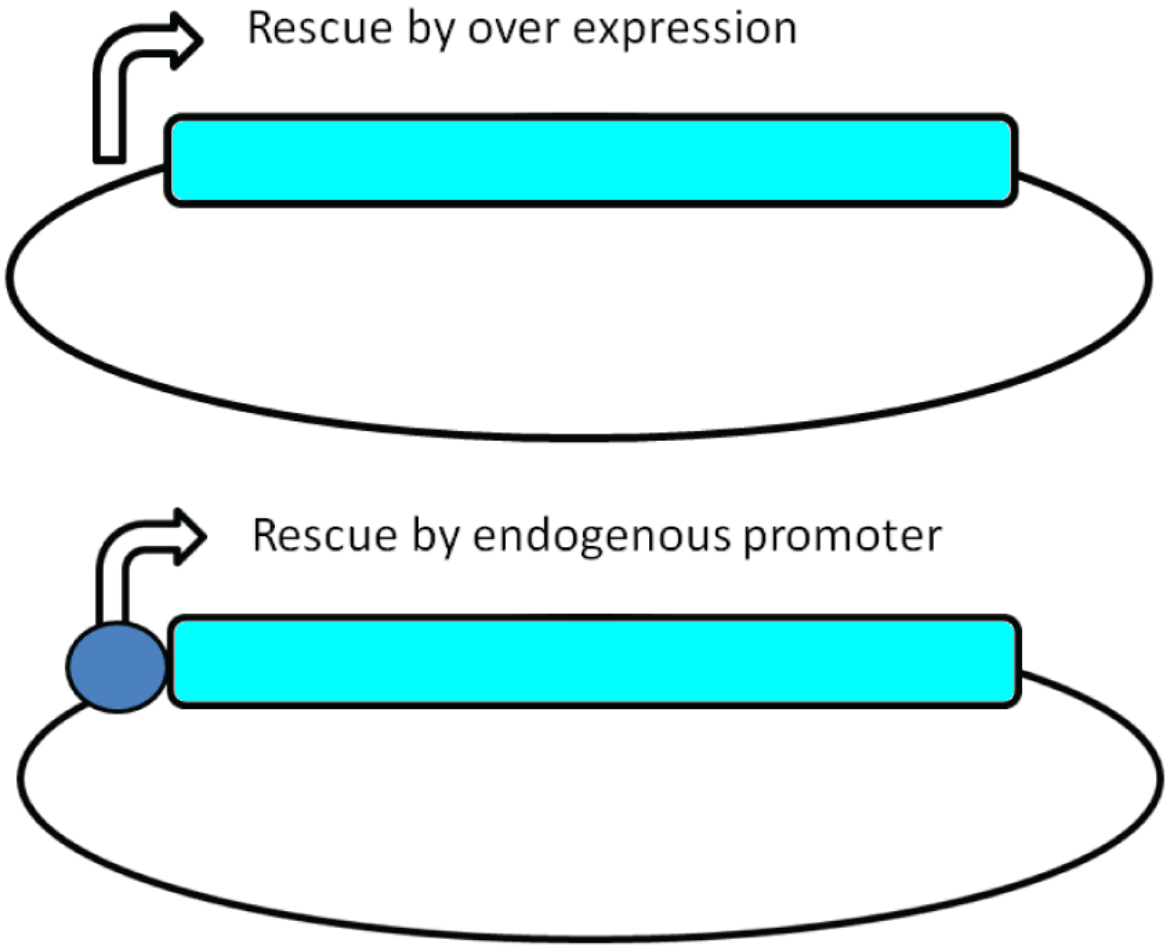

Figure 10. Expression of McmD under constitutive and endogenous promoters

A schematic diagram showing the expression of McmD under a strong constitutive promoter included on vector (Top) and an endogenous promoter (Bottom). Re-cloning McmD downstream of an endogenous promoter derived from the *M. Maripaludis* genome was proposed for the expression of McmD (& $McmD_{20}$) under typical cellular levels. Figure includes representation of McmD ORF (shown in pale blue), the arrows illustrate direction of expression, endogenous promoter is illustrated by the blue dot.

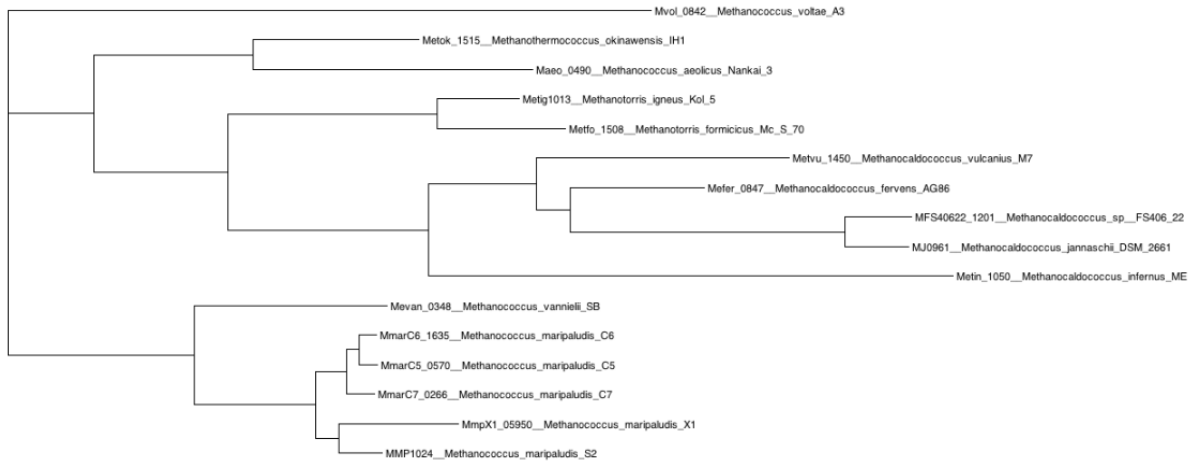

Figure 11. Phylogenetic tree showing relatedness of MMP1024 in the Methanococcales

Homologous genes identified for MMP1024 (using NCBI BLAST), sequences then aligned using ClustalX with default parameters and tree visualised.

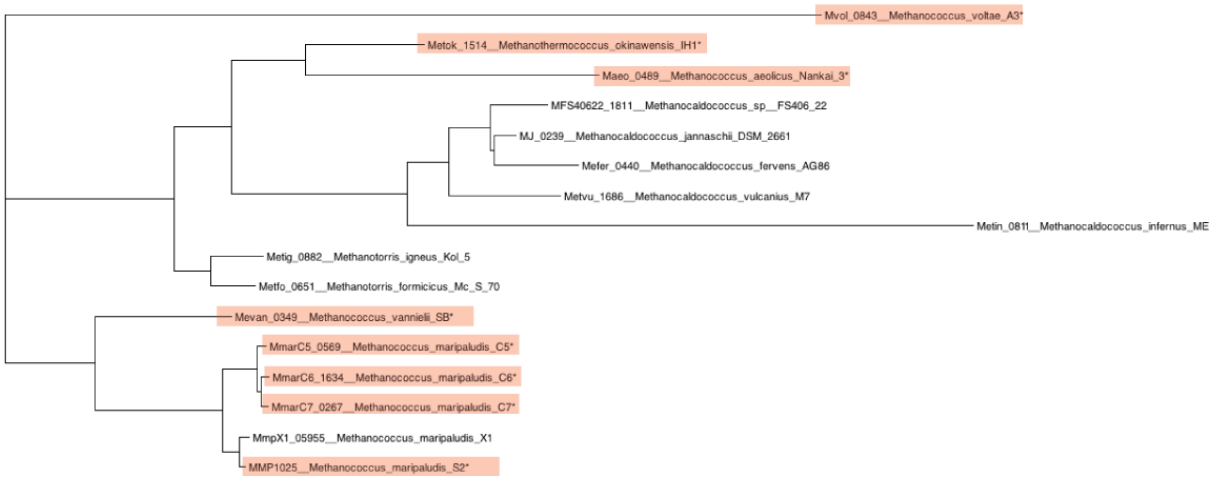

Figure 12. Phylogenetic tree showing relatedness of MMP1025 in the Methanococcales

Species highlighted in red denote MMP1024 / MMP1025 operonic arrangement. Homologous genes identified for MMP1025 (using NCBI BLAST), sequences then aligned using ClustalX with default parameters and tree visualised.

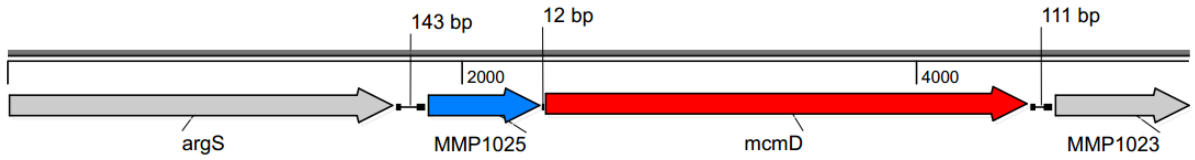

Figure 13. Map of ORFs MMP1025 and MMP1024

Map of ORFs MMP1025 and MMP1024 (red) on the *M. maripaludis* genome; including putative 143bp upstream operonic promoter. Sequences obtained from NCBI; complete genome sequence of *M. maripaludis* S2 (Hendrickson et al., 2004), specific region prepared in Lasergene (from DNA Star).

Tool generation (Technical development)

Multiple knockout strains generated (Southern Blots)

Previous work has found it is possible to produce single knockout strains of *M. maripaludis*- McmBΔ, McmCΔ, and McmDΔ. However, it has not been possible to produce a knockout strain of McmA, so this gene has been assumed as essential. Single, double and a triple mutant of *M. maripaludis* MCM strains have been produced in order to investigate synthetic effects of MCM deletion. Figure 14. is a Southern blot showing McmC deletion in strain McmBΔ, resulting in the production of double mutant McmBCΔ. Figure 15. shows a Southern Blot of McmB deletion in strain McmCDΔ resulting in the production of triple mutant McmBCDΔ. Figure 16. shows a Southern Blot of double mutant McmCDΔ (Hirst, 2011). Mutant strain McmBDΔ (Noble, 2011) was produced, however initial results were deemed unclear and could not be included in this report. The original gel did not run as expected and as a result was omitted from presentation. Deletions were possible by virtue of a knock-in knockout construct whereby flanking regions of the gene of interest were positively selected, incorporated onto the genome resulting in polyploidy, with a subsequent selection driving an estimated result of a 50:50 ratio of mutant or wildtype.

Fusion Protein (*M. maripaludis* McmB N-terminal and McmD C-terminal)

2. To date biochemical analysis with recombinant McmD has proven problematic due to technical issues during expression and purification steps. Previous work has demonstrated it is possible to express McmD recombinantly in *E. coli.* However, the protein did not appear to fold in the correct way resulting in the production of insoluble protein fractions, most likely consisting of mis-folded aggregates (Walters, 2009). Recovered protein levels were insufficient for further biochemical analysis. In order to present a solution to this problem a domain swap was proposed between the N-terminal of McmB and the C-terminal of McmD. McmB recombinant protein has previously be shown to yield good workable recovery levels and has proven the most reliably expressed, purified and active recombinant *M. maripaludis* MCM (Walters, 2009). A domain swap between these two proteins would potentially provide the study with an extremely useful tool which in theory combined the biochemical tractability of McmB with a focus on thus far biochemically intractable McmD further to this, the C-terminal of McmD only, the domain of interest.

Figure 17. shows a schematic diagram of the fusion protein. The C-terminal only of McmD is included. The junction between McmB and McmD is shown in more detail in Figure 21. The strategic approach taken for the preparation of the fusion protein construct is illustrated in Figure 18. Parent plasmids including McmB and McmD coding regions were amplified by PCR across regions of interest producing inserts required for fusion protein. Inserts were gel purified and ligated in commercial

expression vector pProEXHTa (using Fusion ligation kit Clontech), producing construct pJB004. Following cloning strategy for fusion construct pJB004 (Figure 18.), 7 clones containing putative pJB004 constructs were picked into LB nutrient broth, grown for 24 hours at 37^{o}C, plasmids were harvested and purified. All band patterns were present in all samples analysed. The presence of correct inserts were checked on DNA agarose gel (Figure 19.) and subsequently submitted for sequence analysis (University of York Genomics). Correct sequences were confirmed using sequencing primers 12., 13. and 14. (See Table 3.) designed to sequence over target junctions.

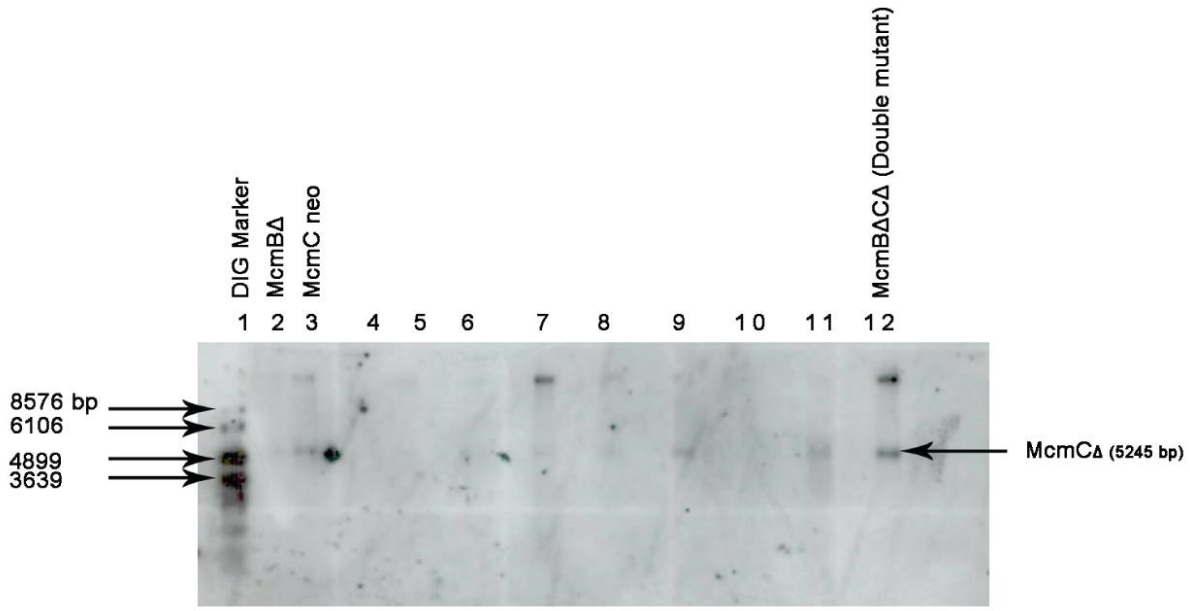

Figure 14. Southern blot of CΔ construct in BΔ strain (double mutant)

Genomic DNA from eight transformants (lanes 4-12) was digested, (using restriction enzymes SacI & PvuII, NEB) blotted and hybridised with a labelled probe specific for McmC flanking regions. Expected fragment sizes were 7245bp and 5245bp for BΔ (WT) and McmCΔ respectively. Not all sample strains produced visible bands. Strain 4 (lane 7), strain 9 (lane 9), strain 18 (lane 11) and strain 19 (lane 12 – putative double mutant), show expected band patterns for mutant. BΔ control (lane 2), Neomycin control (lane 3). Some protein remained undigested in lanes 7 and 12.

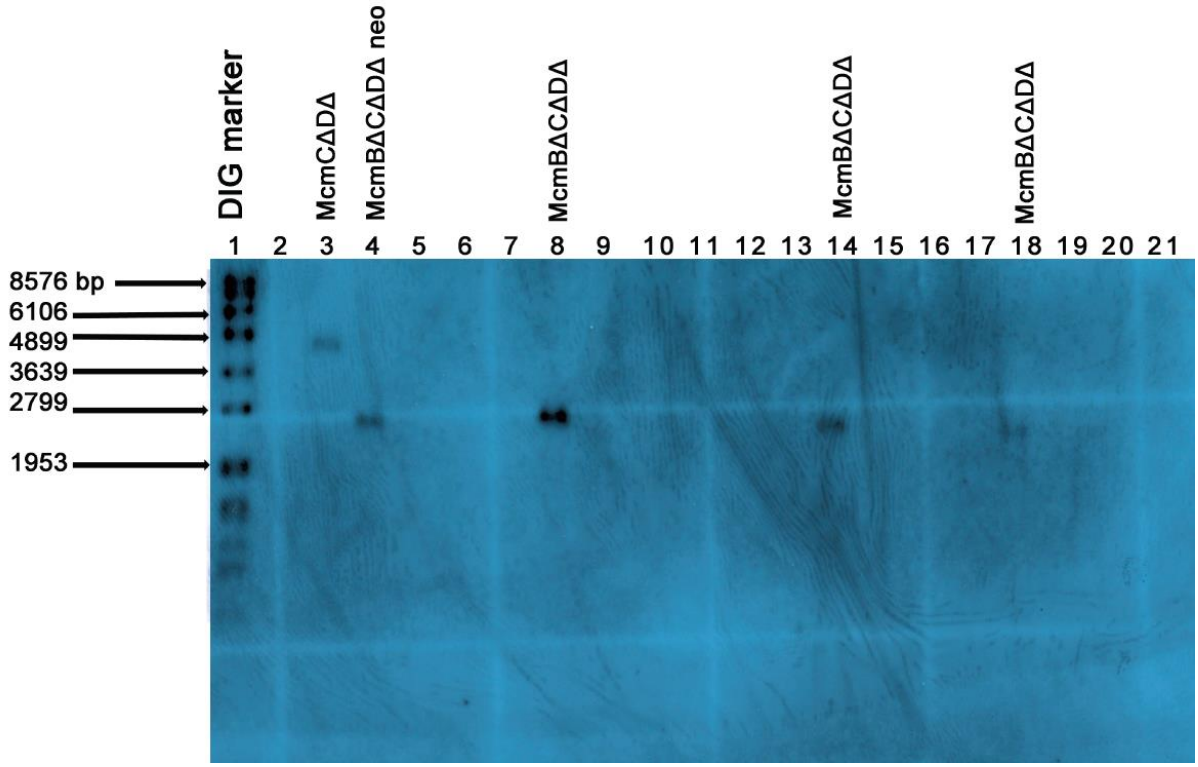

Figure 15. Southern blot of BΔ construct in strain CΔDΔ (triple mutant)

Genomic DNA from eight transformants (lanes 5-20) was digested, (using restriction enzymes SacI & PvuII, NEB) blotted and hybridised with a labelled probe specific for McmB flanking regions. Expected fragment sizes were 4567bp and 2539bp for CΔDΔ (WT) and McmBΔ respectively. Not all sample strains produced visible bands. Strain 2 is the putative triple mutant visible in lane 8. Bands are visible for Strain 5 (lane 14), and Strain 7 (lane 18), CΔDΔ (lane 3), Neomycin Control (lane 4).

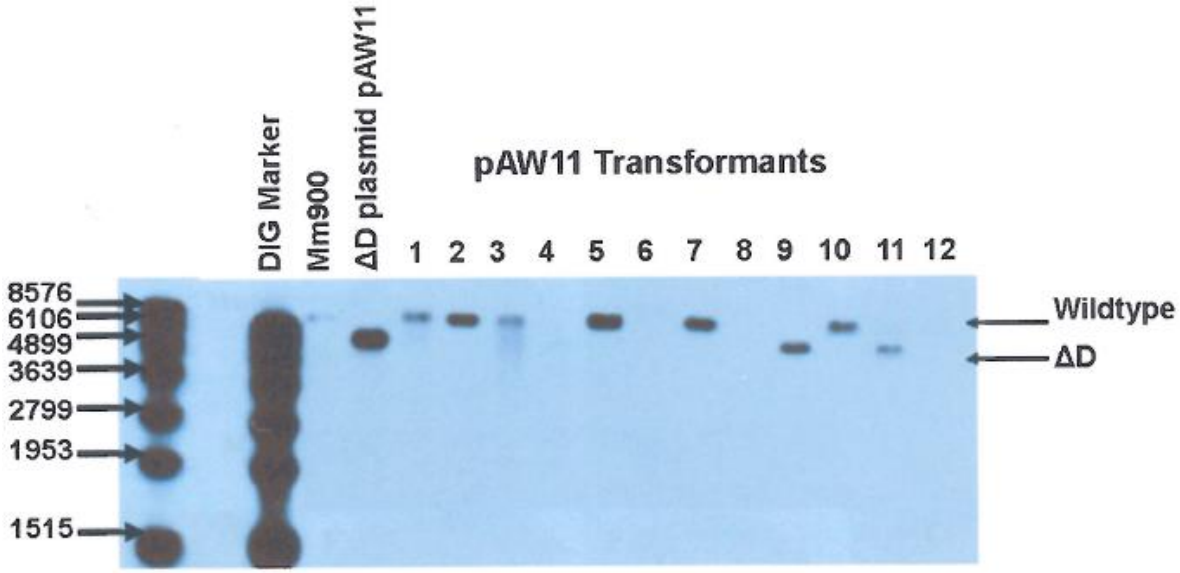

Figure 16. A Southern Blot of DΔ construct in CΔ mutant (double mutant)

Genomic DNA from twelve transformants was digested (using restriction enzymes SacI & PvuII, NEB) blotted and hybridised with a labelled probe specific for McmD flanking regions. Expected fragment sizes were 6518 bp and 4428 bp for CΔ (WT) and DΔ respectively. Transformants in lanes 9 and 11 show a positive result for DΔ. Adapted from (Hirst, 2011), final year undergraduate dissertation.

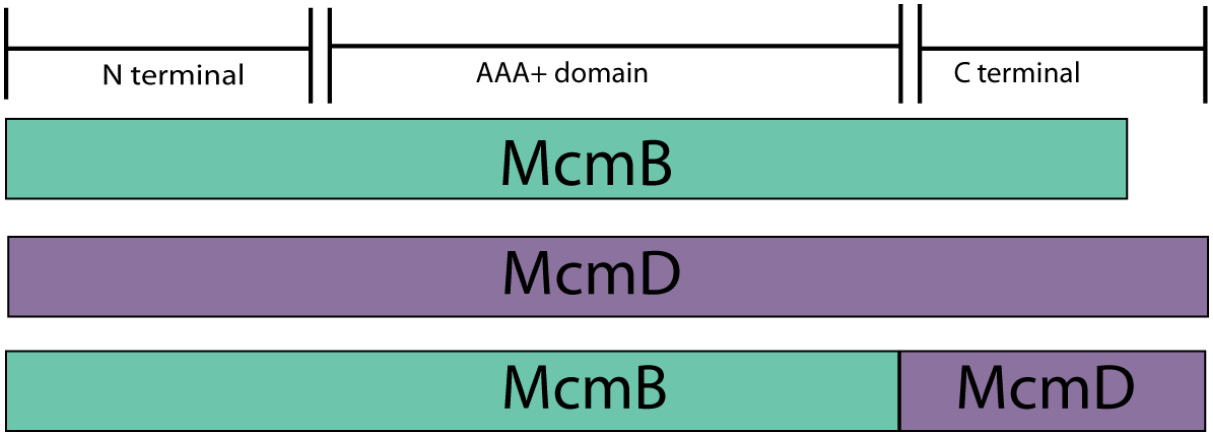

Figure 17. *M. maripaludis* fusion protein

The above schematic diagram shows the theoretical *M. maripaludis* fusion protein. McmB N-terminal domain is fused with McmD C-terminal domain.

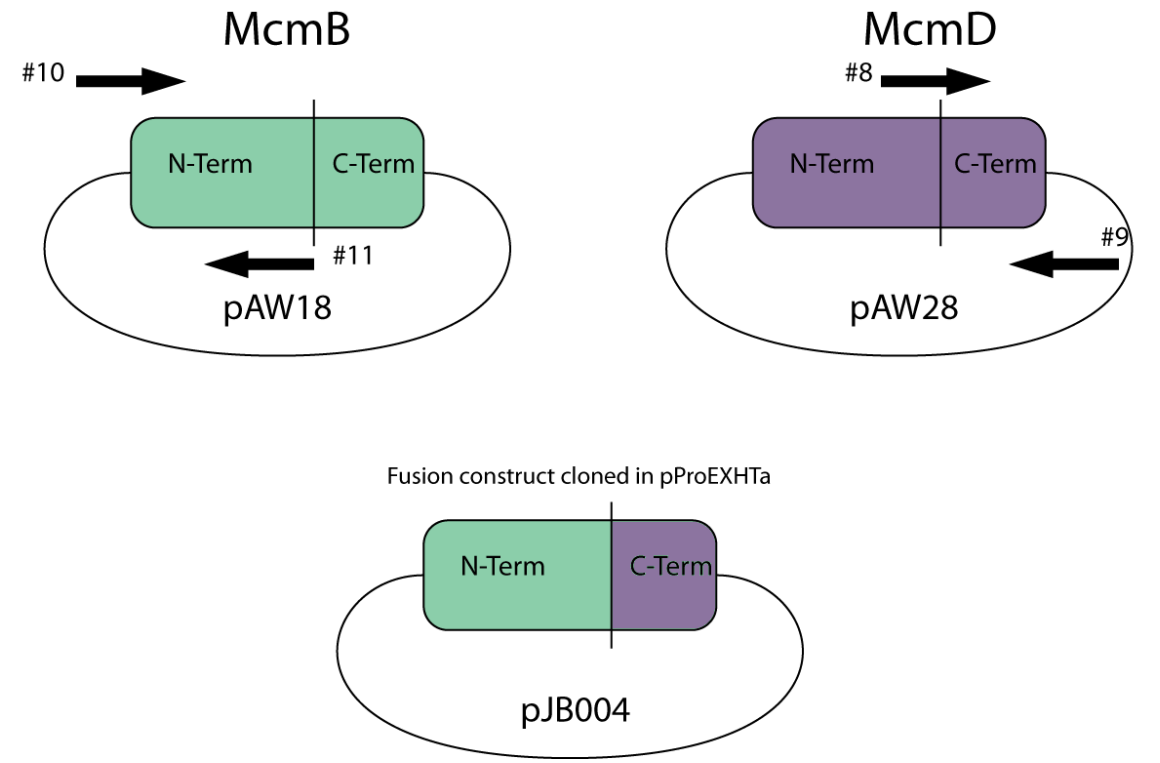

Figure 18. Cloning strategy for fusion construct pJB004

Step 1. Holding plasmids pAW18 (McmB His, pProEXHTa, Invitrogen) and pAW28 (McmD His, p-ET28a) (see Table 1.) (Walters, 2009), were used to amplify DNA fragments; McmB-His N-terminal domain, and McmD C-terminal domain) by PCR. Primers 8., 9., 10. & 11. see Table 3.), were designed engineering overhangs on amplified fragments complementary to final insertion in pProEXTHa producing fusion construct pJB004. The In-Fusion® HD Cloning kit (Clontech) used to anneal both inserts in the vector employed the used of LIC overhangs. Primer 8. included 15bp overhangs complementary with McmB N-term, primers 9. & 10., Included 15bp overhangs complementary to pProEXHTa and primer 11. included 15bp overhangs complementary to McmD C-term.

Step 2. Bacterial expression of pJB004 and purification of fusion protein (represented by linear bar), including a single His6 tag at the amine terminus.

Key:
McmB-His ORF shown in green, McmD-His ORF shown in violet, #8-#11 denotes primer number, (see Table 3.), direction of DNA amplification by PCR denoted by arrows.

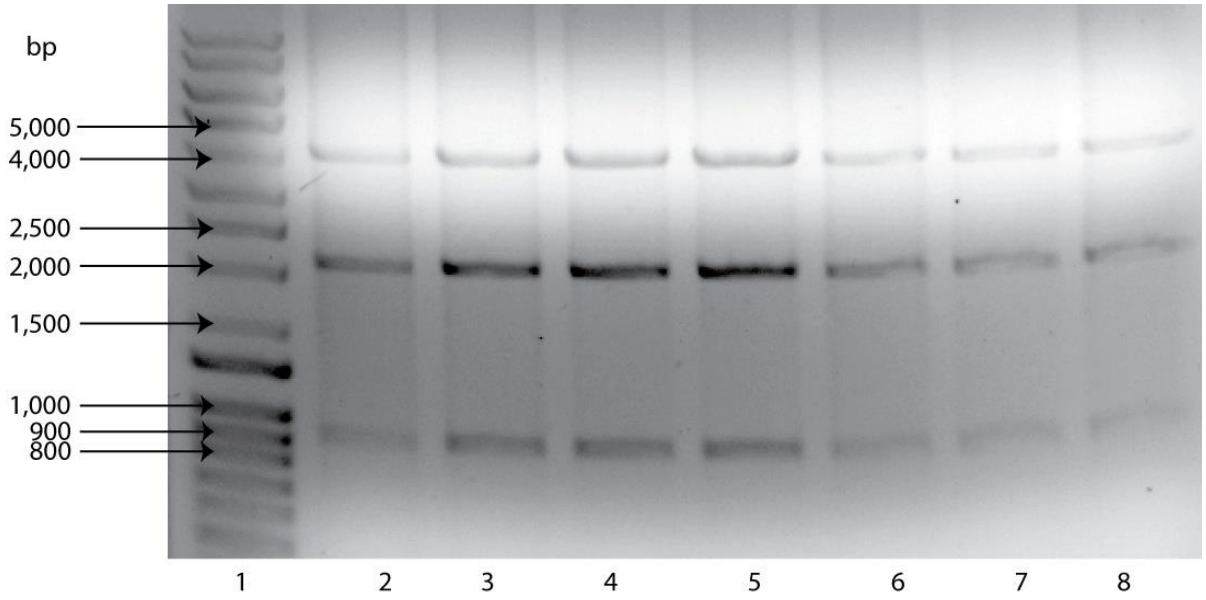

Figure 19. pJB004 screen

Fusion cloning reaction of McmB N-terminal , McmD C-terminal and pProEXHTa (Invitrogen) (commercial expression vector) in restriction digest with NcoI, KpnI and PspOMI. Expected bands were 2088 bp for McmB N-terminal and McmD C-terminal construct, and 4372 bp for vector (pProEXHTa). An additional fragment of 887bp was also expected, enzyme PspOMI was added to digest reaction in error. Lane 1 Q-Step 4 Ladder (Yorkshire Biosciences Ltd), lanes 2 to 8 all contain restriction digest samples of putative pJB004 clones. All lanes contained expected band patterns. The sample in lane 5 was submitted for sequencing, correct sequence was confirmed (University of York Genomics).

Figure 20. Alignment of the four *M. maripaludis* MCM sequences

All four *M. maripaludis* MCM sequences were aligned (sequences obtained from NCBI, aligned in ClustalX). All additional inserts included in the McmD sequence are highlighted blue, longest insert highlighted in pink.

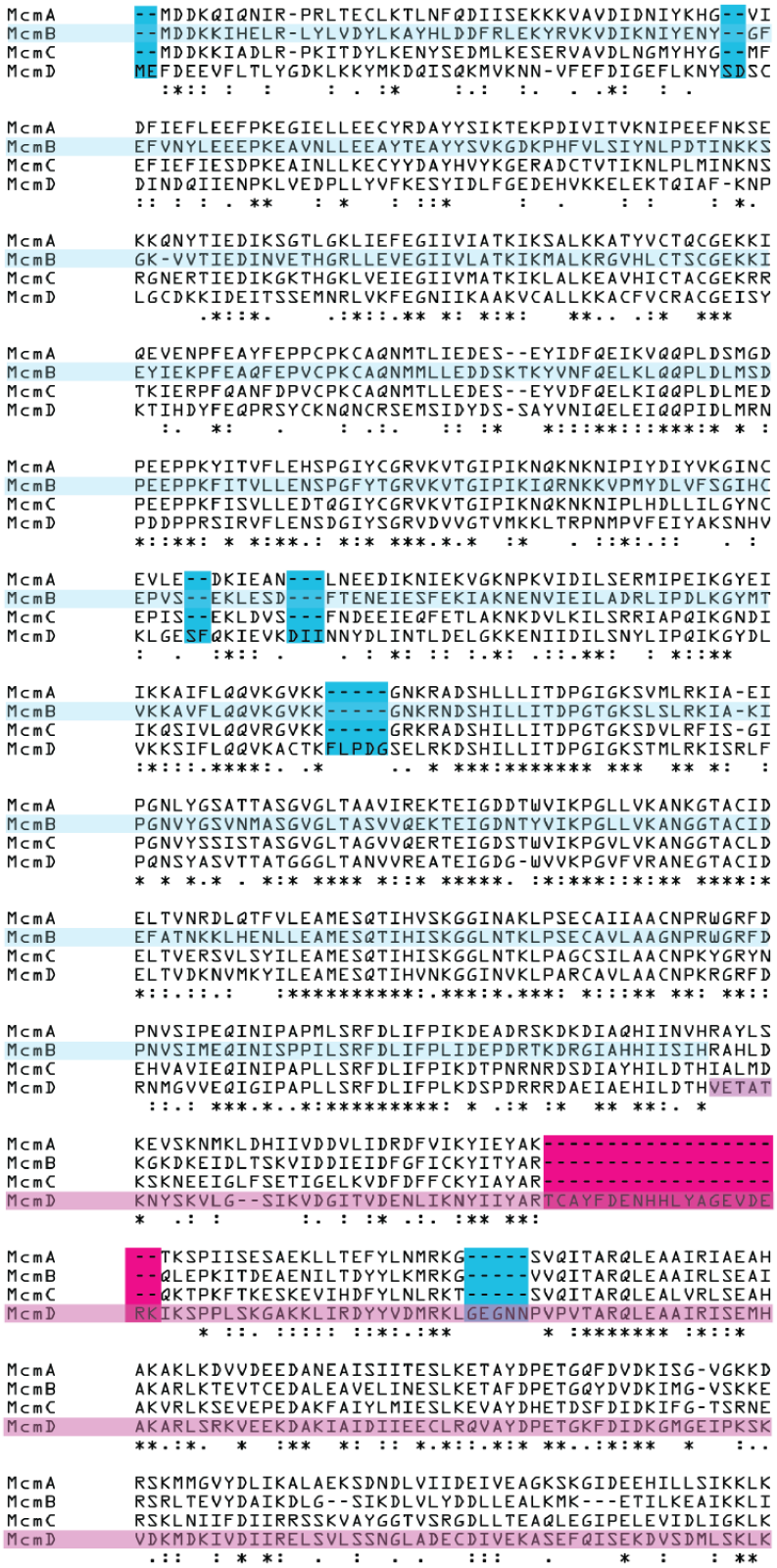

Figure 21. as Figure 20, including fusion protein amino acid sequence

Fusion protein sequence is highlighted in pale blue and pale pink. The junction between McmB N-terminal and McmD C-terminal is located at residue 495,496, joining sequence SIH-VET

Chapter 4

Discussion

Overview

Work performed as part of this project included the manipulation of genomic and plasmid DNA in model organisms *E. coli* and *M. maripaludis*. In the interests of simplicity, flexibility and reproducibility the majority of genetic manipulation was performed in *E. coli* and subsequently introduced in *M. maripaludis* test strains. Also a complete and comprehensive range of genetic tools to rival those currently available for *E. coli* cannot be as yet, matched for archaeal model species.

This study aimed to investigate the four *M. maripaludis* MCM genes, building on previous work (Walters and Chong, 2010); (Walters et al., 2011). Unpublished data from our laboratory has led to an interest in damage sensitive phenotype *M. maripaludis* McmDΔ and the involvement of McmD protein in DNA damage signalling and / or repair. The main focus was centred on McmD and the C-terminal of McmD and the possible connection with DNA damage repair under UV light exposure. Investigations with overexpression of McmD via cloned constructs pJB002 and pJB003 in *M. maripaludis* strain McmDΔ, unfortunately were not successful leading to the requirement for a revised strategy, influencing the direction of study.

Co-expression of MMP1024 and MMP1025

It is possible McmD expression, translation and post-translational modification would be enhanced with the co-expression of MMP1025, given that we have found both genes are expressed in an operon. It is possible MMP1025 acts as a chaperone or protein partner to MMP1024, facilitating the production of active McmD protein. In order to investigate this hypothesis, planned experiments would involve the production of a construct which would allow simultaneous expression of both proteins in high yielding active forms so that purification and biochemical characterisation would be possible.

Following difficulties presented by the 'rescue' explained previously in the results chapter, rather than re-clone pJB002 and pJB003 under an endogenous promoter, MMP1024 and MMP1025 would be cloned in a commercially obtained vector designed for the co-expression of proteins. If successful, co-expression and purification would be possible. Previous problems with the production of active recombinant McmD protein would consequently have a compelling solution.

Initial searches on ORF MMP1025 has revealed no additional information regarding analysis (bioinformatics or otherwise), therefore further analysis of this ORF would be valuable.

Cloning strategy: co-expression construct

The co-expression construct (ORFs MMP1024, MMP1025 in vector pRSF -EMD Biosciences) was in the developmental stages towards the end of the project, and therefore the production of a correct and sequenced construct was not seen through to completion. However, inserts and vector samples for use at the annealing step had been prepared.

MMP1024 and MMP1025 open reading frames were amplified from plasmid and genomic samples, purified and cloned in commercial vector LIC Duet (minimal adaptor) Novagen. LIC Duet adaptors enable cloning of two ORFs simultaneously into one plasmid for co-expression in bacteria. The protocol employed ligation independent cloning (LIC) without the requirement for restriction enzyme digestion or ligation reactions. The adaptors bridged the 3' end of the first prepared PCR product and the 5' end of the second prepared PCR product. The LIC overhangs annealed with 5' end of the first PCR product and the 3' end of the second PCR product. Therefore, the first target protein is fused to an N-terminal tag encoded by the LIC adaptor. LIC Duet possessed a T7lac promoter, the minimal adaptor included the ATG start codon. PCR products were treated with T4 DNA polymerase in the presence of dATP to generate overhangs compatible with LIC vector and LIC Duet adaptor. Once annealing reaction was complete plasmid could be transformed in competent *E. coli* and used for bacterial expression.

Fusion 'domain swap' protein

The study aimed to address previous problems with McmD recombinant expression and purification. Initial strategies outlined work with the domain swap protein whereby a fusion protein of McmB and McmD was produced to improve active protein recovery for ongoing biochemical characterisation and construct pJB004 was cloned as part of the sub-project. As the study progressed we had evidence to suggest that another avenue should be investigated. Bioinformatic analysis showed that the operonic arrangement of MMP1024 and MMP1025 could be the key to effective production of active recombinant McmD protein, which would be prioritised in planned future work. Previous work has provided biochemical characterisation for *M. maripaludis* MCMs, A, B and C, this study aimed to complete these data by the addition of McmD characterisation, in the form of helicase, ATPase and DNA binding assays.

Knockout Mutants

Single knockout mutants; McmBΔ, -CΔ, -DΔ, were produced as part of a preceding study. McmA knockout mutant could not be produced so McmA was assumed essential. Previous unpublished data showed that the McmDΔ strain showed pronounced sensitivity to UV induced DNA damage in comparison with the other single knockout strains. Additionally flow cytometry with the McmDΔ

strain showed an unusual cell cycle progression where cell division may have been occurring in the absence of complete genome replication, suggesting replication checkpoints may have been compromised. This would lead to an inevitable occurrence of lethality and apoptosis.

During this study knockout strains; McmBCΔ, -McmCDΔ, McmBDΔ (double mutants) and McmBCDΔ (triple mutant) were prepared confirming these proteins are non- essential. Future work with these strains was planned in the form of comparative analysis alongside the McmDΔ strain in similar UV exposure / flow cytometry experiments. The sub project also aimed to investigate the synthetic effects of deleting multiple MCMs.

UV experiments comparing McmBCΔ, McmBCDΔ, McmDΔ and Mm900 (wildtype) were trialled on a number of occasions. These particular strains were initially selected as the data provided would show a comprehensive comparison of all McmDΔ strains and therefore synthetic effects could be analysed. Also strain McmBCΔ (deemed most stable, possessing McmD and closest to wildtype) was included. Due to technical issues with gas stability in compression canisters and varying growth rates of colonies on plates, a reliable data set remained elusive. Repeated UV experiments with these strains would be useful in the future under optimised conditions. Flow cytometry analysis with the double and triple mutants was also planned in ongoing projects.

***M. maripaludis* MCM expression and protein purification**

The expression and purification of recombinant MCM was trialled toward the end of this study. Fusion construct pJB004 was prepared with the foresight of producing a fusion protein where the N-terminal of McmB is fused with the C-terminal of McmD, allowing further biochemical characterisation of the C-terminal of McmD. Purification of recombinant McmB was attempted in the early stages of this sub-project in order to optimise the protocol, as previous studies showed that McmB was the most readily purified MCM protein in *vitro* with relatively high recoveries (Walters, 2009). However, McmB protein levels were detected, but only in very low concentrations by Western blotting (data not shown), further optimisation of the process was therefore required. The original Western blot was of insufficient quality for presentation and was not included as advised. Further work would ideally include a repeat of the McmB expression and multi-step protein purification using more robust automated systems including e.g. HisTrap™FF crude 5 mL column, FPLC –AKTA systems, (GE Healthcare) including size exclusion chromatography.

Oligo description			Sequence 5' > 3'
Amplification of McmD / McmD$_{20}$ from pAW22 / pAW23			
1.	McmD	Fwd.	AAAAGGGCCCATGGAATTTGATGAAGAAGTTTTTTTGACCC
		Rev.	AAAAGGCGCGCCTTAAGTTAATCTATAATATCCATATTTTGGGCTGAATAC
2.	McmD$_{20}$	Fwd.	AAAAGGGCCCATGGAATTTGATGAAGAAGTTTTTTTGACCC
		Rev.	AAAACCTAGGTTAAGTTAATCTATAATATCCATATTTTGGGCTGAATAC
G^{V}GGCCC PspOMI (NEB) restriction site GG^{V}CGCGCC AscI (NEB) restriction site C^{V}CTAGG AvrII (NEB) restriction site ^{V}*Denotes restriction enzyme cleavage site			
Sequencing primer: pAW42 multiple cloning site junction			
6.	Fwd.	CTGATCGATCAAAATATAACATAAATAACATAGG	
7.	Rev.	CCTCTAGAGGATGATTAATTTTAAGAGAG	
Amplification of fusion construct			
8.	McmB N-terminal / McmD C-terminal	Fwd.	CACATTTTAGACACACACGTTGAAACCGCAACC
9.	pAW28 /McmD C-terminal	Rev.	GCCAAGCTTGGTACCCTCGAGTTAAGTTAATCTATAATATCC
10.	pAW18 / McmB N-terminal	Fwd.	TGTATTTTCAGGGCGCCATGGATGATAAAAAAATAGCCGA
11.	McmD C-terminal McmB N-terminal	Rev.	GGTTGCGGTTTCAACGTGTGTGTCTAAAATGGATATGC
Sequencing of fusion construct			**Sequence 5' > 3'**
12.	84bp upstream of McmB N-term in pPROEXHTa		GAGCGGATAACAATTTCACACAGG
13.	Junction between McmB N-term / McmD C-term		GCAACAGGTCAGAGGCG
14.	Rev. in pPROEXHTa 711bp from vector / McmD C-term junction		GGTCCCACCTGACCCC
Amplification of MMP1024 & MMP0125 for combined cloning in LIC Duet pRSF			
15.	MMP1024 (McmD)	Fwd.	GCGGGCCCGGCCTTGATGGAATTTGATGAAGAAGTTTTTTTGACC
		Rev.	GAGGAGAAGCCCGGTTTAAGTTAATCTATAATATCCATATTTTGGTCAGAA
16.	MMP0125	Fwd.	GACGACGACAAGATGGACGTTTATGATATTTTATTCTTAAAATGC
		Rev.	CGCGGGCGGCCGTTAATTTTGTAAATTAAAATCGTATTTTTTAAGTTTTGAAATTATTTCTTCAAA

Table 3. Table of primers

Strain	**Genotype**
Mm900	*Δhprt*
S0001	*ΔORF1* from pURB500 integrated on genome
S0001DΔ	As above, *Δhprt* McmDΔ
McmBΔ	*Δhprt* McmBΔ
McmCΔ	*Δhprt* McmCΔ
McmDΔ	*Δhprt* McmDΔ
McmBCΔ	*Δhprt* McmBCΔ
McmBDΔ	*Δhprt* McmBDΔ
McmCDΔ	*Δhprt* McmCDΔ
McmBDCΔ	*Δhprt* McmBCDΔ

Table 4. *M. maripaludis* S2 strains used in this study

Strain	**Genotype**
Novablue	*endA1 hsdR17* ($r_{K12}^{-}m_{K12}^{+}$) *supE44 thi- 1 recA1 gyrA96 relA1 lac* F'[*proA+B+ lacI*^{q}ZD*M15::Tn10*]Tet^{R}
Rosetta	F^{-} *ompT hsdS*$_B$($r_B^{-}m_B^{-}$) *gal dcm* (DE3) pRARE2 (Cam^{R})
Arctic Express RIL (Stratagene)	*E. coli* B F^{-} *ampT hsdS*($r_B^{-}m_B^{-}$) *dcm*$^{+}$ *Tet*^{r} *gal endA* Hte *[cpn10 cpn60* Gent^{r}*] [argU ileY leuW Str*^{r}*]*
XL 10 gold (Stratagene)	Tet^{r}*Δ(mcrA)Δ*(*mcrCB-hsdSMR-mrr*) *173 endA1 supE4 thi-1 recA1 gyrA96 relA1 lac Hte* [F′ *proAB lacIqZDM15* Tn*10 (Tet*^{r}*)* Amy Cam^{r}]

Table 5. *E.coli* strains used in this study

Appendices

Appendix I. Amplification of ORFs McmD / $McmD_{20}$

Amplification of *M. maripaludis* ORFs McmD / $McmD_{20}$ from pAW22 / pAW23

PCR Component	Volume	Final Concentration
10X *Pfx* Amplification buffer	5 µl	1X
10 mM dNTP mixture	1.5 µl	0.3 mM each
50 mM $MgSO_4$	1 µl	1 mM
Primer mix (10 µM each) (primers 1 & 2)	0.5 µl of each	0.3 µM each
Template plasmid DNA	1 µl	50 ng
Platinum ® *Pfx* DNA Polymerase	0.4 µl	1 unit
Autoclaved, distilled water	to 50 µl	

Appendix II. Thermocycler programme for use with primers 1. & 2.

Thermocycler programme for use with primers 1. & 2. / Platinum ® *Pfx* DNA Polymerase (amplification of McmD / $McmD_{20}$)

Temperature ^oC	Step	Time
94	Initial denaturation	2 min
94	Denaturation	30 sec
60	annealing	30 sec
68	elongation	2.5 min (1 min = 1Kbp)
68	Final elongation	5 min
~ Hold at 4 ^oC		
30 cycles		

Appendix III. A-tailing reaction of blunt-ended PCR product (McmD / $McmD_{20}$)

Component	Volume
PCR product (post clean-up)	5 µl
Taq 10X buffer (NEB)	1 µl
dATP 0.2mM	1 µl
Taq DNA polymerase (NEB)	1 µl
Autoclaved, distilled water	to 10µl
Incubated at 70^oC for ~30 min	

Standard ligation reaction for manipulation in *E. coli*

Standard ligation reaction for manipulation in *E. coli*; Ligation reaction of A-tailed PCR product (McmD / $McmD_{20}$) for cloning in pGEM®-T easy vector systems, Promega

Component	Volume
T4 DNA ligase (NEB)	1 µl
PCR product (A-tailed)	2 µl
pGEM®T –easy (Promega)	1 µl
10X ligase buffer (NEB)	1 µl
Autoclaved, distilled water	to 10µl
Incubated at 4-10^{o}C overnight	

Appendix IV. In –fusion HD cloning; sample preparation.

In –fusion HD cloning kit sample preparation; for the cloning of McmB N-terminal / McmD C-terminal in pProEXTHa (Invitrogen).

Component	Volume / final concentration
Purified PCR fragments (McmB N-term / McmD C-term)	75 ng
Linearised vector (pProEXHTa)	75 ng
5X In-fusion HD Enzyme premix	2 µl
Autoclaved, distilled water	to 10 µl

Appendix V. Amplification of ORFs MMP1024 / MMP1025

Amplification of MMP1024 / MMP1025 ORFs from *M. maripaludis* genomic DNA

PCR Component	Volume	Final Concentration
5X buffer primeSTAR GXL (Takara)	10 µl	1X
2.5 mM dNTP mixture	4 µl	0.2 mM each
25mM MgCl	2 µl	1 mM
Primer mix (10 µM each) (primers 12 &13)	0.5 µl of each	0.3 µM each
Template genomic DNA	2 µl	100 ng
primeSTAR GXL DNA Polymerase (Takara)	0.4 µl	1 unit
Autoclaved, distilled water	to 50 µl	

Appendix VI. Thermocycler programme for use with primers 15. & 16.

Thermocycler programme for use with primers 15. & 16. / primeSTAR GXL DNA Polymerase (Takara) (amplification of MMP1024 / MMP1025)

Temperature ^{o}C	**Step**	**Time**
98	Initial denaturation	30 sec
98	Denaturation	10 sec
60	annealing	15 sec
68	elongation	10 sec per Kbp (7 secs for 658 and 536bp fragments)
68	Final elongation	5 min
~ Hold at 4 ^{o}C		
30 cycles		

Appendix VII. Treatment of LIC inserts with T4 DNA Poly., creating LIC overhangs

Treatment of LIC (LIC Duet ™ Adaptor kit, Novagen) inserts (amplified and purified) ORFs MMP1024 / MMP1025 with T4 DNA Polymerase for the creation of LIC overhangs.

Component	**Volume / final concentration**
Purified PCR product (stored in 10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0).	0.2 pmol
10X T4 DNA Polymerase Buffer (500 mM Tris-HCl pH 8.0, 100 mM $MgCl_2$, 0.5mg/ml acetylated BSA)	2 µl
25mM dATP	2 µl
100mM DTT	1 µl
2.5 U/µl T4 DNA Polymerase (LIC Duet ™ Adaptor kit, Novagen)	0.4 µl
Autoclaved, distilled water to 20 µl (Final concentration of insert is 0.01 pmol/µl)	

Appendix VIII. Standard restriction digest reaction

Component	**Volume / final concentration**
Template DNA	1-5µg
10 Units of each restriction endonuclease (NEB)	1-2 µl
10X buffer (NEB)	5 µl
100X BSA	0.5 µl
Autoclaved, distilled water to 50 µl , incubated at 37^{o}C for 1 hour	

Appendix IX. Annealing Vector (pRSF), minimal adaptor, and Inserts

Annealing Vector (pRSF), minimal adaptor, (LIC Duet ™ Adaptor kit, Novagen) and Inserts (MMP1024 / MMP1025)

Component	**Volume / final concentration**
pRSF	1 µl
T4 DNA Polymerase-treated target ORF-1(MMP1025) insert (0.02 pmol)	2 µl
LIC Duet minimal adaptor	0.02 pmol
T4 DNA Polymerase-treated target ORF-2 (MMP1024) insert (0.02 pmol)	2 µl
Incubate at 70°C for 30 sec, place at room temperature for 2 min to cool, then add:	
25mM EDTA	1 µl
Mix by stirring with the pipet tip and incubate at 22°C for 5 min.	

Abbreviations

A_{260}	Absorbance measured at 260 nm
AAA+	ATPases associated with diverse cellular activities
Amp	Ampicillin
AC	Anaerobic chamber
ATP	Adenosine triphosphate
ATPase	ATP hydrolysis (activity or domain
ARS	Autonomously replicating sequences (or A-element)
BLAST	Basic Local Alignment Search Tool
bp	DNA base pair(s)
C-term	C-terminal (carboxyl-terminus, C-terminal tail, C-terminal end, COOH-terminus)
DnaA	A trans-acting protein that activates initiation of DNA replication in prokaryotes
DnaA box	cis-acting element, bound to by DnaA
DnaB	Bacterial replicative helicase
DIG	Digoxigenin
dNTPs	Deoxynucleoside triphosphates
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
GINS	Eukaryotic replication factor
H2i	Helix-2 insert
His-tag	Histidine tag
HTH	Helix Turn Helix
IPTG	Isopropyl
LB	Luria-Bertani medium
LIC	Ligation Independent Cloning
MCM	Minichromosome maintenance
*Mth*MCM	*Methanothermobacter thermautotrophicus* MCM protein
13-mer	A repeated unit (DNA base pairs in this context)
min	Minute(s)
NEB	New England BioLabs® *Inc*.
NCBI	National Center for Biotechnology Information
N-term	N terminal (amino terminus, NH_2-terminus, N-terminal end or amine-terminus)
NTPs	Nucleoside Triphosphates
OD_{600}	Optical Density measured at 600 nm
ORB	Origin Recognition Box
ORC	Origin Recognition Complex
ORF	Open Reading Frame
Ori	Origin of DNA replication
OriC	Bacterial origin of replication
PEG	Polyethylene glycol
pGEMT	pGEM®-T easy vector systems, Promega
pPROEXHTa	pPROEX HTa, bacterial expression vector, Invitrogen
preRC	Pre-replication complex
pRSF	pRSFDuet-1 bacterial expression and coexpression vector, EMD Biosciences
PCR	polymerase Chain Reaction
PMSF	Phenylmethylsulphonyl Fluoride
PS1BH	Pre-sensor 1 beta-hairpin
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography

SSC	Saline Sodium Citrate Buffer
TE	Tris EDTA
TEMED	Tetramethylethylenediamine
UV	Ultra Violet
wHTH	winged Helix Turn Helix
WT	Wild Type
w/v	weight/volume
5', 3'	5 prime, 3 prime (directional orientation of DNA molecule)

References

BAPTESTE, E., BROCHIER, C. & BOUCHER, Y. 2005. Higher-level classification of the Archaea: evolution of methanogenesis and methanogens. *Archaea,* 1**,** 353-63.

BARRY, E. R. & BELL, S. D. 2006. DNA replication in the archaea. *Microbiol Mol Biol Rev,* 70**,** 876-87.

BARRY, E. R., MCGEOCH, A. T., KELMAN, Z. & BELL, S. D. 2007. Archaeal MCM has separable processivity, substrate choice and helicase domains. *Nucleic Acids Res,* 35**,** 988-98.

BELL, S. P. 2002. The origin recognition complex: from simple origins to complex functions. *Genes Dev,* 16**,** 659-72.

BELL, S. P. & DUTTA, A. 2002. DNA replication in eukaryotic cells. *Annu Rev Biochem,* 71**,** 333-74.

BOCHMAN, M. L. & SCHWACHA, A. 2007. Differences in the single-stranded DNA binding activities of MCM2-7 and MCM467: MCM2 and MCM5 define a slow ATP-dependent step. *J Biol Chem,* 282**,** 33795-804.

BOCHMAN, M. L. & SCHWACHA, A. 2008. The Mcm2-7 complex has in vitro helicase activity. *Mol Cell,* 31**,** 287-93.

BOCHMAN, M. L. & SCHWACHA, A. 2009. The Mcm complex: unwinding the mechanism of a replicative helicase. *Microbiol Mol Biol Rev,* 73**,** 652-83.

BOCHMAN, M. L. & SCHWACHA, A. 2010. The Saccharomyces cerevisiae Mcm6/2 and Mcm5/3 ATPase active sites contribute to the function of the putative Mcm2-7 'gate'. *Nucleic Acids Res,* 38**,** 6078-88.

BOONE, B. R., CASTENHOLZ, R. W., STALEY, J.T. 2001. *The Archaea and the Deeply Branching and Phototrophic Bacteria,* New York, USA, Springer-Verlag.

BOULIKAS, T. 1996. Common structural features of replication origins in all life forms. *J Cell Biochem,* 60**,** 297-316.

BREWSTER, A. S., WANG, G., YU, X., GREENLEAF, W. B., CARAZO, J. M., TJAJADI, M., KLEIN, M. G. & CHEN, X. S. 2008. Crystal structure of a near-full-length archaeal MCM: functional insights for an AAA+ hexameric helicase. *Proc Natl Acad Sci U S A,* 105**,** 20191-6.

BULT, C. J., WHITE, O., OLSEN, G. J., ZHOU, L., FLEISCHMANN, R. D., SUTTON, G. G., BLAKE, J. A., FITZGERALD, L. M., CLAYTON, R. A., GOCAYNE, J. D., KERLAVAGE, A. R., DOUGHERTY, B. A., TOMB, J. F., ADAMS, M. D., REICH, C. I., OVERBEEK, R., KIRKNESS, E. F., WEINSTOCK, K. G., MERRICK, J. M., GLODEK, A., SCOTT, J. L., GEOGHAGEN, N. S. & VENTER, J. C. 1996. Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii. *Science,* 273**,** 1058-73.

CARPENTIERI, F., DE FELICE, M., DE FALCO, M., ROSSI, M. & PISANI, F. M. 2002. Physical and functional interaction between the mini-chromosome maintenance-like DNA helicase and the single-stranded DNA binding protein from the crenarchaeon Sulfolobus solfataricus. *J Biol Chem,* 277**,** 12118-27.

CHABAN, B., NG, S. Y. M ., JARRELL, K.F. 2006. Archaeal habitats — from the extreme to the ordinary. *Canadian Journal of Microbiology,* 52**,** 73-116.

CHONG, J. P. 2005. Learning to unwind. *Nat Struct Mol Biol,* 12**,** 734-6.

CHONG, J. P., HAYASHI, M. K., SIMON, M. N., XU, R. M. & STILLMAN, B. 2000. A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. *Proc Natl Acad Sci U S A,* 97**,** 1530-5.

CORTEZ, D., GLICK, G. & ELLEDGE, S. J. 2004. Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 10078-10083.

COSTA, A., PAPE, T., VAN HEEL, M., BRICK, P., PATWARDHAN, A. & ONESTI, S. 2006. Structural studies of the archaeal MCM complex in different functional states. *J Struct Biol,* 156**,** 210-9.

DAVEY, M. J. & O'DONNELL, M. 2003. Replicative helicase loaders: ring breakers and ring makers. *Current Biology,* 13**,** R594-R596.

EDGELL, D. R. & DOOLITTLE, W. F. 1997. Archaea and the origin(s) of DNA replication proteins. *Cell,* 89**,** 995-8.

FANGMAN, W. L. & BREWER, B. J. 1991. Activation of replication origins within yeast chromosomes. *Annu Rev Cell Biol,* 7**,** 375-402.

FORSBURG, S. L. 2004. Eukaryotic MCM proteins: beyond replication initiation. *Microbiol Mol Biol Rev,* 68**,** 109-31.

FREEMAN, A., MORRIS, L. S., MILLS, A. D., STOEBER, K., LASKEY, R. A., WILLIAMS, G. H. & COLEMAN, N. 1999. Minichromosome maintenance proteins as biological markers of dysplasia and malignancy. *Clin Cancer Res,* 5**,** 2121-32.

GAMBUS, A., JONES, R. C., SANCHEZ-DIAZ, A., KANEMAKI, M., VAN DEURSEN, F., EDMONDSON, R. D. & LABIB, K. 2006. GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol,* 8**,** 358-66.

GARRETT, R. A., KLENK, H. P. 2007. *Archaea: Evolution, Physiology, and Molecular Biology*, Blackwell publishing.

GAUDIER, M., SCHUWIRTH, B. S., WESTCOTT, S. L. & WIGLEY, D. B. 2007. Structural basis of DNA replication origin recognition by an ORC protein. *Science,* 317**,** 1213-6.

GILLE, H. & MESSER, W. 1991. Localized DNA melting and structural pertubations in the origin of replication, oriC, of Escherichia coli in vitro and in vivo. *EMBO J,* 10**,** 1579-84.

GOMEZ-LLORENTE, Y., FLETCHER, R. J., CHEN, X. S., CARAZO, J. M. & SAN MARTIN, C. 2005. Polymorphism and double hexamer structure in the archaeal minichromosome maintenance (MCM) helicase from Methanobacterium thermoautotrophicum. *J Biol Chem,* 280**,** 40909-15.

GONZALEZ, M. A., TACHIBANA, K. E., LASKEY, R. A. & COLEMAN, N. 2005. Control of DNA replication and its potential clinical exploitation. *Nat Rev Cancer,* 5**,** 135-41.

HAMLIN, J. L., MESNER, L. D., LAR, O., TORRES, R., CHODAPARAMBIL, S. V. & WANG, L. 2008. A revisionist replicon model for higher eukaryotic genomes. *J Cell Biochem,* 105**,** 321-9.

HENDRICKSON, E. L., KAUL, R., ZHOU, Y., BOVEE, D., CHAPMAN, P., CHUNG, J., CONWAY DE MACARIO, E., DODSWORTH, J. A., GILLETT, W., GRAHAM, D. E., HACKETT, M., HAYDOCK, A. K., KANG, A., LAND, M. L., LEVY, R., LIE, T. J., MAJOR, T. A., MOORE, B. C., PORAT, I., PALMEIRI, A., ROUSE, G., SAENPHIMMACHAK, C., SOLL, D., VAN DIEN, S., WANG, T., WHITMAN, W. B., XIA, Q., ZHANG, Y., LARIMER, F. W., OLSON, M. V. & LEIGH, J. A. 2004. Complete genome sequence of the genetically tractable hydrogenotrophic methanogen Methanococcus maripaludis. *J Bacteriol,* 186**,** 6956-69.

HIRST, C. 2011. MCM gene deletion in *Methanococcus maripaludis* S2 (McmD knockout in McmCΔ strain). University of York.

ISHIMI, Y. 1997. A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *J Biol Chem,* 272**,** 24508-13.

JARRELL, K. F., WALTERS, A. D., BOCHIWAL, C., BORGIA, J. M., DICKINSON, T. & CHONG, J. P. 2011. Major players on the microbial stage: why archaea are important. *Microbiology,* 157**,** 919-36.

JENKINSON, E. R. & CHONG, J. P. 2006. Minichromosome maintenance helicase activity is controlled by N- and C-terminal motifs and requires the ATPase domain helix-2 insert. *Proc Natl Acad Sci U S A,* 103**,** 7613-8.

KASIVISWANATHAN, R., SHIN, J. H., MELAMUD, E. & KELMAN, Z. 2004. Biochemical characterization of the Methanothermobacter thermautotrophicus minichromosome maintenance (MCM) helicase N-terminal domains. *J Biol Chem,* 279**,** 28358-66.

KELMAN, L. M. & KELMAN, Z. 2003. Archaea: an archetype for replication initiation studies? *Mol Microbiol,* 48**,** 605-15.

KELMAN, Z., LEE, J. K. & HURWITZ, J. 1999. The single minichromosome maintenance protein of Methanobacterium thermoautotrophicum DeltaH contains DNA helicase activity. *Proc Natl Acad Sci U S A,* 96**,** 14783-8.

KUIPERS, M. A., STASEVICH, T. J., SASAKI, T., WILSON, K. A., HAZELWOOD, K. L., MCNALLY, J. G., DAVIDSON, M. W. & GILBERT, D. M. 2011. Highly stable loading of Mcm proteins onto chromatin in living cells requires replication to unload. *J Cell Biol,* 192**,** 29-41.

LABIB, K., TERCERO, J. A. & DIFFLEY, J. F. 2000. Uninterrupted MCM2-7 function required for DNA replication fork progression. *Science,* 288**,** 1643-7.

LEI, M. 2005. The MCM Complex: Its Role in DNA Replication and Implications for Cancer Therapy. *Current Cancer Drug Targets.* Bentham Science publishers.

LEI, M. & TYE, B. K. 2001. Initiating DNA synthesis: from recruiting to activating the MCM complex. *J Cell Sci,* 114**,** 1447-54.

LIN, D. I., AGGARWAL, P. & DIEHL, J. A. 2008. Phosphorylation of MCM3 on Ser-112 regulates its incorporation into the MCM2-7 complex. *Proc Natl Acad Sci U S A,* 105**,** 8079-84.

LIU, Y. 2010. Methanococcales. *In:* TIMMIS, K. N. (ed.) *Handbook of Hydrocarbon and Lipid Microbiology.* Springer Berlin Heidelberg.

LUO, H., SUN, Z., ARNDT, W., SHI, J., FRIEDMAN, R. & TANG, J. 2009. Gene order phylogeny and the evolution of methanogens. *PLoS One,* 4**,** e6069.

MAINE, G. T., SINHA, P. & TYE, B. K. 1984. Mutants of S. cerevisiae defective in the maintenance of minichromosomes. *Genetics,* 106**,** 365-85.

MCGEOCH, A. T., TRAKSELIS, M. A., LASKEY, R. A. & BELL, S. D. 2005. Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. *Nat Struct Mol Biol,* 12**,** 756-762.

MESSER, W. 2002. The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. *FEMS Microbiol Rev,* 26**,** 355-74.

MILLER, T. L., M. J. WOLIN, E. A. KUSEL. 1986. Isolation and characterization of methanogens from animal feces. . *Syst. Appl. Microbiol. ,* 7**,** 7.

MOORE, B. C. & LEIGH, J. A. 2005. Markerless mutagenesis in Methanococcus maripaludis demonstrates roles for alanine dehydrogenase, alanine racemase, and alanine permease. *J Bacteriol,* 187**,** 972-9.

MUSAHL, C., HOLTHOFF, H. P., LESCH, R. & KNIPPERS, R. 1998. Stability of the Replicative Mcm3 Protein in Proliferating and Differentiating Human Cells. *Experimental Cell Research,* 241**,** 260-264.

NOBLE, S. 2011. MCM gene deletion in *Methanococcus maripaludis* S2 (McmD knockout in McmBΔ strain). University of York.

PEARSON, C. E., ZORBAS, H., PRICE, G. B. & ZANNIS-HADJOPOULOS, M. 1996. Inverted repeats, stem-loops, and cruciforms: significance for initiation of DNA replication. *J Cell Biochem,* 63**,** 1-22.

POPLAWSKI, A., GRABOWSKI, B., LONG, S. E. & KELMAN, Z. 2001. The zinc finger domain of the archaeal minichromosome maintenance protein is required for helicase activity. *J Biol Chem,* 276**,** 49371-7.

PROKHOROVA, T. A. & BLOW, J. J. 2000. Sequential MCM/P1 subcomplex assembly is required to form a heterohexamer with replication licensing activity. *J Biol Chem,* 275**,** 2491-8.

ROBINSON, N. P., DIONNE, I., LUNDGREN, M., MARSH, V. L., BERNANDER, R. & BELL, S. D. 2004. Identification of two origins of replication in the single chromosome of the archaeon Sulfolobus solfataricus. *Cell,* 116**,** 25-38.

SHECHTER, D. F., YING, C. Y. & GAUTIER, J. 2000. The intrinsic DNA helicase activity of Methanobacterium thermoautotrophicum delta H minichromosome maintenance protein. *J Biol Chem,* 275**,** 15049-59.

SHEU, Y. J. & STILLMAN, B. 2006. Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. *Mol Cell,* 24**,** 101-13.

SHI, Y., DODSON, G. E., MUKHOPADHYAY, P. S., SHANWARE, N. P., TRINH, A. T. & TIBBETTS, R. S. 2007. Identification of carboxyl-terminal MCM3 phosphorylation sites using polyreactive phosphospecific antibodies. *J Biol Chem,* 282**,** 9236-43.

STOEBER, K., TLSTY, T. D., HAPPERFIELD, L., THOMAS, G. A., ROMANOV, S., BOBROW, L., WILLIAMS, E. D. & WILLIAMS, G. H. 2001. DNA replication licensing and human cell proliferation. *J Cell Sci,* 114**,** 2027-41.

TAKEI, Y., ASSENBERG, M., TSUJIMOTO, G. & LASKEY, R. 2002. The MCM3 acetylase MCM3AP inhibits initiation, but not elongation, of DNA replication via interaction with MCM3. *J Biol Chem,* 277**,** 43121-5.

THAUER, R. K., KASTER, A. K., SEEDORF, H., BUCKEL, W. & HEDDERICH, R. 2008. Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat Rev Microbiol,* 6**,** 579-91.

TSAO, C.-C., GEISEN, C. & ABRAHAM, R. T. 2004. Interaction between human MCM7 and Rad17 proteins is required for replication checkpoint signaling. *The EMBO Journal,* 23**,** 4660-4669.

TSUJI, T., FICARRO, S. B. & JIANG, W. 2006. Essential role of phosphorylation of MCM2 by Cdc7/Dbf4 in the initiation of DNA replication in mammalian cells. *Mol Biol Cell,* 17**,** 4459-72.

TUMBULA, D. L., BOWEN, T. L. & WHITMAN, W. B. 1997. Characterization of pURB500 from the archaeon Methanococcus maripaludis and construction of a shuttle vector. *J Bacteriol,* 179**,** 2976-86.

WALTERS, A. D. 2009. *The role of multiple maitenance proteins in Methanococcus maripaludis.* PhD, University of York.

WALTERS, A. D. & CHONG, J. P. 2010. An archaeal order with multiple minichromosome maintenance genes. *Microbiology,* 156**,** 1405-14.

WALTERS, A. D., SMITH, S. E. & CHONG, J. P. 2011. Shuttle vector system for Methanococcus maripaludis with improved transformation efficiency. *Appl Environ Microbiol,* 77**,** 2549-51.

WEISS, D. S., AND THAUER, R. K. 1993. Methanogenesis and the unity of biochemistry. *Cell,* 72**,** 819-822.

WEISS, D. S. & THAUER, R. K. 1993. Methanogenesis and the unity of biochemistry. *Cell,* 72**,** 819-22.

WOESE, C. R. & FOX, G. E. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A,* 74**,** 5088-90.

XIA, Q., HENDRICKSON, E. L., ZHANG, Y., WANG, T., TAUB, F., MOORE, B. C., PORAT, I., WHITMAN, W. B., HACKETT, M. & LEIGH, J. A. 2006. Quantitative proteomics of the archaeon Methanococcus maripaludis validated by microarray analysis and real time PCR. *Mol Cell Proteomics,* 5**,** 868-81.