Diversity, Mutagenesis and Recombinant Expression of the Soluble Methane Monooxygenase

Marc G. Dumont

A thesis submitted to the Department of Biological Sciences
in fulfilment of the requirements for the degree of Doctor of Philosophy

July 2004

University of Warwick
Coventry, UK

Table of Contents

List of Tables

List of Figures

Abbreviations

AMO	alkene monooxygenase (Chapter 3); ammonia monooxygenase
AMS	ammonium mineral salts
Ap	ampicillin
ATP	adenosine triphosphate
BAC	bacterial artificial chromosome
bp	nucleotide base pairs
β-RFAP	β-ribofuranosylaminobenzene 5'-phosphate synthetase
Cm	chloramphenicol
CTAB	cetyltrimethylammonium bromide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleotide triphosphate
E.	*Escherichia*
EBP	enhancer binding protein
EDTA	ethylenediamine tetraacetic acid
EtBr	ethidium bromide
FISH	fluorescence *in situ* hybridisation
g	gram
GFP	green fluorescent protein
Gm	gentamycin
GSC	Gisburn soil clone
h	hour
ICM	intracytoplasmic membrane
IPTG	isopropyl β-D-thiogalactopyranoside
kbp	nucleotide kilobase pairs
kDa	kilo Daltons
Km	kanamycin
l	litre
LB	Luria Bertani medium
M	molar
mA	milliamp
Mclla.	*Methylocella*
MDH	methanol dehydrogenase
mg	milligram
min	minute
ml	millilitre
µg	microgram
µl	microlitre
mM	millimolar
mRNA	messenger RNA
Ms.	*Methylosinus*
NAD^+	oxidised nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
$NADP^+$	oxidised nicotinamide adenine dinucleotide phosphate
NADPH	reduced nicotinamide adenine dinucleotide phosphate
ng	nanogram

nm	nanometre
nM	nanomolar
NMS	nitrate mineral salts
OD_{540}	optical density at 540 nm
OTU	operational taxonomic unit
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PFGE	pulsed-field gel electrophoresis
PLFA	phospholipid fatty acid
pMMO	particulate methane monooxygenase
PMSF	phenylmethylsulfonylfluoride
ppmv	parts per million volume
PQQ	pyrroloquinoline quinone
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNAP	RNA polymerase
RNase	ribonuclease
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
RT-PCR	reverse transcription-PCR
RuMP	ribulose monophosphate
s	second
SDS	sodium dodecyl sulphate
SIP	stable isotope probing
Sm	streptomycin
sMMO	soluble methane monooxygenase
sp.	specie
spp.	species
Sp	spectinomycin
SSU rRNA	small subunit ribosomal ribonucleic acid
TAE	Tris acetate EDTA
Taq	*Thermus aquaticus*
TBE	Tris borate EDTA
TCE	trichloroethylene
TE	tris EDTA
TEMED	N, N, N', N' –tetramethylethylenediamine
Tg	teragram
Tris	tris-(hydroxymethyl)-nethylamine
UV	ultraviolet
V	volts
v/v	volume to volume
w/v	weight to volume
x *g*	centrifugal force

Acknowledgements

I am grateful to both Colin Murrell and Carlos Miguez for their assistance in arranging my enrolment at the University of Warwick. Colin has been an excellent supervisor and mentor and I feel privileged to have worked in his laboratory. I am very grateful to Carlos for his encouragement and support during my time "overseas".

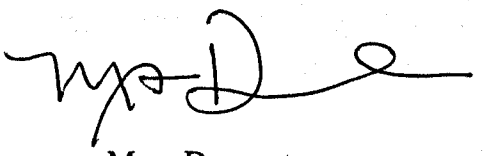
I must thank Stefan Radajewski for his invaluable assistance during the last three years; I never would have cloned any ^{13}C-DNA without his help. Additional thanks to Stefan and to Hendrik Schaefer for comments on the thesis. Thanks to the other members (past and present) of Micro-1; they have been valuable colleagues and many have become good friends.

The most important word of thanks goes to my parents for their tremendous support. Also, I am grateful to my sisters who were always there to help.

I acknowledge the financial support of the FCAR (Quebec, Canada), the Biotechnology Research Institute (CNRC, Canada) and the EU-BOMBBE project.

Declaration

I declare that the work presented in this thesis was conducted by me under the direct supervision of Professor J. Colin Murrell, with the exception of those instances where the contribution of others has been specifically acknowledged. None of the work presented has been previously submitted for any other degree. Some of the data presented in Chapter 4 was published as part of a manuscript (**Hutchens, E., Radajewski, S., Dumont, M. G., McDonald, I. R. & Murrell, J. C. (2004).** Analysis of methanotrophic bacteria in Movile Cave by stable isotope probing. *Environmental Microbiology* **6**, 111-120).



Marc Dumont

Abstract

Methanotrophic bacteria convert methane to methanol using a methane monooxygenase enzyme (MMO). Two types of MMO exist: a membrane bound enzyme (pMMO) and a cytoplasmic enzyme (sMMO). A system for the site-directed mutagenesis of residues in the active site of sMMO has recently been developed that uses a sMMO-minus strain of *Methylosinus trichosporium* OB3b as the expression host (Smith *et al.*, 2002. *Appl. Environ. Microbiol.* 68:5265-5273); this strain, designated Mutant F, was created by disrupting the *mmoX* gene by marker-exchange mutagenesis. In this study a *Ms. trichosporium* OB3b strain was created in which all the sMMO structural genes were deleted or disrupted. This mutant was designated *Ms. trichosporium* SMDM. The recombinant expression of sMMO was performed in *Ms. trichosporium* SMDM using the same sMMO expression plasmid used for *Ms. trichosporium* Mutant F. The effect of sMMO expression in the absence of the enigmatic *mmoD* gene was investigated. Preliminary results indicate that *mmoD* is required for active expression of sMMO. The sMMO genes from *Methylocella silvestris* $BL2^T$ were sequenced and conjugated into *Ms. trichosporium* SMDM on a broad host range plasmid. No expression of *Methylocella silvestris* $BL2^T$ sMMO was detected in *Ms. trichosporium* SMDM.

A new system for the mutagenesis of the *Ms. trichosporium* OB3b sMMO α-subunit was created. Chimaeric sMMO mutants were created by introducing gene sequence from the alkene monooxygenase enzyme of *Rhodococcus corallinus* into the *mmoX*. The chimaeric sMMO enzymes appeared to be unstable in *Ms. trichosporium* Mutant F. An attempt was made to improve the stability of sMMO mutants in *Ms. trichosporium* Mutant F by disrupting the gene encoding the Lon protease. The *Ms. trichosporium* OB3b *lon* gene was cloned and sequenced and attempts were made to disrupt the *Ms. trichosporium* OB3b *lon* by marker-exchange mutagenesis. A mutant was not obtained, suggesting that Lon may be essential for vegetative growth of *Ms. trichosporium* OB3b.

The diversity of sMMO in several environmental samples was investigated using PCR. The objective was to isolate novel *mmoX* sequences from uncultivated methanotrophs that could be used to design sMMO mutagenesis experiments. New PCR primers targeting the *mmoX* were developed. The primers were used to generate libraries from a blanket bog peat (UK) and from cave water (Romania). A group of sequences that did not cluster with the *mmoX* of any cultivated methanotroph was obtained from the cave water. The use of a PCR independent approach to clone methanotroph genes from environmental samples was also investigated. This was performed by developing a method to clone ^{13}C-DNA from a stable isotope probing experiment with $^{13}CH_4$ into a BAC vector. A library of 2300 clones was generated. Greater than 95 % of plasmids analysed contained inserts, which ranged in size from approximately 10 – 30 kb. The library was screened for *mxaF*, *mmoX* and *pmoA* by colony hybridization. A clone (15 kb) containing *pmoA* was completely sequenced. Other genes encoding proteins with (potential) roles in methylotrophy were contained on the clone, including *pmoC, pmoB, folP, folK, mptG* and *moxF*.

Chapter 1

Introduction

1.1 Introduction

The soluble methane monooxygenase (sMMO) is an enzyme than catalyses the oxidation of methane to methanol according to the following equation:

$$CH_4 + NADH + H^+ + O_2 \rightarrow CH_3OH + NAD^+ + H_2O$$

The sMMO is one of two known forms of methane monooxygenase possessed by methanotrophic bacteria and performs the first step in the metabolic pathway that enables these organisms to use methane as a sole source of carbon and energy. The oxidation of methane by methanotrophs plays a significant role in the global carbon cycle and has a major impact on the atmospheric chemistry of the planet (Section 1.6.1). In addition, there is potential in utilising the sMMO in bioremediation and biotechnological processes since the enzyme catalyses the fortuitous oxidation of many organic compounds (Section 1.7). Although considerable research has focussed on the sMMO, the biology of this enzyme is not fully understood and the enzyme has yet to be exploited in commercial applications. This chapter contains an introduction to the biology of methanotrophic bacteria and reviews vital literature relating to the sMMO enzyme.

1.2 Methanotrophic bacteria

1.2.1 Basic characteristics

Methanotrophs are aerobic bacteria that grow using methane, and all representatives can use methane as a sole source of carbon and energy (Anthony, 1982). They are a subgroup of the methylotrophic bacteria, which are a phylogenetically diverse group of organisms capable of growth on one-carbon compounds (Bratina *et al.*, 1992; Bulygina *et al.*, 1990; Bulygina *et al.*, 1993; Murrell & McDonald, 2000). The remarkable ability of methanotrophs to oxidise methane to methanol, using either the particulate MMO (pMMO) enzyme (Section 1.4) or sMMO enzyme (Section 1.5), is what distinguishes the methanotrophs from other Gram-negative methylotrophs (Anthony, 1986).

Most methanotrophs are obligate methylotrophs, growing only on methane and sometimes other one-carbon compounds (Bowman *et al.*, 1993; Hanson & Hanson, 1996). It is easy to see why methanotrophy is a highly successful lifestyle: methane is abundant in many natural habitats and there is no competition from other groups of organisms for this carbon. Reports of facultative methanotrophs reputed to grow on multicarbon sources are present in the literature; however these strains have never been thoroughly studied, and their existence is still questioned (Hanson & Hanson, 1996). Putative facultative methanotrophs include the organisms referred to as *Methylobacterium organophilum* XX (Patt *et al.*, 1974; Patt *et al.*, 1976) and *Methylomonas* sp. strain 761 (Zhao & Hanson, 1984). Most recently *Methylocella silvestris* sp. strain BL2^T has been reported to grow on ethanol and acetate (Dunfield & Dedysh, 2003) and this has been authenticated in our laboratory (Andreas Theisen, personal communication).

Most methanotrophs possess either pMMO or both pMMO and sMMO. The complete metabolism of methane to carbon dioxide (through methanol, formaldehyde and formate intermediates) generates reductant for the cell, and this biochemistry is described in Section 1.3. Methanotrophs incorporate carbon using either the serine or RuMP pathway for formaldehyde fixation. A distinguishing characteristic of methanotrophs, although similar features are also found in autotrophic nitrifiers, are extensive intracytoplasmic membranes (ICMs) visible by electron microscopy (Dedysh *et al.*, 2002; Green, 1992; Whittenbury *et al.*, 1970). The ICMs may serve as

membrane area for the pMMO enzyme, and in *Methylococcus capsulatus* (Bath) the ICMs are absent in cells grown under sMMO expressing conditions (Prior & Dalton, 1985b). Other specific characteristics of methanotrophs and individual methanotroph genera are discussed in the following section.

1.2.2 Taxonomy

The first detailed isolation and cataloguing of methanotrophs was performed by Whittenbury and co-workers (1970). Over 100 methane utilising organisms were isolated and classified based on physiological characteristics, morphological differences, types of resting stages formed and the fine structure of their ICMs. Five genera were proposed: *Methylomonas*, *Methylobacter*, *Methylocystis*, *Methylosinus* and *Methylococcus*. These genera remain valid today, although the genus *Methylomicrobium*, which were previously classified as *Methylobacter* species, has been added together with several new methanotroph genera (see Table 1.1 and references therein).

Whittenbury and co-workers (1970, 1981 & 1984) separated the genera into three groups: type I (*Methylomonas* and *Methylobacter*), type II (*Methylocystis* and *Methylosinus*) and type X (*Methylococcus*). Type I methanotrophs used the RuMP pathway for formaldehyde assimilation and had intracytoplasmic membranes arranged in bundles of vesicular discs; type II methanotrophs used the serine pathway for formaldehyde assimilation and had intracytoplasmic membranes arranged in pairs around the periphery of the cell; type X methanotrophs were similar to type I, but grew at elevated temperatures and possessed some serine pathways enzymes and the RuBP carboxylase (Whittenbury *et al.*, 1970; Whittenbury & Dalton, 1981; Whittenbury & Krieg, 1984). Methanotrophs are still often referred to as type I, II or X.

There are currently 11 recognised methanotroph genera (Table 1.1). Phylogenetic classification based on rRNA gene sequence has shown that the type I methanotrophs group within the γ–subclass of the *Proteobacteria*, and have been assigned the family *Methylococcaceae* (Bowman, 2000; Whittenbury & Krieg, 1984; Wise *et al.*, 2001). The type X methanotroph, *Methylococcus capsulatus*, is very closely related to the type I organisms, and *Methylococcus* is the type genus of the *Methylococcaceae* family (Bowman *et al.*, 1993). For this reason *Methylococcus* spp. are sometimes referred to as type I rather than type X. The type II methanotrophs

Table 1.1 Methanotroph genera and their characteristics.

Genus name	Phylogeny	MMO type	C-1 assimilation	ICM type	N_2 fixation	DNA GC content (mol %)	Major PLFA	Trophic niche	Reference
Methylobacter	γ–*Proteobacteria*	pMMO	RuMP	type I	no	49-54	16:1	some psychrophilic	1, 2, 3, 4
Methylococcus	γ–*Proteobacteria*	pMMO + sMMO	RuMP/Serine	type I	yes	59-66	16:1	thermophilic	1, 5
Methylocaldum	γ–*Proteobacteria*	pMMO	RuMP/Serine	type I	no	57	16:1	thermophilic	6
Methylomicrobium	γ–*Proteobacteria*	pMMO +/– sMMO	RuMP	type I	no	49-60	16:1	some halotolerant some alkaliphilic	7, 8, 9 10, 11, 12, 13
Methylomonas	γ–*Proteobacteria*	pMMO +/– sMMO	RuMP	type I	some strains	51-59	16:1	some psychrophilic	1, 2, 14, 15, 16
Methylosarcina	γ–*Proteobacteria*	pMMO	RuMP	type I	no	54	16:1	not extreme	17
Methylosphaera	γ–*Proteobacteria*	pMMO	RuMP	ND[a]	yes	43-46	16:1	psychrophilic	18
Methylocystis	α–*Proteobacteria*	pMMO +/– sMMO	Serine	type II	yes	62-67	18:1	not extreme	1, 2
Methylosinus	α–*Proteobacteria*	pMMO + sMMO	Serine	type II	yes	63-67	18:1	not extreme	1, 2
Methylocella	α–*Proteobacteria*	sMMO	Serine	NA[b]	yes	60-61	18:1	acidophilic	19, 20
Methylocapsa	α–*Proteobacteria*	pMMO	Serine	type III	yes	63.1	18:1	acidophilic	21

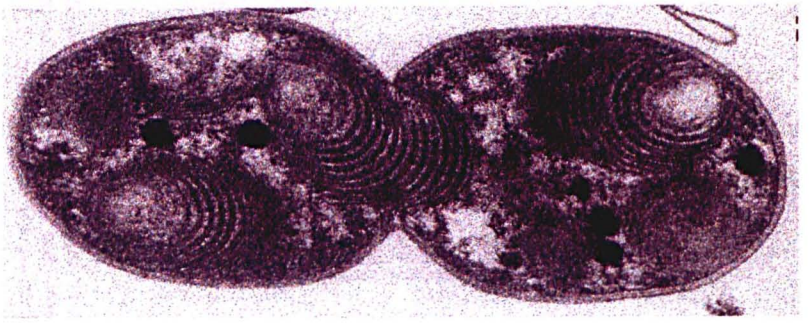
a. ND, not determined; b. NA, not applicable; ICMs are very limited in this genus. 1. (Whittenbury *et al.*, 1970); 2. (Bowman *et al.*, 1993); 3. (Omelchenko *et al.*, 1996); 4. (Kalyuzhnaya *et al.*, 1998); 5. (Malashenko *et al.*, 1975); 6. (Bodrossy *et al.*, 1997); 7. (Bowman *et al.*, 1995); 8. (Sieburth *et al.*, 1987); 9. (Fuse *et al.*, 1998); 10. (Khmelenina *et al.*, 1997) 11. (Kalyuzhnaya *et al.*, 1999b); 12. (Sorokin *et al.*, 2000); 13. (Kaluzhnaya *et al.*, 2001); 14. (Whittenbury & Krieg, 1984); 15. (Omelchenko *et al.*, 1996); 16. (Kalyuzhnaya *et al.*, 1999a) 17. (Wise *et al.*, 2001); 18. (Bowman *et al.*, 1997); 19. (Dedysh *et al.*, 2000); 20. (Dunfield *et al.*, 2003); 21. (Dedysh *et al.*, 2002).

cluster within the α–subclass of the *Proteobacteria* and were assigned the family *Methylocystaceae* (Bowman, 2000). The PLFA profiles of methanotrophs can be used to distinguish type I and II methanotrophs. The membranes of type I methanotrophs contain predominantly 16:1ω7c, 16:1ω5t and 14:0 PLFA molecules; in type X organisms the 16:1ω7c and 16:0 PLFAs dominate. In type II methanotrophs the 18:1ω8c PLFA molecules are most abundant (Bowman *et al.*, 1991). Methanotroph PLFA profiles have been used in molecular ecology studies to detect methanotrophs (Section 1.6.2.4).

Recently, two new genera, *Methylocapsa* and *Methylocella*, have been isolated using an acidic, low nutrient medium (Dedysh *et al.*, 2000; Dedysh *et al.*, 2002). *Methylocapsa* and *Methylocella* are monophyletic α–*Proteobacteria,* but are phylogenetically more closely related to the non-methane-utilising acidophile *Beijerinckia indica* than to *Methylosinus/Methylocystis*. Like other α–*Proteobacteria* methanotrophs, both *Methylocapsa acidiphila* and *Methylocella* spp. incorporate formaldehyde using the serine pathway. Unlike all other characterised methanotrophs, the *Methylocella* species examined appear to lack a pMMO and possess only sMMO (Dedysh *et al.*, 2000; Dedysh *et al.*, 2004; Dunfield & Dedysh, 2003). The methanotroph ICMs are also diminished in the species examined. *Methylocapsa acidiphila* does possess a pMMO. The intracytoplasmic membranes in *Methylocapsa acidiphila*, which the authors have termed type III, are unusual and appear to resemble a hybrid of the type I and type II membrane structure (Figure 1.1) (Dedysh *et al.*, 2002).

With the discovery of the *Methylocella/Methylocapsa* methanotroph genera that do not fit the type I and II classification scheme, in lieu of designating these genera as type III, methanotrophs will be referred to as either α– or γ–*Proteobacteria* methanotrophs.

Type I (*Methylomonas methanica*)



Type II (*Methylocystis parvus*)

Type III (*Methylocapsa* sp. B2)

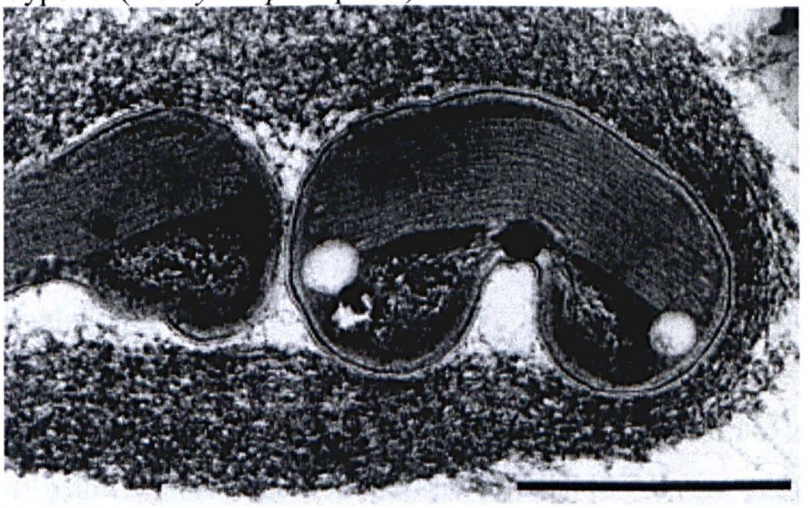


Figure 1.1 Scanning electron micrographs of Type I, II and III ICM membranes. Types I and II taken from Green (1992); Type III taken from Dedysh *et al.* (1992).

1.3 Biochemistry of methane oxidation

Methanotrophs oxidise methane to CO_2 via the intermediates methanol, formaldehyde and formate and incorporate carbon into biomass using the serine or RuMP pathway for formaldehyde fixation (Figure 1.2). A detailed description of the biochemistry of methanotrophs is available (Anthony, 1982). Approximately 50 % of the methane is incorporated into cell carbon and the remainder is oxidised completely to CO_2 which generates reducing power that drives cell metabolism. The oxidation of methane to methanol is performed by MMO (either pMMO or sMMO). Methanol dehydrogenase (MDH) performs the second step in the oxidation pathway. In Gram-negative methylotrophs, which include all known methane-oxidisers, the MDH is a periplasmic pyrroloquinoline quinone (PQQ)-dependent enzyme and consists of a tetrameric dimer ($\alpha_2\beta_2$) of large and small subunits (Anthony & Dales, 1996). The large subunit is encoded by *mxaF* which has been used extensively as a phylogenetic marker for methylotrophs in ecological studies (Section 1.6.2.3). The X-ray crystal structure of MDH has been solved for the enzymes from *Methylobacterium extorquens* (Ghosh *et al.*, 1995) and *Methylophilus* W3A1 (Xia *et al.*, 1996).

Formaldehyde is the branch point intermediate between the energy generating dissimilatory reactions of the methane oxidation pathway and the carbon assimilation pathway. There are several pathways for formaldehyde metabolism in methylotrophs and these serve the purpose of formaldehyde detoxification, formaldehyde dissimilation and energy production, or carbon assimilation (Lidstrom, 2001). Two folate cofactors are involved in formaldehyde metabolism: tetrahydrofolate (H_4F) and tetrahydromethanopterin (H_4MPT). The H_4F-linked pathway is only found in serine pathway methylotrophs (Vorholt *et al.*, 1999). Formaldehyde reacts spontaneously with H_4F to form a methylene-H_4F derivative that can transfer the C-1 unit directly to the serine pathway or can be oxidised to generate formate and NADH. Both serine and RuMP pathway methylotrophs possess the H_4MPT-linked pathway in which a methylene-H_4MPT derivative undergoes several reactions to generate formate and energy (Lidstrom, 2001; Vorholt *et al.*, 1999). This pathway is also present in methanogenic and sulfate-reducing archaea (Chistoserdova *et al.*, 1998). In serine cycle methylotrophs, such as *Methylobacterium extorquens,* the H_4MPT-linked

pathway is the major route for formaldehyde dissimilation, whereas in RuMP pathway methylotrophs, the role

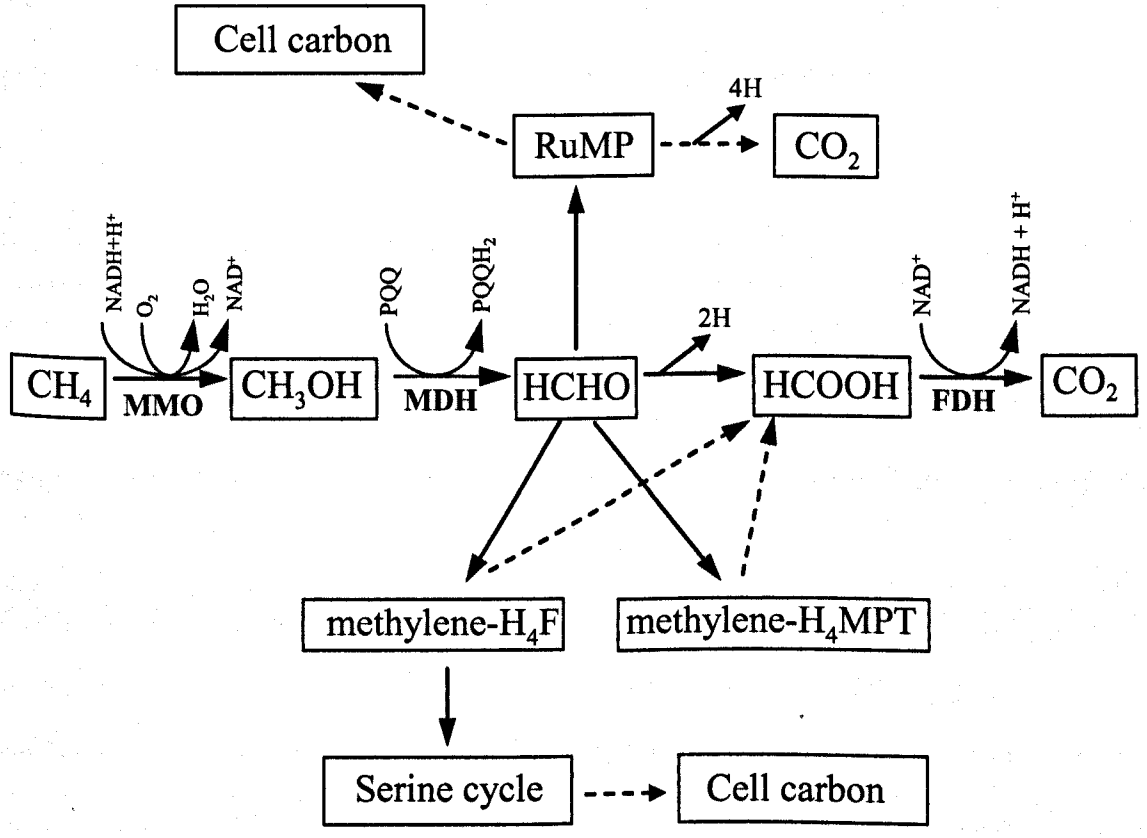


Figure 1.2 The methane oxidation pathway in RuMP and serine pathway methanotrophs.

appears to be minor and contributes mainly to formaldehyde detoxification (Chistoserdova *et al.*, 2000). In RuMP pathway methylotrophs, most of the formaldehyde appears to bind ribulose monophosphate and follows the assimilatory pathway or a cyclic dissimilatory pathway (Lidstrom, 2001).

1.4 Particulate methane monooxygenase (pMMO)

The particulate methane monooxygenase (pMMO) is responsible for the membrane associated methane oxidation activity that was first described in the 1970s (Colby *et al.*, 1975; Ferenci *et al.*, 1975; Tonge *et al.*, 1975). Compared with the sMMO, the pMMO appears to be the "favoured" methane oxidation enzyme in that it is possessed by nearly all methanotrophs and is preferentially expressed in organisms with both pMMO and sMMO (Murrell *et al.*, 2000). Copper is absolutely required for pMMO activity and not only enhances transcription of the genes encoding the enzyme (Choi *et al.*, 2003; Nielsen *et al.*, 1996; Nielsen *et al.*, 1997; Stolyar *et al.*, 2001), but stimulates enzyme production and activity (Basu *et al.*, 2003; Choi *et al.*, 2003; Lieberman *et al.*, 2003; Nguyen *et al.*, 1998; Takeguchi *et al.*, 1998; Yu *et al.*, 2003). Copper also enhances the synthesis of the intracytoplasmic membranes (ICMs) in which the enzyme is found (Choi *et al.*, 2003; Prior & Dalton, 1985b). The optimal copper concentration for pMMO expression is dependent on the cell density of the culture, and has been optimised for fermentor cultures (Choi *et al.*, 2003; Yu *et al.*, 2003).

Unfortunately, the pMMO has been recalcitrant to experimental study since the purified enzyme is unstable. However there have been recent technical advances, such as the use of dodecyl-β-D-maltoside detergent to solubilize the enzyme (Smith & Dalton, 1989), that have led to highly purified preparations of active pMMO (Choi *et al.*, 2003; Lieberman *et al.*, 2003; Nguyen *et al.*, 1998; Takeguchi *et al.*, 1998; Yu *et al.*, 2003; Zahn & DiSpirito, 1996). Despite recent advances, the enzyme is still poorly understood and researchers disagree on some basic biochemistry, including the role(s) of copper, the involvement of iron and the source of reductant. The literature is summarized in this section.

1.4.1 Biochemistry of the particulate methane monooxygenase

The pMMO is composed of three subunits, α (PmoB) β (PmoA) γ (PmoC) of approximately 45, 27 and 25 kDa respectively. Active pMMO enzyme preparations are approximately 200 kDa (Basu *et al.*, 2003; Lieberman *et al.*, 2003; Yu *et al.*,

2003), which suggests the enzyme may be an $(\alpha\beta\gamma)_2$ dimer (Lieberman *et al.*, 2003), or the additional mass could be detergent molecules associated with the monomer (Yu *et al.*, 2003). The 27 kDa protein is believed to contain the active site since radiolabelling studies have shown that the suicide substrate acetylene binds primarily to this subunit (Prior & Dalton, 1985a; Zahn & DiSpirito, 1996).

It is generally accepted that copper is present in the active site of the pMMO and is necessary for catalysis. Active preparations of purified enzyme contain at least two molar equivalents of copper and the EPR spectra indicate that it most resembles a type II copper centre (Basu *et al.*, 2003; Lieberman *et al.*, 2003). The involvement of iron in pMMO catalysis is also debated. One (Basu *et al.*, 2003; Lieberman *et al.*, 2003) or two (Choi *et al.*, 2003) molar equivalents of iron have been found in active preparations of pMMO and are suggested to be essential for catalysis. Nguyen and coworkers (1998) attributed the iron to contamination and this group has recently reported the isolation of active pMMO that contained virtually no iron (Yu *et al.*, 2003).

In addition to the copper associated with catalysis, pMMO preparations also contain a significant amount of less tightly bound copper that can be removed by repeated washes with buffer (Basu *et al.*, 2003; Choi *et al.*, 2003; Lieberman *et al.*, 2003; Zahn & DiSpirito, 1996). The role of the loosely bound copper is unclear, although the removal does appear to compromise the specific activity of the enzyme. At least some of this copper appears to be associated with copper-binding compounds (CBCs). CBCs are low-molecular-mass peptides that are found both in the cytoplasmic membranes and in the growth medium of *Mc. capsulatus* (Bath) and *Ms. trichosporium* OB3b (DiSpirito *et al.*, 1998; Tellez *et al.*, 1998; Zahn & DiSpirito, 1996). In addition to a role in copper acquisition, the Cu-CBCs may have a function in pMMO activity (DiSpirito *et al.*, 1998; Zahn & DiSpirito, 1996) and have been shown to exhibit superoxide dismutase-like activity that may protect the pMMO from oxidative damage (Choi *et al.*, 2003). Chan and colleagues believe that some of this copper is associated directly with the pMMO enzyme and is necessary for electron transfer (Chan *et al.*, 2004; Nguyen *et al.*, 1998). The authors have found that each $\alpha\beta\gamma$ monomer contains 15 copper atoms and that these are arranged in two trinuclear catalytic centres (designated "C-clusters") and three trinuclear electron transfer centres (designated E-clusters") (Chan *et al.*, 2004).

The biological reductant for pMMO has not been well defined. Several research groups believe that the enzyme is coupled to the electron transport chain at cytochrome bc_1 (reviewed in Choi *et al.*, 2003). Duroquinol is an effective artificial reductant for pMMO (reviewed in Chan *et al.*, 2004), and the possibility exists that an *in vivo* pool of a plastoquinol could be the source of reducing equivalents (Shiemke *et al.*, 1995). The Chan group believes that NADH provides reducing equivalents directly to the pMMO via the E-cluster coppers, which can be bypassed using duraquinol (Chan *et al.*, 2004; Nguyen *et al.*, 1998; Yu *et al.*, 2003). However, DiSpirito and colleagues suggest that NADH probably does not provide reducing equivalents directly to the pMMO and that the observed NADH-driven activity with purified pMMO is mediated through contamination with NADH:quinone oxidoreductase (Choi *et al.*, 2003).

The pMMO and the ammonia monooxygenase (AMO) of autotrophic nitrifiers are thought to be evolutionarily related (Holmes *et al.*, 1995) and possess many biochemical similarities. Like the pMMO, the AMO is a copper containing enzyme and is composed of three subunits (Ensign *et al.*, 1993). The genes encoding the AMO are arranged in a similar manner to those encoding pMMO, and this is described in the following section. Acetylene is a suicide substrate of both pMMO and AMO and binds the 27 kDa subunit of both enzymes (Hyman & Arp, 1992; McTavish *et al.*, 1993; Prior & Dalton, 1985a). The enzymes also have broadly similar substrate specificities (Arp *et al.*, 2002; Hooper *et al.*, 1997).

1.4.2 Genetic organisation of genes encoding the particulate methane monooxygenase

Genes encoding components of the pMMO were first cloned from *Mc. capsulatus* (Bath) (Semrau *et al.*, 1995). The genes were detected by Southern blotting using an oligonucleotide probe derived from the N-terminal sequence of the 45 kDa band (PmoB) of *Mc. capsulatus* (Bath) and *Methylomicrobium album* BG8. The sequence of the *amoA* and *amoB* genes from *Nitrosomonas europaea* had previously been determined (McTavish *et al.*, 1993) and Southern blots of methanotroph DNAs were also probed using an *amoA* gene probe. Both probes hybridized with the methanotroph DNA and a 2.1 kb chromosomal fragment containing both *pmoA* and *pmoB* from *Mc. capsulatus* (Bath) was cloned and

sequenced. The probing performed in these studies indicated that additional copies of the genes were present in the genome.

The other *pmo* genes of *Mc. capsulatus* (Bath) were sequenced (Stolyar *et al.*, 1999). A second copy of the *pmoCAB* operon and a third lone *pmoC* gene were found. Sequence analysis indicated that the two *pmoCAB* operons were very similar and differed in only in 13 nucleotide residues over the 3,183 bp of coding region. Most of the changes were synonymous and the derived aminoacyl sequences differed at only one residue in each polypeptide. The third copy of *pmoC* contained several differences to the other copies and the amino acid changes occurred primarily at the N- and C-termini of the derived protein sequence. In an effort to assess the roles of the multiple *pmo* gene copies in this organism, Stolyar and coworkers (1999) disrupted the genes by allelic-exchange mutagenesis. Mutating the *pmoCAB2* genes had a greater effect on reducing methane oxidation rates than did mutating the *pmoCAB1* genes. The most interesting findings were that they could not obtain a mutant in the lone *pmoC* gene, nor were they able to obtain a double mutant in both a *pmoCAB1* and *pmoCAB2* gene, even under sMMO expressing growth conditions.

The *pmoCAB* genes have been sequenced from several other methanotrophs in addition to *Mc. capsulatus* (Bath). Two copies of the operon are present in the other species examined, but unlike *Mc. capsulatus* (Bath), a third copy of *pmoC* has not been observed in other organisms. Two copies of *pmoCAB* are present in both *Ms. trichosporium* OB3b and *Methylocystis* sp. strain M and one of the copies from each organism was sequenced (Gilbert *et al.*, 2000). Although the other *pmoCAB* genes were not sequenced the authors believed them to be very similar. Two very similar copies of *pmoCAB* have been detected and partially sequenced from *Methylocaldum szegediense* strain B1 (Smith, 2004). Several α-*Proteobacteria* methanotrophs, such as *Methylocystis* sp. strain SC2, contain two *pmoCAB* operons, one of which is very similar to those of *Ms. trichosporium* OB3b and *Methylocystis* sp. strain M, and another which is significantly different (Dunfield *et al.*, 2002; Tchawa Yimga *et al.*, 2003). The two copies of the *pmoCAB* operons were cloned and sequenced from *Methylocystis* sp. strain SC2 (Ricke *et al.*, 2004) and RT-PCR analysis showed that both copies were transcribed (Tchawa Yimga *et al.*, 2003). The transcriptional start sites for *pmoCAB* operons have been mapped in *Mc. capsulatus* (Bath) (Stolyar *et al.*, 2001) *Ms. trichosporium* OB3b (Gilbert *et al.*, 2000) and *Methylocystis* sp. strain SC2

(Ricke *et al.*, 2004), and a consensus sequence resembling that of σ^{70}-like promoters were identified. Northern blot analysis indicated that transcription of *pmoCAB* was polycistronic (Nielsen *et al.*, 1996; Nielsen *et al.*, 1997). The transcriptional regulation of *pmoCAB* is described in Section 1.5.3.

1.5 The soluble methane monooxygenase

The sMMO is possessed by some methanotrophs, usually in addition to a pMMO. The enzyme is cytoplasmic and can be easily purified and assayed. It is a member of the non-haem diiron monooxygenases that include, amongst others, the alkene monooxygenase of *Rhodococcus corallinus* (Saeki & Furuhashi, 1994), toluene-*o*-monooxygenase from *Burkholderia cepacia* G4 (Newman & Wackett, 1995), toluene 4-monooxygenase from *Pseudomonas mendocina* KR1 (Pikus *et al.*, 1996) and the phenol hydroxylase from *Pseudomonas* sp. strain CF600 (Nordlund *et al.*, 1990); a recent study has reviewed the diversity of this enzyme family and some deductions relating to the evolution of the enzymes have been proposed (Leahy *et al.*, 2003). The closest known relative of the sMMO appears to be the butane monooxygenase from "*Pseudomonas butanovora*" (Sluis *et al.*, 2002). The sMMO, and in particular those of *Mc. capsulatus* (Bath) and *Ms. trichosporium* OB3b, is the best characterised of the diiron monooxygenases. The biochemistry of sMMO has been reviewed extensively (Lipscomb, 1994; Merkx *et al.*, 2001).

1.5.1 Biochemistry of the soluble methane monooxygenase

The sMMO is a three-component enzyme consisting of a hydroxylase subunit, a reductase and a coupling protein (Figure 1.3). The hydroxylase is composed of three subunits in an $\alpha_2\beta_2\gamma_2$ configuration of approximately 250 kDa (Fox *et al.*, 1989). The reductase (MmoC) is a single protein of approximately 39 kDa that obtains reducing equivalents from NADH. The coupling protein (MmoB) is a 16 kDa protein that shuttles electrons from MmoC to the active site. Each of these sMMO components is described separately in this section.

1.5.1.1 The sMMO hydroxylase

The activation of oxygen and the reaction with methane to produce methanol occurs in the hydroxylase. The hydroxylase is composed of three subunits, designated α, β and γ, of 60, 45 and 20 kDa, respectively. The crystal structures of the

hydroxylase from both *Mc. capsulatus* (Bath) (Rosenzweig *et al.*, 1993) and *Ms. trichosporium* OB3b (Elango *et al.*, 1997) have been solved to 1.7 Å and 2.0 Å, respectively. The structures of the hydroxylases from both organisms were found to be very similar (Elango *et al.*, 1997). The Lippard group at MIT has generated additional crystals of the *Mc. capsulatus* (Bath) hydroxylase in various forms, redox states and in the presence of various substrates and products; the results of these experiments have been reviewed (Merkx *et al.*, 2001).

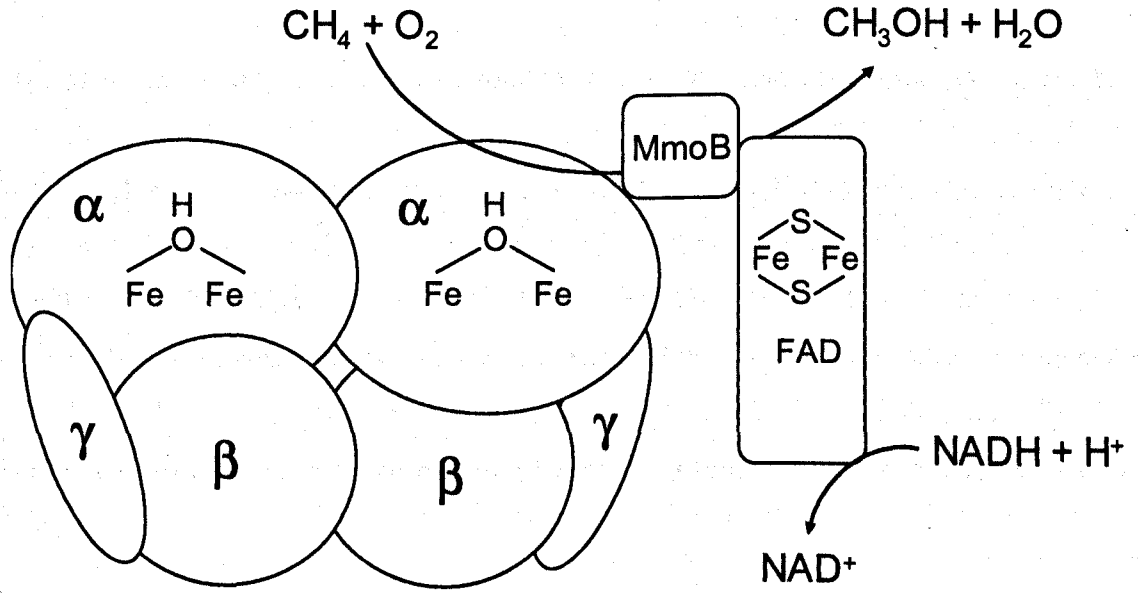


Figure 1.3 Schematic representation of the sMMO. The components are described in the text.

The overall structure of the hydroxylase is that of a "heart-shaped" dimer of $\alpha\beta\gamma$ protomers (Figure 1.4). The protomers are arranged across a two-fold symmetry axis between which a wide canyon is formed. Most of the interaction between the protomers is between helices of the β–subunits. The γ–subunits flank the sides of the hydroxylase and have no contribution to dimerization. Ten of the twelve helices of the β–subunit contain virtually identical folds to those of the α-subunit and it has been suggested that they have evolved from a common ancestor (Leahy *et al.*, 2003). The γ–subunit, with the exception of the butane monooxygenase, is unique to sMMO and its function is not clear.

The highly helical α-subunit contains the active site of the enzyme. The core region contains a non-haem bis-μ-hydroxo-bridged diiron, which is coordinated between four helix bundles by two glutamate-X-X-histidine motifs (Glu144 & His147; Glu243 & His246) and two glutamate residues (Glu114 & Glu209) (Figure 1.5). Similar diiron centres occur in other enzymes, including hemerythrin, rubrerythrin, class 1 ribonucleotide reductase and stearoyl-ACP Δ^9 desaturase (deMare *et al.*, 1996; Wallar & Lipscomb, 1996). The Glu243 in the sMMO hydroxylase, and the analogous glutamate in the related enzymes, undergoes a so-called "carboxylate-shift" upon reduction of the diiron centre from the diferric (Fe^{II}-Fe^{II}) state to the diferrous state (Fe^{III} - Fe^{III}). In the oxidised state, one oxygen atom of the carboxyl ligates one of the irons (Fe2), but shifts upon reduction such that it displaces the bridging solute and coordinates between both irons (Figure 1.5). The other carboxylate oxygen atom provides an additional bridge to the diferrous Fe2 atom.

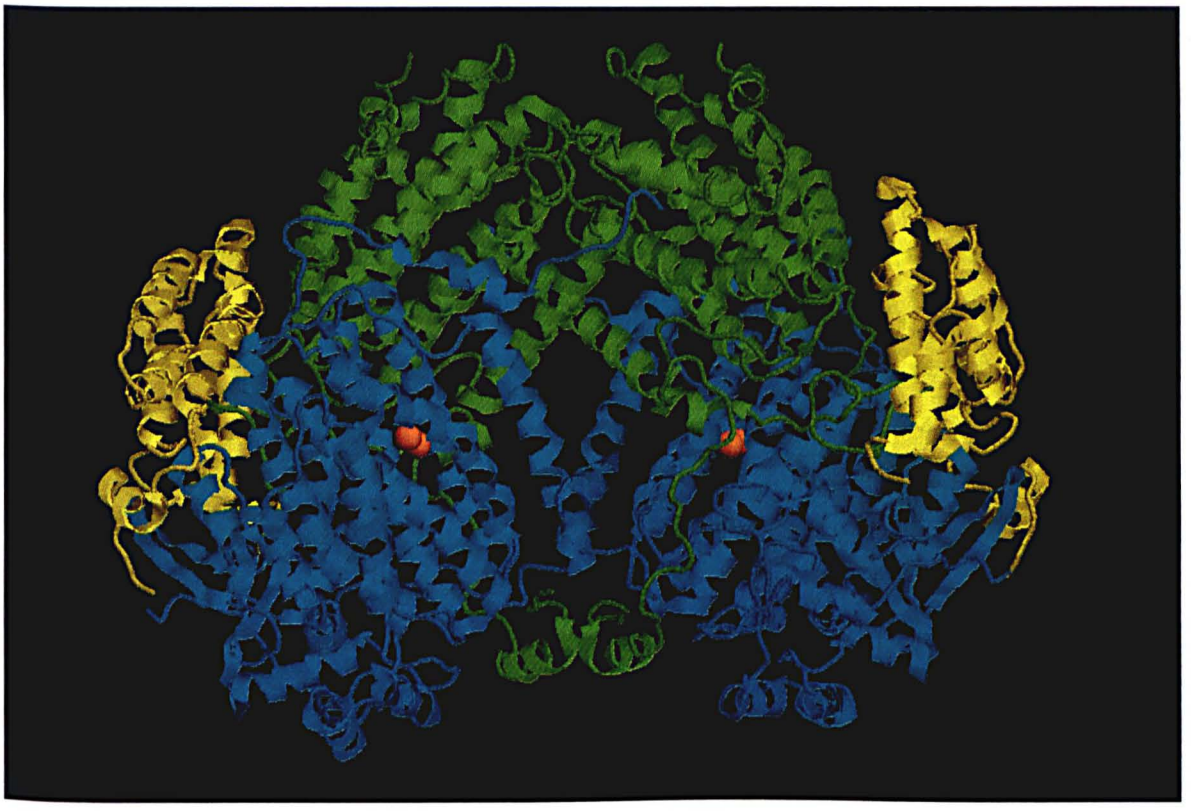


Figure 1.4 Model of the structure of the hydroxylase. The α-subunit is shown in blue, the β-subunit in green and the γ-subunit in yellow. The iron atoms are represented as orange spheres. Taken from Rosenzweig *et al.* (1993).

The active site of the hydroxylase, designated cavity 1, is located between the four long helices of the α-subunit and is lined with residues 97-127, 131-161, 197-226 and 231-257 (Elango *et al.*, 1997). Thr213 and Cys151 possess the only available protonated side chains in the active site and therefore it has been suggested that they play key roles in catalysis. In addition, the Cys151 occupies the position corresponding to the catalytically essential tyrosyl radical in the R2 protein of ribonucleotide reductase and therefore it was proposed that the Cys151-sulfhydryl-group may have a similar redox role in sMMO (Rosenzweig *et al.*, 1993); however, site-directed mutagenesis of this residue has shown this is unlikely (see Section 1.9).

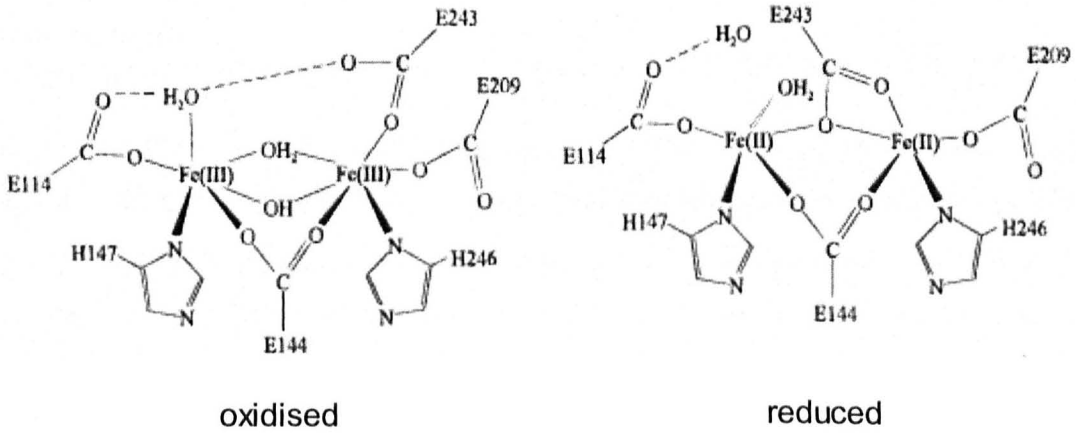


Figure 1.5 Structural representations of the sMMO hydroxylase diiron centre in the oxidised and reduced states. The carboxylate shift of the Glu243 (E243) can be clearly seen. The oxidised structure is based on the *Ms. trichosporium* sMMO hydroxylase crystal structure (Elango *et al.*, 1997) and the reduced structure is that of *Mc. capsulatus* (Bath) (Rosenzweig *et al.*, 1993). Adapted from Wallar & Lipscomb (1996).

In addition to cavity 1 which constitutes the active site, the hydroxylase also contains additional hydrophobic pockets (Rosenzweig *et al.*, 1997). Several of the cavities may be involved in channeling substrate molecules from the cytoplasm to the catalytic centre. The pocket nearest cavity 1, designated cavity 2, is completely hydrophobic and is lined with residues Val105, Val106, Phe109, Leu110, Met184, Phe188, Leu216, Phe282, Val285, Leu286 and Leu289. In the oxidised hydroxylase, the Leu110 and Phe188 side chains form a barrier between cavities 1 and 2. Upon reduction of the hydroxylase, the Leu110 rearranges such that cavities 1 and 2 are connected. Therefore, the Leu110 (and possibly Phe188) may act as a gate to regulate access of substrate and solvent into the active site. This effect is believed to be more pronounced in the presence of MmoB (Merkx *et al.*, 2001; Rosenzweig *et al.*, 1997), which would account for the some of the effects of MmoB on the sMMO oxidations (Section 1.5.1.2).

1.5.1.2 MmoB

MmoB is a sMMO effector protein and is a strong regulator of sMMO enzyme activity. It is comprised of a single protein (16 kDa) that contains no prosthetic groups (Green & Dalton, 1985). Not only does MmoB enhance the rate of electron transfer between the reductase and the hydroxylase (Brazeau & Lipscomb, 2003; Froland *et al.*, 1992; Wallar & Lipscomb, 2001), but it regulates the redox state of the hydroxylase (Kazlauskaite *et al.*, 1996; Liu & Lippard, 1991; Paulsen *et al.*, 1994), regioselectivity of oxidations (Fox *et al.*, 1991; Froland *et al.*, 1992), and effects a 1000-fold increase on the rate of oxygen binding to the reduced hydroxylase (Brazeau *et al.*, 2001b). MmoB binds to the hydroxylase which is believed to alter the geometry of the active site (Pulver *et al.*, 1997; Rosenzweig *et al.*, 1997; Wallar & Lipscomb, 2001). The steady-state hydroxylation rate of sMMO is optimal at a concentration ratio of two moles MmoB per mole of hydroxylase (Fox *et al.*, 1991; Liu *et al.*, 1995).

The structure of MmoB has been solved by NMR spectroscopy (Chang *et al.*, 1999; Walters *et al.*, 1999). The exact residues in MmoB and the hydroxylase that are involved in binding have not been determined, but it has been proposed to involve the N-terminus of MmoB and occurs in the canyon region between the $\alpha\beta\gamma$ dimers of the

hydroxylase (MacArthur *et al.*, 2002; Walters *et al.*, 1999). Positively charged residues on the surface of the hydroxylase have been shown to be essential for the interaction and electron transfer from the reductase (Balendra *et al.*, 2002). The MmoB from *Mc. capsulatus* (Bath) is susceptible to proteolysis which inactivates the component and therefore may constitute an important regulatory feature of the system (Lloyd *et al.*, 1997b).

1.5.1.3 MmoC

The three-dimensional structure of the reductase component of the sMMO (MmoC) has not been determined; however, the protein has been purified and extensively characterized. The MmoC is a single protein (38.5 kDa) and contains one [2FE-2S] cluster and an FAD cofactor which mediate electron transfer from NADH to the hydroxylase. The [2FE-2S] cluster is similar to that found in other reductases (Merkx *et al.*, 2001). The MmoC is present at 10 % of the molar concentration of the other sMMO components, perhaps to limit the production of reactive oxygen species (Fox *et al.*, 1991). Like MmoB, the MmoC interacts with the hydroxylase and there is evidence that MmoC may interact with the β-subunit (Fox *et al.*, 1991) and the N-terminal region of the α-subunit (Kopp & Lippard, 2002).

1.5.1.4 MmoD

The role of MmoD (also known as OrfY) is not known. The derived aminoacyl sequence of MmoD is the least conserved of the sMMO components and the only other non-haem diiron monooxygenase that appears to have a protein evolutionarily-related to MmoD is the butane monooxygenase of '*Pseudomonas butanovora*' (Sluis *et al.*, 2002). The MmoD from *Mc. capsulatus* (Bath) was studied by Western blot analysis. This showed that it was expressed at 1-2 % of the level of the sMMO hydroxylase proteins (Merkx & Lippard, 2002). In the same study, the MmoD was over-expressed in *E. coli* and added to purified sMMO *in vitro*. The MmoD protein appeared to bind the hydroxylase and inhibit propene oxidation. A role in assembly of the diiron centre was proposed, a function analogous to that of the

DmpK protein of phenol hydroxylase from *Pseudomonas* sp. CF600 (Powlowski *et al.*, 1997).

1.5.1.5 The catalytic cycle of sMMO

Several of the intermediates of the catalytic cycle of sMMO have been determined using a range of spectroscopic techniques (Figure 1.6); only a brief summary is given in this section but the subject has been reviewed extensively (Kopp & Lippard, 2002; Merkx *et al.*, 2001; Wallar & Lipscomb, 1996). The first step is the reduction of the diferric hydroxylase (H^{ox}) to the reduced state (H^{red}) by the MmoC and MmoB components. At this step the Glu243 side chain undergoes a carboxylate shift, as described in Section 1.5.1.1. There is evidence for an intermediate (compound P*, or 'superoxo') between the reduced hydroxylase and the formation of the peroxo intermediate (compound P) (Lee & Lipscomb, 1999; Merkx *et al.*, 2001). Compound Q, the form that reacts with the substrate, is perhaps the most powerful oxidant in biology. The precise mechanism of carbon activation is not known, but some insights have been obtained by analysing kinetic isotope effects (KIEs) and by the hydroxylation of probe substrates such as norcarane (Brazeau *et al.*, 2001a; Kopp & Lippard, 2002; Merkx *et al.*, 2001; Nesheim & Lipscomb, 1996; Newcomb *et al.*, 2002).

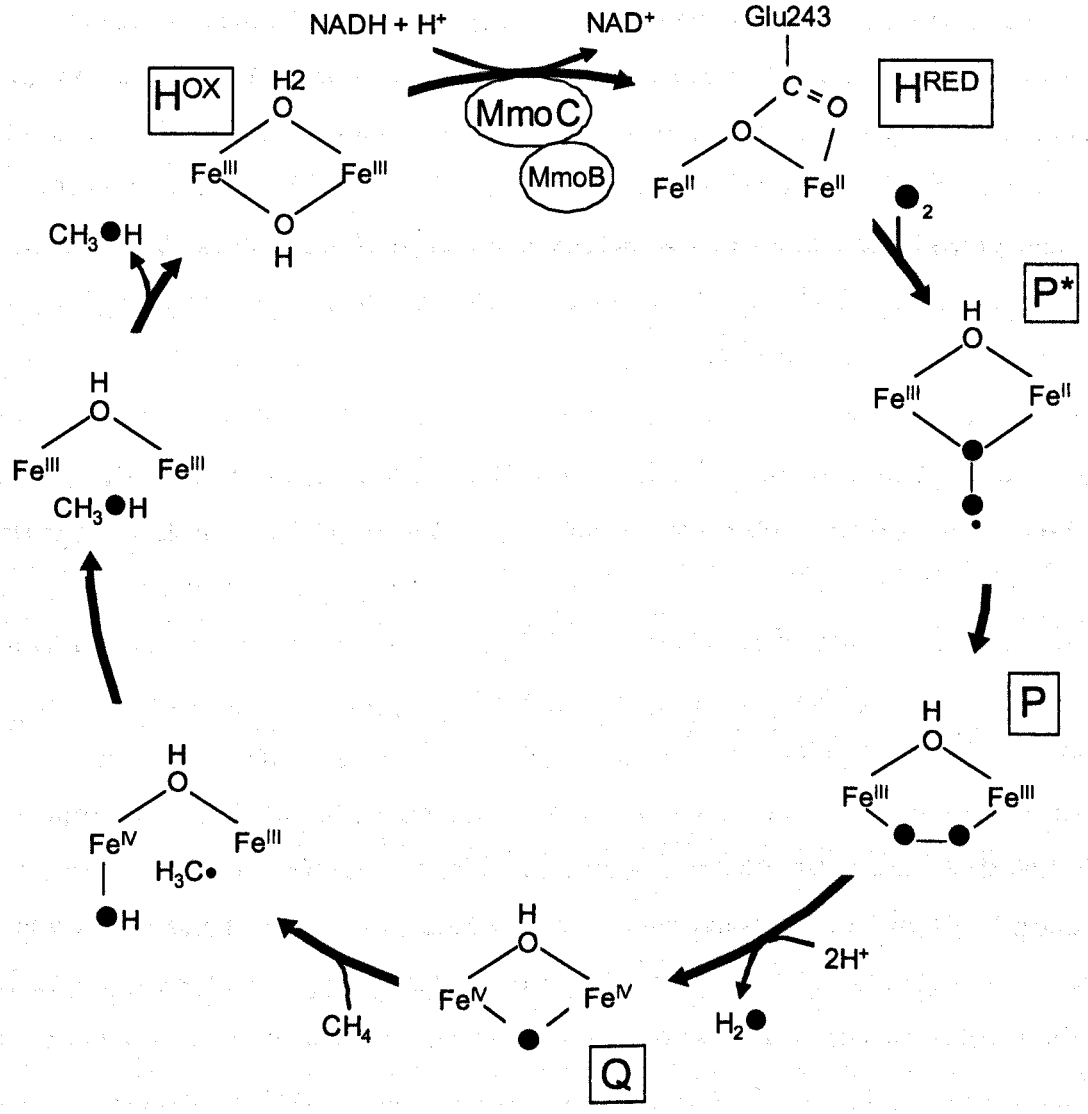


Figure 1.6 The proposed catalytic cycle of sMMO. Oxygen atoms derived from O_2 are filled. The intermediates are described in the text. Adapted from Elango *et al.* (1997); the configuration of the oxygen in P* is that shown in Merckx *et al.* 2001.

1.5.2 Molecular genetics of the soluble methane monooxygenase

The first sMMO structural genes to be cloned and sequenced were from *Mc. capsulatus* (Bath) (Stainthorpe *et al.*, 1989; Stainthorpe *et al.*, 1990), followed by those of *Ms. trichosporium* OB3b (Cardy *et al.*, 1991a; Cardy *et al.*, 1991b), *Methylocystis* sp. strain M (McDonald *et al.*, 1997) *Methylocystis* sp. strain WI14 (Grosse *et al.*, 1999), and *Methylomonas* species KSWIII and KSPIII (Shigematsu *et al.*, 1999). The sMMO genes from *Methylocella silvestris* $BL2^T$ were sequenced in this study (Chapter 9). The organisation of the genes in the order *mmoXYBZDC* is identical in all methanotrophs examined; *mmoX*, *mmoY* and *mmoZ* encode the α, β and γ-subunits of the hydroxylase (Section 1.5.1.1), *mmoB* encodes the effector protein (Section 1.5.1.2), *mmoC* the reductase (Section 1.5.1.3) and *mmoD* the enigmatic MmoD protein (Section 1.5.1.4) (Figure 1.7). The *mmoX* genes are the most conserved, sharing a minimum of 72 % nucleotide identity, and *mmoD* is the least conserved (Merkx & Lippard, 2002; Shigematsu *et al.*, 1999).

A consistent feature of the operons encoding non-haem diiron monooxygenases is that the reductase gene is present at the end of the cluster, which may indicate that the expression of the reductase in the absence of the hydroxylase is toxic to the cell (Leahy *et al.*, 2003). The expression of MmoC in *Mc. capsulatus* (Bath) appears to be approximately 10 % of the level of the hydroxylase proteins (Section 1.5.1.3), which is consistent with this theory. The *mmoC* genes in all the sMMO sequences have weak ribosomal binding sites which may account for the relatively poor expression.

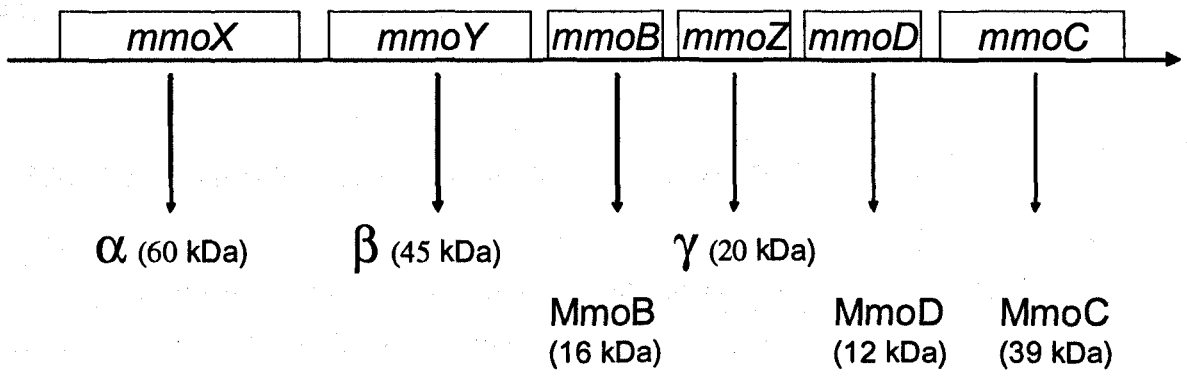


Figure 1.7 Organisation of the sMMO structural genes.

1.5.3 The copper switch

In methanotrophs that possess both sMMO and pMMO, the expression of the genes encoding these enzymes is regulated by the copper-to-biomass ratio, where sMMO is expressed at low copper-to-biomass ratios (less than 0.89 μmol copper per g dry weight cells (Hanson & Hanson, 1996) and pMMO is expressed at higher copper-to-biomass ratios (Dalton *et al.*, 1984; Stanley *et al.*, 1983). Northern blot and primer extension analysis of *pmoCAB* and *mmo* genes from *Mc. capsulatus* (Bath) demonstrated that this phenomenon originated at the level of gene transcription (Nielsen *et al.*, 1996; Nielsen *et al.*, 1997). The results of the gene analysis were consistent with the phenotype and showed that *mmo* gene transcription occurred in low copper medium and *pmo* transcription occurred in medium containing elevated levels of copper. The *mmo* transcripts disappeared within 15-30 minutes of $CuSO_4$ addition to low-copper growth medium.

Choi *et al.* (2003) used quantitative RT-PCR to measure the amount of *mmoX* and *pmoA* transcript present in cells growing in a fermentor culture (OD_{600} 1.8 – 2.0) with varying amounts of copper. With no added copper, the *pmoA* transcript level corresponded to 0.47 mole mRNA per μg total RNA. At 1 μM copper the transcript increased to 1.8 mole μg^{-1}, at 5 μM copper there was 2.9 mole μg^{-1}, at 25 μM copper

there was 22 mole μg^{-1} and finally at 55 μM copper there was 34 mole μg^{-1}. These results showed that the level of *pmo* gene transcription gradually increased with increases in copper concentration, but *mmo* transcription appeared to be an on-off switch.

1.5.3.1 Factors involved in *mmoXYBZDC* gene expression

The Northern blot and primer extension analysis performed by Nielsen and coworkers (1996 & 1997) showed that transcription of *mmoXYBZDC* occurs from a σ^{54}-type promoter 5' of *mmoX* in *Mc. capsulatus* (Bath)[1]. Transcription occurred from a similar σ^{54} promoter 5' of *mmoX* in *Ms. trichosporium* OB3b; however there was evidence of additional transcription from a second σ^{70}-type promoter between *mmoX* and *mmoY* (Figure 1.8). Stafford *et al.* (2003) provided conclusive evidence that the sMMO operon is transcribed from the σ^{54} promoter by disrupting the gene that encodes σ^{54} *(rpoN)* in *Ms. trichosporium* OB3b and showing that it abolished sMMO expression. The -24, -12 recognition motif (YTGGCACG-N_4-TTGCW) characteristic of σ^{54} promoters (Merrick, 1993) has been identified upstream of every sequenced *mmoXYBZDC* operon (see Chapter 9).

The σ^{54} is sometimes referred to as the alternative sigma factor and is responsible for transcription of relatively few genes in bacterial genomes, such as those involved in nitrogen fixation and flagellar biosynthesis (Reitzer & Schneider, 2001; Studholme & Buck, 2000; Studholme *et al.*, 2000). Unlike promoter-bound σ^{70}-RNA polymerase (RNAP) that can spontaneously form an open-complex and initiate transcription, σ^{54}-RNAP requires activation by an enhancer binding protein (EBP). An EBP binds its corresponding enhancer which can be positioned anywhere in the genome in either orientation, but is usually present 100-200 bp 5' of the σ^{54} promoter (Morett & Segovia, 1993). Activation of the EBP causes ATP hydrolysis and then interacts with the σ^{54}-RNAP to initiate transcription. Unlike σ^{70} promoters that often exhibit basal levels of transcription when repressed, transcription from σ^{54} controlled genes can be completely shut-off. For this reason, σ^{54} promoters may be important for controlling the expression of genes that encode proteins that are only

[1] The original sequence of the promoter region from *Mc. capsulatus* (Bath) contained sequencing errors in the -24 domain (Csáki *et al.*, 2003) which was confusing.

required under certain conditions and that have energy expensive activities, such as nitrogenase.

The gene encoding the *mmo* σ^{54} EBP, designated *mmoR*, has been characterised in *Ms. trichosporium* OB3b (Stafford *et al.*, 2003) and *Mc. capsulatus* (Bath) (Csáki *et al.*, 2003) and its position relative to the *mmoXYBZDC* operon in other methanotrophs has also been determined (Figure 1.8). The gene was disrupted by marker-exchange mutagenesis in *Ms. trichosporium* OB3b (Stafford *et al.*, 2003) and *Mc. capsulatus* (Bath) (Csáki *et al.*, 2003) which resulted in the absence of sMMO expression in both methanotrophs. Since the promoter has been shown to be regulated by copper, it is likely that copper directly or indirectly affects MmoR. The MmoR does not possess an identifiable regulatory domain, such as a GAF or PAS[2] domain found in many EBPs (Studholme & Dixon, 2003), and there is no obvious copper binding motif in MmoR. In addition, the *Methylocella silvestris* BL2^T contains *mmoR* (see Chapter 9), but its sMMO is not regulated by copper (Andreas Theisen, unpublished results) suggesting that MmoR may not have evolved to bind copper directly but instead responds to a separate sensor protein. An ORF positioned 3' of *mmoR*, designated *mmoR2*, is present in *Ms. trichosporium* OB3b, *Ms. sporium* strain 5 (Hanif Ali, unpublished data) and *Methylocella silvestris* BL2^T (Chapter 9). *mmoR2* has a possible ribosome binding site and the codon usage is similar to that of other methanotroph genes, which suggests that it may be a *bona fide* gene. The derived peptide sequence has very low similarity to a kinase, and therefore it may be part of a sensory pathway that activates MmoR by phosphorylation.

The *mmoG* gene, which encodes a GroEL-like chaperonin, is also found in the vicinity of *mmoXYBZDC* operons and has been identified both 3' and 5' of the operon in various methanotrophs (Figure 1.8). Originally it was thought that MmoG could be involved in the assembly of the sMMO hydroxylase, and although this has not been ruled-out, Δ*mmoG* mutants do not transcribe *mmoX* and therefore MmoG probably has a role in the assembly of the transcription machinery (Csáki *et al.*, 2003; Stafford

---

[2]GAF is an acronym for cGMP-specific and -stimulated phosphodiesterases, *Anabaena* adenylate cyclases and *Escherichia coli* FhlA (**Aravind, L. & Ponting, C. P. (1997).** The GAF domain: an evolutionary link between diverse phototransducing proteins. *Trends Biochem Sci* **22**, 458-459.); PAS is an acronym for Per, ARNT and Sim (**Ponting, C. P. & Aravind, L. (1997).** PAS: a multifunctional domain family comes to light. *Curr Biol* **7**, 674-677.).

et al., 2003). The genes encoding a two-component sensor-regulator, designated *mmoS* and *mmoQ*, were identified 3' of the *mmoXYBZDC* operon in *Mc. capsulatus* (Bath). The genes were disrupted by marker-exchange mutagenesis and this had little effect on sMMO expression (Csáki *et al.*, 2003). In light of this and the fact that similar genes have not been found in the vicinity of the *mmoXYBZDC* operon in other methanotrophs, MmoS and MmoQ may not be involved in *mmo* gene transcription. A proposed model for the expression of *mmoXYBZDC* is shown in Figure 1.9.

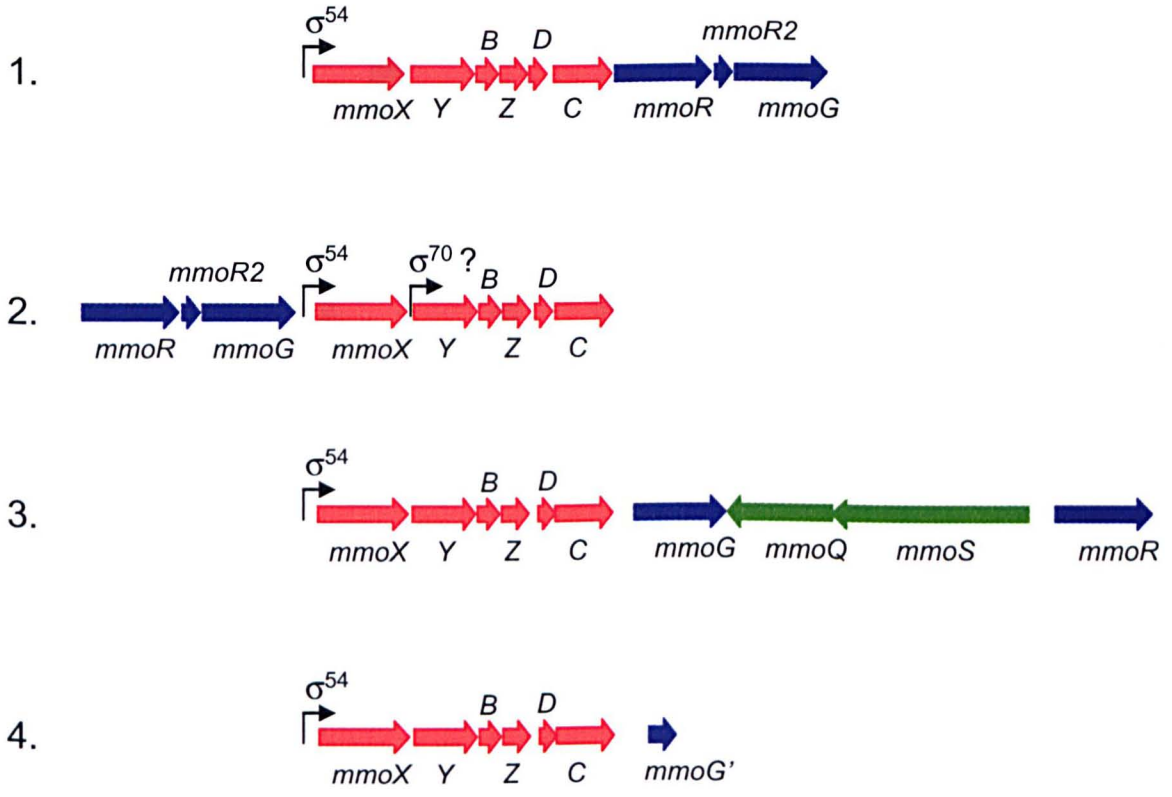


Figure 1.8 The organisation of *mmo* structural (red) and transcription-related (blue) genes from various methanotrophs: 1. *Methylocella silvestris* BL2^T, 2. *Ms. trichosporium* OB3b, 3. *Methylococcus capsulatus* (Bath), 4. *Methylomonas* sp. strain KSWIII. The genes encoding the possible two-component regulatory genes are indicated in green. The *Methylocystis* sp. strain M and WI14 gene clusters are not shown since the regulatory and accessory gene sequences are not available. The genes from *Methylocystis sporium* have been cloned and partially sequenced (Hanif Ali, personal communication) and the positions of *mmoR*, *mmoR2* and *mmoG* are identical to that of *Ms. trichosporium* OB3b.

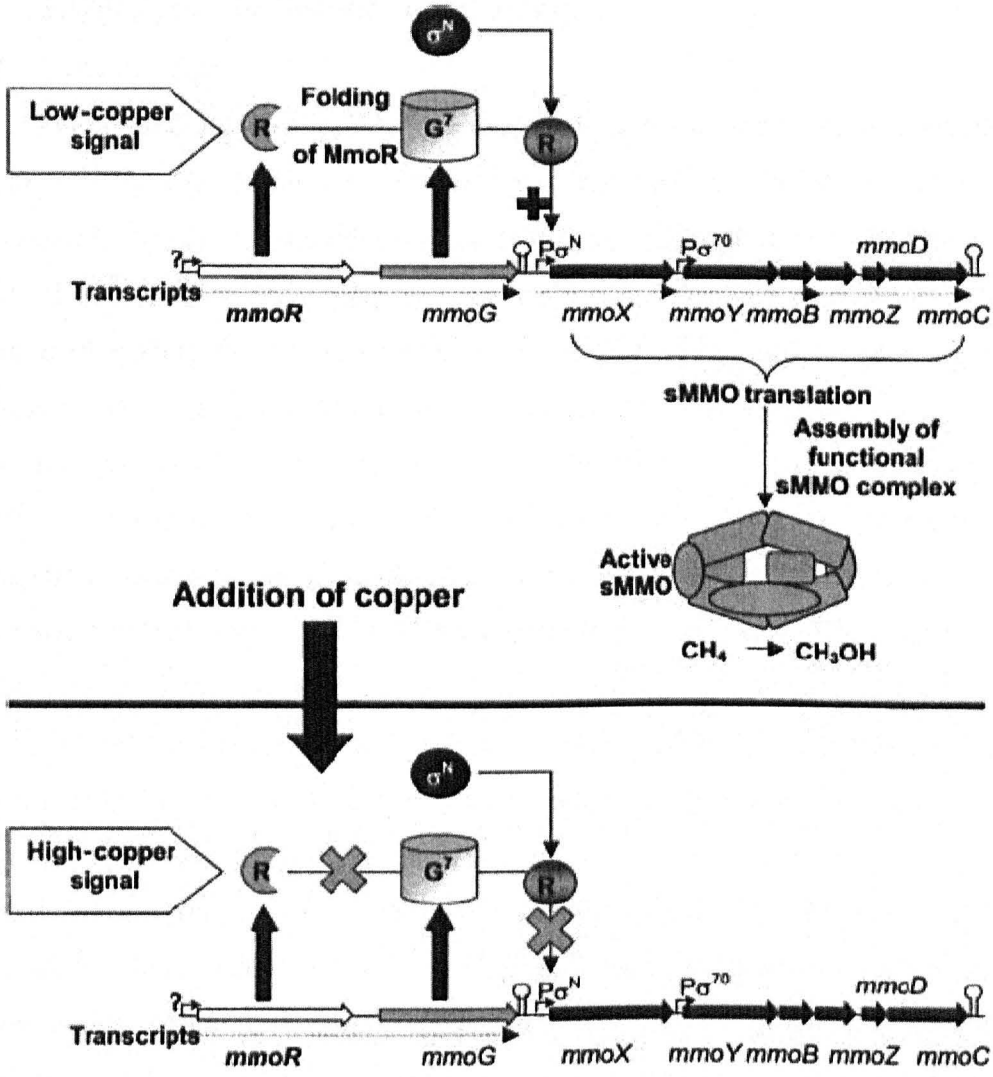


Figure 1.9 A hypothetical model for the regulation of the *mmoXYBZDC* operon under high and low copper conditions in *Ms. trichosporium* OB3b. Both MmoG (G^7) and MmoR (R) are necessary for activating the transcription of the *mmo* operon. MmoG may directly or indirectly activate MmoR, which then initiates transcription via σ^N (σ^{54}). In the presence of high levels of copper, the transcriptional activation of sMMO that is mediated by MmoR and σ^{54} is abolished and the transcription of *pmoCAB* is up-regulated by an unknown mechanism. Taken from Stafford *et al.* (2003).

1.6 Ecology of methanotrophic bacteria

1.6.1 Introduction

Methane is an important greenhouse gas and the atmospheric concentration has been rising at approximately 0.8-1 % per annum (Cicerone & Oremland, 1988; Steele *et al.*, 1992). Although methanotrophs convert methane to carbon dioxide, which is also a greenhouse gas, methane has 26 times the greenhouse warming potential of carbon dioxide (Lelieveld *et al.*, 1993). The total sources of methane contribute 525 Tg of methane to the atmosphere per annum (IPCC, 1995). Methanotroph activity, primarily in wetland environments where methanogenesis is substantial, mitigates the release of an additional 700 Tg y^{-1} of methane to the atmosphere (Reeburgh *et al.*, 1993). Chemical oxidation of methane by hydroxyl radicals in the troposphere accounts for approximately 90 % (420-520 Tg y^{-1}) of the atmospheric methane sink. In addition, upland soils are estimated to consume 15-45 Tg y^{-1} of atmospheric methane (IPCC, 1995). Therefore methanotrophs are responsible for an important component of the carbon cycle and their activity has an impact on the Earth's atmospheric chemistry.

Methanogenesis occurs in anaerobic environments (Garcia *et al.*, 2000) such as wetlands, lakes, landfills and rice paddies, and these are referred to as high methane environments. On a global scale, most of the methane oxidation by methanotrophs occurs in these high methane environments and the activity is generally located at the oxic-anoxic interface as the methane diffuses from the deep anoxic layer (Hanson & Hanson, 1996). The methanotrophs in culture collections and those with similar physiology are believed to be responsible for this methane oxidation. Atmospheric methane oxidation occurs in upland soils, in which the organic layer is aerobic and methanogenesis rarely, if ever, occurs. This activity is different from the methane oxidation that occurs in high methane environments and the organisms responsible for atmospheric methane oxidation have not been identified (Hanson & Hanson, 1996). The methane half-saturation constants of upland soils are typically below 100 nM (Bender & Conrad, 1993) whereas the values for cultivated methanotrophs are in the order 1 μM (Bédard & Knowles, 1989).

1.6.2 Methods for detecting methanotrophs

The diversity of methanotrophs, like all microorganisms, was traditionally studied by laboratory-based culturing. It is estimated that < 5 % of soil microorganisms have been cultivated and that laboratory strains only account for a small fraction of the phylogenetic and physiological diversity of bacteria (Amann *et al.*, 1995; Felske *et al.*, 1999; Hugenholtz *et al.*, 1998). Many culture independent methods for studying microbial diversity have been developed over the last fifteen years (Gray & Head, 2001) and several have been applied to the detection of methanotrophs (Murrell *et al.*, 1998; Murrell & Radajewski, 2000).

1.6.2.1 Laboratory enrichments

Methanotrophs have been isolated from many habitats and the culture collection now contains organisms demonstrating a wide-range of physiological adaptations, including moderately halophilic, alkaliphic, thermophilic, psychrophilic and acidophilic representatives (Table 1.1). Successful isolation of many of the "extremophilic" methanotrophs has only been achieved in the last few years. This is mostly the result of an increased effort to isolate methanotrophs from extreme environments, but modifications to the conventional culturing methods, such as the discovery that the acidophilic *Methylocella* species do not tolerate the mineral salts concentrations traditionally used in methanotroph growth media (Dedysh *et al.*, 2000), have contributed to the isolation of new genera.

Culturing methods favour the isolation of the fastest grower. Conditions known to select the growth of various methanotrophs include temperature, pH, salt concentration, nitrogen source (N_2 or fixed) and the use of "copper-free" medium (Hanson & Hanson, 1996). In addition, methane and oxygen concentrations appear to influence the competition between α- and γ-*Proteobacteria* methanotrophs. Competition experiments between *Methylomicrobium album* BG8 and *Methylosinus trichosporium* OB3b in a chemostat showed that the γ-*Proteobacteria* methanotroph (*Methylomicrobium album* BG8) dominated under a low methane-oxygen mixing ratio (Graham *et al.*, 1993). This phenomenon was also observed in agarose diffusion columns used to enrich for methanotrophs by inoculating the agarose with a soil

sample and providing the column with methane at the base and oxygen at the top (Amaral & Knowles, 1995). A band of methanotroph growth appeared within the column after several days of incubation and always consisted of γ-*Proteobacteria* methanotrophs. A small amount of oxygen leakage at the bottom of the column occasionally resulted in the development of a second band of growth, which always consisted of α-*Proteobacteria* representatives. The authors suggested that the apparent favoured growth of γ-*Proteobacteria* methanotrophs near the top of the column was a result of the low methane-oxygen ratio in this part of the gradient and the favoured growth of α-*Proteobacteria* methanotrophs at the base of the column was the result of the high methane-oxygen ratio. There is evidence that this phenomenon also occurs in natural environments: an experiment with rice paddy soil found that most methane uptake was performed by α-*Proteobacteria* methanotrophs when the soil was incubated with high concentrations of methane (10,000 ppmv) and γ-*Proteobacteria* methanotrophs when incubated with low concentrations of methane (1000 ppmv) (Henckel *et al.*, 2000b).

1.6.2.2 Phylogenetic probes

PCR primers designed to anneal specifically to the 16S rRNA gene of methanotrophs have been designed and used to detect methanotrophs in environmental samples (Murrell *et al.*, 1998). The phylogeny of methanotrophs is diverse and therefore the design of primers that are specific to groups of methanotrophs, such as all α-*Proteobacteria* methanotrophs or all γ-*Proteobacteria* methanotrophs, is not possible. To properly target methanotroph 16S rRNA genes, it is now necessary to design primers for more narrow groups, such as methanotroph families and individual genera (Gulledge *et al.*, 2001). It is still necessary to modify existing primer sequences or design new primers after new methanotroph isolates are obtained and their 16S rRNA gene sequences are deposited in sequence databases.

Fluorescence *in situ* hybridization (FISH) is a powerful technique to visualize target populations in their natural environment using fluorescence-labelled 16S rRNA oligonucleotide probes (Amann *et al.*, 1995). FISH has been used to visualize and quantify groups of methanotrophs in enrichment cultures (Bourne *et al.*, 2000; Holmes *et al.*, 1996; Tsien *et al.*, 1990) and to visualize and count acidophilic

methanotrophs in a peat sample (Dedysh *et al.*, 2001). The peat soil study indicated that methanotrophs were present at 3 x 10^6 cells g^{-1} of peat and consisted almost entirely of α–*Proteobacteria* representatives.

1.6.2.3 Functional gene probes

The genes that encode key enzymes in methane metabolism are referred to as methanotroph functional genes. The methanotroph functional genes that have been targeted in environmental samples by PCR include *pmoA* (encoding the β-subunit of pMMO), *mmoX* (encoding the α-subunit of sMMO) and *mxaF* (encoding the α-subunit of MDH) (Auman *et al.*, 2000; Hutchens *et al.*, 2004; McDonald *et al.*, 1995; McDonald & Murrell, 1997a; McDonald & Murrell, 1997b; Murrell *et al.*, 1998). Their advantage over 16S rRNA probes is that they correspond to the metabolic function of interest. Also, primers can often be developed that target phylogenetically diverse groups of organisms. Often the sequences of functional genes follow evolutionary patterns that allow assumptions to be made regarding the phylogeny of the gene, and this can tentatively be extrapolated to infer the phylogeny of the organism from which it originated (McDonald & Murrell, 1997b). The *mxaF* gene has been shown to be less reliable than *pmoA* as a phylogenetic marker (Heyer *et al.*, 2002).

There are some considerations that need to be made when constructing methanotroph functional gene libraries. First, the sMMO is only present in some representatives and therefore analysis of *mmoX* excludes some methanotrophs. The only cultivated methanotrophs for which a *pmoA* gene can not be amplified are the *Methylocella* spp. But used together, the *pmoA* primers (Costello & Lidstrom, 1999; Holmes *et al.*, 1995) and the *mmoX* primers (Auman *et al.*, 2000; Hutchens *et al.*, 2004) should detect all known methanotrophs. All known methanotrophs possess *mxaF*. The available *mxaF* PCR primers amplify the gene from both methanotrophs and non-methane-utilising Gram-negative methylotrophs (McDonald & Murrell, 1997a).

There are some particular caveats to consider when amplifying and analysing the *pmoA* gene. *pmoA* is frequently amplified from environmental samples using the PCR primers A189f and A682r which also target the *amoA* gene from the ammonia

monooxygenase (Holmes *et al.*, 1995). Researchers should be aware that the primers were designed when the only sequences available were the *pmoA* of *Mc. capsulatus* (Bath) and the *amoA* of *Nitrosomonas europaea*. Numerous *amoCAB* and several *pmoCAB* operons have since been sequenced and demonstrate extensive mismatching with the A189f/A682r primers. Although the primers function well from individual strains despite the mismatches, they have been shown to be biased with mixed template (Bourne *et al.*, 2001).

The diversity of *amoA/pmoA* genes recovered from environmental samples using the A189f/A682r primers can be very large (Henckel *et al.*, 2000a; Holmes *et al.*, 1999; Morris *et al.*, 2002). Many of the sequences demonstrate no greater similarity to either *pmoA* or *amoA*, which makes it impossible to assign an enzymatic function (if any) to the protein it encodes. Holmes *et al.* (1999) aligned several of the PmoA and AmoA sequences present in the database and looked for residues conserved for each enzyme, and these were designated AMO or pMMO signature residues. Unfortunately many of the proposed pMMO-signature residues are found in the AmoA sequence of the γ-*Proteobacteria* autotrophic nitrifier, *Nitrosococcus oceanus*, suggesting that these residues still reflect phylogeny and not function. An unusual *pmoA* sequence detected in environmental studies has recently been shown to be present in members of the *Methylocystaceae* family (Dunfield *et al.*, 2002; Tchawa Yimga *et al.*, 2003) and appears to a *bona fide* pMMO gene (Ricke *et al.*, 2004).

Denaturing gradient gel electrophoresis (DGGE) is a technique for separating PCR sequences based on their nucleotide base composition (Muyzer *et al.*, 1993) and protocols for the separation of methanotroph 16S rRNA, *pmoA* and *mxaF* genes by DGGE have been developed and used to detect methanotrophs in environmental samples (Fjellbirkeland *et al.*, 2001; Henckel *et al.*, 1999; Jensen *et al.*, 1998). The advantages and limitation of the DGGE technique have been reviewed (Muyzer & Smalla, 1998).

1.6.2.4 Phospholipid fatty acid (PLFA) profiling

Bacterial PLFAs often contain fingerprint molecules that can be used to identify or detect specific organisms in microbial communities (Roslev *et al.*, 1998). PLFA analysis has been used to discriminate between α– and γ–*Proteobacteria*

methanotrophs since the former contain mainly 18 carbon (18C) PLFAs and the latter 14 carbon (14C) and 16 carbon (16C) PLFAs (see Section 1.2.2) (Bowman *et al.*, 1991; Guckert *et al.*, 1991). Incubating an environmental sample with either $^{13}CH_4$ or $^{14}CH_4$ and analysing the isotopically-labelled PLFAs is a powerful approach for analysis of natural populations of methanotrophs. This approach has been used to detect methanotrophs in forest soil (Bull *et al.*, 2000; Holmes *et al.*, 1999; Roslev & Iversen, 1999), lake sediment (Costello *et al.*, 2002) and rice field soil (Henckel *et al.*, 2000b).

PLFA profiling has been used to detect the organisms that incorporate atmospheric concentrations of isotope-enriched methane in upland soil samples (Bull *et al.*, 2000; Holmes *et al.*, 1999; Roslev & Iversen, 1999). All of the studies indicated that organisms with PLFA molecules similar, but not identical, to those of α-*Proteobacteria* methanotrophs incorporated the methane. Holmes and co-workers (1999) combined the PLFA analysis with the construction of *pmoA* gene libraries from forest soils. Interestingly, the library was dominated by a cluster of unusual *pmoA* sequences (designated RA14) that lie on the same evolutionary branch as those of α-*Proteobacteria* methanotrophs. RA14-*pmoA* genes have never been detected in a cultivated methanotroph and are only found in soils that exhibit uptake of atmospheric concentrations of methane.

1.6.2.5 DNA-based stable isotope probing (DNA-SIP)

It is difficult to link genes from uncultivated microorganisms with a specific metabolic function. Stable isotope probing (SIP) is a method of labelling an active population in an environmental sample, thus enabling uncultivated microorganisms to be linked with a metabolic process (Radajewski & Murrell, 2002; Radajewski *et al.*, 2003). DNA-SIP is performed by incubating a sample (ex. soil or sediment) with a ^{13}C-substrate (> 99 % ^{13}C). Most of the natural carbon is ^{12}C and the natural abundance of ^{13}C is about 1 %. Therefore the organisms that grow using the ^{13}C-substrate incorporate the label and become enriched with the heavy carbon isotope. DNA is isolated directly from the sample and the ^{13}C-DNA of the active population can be separated from the community ^{12}C-DNA (ie. the DNA from all the other

organisms that do not incorporate the ^{13}C-substrate into cell carbon) by CsCl-ethidium bromide density gradient centrifugation.

The DNA-SIP technique has been mostly used with the one-carbon growth substrates $^{13}CH_3OH$ and $^{13}CH_4$ (Hutchens *et al.*, 2004; Lin *et al.*, 2004; Lueders *et al.*, 2004b; Morris *et al.*, 2002; Radajewski *et al.*, 2000; Radajewski *et al.*, 2002). In these studies the ^{13}C-DNA was used as a template in PCR targeting the 16S rRNA, *pmoA, mxaF* and/or *mmoX* genes. As expected, the sequences of many of the cloned functional genes were similar to those from cultivated methylotrophs. Unexpected findings include a relatively large proportion of *Acidobacterium* 16S rRNA gene clones from DNA-SIP experiments with $^{13}CH_3OH$ (Radajewski *et al.*, 2000; Radajewski *et al.*, 2002), and a relatively large proportion of β-*Proteobacteria* gene clones from DNA-SIP experiments with $^{13}CH_4$ (Hutchens *et al.*, 2004; Lin *et al.*, 2004). The Hutchens *et al.* (2004) manuscript is present in the appendix and is a detailed example of how a DNA-SIP experiment is performed.

The stable-isotope probing technique has also been developed for the isolation of "heavy" RNA (RNA-SIP). This method was first demonstrated using an industrial bioreactor system supplied with $^{13}C_6$-phenol (Manefield *et al.*, 2002a). The RNA was separated and fractionated using a caesium trifluoro-acetate gradient and the ^{13}C-RNA converted to cDNA by RT-PCR. This experiment yielded the surprising discovery that phenol was being degraded by a *Thauera* species, an organism not identified as a phenol-degrader by conventional methods. RNA-SIP is more sensitive than DNA-SIP since the turn-over of cellular RNA is faster than that of DNA and the isotope incorporation does not require cell division. The incubation time required to isolate ^{13}C-RNA from the highly enriched bioreactor community was only several hours (Manefield *et al.*, 2002b), whereas DNA-SIP incubations with ^{13}C-substrate traditionally take days or weeks.

1.6.3 sMMO diversity and significance

Relatively little is known about the diversity of sMMO enzymes in the environment. Most molecular ecology studies of methanotrophs have only generated *pmoA* libraries by PCR and have ignored *mmoX*. The bias against the use of sMMO genes as functional gene markers probably stems from the fact that it is only found in some methanotrophs. Also, the lack of *mmo* gene sequences in the database made it difficult to design PCR primers for *mmoX*, and therefore more reliable PCR primers targeting *mmoX* have only recently become available (Auman *et al.*, 2000). It is now clear that sMMO is present in many more strains of γ–*Proteobacteria* methanotrophs than had been previously thought.

It is still not known if sMMO has a significant role in methane oxidation in the natural environment. It is believed that the concentrations of copper in most environments are sufficient to repress expression of sMMO in organisms similar to *Ms. trichosporium* OB3b and *Methylococcus capsulatus* (Bath) (Hanson & Hanson, 1996). However, the bioavailability of the copper may be low, particularly in micro-sites and biofilms. Also, copper levels are relatively low in oligotrophic environments such as peat, and therefore there may be significant sMMO expression in these systems (Murrell *et al.*, 2000). Methanotrophs such as *Methylocella* spp. which do not have pMMO and possess an sMMO that does not appear to be repressed by copper, may also contribute substantially to methane oxidation in some environments. It may be possible to screen for sMMO expression by analysing *mmoX* gene expression in environmental samples (Han & Semrau, 2004) and thereby estimate the significance of sMMO.

1.7 Commercial applications

The oxidation of hydrocarbons by sMMO or pMMO to produce commercially valuable compounds is one of the most attractive prospects for commercialisation and has been reviewed recently by Smith and Dalton (2004). The sMMO in particular can oxidise a wide range of hydrocarbons ranging from alkanes to substituded polyphenyls (Burrows *et al.*, 1984; Colby *et al.*, 1977; Smith & Dalton, 2004, in press). The utility of sMMO was investigated for the production of epoxypropane from propene, a reaction which is difficult to achieve chemically but which is readily performed by sMMO expressing methanotrophs (Richards *et al.*, 1994). Richards *et al.* (1994) used *Mc. capsulatus* (Bath) in a two-stage bioreactor to produce the epoxide. The bioconversion was achieved in the first stage; however, epoxypropane was cytotoxic and the cells needed to be transferred to the second stage where the cells were reactivated in the absence of propene and epoxypropane. The system generated 20-30 g epoxide per day over a 3 week period, which was not commercially competitive (Smith & Dalton, 2004, in press).

There is considerable value in the synthesis of enantiopure chemicals for the pharmaceutical industry. Unfortunately neither sMMO nor pMMO oxidations are exceedingly stereospecific. However, it may be possible to increase the stereospecificity of sMMO oxidations by mutagenesis of the active site (Smith *et al.*, 2002). There is currently no expression system that would enable the mutagenesis of pMMO.

The University of Bergen and The Institute for Genome Research are currently sequencing the genome of *Mc. capsulatus* (Bath). The genome sequence will eventually lead to an improved understanding of the biochemistry of the organism that may enable the engineering of novel metabolic pathways and the over-production of various compounds, for example methanol, chiral alcohols or poly-hydroxybutyrate, that could be marketable. Because methane is a cheap growth substrate and therefore methanotrophs are less expensive to culture than most bacteria, methanotrophs may be valuable as hosts for the over-expression of commercially relevant recombinant proteins.

The ability of methanotrophs to co-metabolise many halogenated hydrocarbons has made them attractive for use in bioremediation. The degradation of trichloroethene (TCE) by methanotrophs has been extensively studied and is much

more effective by sMMO-expressing cells than pMMO-expressing cells (Alvarez Cohen *et al.*, 1992; Fox *et al.*, 1990; Tsien *et al.*, 1989). The increased removal of TCE from a contaminated river site by stimulating methanotroph growth has been demonstrated in a field experiment (Pfiffner *et al.*, 1997). Because the level of sMMO expression in environmental samples may be low, the bioremediation of contaminants metabolised by sMMO and not pMMO could be ineffective *in situ.* Because the sMMO of *Methylocella* sp. is not repressed by copper, the potential to exploit this genus for bioremediation should be investigated.

1.8 Recombinant expression of the sMMO

There is significant interest in expressing sMMO in a non-methane-using organism for potential use in biotechnological and bioremediation processes. There are several reasons: 1. Genetic engineering of a methanotroph to improve its function in a bioreactor is not feasible since they are difficult to manipulate using molecular biology techniques (Murrell, 1994; Murrell *et al.*, 2000). 2. Since all methanotrophs are obligate methylotrophs and often grow most effectively with methane, the target substrate must compete for the active site of the sMMO. The expression of sMMO in a non-methanotroph would allow growth on an alternate carbon source such that the sMMO would be fully available for the bioconversion process. 3. Methanotrophs are unable to grow using complex carbon and may not have optimal uptake systems for many organic compounds targeted for bioconversion. 4. Depending on the system, it may be difficult or impossible to reduce the copper concentration to levels where sMMO is fully expressed. Once in a heterologous host, the sMMO could be placed under the control of an alternative regulatory system such that it is transcribed when required.

The recombinant expression of active sMMO has proved difficult (Murrell *et al.*, 2000). Active MmoB (Brandstetter *et al.*, 1999; West *et al.*, 1992), MmoC (Blazyk & Lippard, 2002; West *et al.*, 1992) and MmoD (Merkx & Lippard, 2002) can be achieved in *E. coli*. The expression of a fully active sMMO does not occur in *E. coli, Agrobacterium tumefaciens*, nor *Sinorhizobium meliloti* (Murrell *et al.*, 2000). Native-PAGE and Western blot analysis indicate that the hydroxylase may not be assembled (Lloyd, 1997). The expression in *Pseudomonas putida* F1 did show some promise. Recombinant sMMO expression under the control of a foreign promoter resulted in increased degradation of chloroform by the host organism (Jahng & Wood, 1994). Oddly, the system did not oxidise propene which is the standard assay for sMMO activity. Lloyd (1997) repeated the experiments and was also unable to detect any specific sMMO activity in recombinant *Pseudomonas putida*. Native-PAGE and Western blot analysis indicated that, unlike *E. coli,* intact hydroxylase was assembled in *Pseudomonas putida*. Lloyd (1997) was unable to purify the enzyme from the expression system, suggesting that the enzyme was synthesized only at very low levels.

Low level expression of recombinant sMMO from *Ms. trichosporium* OB3b was achieved in *Methylomicrobium album* BG8 and *Methylocystis parvus* OBBP, which are two methanotrophs that only possess pMMO (Lloyd *et al.*, 1999a). The expression was achieved on a broad-host-range plasmid containing the *mmoRGXYBZDC* genes and the native *mmoX* σ^{54} promoter. The expression appeared to be copper sensitive which suggests that system possessed the factors required for repression of transcription under elevated copper. Another remarkable aspect is that this *mmoRGXYBZDC* fragment does not complement Δ*mmoR* or Δ*mmoG* mutants of *Ms. trichosporium* OB3b (Julie Scanlan, personal communication), presumably since the *mmoR* and *mmoG* promoter is absent from the expression plasmid. Either *Methylomicrobium album* BG8 and *Methylocystis parvus* OBBP contain *mmoR* and *mmoG* homologues or there was some expression of the *Ms. trichosporium* OB3b *mmoRG* genes from the expression plasmid used in the study.

The highest levels of recombinant sMMO expression occurs in an *Ms. trichosporium* OB3b sMMO$^-$ mutant (Lloyd *et al.*, 1999b). The mutant, designated *Ms. trichosporium* Mutant F (or simply "Mutant F"), contains a kanamycin gene cassette which has been inserted into the *mmoX* gene on the chromosome of *Ms. trichosporium* OB3b (Martin & Murrell, 1995). The conjugation of *mmoRGXYBZDC* genes contained on a broad-host-range plasmid into *Ms. trichosporium* Mutant F resulted in excellent expression of sMMO at levels at least as high as in the wild-type organism. Although the *mmoX* gene is disrupted in Mutant F, Western blot analysis indicated that the β– and γ–subunits of the sMMO hydroxylase are still expressed (Lloyd *et al.*, 1999b). This system has been shown to be suitable for the expression of site-directed mutants of the α-subunit (see Section 1.9).

1.9 sMMO mutagenesis

Mutagenesis of the MmoB and MmoC subunits is relatively straightforward since these proteins can be over-expressed and purifed from *E. coli.* Mutagenized MmoB obtained from an *E. coli* expression host is studied *in vitro* with native sMMO hydroxylase purified from *Mc. capsulatus* (Bath) or *Ms. trichosporium* OB3b. Considerable mutagenesis studies have been performed with the MmoB and has demonstrated remarkable effects on sMMO catalysis (see Section 1.5.1.2).

The development of the homologous *Ms. trichosporium* Mutant F system has permitted the generation of site-directed mutants of the α-subunit (Smith *et al.*, 2002). Smith and co-workers described the development of a system that allowed the mutagenesis of residues over a significant portion of the active site. The mutants were made using a four-primer overlap extension PCR method. The PCR product was ligated in to pUC18 and then reintroduced into the *mmoRGXYBZDC* fragment in pTJS176 (pMTL24 derivative) using *Bam*HI and *Nde*I. The insert from pTJS176 containing the mutagenized *mmoX* gene was transferred into the broad-host-range (RK2-derivative) and mobilizable plasmid pTJS140. The resulting plasmid was conjugated into *Ms. trichosporium* Mutant F via the *E. coli* S17-1 donor strain and transconjugants selected on agar plates containing streptomycin and spectinomycin.

Four α-subunit sMMO mutants were created and expressed in Mutant F; the residues targeted were the cysteine at position 151 and the threonine at 213 of the α-subunit. The changes made were cysteine to glutamate (C151E), cysteine to tyrosine (C151Y), threonine to serine (T213S) and threonine to alanine (T213A). The mutant constructs were introduced into Mutant F and analysed by the sMMO-specific naphthalene assay and by SDS-PAGE (Table 1.2). Only the C151S was stable enough for sMMO polypeptides to be visualized by SDS-PAGE and for the enzyme to be purified. The purified T213S displayed an altered activity towards methane, propane and toluene. In particular, the conversion of propene to propylene oxide was dramatically reduced in the T213S mutant compared to the wild-type sMMO.

The reason for poor expression of C151S, C151Y and T213A is unclear. These residues may be important for the stability of the enzyme which resulted in a decrease in the stability of the sMMO enzyme and polypeptides or decreased transcription of the *mmo* genes. Conceivably the mRNA transcript may have become unstable. The loss of C151E naphthalene oxidation activity upon cell breakage

suggests that this enzyme was degraded by proteases in Mutant F; however the addition of the protease inhibitors benzamidine and PMSF to the cell breaking buffer did not improve the recovery of the enzyme.

Table 1.2 A summary of some characteristics of the α-subunit hydroxylase mutants from Smith *et al.* (2002). The wild-type was the Mutant F expressing the wild-type sMMO from the expression plasmid. Stability (in the Mutant F expression system) was derived on the basis of being able to detect sMMO polypeptides by SDS-PAGE.

			Activity (nmol min^{-1} mg^{-1} hydroxylase)		
sMMO	Naphthelene assay	stability	Methane	Propene	Toluene
Wild-type	+	+	235 ± 22	244 ± 7	7.8 ± 0.83
T213S	+	+	169 ± 27	43 ± 3	5.6 ± 0.49
T213A	diminished	-			
C151E	+	-			
C151Y	-	-			

1.10 Project aims

1. To develop a method of performing chimaeragenesis of the sMMO hydroxylase α-subunit. To make chimaeras of sMMO with the alkene monooxygenase of *Rhodococcus corallinus* B-276.

2. To explore the diversity of *mmoX* genes in environmental samples using the PCR. To use any "interesting" *mmoX* gene retrieved from an environmental sample in a chimaeragenesis experiment.

3. To develop a method to clone sMMO genes directly from an environmental sample without using the PCR.

4. To improve the *Ms. trichosporium* Mutant F expression system: i) the plasmid system, ii) the stability of mutagenized sMMO, and iii) increasing the range of structural genes that can be targeted by mutagenesis.

Chapter 2

Materials and Methods

2.1 Bacterial strains and plasmids

Table 2.1 Bacterial strains used in this study.

Strain	Genotype	Source/Reference
Escherichia coli TOP10™	F⁻, *mcrA*, *endA1*, *recA1*, ϕ80*lacZ*ΔM15, Δ*lacZX*74, *deoR*, *araD*139, *galK*, *rpsL* (Str^R), *nupG*, Δ(*mrr-hsdRMS-mcrBC*)	Invitrogen
Escherichia coli S17-1	Tp^R, Sm^R, *thi*, *pro*, *hsdR*, *hsdM*, *recA*, RP4 2-Tc::Mu-Km::Tn7	(Simon *et al.*, 1983)
Escherichia coli BLR[1]	F⁻, *ompT*, *hsdS$_B$* (rB⁻ mB⁻), *gal*, Δ(*srl-recA*)306::Tn*10*(Tc^R)	Novagen
Escherichia coli BLR(DE3)	DE3 version of *E. coli* BLR[2]	Novagen
Escherichia coli TransforMax EC300™	F⁻, *mcrA*, Δ(*mrr-hsdRMS-mcrBC*), f80d*lacZ*ΔM15, Δ*lacZX*74, *recA1*, *endA1*, *araD*139, Δ(*ara*, *leu*)7697, *galU*, *galK*, λ⁻, *rpsL*, *nupG*, *tonA*	Epicentre
Sinorhizobium meliloti 1021		Warwick Culture Collection
Methylosinus trichosporium OB3b		Warwick Culture Collection
Methylosinus trichosporium mutant F	*mmoX*::Kan^R (sMMO⁻)	(Martin & Murrell, 1995)
Methylosinus trichosporium SMDM	*mmoC*, Δ(*mmoX-mmoD*)::Gen^R (sMMO⁻)	This Study

1. *E. coli* B and its derivatives are naturally Lon⁻ and Dcm⁻; 2. λ prophage carrying the T7 RNA polymerase gene controlled by the IPTG-inducible *lacUV5* promoter

Table 2.2 Plasmids used in this study.

Plasmid	Description	Source and/or reference
Clone23	2.6 kb *Ms. trichosporium* OB3b *Bgl*II chromosome fragment (containing *clpX & lon*); pUC19 backbone	This study
Clone28	2.3 kb *Ms. trichosporium* OB3b *Bgl*II chromosome fragment (containing *lon*); pUC19 backbone	This study
p34S-Km	Broad-host-range Tn*Mod*-derived plasposon, 4 kb; Ap^R, Km^R	(Dennis & Zylstra, 1998)
p34S-Gm	Broad-host-range Tn*Mod*-derived plasposon, 3.6 kb; Ap^R, Gm^R	(Dennis & Zylstra, 1998)
pACYC177	*E. coli* cloning vector, 3.9 kb; Ap^R, Km^R	NEB/ (Chang & Cohen, 1978)
pBeloBAC11	*E. coli* oriS replicon (single copy) cloning vector, 7.5 kb; Cm^R, *cos* site, *loxP* site, *lacZ*	NEB/ (Shizuya *et al.*, 1992)
pBBR1MCS5	Broad-host-range cloning vector, 4.8 kb; Mob^+, Gm^R, *lacZ*	(Kovach *et al.*, 1994)
pBBR1MCS5b	pBBR1MCS5, digested with *Cla*I/*Bst*BI (to remove the *Bam*HI site) and re-circularized, 4.3 kb; *lacZ'*	This study
pCC1BAC™	Derivative of pBeloBAC11 containing an inducible oriV, 8.1 kb	Epicentre
pCR2.1	*E. coli* TA cloning vector, 3.9 kb; Ap^R, Km^R, *lacZ*	Invitrogen
pD-del	*mmo* cluster, with *mmoD* cloned-out; pJB3Km1 backbone	This study
pD-del-140	pTJS140, carrying the insert from pD-del	This study
pET-3a	*E. coli* expression vector, 4.6 kb; pBR322 derived, T7 promoter, Ap^R	Novagen
pJB3Km1	Broad-host-range cloning vector, 6.1 kb; Mob^+, Ap^R, Km^R, *lacZ*	(Blatny *et al.*, 1997)
pK18mob	Mobilizable pK18 (*E. coli* plasmid) derivative, 3.8 kb; oriT (RP4-mediated), Km^R, *lacZ*	(Schäfer *et al.*, 1994)
pK18mobsacB	pK18mob containing the *B. subtilis sacB* gene; sucrose sensitivity	(Schäfer *et al.*, 1994)
pLONDEL-mob	Clone23 insert containing a Gm cassette in the *Cla*I site (within the *lon* gene); pK18mob backbone	This study
pLONDEL-mobsacB	pLONDEL-mob insert transferred to pK18mobsacB as a *Hin*dIII- *Apa*LI fragment	This study
pMD1	pACYC177, carrying a 1.5 kb *Bam*HI/*Apa*LI *mmoX* fragment from pTJS176; Ap^S	This study

pMD2	pBBR1MCS5b carrying the 10 kb *Kpn*I insert from pTJS176	This study
pMD3	PCR fragment of *mmoX* and upstream region (*Kpn*I sites in primers), ligated at *Kpn*I in pTJS140	This study
pMD3Km1	pMD3 insert excised with *Kpn*I and ligated into pJB3Km1	This study
pMD3KmN	PCR-amplified neomycin gene promoter ligated upstream of *mmoX* between *Bam*HI-*Bst*BI of pMD3Km1	This study
pMD3N-BBR5	pMD3KmN insert excised with *Bam*HI-*Kpn*I and ligated into pBBR1MCS5	This study
pMD4	pJB3Km1 carrying the 10 kb *Kpn*I insert from pTJS176	This study
pMD13-PCR2.1	*amoC* gene (from pTJS104) amplified by PCR to introduce *Kas*I-*Fse*I sites; TA-cloned into pCR2.1	This study
pMD14-PCR2.1	*amoC* gene (from pTJS104) amplified by PCR to introduce *Kas*I-*Nde*I sites; TA-cloned into pCR2.1	This study
pMD15-PCR2.1	*amoC* gene (from pTJS104) amplified by PCR to introduce *Fse*I-*Nde*I sites; TA-cloned into pCR2.1	This study
pMD16-PCR2.1	*amoC* gene (from pTJS104) amplified by PCR to introduce *Fse*I-*Bgl*II sites; TA-cloned into pCR2.1	This study
pMD17-PCR2.1	*amoC* gene (from pTJS104) amplified by PCR to introduce *Fse*I- *Apa*LI sites; TA-cloned into pCR2.1	This study
pMD18-PCR2.1	*amoC* gene (from pTJS104) amplified by PCR to introduce *Nde*I- *Bgl*II sites; TA-cloned into pCR2.1	This study
pMD19-PCR2.1	*amoC* gene (from pTJS104) amplified by PCR to introduce *Nde*I - *Apa*LI sites; TA-cloned into pCR2.1	This study
pMD20-PCR2.1	*amoC* gene (from pTJS104) amplified by PCR to introduce *Bgl*II - *Apa*LI sites; TA-cloned into pCR2.1	This study
pMD13-MD1	Insert from pMD13-pCR2.1 released with *Kas*I-*Fse*I and ligated into *Kas*I-*Fse*I sites of pMD1	This study
pMD14- MD1	Insert from pMD14-pCR2.1 released with *Kas*I-*Nde*I and ligated into *Kas*I-*Nde*I sites of pMD1	This study
pMD17- MD1	Insert from pMD17-pCR2.1 released with *Fse*I- *Apa*LI and ligated into *Fse*I- *Apa*LI sites of pMD1	This study
pMD19- MD1	Insert from pMD19-pCR2.1 released with *Nde*I - *Apa*LI and ligated into *Nde*I - *ApaLI* sites of pMD1	This study
pMD13-MD2	Fragment from pMD13-MD1released with *Bam*HI - *Apa*LI and transferred into *BamHI*- *Apa*LI sites of pMD2	This study
pMD14-MD2	Fragment from pMD14-MD1released with *Bam*HI - *Apa*LI and transferred into *BamHI*- *Apa*LI sites of pMD2	This study
pMD17-MD2	Fragment from pMD17-MD1released with *Bam*HI - *Apa*LI and transferred into *BamHI*- *Apa*LI sites of pMD2	This study
pMD19-MD2	Fragment from pMD19-MD1released with *Bam*HI -*Apa*LI and transferred into *BamHI*- *Apa*LI sites of pMD2	This study
pMD30	*mmo* driven by neomycin promoter; *mmoX'YBZDC* introduced into pMD3N-BBR5 as an *Apa*LI -*Kpn*I fragment	This study
pMD38	*mmoG* driven by the T7 promoter; *Xba*I-*Bam*HI fragment from pMD2 in pET-3a	This study

pMMOC	For construction of pD-del; PCR fragment introducing *Not*I site 5' of *mmoC*, cloned into pCR2.1	This study
pMMODEL-pUC	*mmoG* (*Xba*I-*Bam*HI), *mmoC* (*Bam*HI-*Kpn*I) in pUC19 ligated at *Xba*I-*Kpn*I	This study
pMMODEL-pUC-Gm	Gentamycin cassette from p34S-Gm introduced at *Bam*HI site of pMMODEL-pUC	This study
pMMODEL-pK18-Gm	Insert from pMMODEL-pUC-Gm transferred to pK18mobsacB with *Hin*dIII	This study
pMTL24	Cloning vector, 2.4 kb; Ap^R, *lacZ*	(Chambers *et al.*, 1988)
pSMR1	11 kb *Methylocella silvestris* $BL2^T$ *Sst*I chromosome fragment (Clone 7) containing *mmoXYBZDC*; pUC19	Radajewski, unpublished
pSMR1-140	Insert from pSMR1 transferred to pTJS140 broad-host-range vector	This study
pSMR2	8.6 kb *Methylocella silvestris* $BL2^T$ *Hin*dIII chromosome fragment (Clone 95) containing *mmoRG*; pUC19	Radajewski, unpublished
pTJS140	Sp^R, Sm^R derivative of pJB3km1, 7.5 kb; Km^S	(Smith *et al.*, 2002)
pTJS104	pET-3a with a 4.0 kb insert containing *amoABCD of Rhodococcus corallinus* B-276, Ap^R	(Smith *et al.*, 1999)
pTJS175	pTJS140, carrying a 10 kb *Kpn*I insert containing *mmoRGXYBZDC*[1] of *Ms. trichosporium* OB3b	(Smith *et al.*, 2002)
pTJS176	pMTL24, carrying a 10 kb *Kpn*I insert containing *mmoRGXYBZDC*[1] of *Ms. trichosporium* OB3b	(Smith *et al.*, 2002)
pUC19	*E. coli* cloning vector, 2.7 kb; pMB1 ori (from pBR322) Ap^R, *lacZ*	Invitrogen

1. A *Bam*HI site was introduced upstream of *mmoX* in pTJS175 (described in the source reference).

2.2 Chemicals

Most general purpose chemicals were analytical grade and obtained from Fisher Scientific or BDH. Fine chemicals and reagents were obtained from Sigma Aldrich unless otherwise stated. Water was deionized to 18 MOhm using the SuperQ system (Millipore).

2.3 Growth conditions

2.3.1 *Escherichia coli*

E. coli was routinely cultured using LB medium (Sambrook & Russell, 2001) in conical flasks on a rotary shaker (200 rpm) at 37 °C. Antibiotics were added to the medium at the concentrations indicated in Table 2.3 when appropriate. Long-term storage of cells was achieved by freezing at –80 °C after the addition of sterile glycerol to a final concentration of 15 % (v/v). Solid media were prepared by the addition of 1.5 % (w/v) Bacto Agar (Difco) prior to autoclaving.

Table 2.3 Antibiotics used in this study. All antibiotics, with the exception of cycloheximide, were filter sterilized with a 0.22 μm filter and added aseptically to the cooled autoclaved growth medium. Cycloheximide was added to the medium before autoclaving. All antibiotics were stored at –20 °C except tetracycline dissolved in water, which was made fresh.

Antibiotic	Dissolved in	Stock concentration ($mg\ ml^{-1}$)	Working concentration ($\mu g\ ml^{-1}$)
Ampicillin	water	100	50 – 100
Chloramphenicol	ethanol	25	12.5
Cycloheximide	DMSO	50	50
Gentamycin	water	5	5
Kanamycin	water	50	12.5 – 50
Nalidixic acid	water	1	10
Spectinomycin	water	20	20
Streptomycin	water	20	20
Tetracycline	water/ethanol[a]	25	12.5

[a] tetracycline was dissolved in water to 10 mg ml^{-1} for selection of Tet^R plasmids in *Ms. trichosporium* to prevent ethanol toxicity.

2.3.2 *Ms. trichosporium*

Ms. trichosporium was routinely cultured in Nitrate Mineral Salts medium (NMS), which includes 9.9 mM KNO_3, 3.9 mM phosphate and 0.8 μM Cu^{2+}.

NMS medium composition and preparation:

Salts solution (x 10 stock):

• Potassium nitrate	(KNO_3)	10 g
• Magnesium sulfate	($MgSO_4 \cdot 7H_2O$)	10.8 g
• Calcium chloride	($CaCl_2 \cdot 2H_2O$)	2.65 g

Dissolved in the above order in 700 ml of deionized water and diluted to 1 l in a graduated cylinder.

Iron solution (x 10,000 stock):

• Ferric-EDTA	(Fe-EDTA)	3.8 g

Dissolved in deionized water to a final volume of 100 ml.

Molybdate solution (x 1000 stock):

• Sodium molybdate	($NaMO_3 \cdot 2H_2O$)	0.31 g

Dissolved in deionized water to a final volume of 1 l.

Trace elements solution (x 1000 stock):

• Copper sulfate	($CuSO_4 \cdot 5H_2O$)	0.050 g
• Ferric sulfate	($FeSO_4 \cdot 7H_2O$)	0.125 g
• Zinc sulfate	($ZnSO_4 \cdot 7H_2O$)	0.100 g
• Orthoboric acid	(H_3BO_3)	0.004 g
• Cobalt chloride	($CoCl_2 \cdot 6H_2O$)	0.013 g
• Disodium EDTA	($Na_2EDTA \cdot 2H_2O$)	0.070 g
• Manganese chloride	($MnCl_2 \cdot 4H_2O$)	0.005 g
• Nickel sulfate	($NiSO_4 \cdot 7H_2O$)	0.003 g

Chemicals were dissolved in the above order in deionized water to a final volume of 250 ml.

Phosphate buffer (x 100 stock):

- Disodium orthophosphate ($Na_2HPO_4 \cdot 12H_2O$) 71.6 g
- Potassium orthophosphate (KH_2PO_4) 26 g

The chemicals were dissolved in 800 ml of deionized water in the above order and the pH was adjusted to 6.8. The volume was then adjusted to 1 l.

NMS was prepared by diluting 100 ml of the salts solution to 1 l and adding 1 ml of molybdate solution, 1 ml trace elements solution and 0.1 ml ferric-EDTA solution. 10 ml of phosphate buffer were added per litre of cooled medium (< 50 °C) after autoclaving. For the preparation of solid medium, 15 g l^{-1} Bacto agar (Difco) was added prior to autoclaving. When appropriate, antibiotic was added to the cooled medium after autoclaving (Table 2.3). The occasional contamination with a fungus was eliminated by subculturing several times on NMS plates supplemented with cycloheximide (Table 2.3).

Low-copper medium was prepared using a trace elements solution lacking copper. All glassware was acid-washed (1 M HCl) and rinsed with deionized water. Low-copper NMS medium was solidified by the addition of 20 g l^{-1} Noble agar (Difco) prior to autoclaving.

10 ml NMS cultures of *Ms. trichosporium* were routinely cultivated in 125 ml serum vials sealed with grey butyl stoppers. The vial headspace was supplemented with 10 ml of CH_4/CO_2 (95 %/ 5 % (v/v) respectively) by injection with a needle and syringe. 50 to 100 ml batch cultures were cultured in 250 ml or 1 l Quickfit™ flasks sealed with Subaseals (William Freeman Ltd) and the atmosphere supplemented with 60 ml of CH_4/CO_2. Liquid cultures were incubated at 30 °C on a 200 rpm shaker. The cultures typically reached stationary phase within 3 – 5 days of incubation, normally at an OD_{540} of ~ 0.8. NMS agar plates were incubated at 30 °C in gas-tight containers containing 20 % to 70 % CH_4/ CO_2 (95 %/ 5 % (v/v) respectively) in air. The atmosphere was replenished daily and colonies typically formed in 7 to 10 days.

Purity checks were routinely performed by streaking the culture on nutrient agar (NA) (Sambrook & Russell, 2001) and incubating the plate at 30 °C for 3 days. Since *Ms. trichosporium* does not grow on NA in the absence of CH_4, growth indicated contamination. The presence of contamination was also screened by phase contrast microscopy (Section 2.4).

2.4 Light microscopy

Microscopic examination of cultures was done using a Kyowa Unilux-11 binocular microscope at 1000 x magnification.

2.5 General purpose buffers/ solutions

Agarose gel sample loading buffer (6X)

- Bromophenol blue 0.0125 g
- Xylene cyanol 0.0125 g
- Ficoll (Type 400) 0.75 g
- Deionized water 5 ml

TE buffer (pH 8.0)

- Tris-HCl 10 mM
- Na_2EDTA 1 mM

Prepared from 1 M Tris-HCl (pH 8.0) and 0.5 M Na_2EDTA (pH 8.0)

TBE electrophoresis buffer (10X)

- Tris base 108 g
- Orthoboric acid 55 g
- Na_2EDTA 40 ml 0.5 M solution (pH 8.0)

The volume was adjusted to 1 l with deionized water

TAE electrophoresis buffer (50X)

- Tris 242 g
- Glacial acetic acid 57.1 ml
- Na_2EDTA 100 ml of a 0.5 M solution (pH 8.0)

The Tris was dissolved in ~ 700 ml of water and the volume adjusted tol 1 after the addition of acetic acid and EDTA.

SSC (20X)

- NaCl 173.3 g l^{-1}
- Tri-sodium citrate 88.2 g l^{-1}

The pH was adjusted to 7.0 with 10 M HCl

2.6 DNA extractions

2.6.1 DNA extraction from *Ms. trichosporium*

DNA was isolated from *Ms. trichosporium* as described previously (Oakley & Murrell, 1988). In some instances the CsCl/ ethidium bromide ultracentrifugation was omitted and the crude DNA was purified by an extraction with phenol, chloroform, isoamyl alcohol and precipitated with ethanol (Section 2.8.1). Sometimes the DNA preparations contained considerable amounts of polysaccharide "gloop" which was not removed by repeated CsCl/ ethidium bromide gradient centrifugations or by phenol chloroform extractions. Purification by binding to a silica matrix proved effective (Section 2.8.1).

2.6.2 Small-scale plasmid extractions (minipreps) from *E. coli*

For routine analyses, plasmids were purified from *E. coli* by the alkaline lysis miniprep method (Sambrook & Russell, 2001). Plasmids used as template for DNA sequencing were prepared using the QIAprep™ Miniprep Kit (Qiagen) according to the manufacturer's instructions.

2.6.3 Large-scale plasmid extractions (maxipreps) from *E. coli*

Large scale plasmid preparations were obtained by either alkaline lysis followed by CsCl/ ethidium bromide centrifugation (Sambrook & Russell, 2001) or with the QIAprep® Maxiprep Kit (Qiagen). High-copy number plasmids were prepared from 500 ml *E. coli* cultures (grown for 16 h) and low-copy number plasmids from 0.5 to 2 l cultures.

2.7 Bacterial transformations

2.7.1 Transformation of *Ms. trichosporium* by conjugation

The method for conjugal transfer of RP4-*mob*-containing plasmids into *Ms. trichosporium* was based on the method described previously (Martin & Murrell, 1995). 10 ml of a 16 h culture of *E. coli* S17-1 carrying the plasmid was collected on a sterile 47 mm, 0.2 μm pore-size, nitrocellulose filter (Millipore). The cells were

washed with 50 ml NMS. A fresh 50 ml culture of the *Ms. trichosporium* recipient grown to an OD_{540} of 0.2 – 0.4 (mid-exponential phase of growth) was collected on the same filter as the *E. coli* S17-1 host cells. The cells were washed with 50 ml NMS and the filter placed on an NMS agar plate supplemented with 0.02 % (w/v) Proteose Peptone (Difco) and incubated for 24 – 48 h at 30 °C in the presence of 10 – 50 % CH_4 (v/v) in air. After incubation, the cells from the conjugation plate were washed from the filter with 10 ml of NMS, pelleted by centrifugation at 7,000 x *g* in a Mistral 100 swing rotor (MSE), and resuspended in 1 ml of NMS. 100 µl aliquots were spread onto selective NMS plates and incubated at 30 °C (as described in Section 2.3.2). Colonies typically took 2 to 3 weeks to form. The transconjugants were purified from the *E. coli* S17-1 donor by streaking onto NMS plates containing 10 µg ml^{-1} nalidixic acid, to which *Ms. trichosporium* is naturally resistant, in addition to the appropriate antibiotics for plasmid selection.

2.7.2 Preparation and transformation of chemically competent *E. coli*

E. coli was made chemically competent by treatment with $CaCl_2$ according to the method of Sambrook & Russell (2001). Glycerol or DMSO was used as a cryoprotectant and 100 µl aliquots were frozen in liquid nitrogen. The cells were stored at –80 °C in 2 ml microcentrifuge tubes and were found to retain adequate competence after several months of storage.

A maximum of 5 µl of plasmid DNA was added to the cells which had been thawed on ice. The solution was mixed gently and incubated for 15 min on ice. The tube was placed in a 42 °C water bath for 1 min and then immediately placed on ice for a minimum of 2 min. 900 µl of sterile SOC (Sambrook & Russell, 2001) was added and the tube placed on a 200 rpm shaker at 37 °C for 1 h. Aliquots of the cells were spread on LB agar plates containing the appropriate antibiotic for plasmid selection. The plates were incubated at 37 °C for 16 – 20 h. 0.4 mM IPTG (when necessary) and 40 µg ml^{-1} X-gal were added to the LB plates for blue/ white selection when appropriate.

2.7.3 Preparation and transformation of electrocompetent *E. coli*

A 500 ml culture of *E. coli* was grown to an OD_{600} of 0.5 – 0.7, placed on ice for 15 min and then centrifuged at 4,000 x *g* for 15 min at 4 °C. The supernatant was decanted and the cells were resuspended in 500 ml cold deionized water and centrifuged as before. The supernatant was again decanted and the cells resuspended in 250 ml cold deionized water and centrifuged. The cells were resuspended in 10 ml of cold 10 % (v/v) glycerol and centrifuged. Finally, the cells were suspended in 1 ml of cold 10 % (v/v) glycerol and 100 µl aliquots in 0.5 ml microcentrifuge tubes were frozen in liquid nitrogen and stored at –80 °C. The cells could be stored for several months without excessive loss of competence.

The cells were quickly thawed between thumb and finger and placed on ice. A maximum of 5 µl of solution containing no greater than 100 ng of DNA was added to the cell suspension; if the DNA solution contained significant salt concentrations, for example 1 x ligation buffer, the DNA was first desalted by drop dialysis (Section 2.8.9). The DNA and cells were mixed and placed in a cold 0.2 cm electroporation cuvette. Using a Bio-Rad GenePulser™, an electric field strength of 12.5 kV cm^{-1} at 25 µF and 200 Ω was applied. The cells were immediately suspended in 900 µl of SOC medium (Sambrook & Russell, 2001) at 20 °C and then incubated for 1 h on a rotary shaker (200 rpm) at 37 °C. The transformants were identified as described for chemical transformation (Section 2.7.2). Electrotransformation was always used for *E. coli* S17-1.

2.8 DNA manipulation techniques

2.8.1 DNA purification

DNA was occasionally purified by extraction with phenol (pH 7.6), chloroform, isoamyl alcohol (25:24:1) (Sigma). This was particularly useful for removing protein from DNA solutions. A volume of phenol/chloroform/isoamyl alcohol equal to that of the DNA solution was added and the layers mixed. The aqueous phase was removed and precipitated (Section 2.8.2).

Alternatively, DNA was purified by binding to silica particles using the Geneclean™II Kit (Bio101) according to the manufacturer's instructions. This

method was useful for removing polysaccharides from genomic DNA preparations or for purifying small quantities of DNA (< 100 ng), often lost with methods requiring ethanol precipitation.

2.8.2 DNA precipitation with ethanol

DNA was precipitated from solutions by adding 0.1 volumes of 3 M sodium acetate (pH 5.2) and 2 volumes of ethanol. The solution was mixed and the DNA centrifuged at 20,800 x *g* at 4 °C for 20 min. The supernatant was decanted and the DNA pellet rinsed with 70 % (v/v) ethanol, air dried and dissolved in TE buffer.

2.8.3 DNA quantification

DNA concentrations were estimated by agarose gel electrophoresis (Section 2.8.8) and comparison with a known quantity of 1 kb ladder (Gibco BRL) or *Hin*dIII digested λ DNA.

2.8.4 DNA restriction digests

DNA restriction enzymes were supplied by Gibco BRL, New England Biolabs and MBI Fermentas (Helena Biosciences Ltd) and used according to the manufacturers' recommendations.

2.8.5 Dephosphorylation

DNA ends were dephosphorylated when appropriate using calf intestinal alkaline phosphatase (Roche) according to the manufacturer's instructions.

2.8.6 DNA ligations

DNA ligations were performed using T4 DNA ligase and ligation buffer from Gibco BRL or MBI Fermentas. Vector molecules were digested with the appropriate restriction enzyme(s), dephosphorylated (Section 2.8.5) when necessary, and agarose gel purified (Section 2.8.9). The insert DNA was digested with the appropriate enzyme and agarose gel purified. Approximately 50 ng of purified vector and an equimolar quantity of insert DNA were combined in a minimum reaction volume

(usually 10 – 20 µl), containing 1 x ligation buffer, 1 mM dATP and 1 unit of T4 DNA ligase. The reaction mixture was incubated at 16 °C overnight.

2.8.7 TA cloning PCR fragments

PCR products (Section 2.9) were cloned using the TA cloning kit (Invitrogen) according to the manufacturer's instructions. PCR products generated with *Pfu*Turbo™ (Stratagene) were gel purified (Section 2.8.9) and incubated at 72 °C for 15 min in 1 x PCR buffer (Gibco-BRL) containing 1.5 mM $MgCl_2$, 200 µM dATP and 1 unit of *Taq* DNA polymerase (Gibco-BRL) prior to cloning.

2.8.8 Agarose gel electrophoresis of DNA

Routine agarose gel electrophoresis was performed using a Flowgen Minigel Systems electrophoresis unit. Gels were prepared with 1 x TBE electrophoresis buffer containing between 0.8 % and 2 % (w/v) agarose and 0.5 µg ml^{-1} ethidium bromide. Gels were visualized on a UV transilluminator and photographed using a Polaroid camera (CU5 Land Camera) with Polaroid 665 black and white film. Alternatively, gels were photographed using a UVP GD8000 gel documentation system (Ultra-Violet Products Ltd).

2.8.9 DNA recovery from agarose gels

Routine DNA purifications from both TBE and TAE agarose gels were performed using the Geneclean™II Kit (Bio101) according to the manufacturer's instructions. DNA fragments between 70 and 200 bp were purified using the QIAquick™ Gel Extraction Kit (Qiagen).

High molecular mass DNA was recovered from agarose gels by electroelution. Dialysis tubing was prepared according to Sambrook and Russell (2001). The gel segment containing the DNA to be recovered was excised with a scalpel and clamped inside the dialysis tubing with approximately 250 µl of 1 x TAE buffer. The bag was submerged in 1 x TAE inside an electrophoresis unit and 5 volts cm^{-1} applied for 45 min. The current was reversed for 1 min to free DNA retained on the tubing wall. To minimize mechanical shearing of the DNA, the solution was removed from the tubing using wide bore 200 µl Gilson pipette tips constructed by cutting ~ 3 mm from the tip

end. The DNA was purified by drop dialysis against 0.5 x TE buffer using VSWP membranes (Millipore) according to the method of Sambrook and Russell (2001).

2.9 Polymerase chain reaction (PCR)

PCR amplifications were performed in a total volume of 50 µl in 0.5 ml microcentrifuge tubes using a Hybaid Touchdown™ Thermal Cycling System. Routine PCR amplifications used 2.5 units of *Taq* DNA polymerase from Gibco BRL or MBI Fermentas with the reagents supplied by the manufacturer at the recommended concentrations. Unless otherwise indicated, a concentration of 1.5 mM $MgCl_2$, 200 µM each dNTP and 200 nM of each primer (equivalent to 10 pmol per 50 µl reaction) was used. Custom made primers were obtained from Invitrogen. Unless otherwise indicated, cycling conditions were as follows: a hot start was performed by adding the *Taq* DNA polymerase after 5 min at 94 °C, followed by 30 cycles of: 1 min at 92 °C, 1 min at 60 °C and 1 min at 72 °C. A final extension of 72 °C for 10 min was performed.

PCR amplifications requiring greater polymerase accuracy were performed using *Pfu*Turbo™ (Stratagene) proofreading polymerase according to the manufacturer's instructions. Primers were added to the reactions immediately before cycling to minimize 5' to 3' primer degradation by the *Pfu*Turbo™ enzyme.

2.10 DNA blotting, hybridization and detection

2.10.1 Southern transfer of DNA

Approximately 4 µg of restriction digested genomic DNA samples were separated by electrophoresis through a 0.8 % (w/v) agarose gel on a BRL Model H4 Horizontal Gel Systems unit (Bethesda Research Laboratories) for 16 h at 70 volts in 1 x TAE. The DNA was visualized and the gel photographed after staining for 30 min in 1 x TAE containing 0.5 µg ml^{-1} ethidium bromide; a ruler was placed alongside the gel so that the position of a signal on an autoradiograph could be correlated to the position in the agarose gel. The DNA was denatured by soaking the gel in Denaturing Solution (0.5 M NaOH, 1.5 NaCl) and transferred to Hybond-N+ nylon membrane

(Amersham) by upward capillary transfer as described in the Hybond-N+ product literature. DNA was fixed to the membrane with a UV Stratalinker (Stratagene).

2.10.2 Colony blots

E. coli colony blots were performed by picking colonies onto a grid-marked Hybond-N+ membrane (Amersham) placed on the surface of a selective LB agar plate. The plates were incubated for 16 to 24 h at 37 °C and then lysed on the membrane by the following procedure:

3MM Whatman paper was cut larger than the size of the membrane and 2-sheet stacks were placed in dishes. One stack of Whatman paper was soaked in 10 % (w/v) SDS, another in Denaturing Solution (0.5 M NaOH, 1.5 NaCl), another in Neutralizing Solution (1 M Tris-HCl (pH 7.4), 1.5 M NaCl) and another in 2 x SSC. The membrane was placed on each wetted Whatman paper for 10 min, in the above order. The membrane was dried for 30 min at room temperature and the DNA was then fixed to the membrane with a UV Stratalinker (Stratagene).

2.10.3 Radiolabelling of DNA by random priming

DNA probes were labelled by random priming (Feinberg & Vogelstein, 1983; Feinberg & Vogelstein, 1984) using hexanucleotide primers and dNTPs from Roche according to the manufacturer's instructions. 25 to 50 ng of agarose gel purified DNA fragment was labelled with 50 μCi of [α-^{32}P]dGTP for 1 h at 37 °C with Klenow polymerase (Invitrogen). Unincorporated label was sometimes removed using a MicroSpin™ Column (Amersham Pharmacia Biotech) according to the manufacturer's instructions. The probes were denatured by the addition of NaOH to a final concentration of 0.4 M and incubated for 2 min at room temperature before adding to the hybridization solution.

2.10.4 DNA hybridizations

Membranes were rolled in mesh and placed in a Hybaid tube containing 20 ml of prewarmed prehybridization solution (0.5 M sodium phosphate (pH 7.2), 7 % (w/v)

SDS, 5 mM EDTA) and incubated in a Hybaid oven for 30 min at 55 °C. The prehybridization solution was discarded and replaced with 20 ml of fresh 55 °C solution, to which the denatured radiolabelled DNA probe was added. Hybridizations were carried out overnight at 55 °C.

The hybridization solution was discarded and the membrane washed twice with 100 ml of 2 x SSC at 55 °C. The membrane was removed and scanned with a Geiger Müller detector to estimate the amount of bound probe. If necessary, the membrane was washed at greater stringency by incrementally decreasing the salt concentration and/or increasing the wash temperature. In most instances, a wash with 0.5 x SSC at 65 °C was adequate to remove "non-homologous" hybridization signals.

2.10.5 Autoradiography

Fuji 100NIF Medical X-ray Film (Fujifilm) was exposed to hybridized membranes between intensifying screens in autoradiography cassettes at –80 °C. The length of exposure varied between 1 and 72 h, based on the signal intensity. Autoradiograph film was developed and fixed in accordance with the manufacturer's instructions. Alternatively, membranes were exposed to a phosphorimaging plate (Fujifilm) and the signal detected using a FLA-5000 phosphorimager (Fujifilm) according to the manufacturer's instructions.

2.11 Measuring soil pH

1 g of soil was suspended in 10 ml of 10 mM $CaCl_2$ and the pH of the suspension measured with a pH meter (Hanna).

2.12 Stable Isotope Probing (SIP) of soil with $^{13}CH_4$

The characteristics of the Gisburn forest soil were described previously (Radajewski *et al.*, 2002). Soil was collected in August 2002 and SIP experiments were performed essentially as described previously. A soil slurry was made by adding 10 ml of ANMS medium (NMS medium (Section 2.3.2) except containing 0.5 g NH_4Cl, 0.5 g KNO_3 and buffered with 4 mM phosphate (pH 3.5)) to 5 g of soil. The slurry was incubated in a 125 ml serum vial sealed with a butyl stopper on a

rotary shaker (~ 100 rpm) at room temperature (20 – 25 °C) and in the dark. The soil DNA was harvested (Section 2.13) after consumption of a total of 50 ml of $^{13}CH_4$ (Linde) added in 10 ml aliquots to the soil slurry microcosm (Radajewski *et al.*, 2002).

2.13 DNA extraction from soil

2.13.1 DNA isolation

The protocol for DNA extraction from soil was based on the method previously described (Zhou *et al.*, 1996), with several modifications. 5 g (wet weight) of soil was placed in a 35 ml Oakridge tube and centrifuged at 7,000 rpm in a JA20 rotor. The soil pellet was suspended in 13.5 ml of extraction buffer (100 mM Tris-HCl (pH 8.0), 100 mM sodium EDTA (pH 8.0), 100 mM sodium phosphate (pH 8.0), 1.5 M NaCl, 1 % (w/v) CTAB) to which 100 µl of fresh Proteinase K (10 mg ml^{-1}) was added; the tube was placed horizontally on a 200 rpm shaker at 37 °C for 30 min. 1.5 ml of 20 % (w/v) SDS was added after the 37 °C incubation and the tube was then placed at 65 °C for 2 h and mixed by inversion every 15 min. The supernatant was decanted into a clean tube after centrifugation for 10 min at 7,000 rpm at 25 °C. The soil pellet was again resuspended in 4.5 ml of extraction buffer and 0.5 ml of 20 % (w/v) SDS, incubated at 65 °C, centrifuged as before and the supernatant added to the first aliquot. The crude extract (~ 20 ml) was gently extracted with chloroform (containing 4 % (v/v) isoamyl alcohol to minimize foaming) and centrifuged at 11,500 rpm for 10 min at 25 °C. The aqueous phase was transferred to a clean tube with a wide bore 5 ml pipette tip, made by cutting ~ 5 mm from the tip end. Care was taken to leave the interface undisturbed. The DNA was precipitated from the aqueous phase by adding 0.6 volumes (~ 10.5 ml) of 2-propanol, mixing gently and incubating for 1 h at room temperature. The DNA was pelleted by centrifugation at 11,500 rpm for 20 min at 20 °C. The DNA pellet was rinsed with 5 ml of 70 % (v/v) ethanol and air dried for 20 min.

2.13.2 Ultracentrifugation

DNA was dissolved in 20 ml of TE buffer at 4 °C for 16 h. The volume of DNA solution was measured and exactly 1 g ml^{-1} CsCl was dissolved by gentle mixing. 500 µl of ethidium bromide (10 mg ml^{-1}) was added and the solution

transferred to a 25 x 89 mm polyallomer Quick-Seal™ ultracentrifuge tube (Beckman). Unfilled tube volume was filled by adding 1 g ml^{-1} CsCl in TE buffer. The tube was centrifuged at 180,000 x *g* (45,000 rpm) for 20 h in a VTi50 rotor at 20 °C and stopped without braking. If the position of the DNA in the gradient could not be seen in visible light, the tube was exposed to long wavelength UV (365 nm) for a minimum time (< 2 sec); it was found that indiscriminate UV exposure subsequently made it impossible to clone the DNA into pBeloBAC11 or pCC1BAC™. The portion of gradient containing the DNA was collected using a 16 gauge needle and 2.5 ml syringe (Sambrook & Russell, 2001). Another 250 µl of ethidium bromide (10 mg ml^{-1}) was added to the DNA solution, which was transferred to a new 25 x 89 mm Quick-Seal™ tube and centrifuged as before. DNA bands were collected with a needle and syringe as before. The DNA was then transferred to a 13 x 51 mm polyallomer Quick-Seal™ tube and centrifuged at 265,000 x *g* (55,000 rpm) in a VTi65 rotor at 20 °C. If the DNA was not from a SIP experiment, the DNA band was collected and purified as described in Section 2.13.3. If the DNA preparation contained light (^{12}C-DNA) and heavy (^{13}C-DNA) fractions, the bands were collected independently (^{13}C-DNA followed by ^{12}C-DNA) and each placed in a new 13 x 51 mm Quick-Seal™ tube and centrifuged as before. DNA was collected as above and then purified further (Section 2.13.3).

2.13.3 DNA purification from CsCl gradients

Ethidium bromide was removed from DNA preparations by repeated extractions with water saturated 1-butanol (Sambrook & Russell, 2001). The CsCl was removed by dialysis against TE buffer (Sambrook & Russell, 2001) and the DNA stored at 4 °C. The DNA often required further purification for use in molecular biology applications. The DNA was purified using the GenecleanII™ Kit (Bio101) for use as template in PCR. Purification for cloning is described in Section 2.14.1.

2.14 Cloning soil DNA

2.14.1 Further purification of DNA for cloning

The desalted DNA was purified by electrophoresis on a 1 % (w/v) low melting point (LMP) agarose gel in 1 x TAE buffer without added ethidium bromide. LMP agarose (Gibco) was used since it has a large pore size which allows the restriction enzyme to penetrate the agarose plug (Sambrook & Russell, 2001) as described in Section 2.14.2. To minimize DNA shearing, 3 mm from the end of the pipette tips were cut-off to increase the bore of the opening for all pipetting of ^{13}DNA. A Nile Blue staining protocol, which forms a complex with DNA visible in white light, was used in place of ethidium bromide to eliminate any further exposure of the DNA to UV radiation (Adkins & Burmeister, 1996; Yang *et al.*, 2000). The gel was stained by soaking overnight in 15 μg ml^{-1} Nile Blue in water. The gel fragment (~ 3 mm thick) containing the DNA was excised with a clean scalpel and stored in TE buffer at 4 °C.

2.14.2 Partial restriction enzyme digestion of DNA

Agarose gel fragments (Section 2.14.1) containing ~ 2 μg DNA were washed three times at room temperature in 20 ml TE buffer on a tube roller. The agarose plugs were suspended in 1 ml of 1 x restriction enzyme buffer and incubated for 1 h at 37 °C. The equilibration in restriction enzyme buffer was repeated with fresh buffer. The plugs were transferred to 2 ml microcentrifuge tubes containing 500 μl of cold 1 x restriction enzyme buffer with *n* units of restriction enzyme. For each experiment, a range of enzyme concentrations was used; for example, with *Bam*HI n = 0, 5, 10, 25, 50, 100, 200 and 500 units. The enzyme was allowed to penetrate the agarose matrix for 1 h on ice. The tubes were then incubated for 1 h in a 37 °C water bath so that the restriction enzyme could cut the DNA molecules.

Immediately after restriction enzyme digestion, the agarose plugs were inserted into the wells of a 20 cm 1 % (w/v) agarose TAE gel. DNA markers of an appropriate size were included on the gel and loaded on both sides of the lanes containing the agarose plugs. Electrophoresis was performed as described in Section 2.10.1. A relatively complex staining protocol was followed to circumvent exposing

the partially digested DNA to UV radiation. The lanes containing the DNA markers were cut from the gel and stained by soaking in 0.5 μg ml^{-1} ethidium bromide (Section 2.10.1). The stained gel was placed on a UV transilluminator and the positions of the DNA bands of the markers were indicated by placing strips of paper on the gel surface. The DNA marker lanes were then replaced alongside the unstained gel. Using the paper strips as a guide, the gel region containing the desired size of partially digested DNA was derived and the regions removed with a scalpel from each lane. The remainder of the gel was stained in ethidium bromide and photographed on a UV transilluminator. The gel slices containing size-selected DNA were identified as those having DNA both ahead and behind the removed region; for *Bam*HI, this was found to be those treated with between 50 and 100 units of enzyme.

2.14.3 Ligation of DNA into pCC1BAC™

The partially digested and size-selected DNA was eluted from the agarose as described in Section 2.8.9. The DNA was ligated into pCC1BAC™ in a 100 μl volume according to the manufacturer's instructions (Epicentre).

2.14.4 Electrotransformation of *E. coli* TransforMax EC300™

Transformations of DNA into *E. coli* TransforMax EC300™ were performed according to the supplier's instructions (Epicentre) and as described in Section 2.7.3. The ligations were desalted by drop dialysis against 0.5 x TE buffer using VSWP membranes (Millipore) according to the method of Sambrook and Russell (2001). The electroporation was performed in 0.1 cm cuvettes at the following settings: 2.5 kV cm^{-1} at 25 μF and 100 Ω on a Bio-Rad GenePulser™.

2.14.5 Construction of a soil DNA clone library

White colonies were picked from the LB (40 μg ml^{-1} X-gal, 0.4 mM IPTG, 12.5 μg ml^{-1} Cm) agar plates using sterile toothpicks and transferred to individual wells of 96 well plates containing 150 μl per well of freezing medium (LB broth, 7.5 % (v/v) glycerol, 12.5 μg ml^{-1} Cm). The plates were incubated for 24 h at 37 °C and stored at –80 °C.

Colony blots were performed essentially as described in Section 2.10.2. The membranes were inoculated from the 96 well plates using a 96-pin replicating device (Boekel). Colony hybridization of the soil library was performed as described in Section 2.10.

2.15 DNA sequencing and analysis

2.15.1 DNA sequencing and data analysis

DNA was sequenced at the University of Warwick Molecular Biology Service using Big Dye Terminator Version 3.1 chemistry and run on a 3100 Genetic Analyser (Applied Biosystems).

Sequences were analysed using the Chromas software programme (http:\\www.technelysium.com.au\chromas.html) and the SeqMan™II programme (DNASTAR Inc). Phylogenetic analyses were performed using the ARB software package (http:\\www.mikro.biologie.tu-muenchen.de).

2.15.2 Shotgun sequencing

Large clones (>10 kb) were sequenced using a shotgun sequencing method. The cloned insert DNA was excised from the vector and gel purified (Section 2.8.9). The purified insert DNA was partially digested for 1 h at 37 °C with *Bsp*143I by incubating with a limiting quantity of enzyme (~ 0.05 units, optimized empirically). The digested fragments were resolved by agarose gel electrophoresis and fragments 0.8 – 1.5 kb were gel purified and ligated into pUC19, that had been linearized with *Bam*HI and dephosphorylated (Sections 2.8.5 & 2.8.6). *E. coli* transformants were isolated, the clones purified by plasmid miniprep (Section 2.6.2) and analysed by RFLP to ensure the inserts were 0.8 – 1.5 kb. Clones were chosen at random and the plasmid DNA sequenced using the M13F primer (Section 2.15.1). Approximately six-fold sequence data were accumulated and analysed using the SeqMan™II software programme (DNASTAR Inc). To join contigs and re-sequence regions of poor data, the original (full length) clone was sequenced using custom designed sequencing primers.

2.16 Naphthalene oxidation assay for sMMO activity

The routine qualitative detection of sMMO activity was performed using the naphthalene oxidation assay (Brusseau *et al.*, 1990; Graham *et al.*, 1992). The assays were performed as originally described, except that the tetrazotized-*o*-dianisidine solution was pipetted onto the colonies rather than sprayed. Colonies on copper-containing NMS plates were grown to > 5 mm diameter before assaying.

2.17 Denaturing polyacrylamide gel electrophoresis (SDS-PAGE)

2.17.1 Whole cell protein preparation

Protein was routinely prepared from *Ms. trichosporium* and *E. coli* by boiling cells from 1-10 ml of culture for 5 min in 100 µl SDS-PAGE Sample Buffer (63 mM Tris-HCl (pH 6.8), 10 % (v/v) glycerol, 5 % (v/v) β-mercaptoethanol, 2 % (w/v) SDS, 0.00125 % (w/v) bromophenol blue). Alternatively, protein was prepared using a Ribolyser (Hybaid). Cells were suspended in 200 µl of 20 mM Tris-HCl (pH 7.4) and bead-beaten 5 times in Ribolyser Blue tubes for 10 seconds at speed setting 6 ms^{-1}. The samples were incubated on ice for 2 min between each round of bead-beating. The supernatant was collected after centrifugation at 13,000 x *g* for 5 min in a microcentrifuge.

2.17.2 Protein quantification

Total protein concentration was determined using the Bio-Rad Protein Assay according to the manufacturer's instructions. Bovine serum albumin was used as the standard.

2.17.3 SDS-PAGE

Protein samples were analysed by SDS-PAGE (Laemmli, 1970) using an X-Cell II™ Mini-Cell system (Novex). A 4 % (w/v) stacking gel and a 10 % or 12 % (w/v) resolving gel was used, prepared as follows:

Stacking gel:

- 0.53 ml 40 % (w/v) acrylamide
- 0.3 ml 2 % (w/v) N,N'-methylenebisacrylamide
- 1.36 ml stacking buffer (0.5 M Tris-HCl, pH 6.8)
- 2.8 ml water
- 50 µl 10 % (w/v) SDS
- 50 µl 10 % (v/v) glycerol
- 0.5 ml fresh 10 % (w/v) ammonium persulfate
- 5 µl TEMED

10 % Resolving gel:

- 2.4 ml 40 % (w/v) acrylamide
- 1.36 ml 2 % (w/v) N,N'-methylenebisacrylamide
- 2.5 ml resolving buffer (3 M Tris-HCl, pH 8.8)
- 2.65 ml water
- 100 µl 10 % (w/v) SDS
- 100 µl 10 % (v/v) glycerol
- 0.5 ml fresh 10 % (w/v) ammonium persulfate
- 5 µl TEMED

A 12 % gel was made as for a 10 % gel except that 3.15 ml of acrylamide and 1.69 ml of 2 % N,N'-methylenebisacrylamide were used and the volume of water decreased accordingly.

Samples were mixed with at least an equal volume of SDS-PAGE Sample Buffer (Section 2.17.1), boiled for 5 min, loaded onto the gel and separated by applying 200 volts for 90 min with 1 x Tris-glycine electrophoresis buffer (14.4 g l^{-1} glycine, 3.0 g l^{-1} Tris, 1.0 g l^{-1} SDS). Gels were calibrated by including the Dalton Mark VII-L protein standard (Sigma).

Proteins were visualized by soaking the gel for 1 h in Coomassie Brilliant Blue staining solution (0.1% Coomassie Brilliant Blue R-250, in 40 % methanol, 10 % acetic acid, 50 % water by volume) and destained in a solution containing 30 % methanol, 10 % acetic acid and 60 % water by volume. Gels were photographed using a UVP GD8000 gel documentation system (Ultra-Violet Products Ltd).

Chapter 3

Chimaeric sMMO Hydroxylase Mutants

3.1 Introduction

Mutagenesis of residues in the active site of the sMMO enzyme became possible with the development of an homologous sMMO expression system (Lloyd *et al.*, 1999b). The first mutagenesis experiments were performed by Smith and co-workers (2002), where the cysteine-151 and threonine-213 residues were modified by site-directed mutagenesis. This study demonstrated the effectiveness of the expression system and proof-of-principle for sMMO mutagenesis.

Combinatorial mutagenesis is a powerful approach to obtain enzymes with enhanced catalytic properties (Crameri *et al.*, 1998). In this approach, chimaeric proteins are obtained by recombining gene sequences of related enzymes, which can result in novel catalytic functions and in some instances in properties not possessed by either parent enzyme (Campbell *et al.*, 1997). The gene shuffling technique is a popular method that involves the reassembly of a gene mixture partially digested with DNAse 1 (Stemmer, 1994). This technique is effective with genes sharing high sequence identity and the generation of chimaeras between enzymes sharing less identity (corresponding to < 65 % amino acid identity) require other techniques (Hiraga & Arnold, 2003; Horton *et al.*, 1989; Kawarasaki *et al.*, 2003). In this study, a simple approach was employed, using a non-random approach to generate chimaeras of sMMO and alkene monooxygenase (AMO) from *Rhodococcus corallinus* B-276.

The AMO is one of the closest relatives of sMMO (Leahy *et al.*, 2003; Saeki & Furuhashi, 1994). The AMO epoxygenase component contains a diiron centre with similar EXXH iron-bridging ligands to those found in the sMMO hydroxylase, but unlike the $(\alpha\beta\gamma)_2$ of sMMO, the AMO epoxygenase has an $\alpha\beta$ configuration (Gallagher *et al.*, 1997; Gallagher *et al.*, 1998; Miura & Dalton, 1995). The derived sequences of the α–subunits of sMMO and AMO share 35 % amino acid identity (Saeki & Furuhashi, 1994) and, based on the X-ray crystal structure of the sMMO hydroxylase and sequence-alignment modelling between AmoC and MmoX (which are the α–subunits of the AMO and sMMO respectively), the tertiary structure of the AMO α–subunit is believed to be similar to that of the sMMO, despite the relatively low sequence identity (Gallagher *et al.*, 1998). A conserved structure between the parent proteins is most often critical in obtaining chimaeras that fold properly (Bernhardt, 2004; Voigt *et al.*, 2002).

A noteworthy difference between the α-subunits of the sMMO and AMO is that the residue corresponding to the cysteine-151 of the sMMO is a glutamate in AMO. The cysteine-151 of sMMO corresponds to tyrosine-122 of ribonucleotide reductase which forms a catalytically essential tyrosyl radical (Nordlund *et al.*, 1992; Ormö *et al.*, 1995). Therefore there is the possibility that an essential sulfhydryl radical at cysteine-151 is essential for sMMO hydroxylations (Rosenzweig *et al.*, 1993). By site-directed mutagenesis, Smith and co-workers (2002) changed the cysteine-151 residue to a glutamate (C151E), as found in the AMO. Whole cells of *Ms. trichosporium* Mutant F expressing C151E-sMMO were still capable of oxidizing naphthalene, indicating that the enzyme was active and therefore a catalytically essential free-radical at this position was not possible. Therefore, the presence of glutamate in the AMO rather than cysteine is perhaps less significant than originally believed.

Catalysis of the AMO and sMMO have several important differences. The substrate specificity of the sMMO is considerably wider than the AMO, which is limited to epoxygenations of C3 and C4 1- and 2-alkenes (Miura & Dalton, 1995; Saeki *et al.*, 1999). The most intriguing difference is that the AMO performs stereospecific oxidations, such as propene to chiral propene oxide (Gallagher *et al.*, 1997). Molecular docking studies suggest that hydrophobic residues alanine-91, alanine-95, phenylalanine-169, leucine-181 and alanine-185 in the putative active site of the AMO create a substrate binding pocket, and that the alanines 91 & 185 correspond to glycines in the sMMO (Gallagher *et al.*, 1998). The slightly bulkier alanine-91 and alanine-185 compared with glycine perhaps help to clamp the substrate in the active site and may account for the difference in stereoselectivity between sMMO and AMO (Gallagher *et al.*, 1998).

In this study, combinatorial mutagenesis was performed by swapping regions of the *mmoX* gene with the corresponding sequence of the *amoC* gene. This was done simply using selected restriction enzyme sites, rather than constructing a random library using gene shuffling techniques. The aim of the study was to determine if active chimaeras could be generated, and if obtained, to compare the catalytic properties of the chimaeras with that of AMO and sMMO.

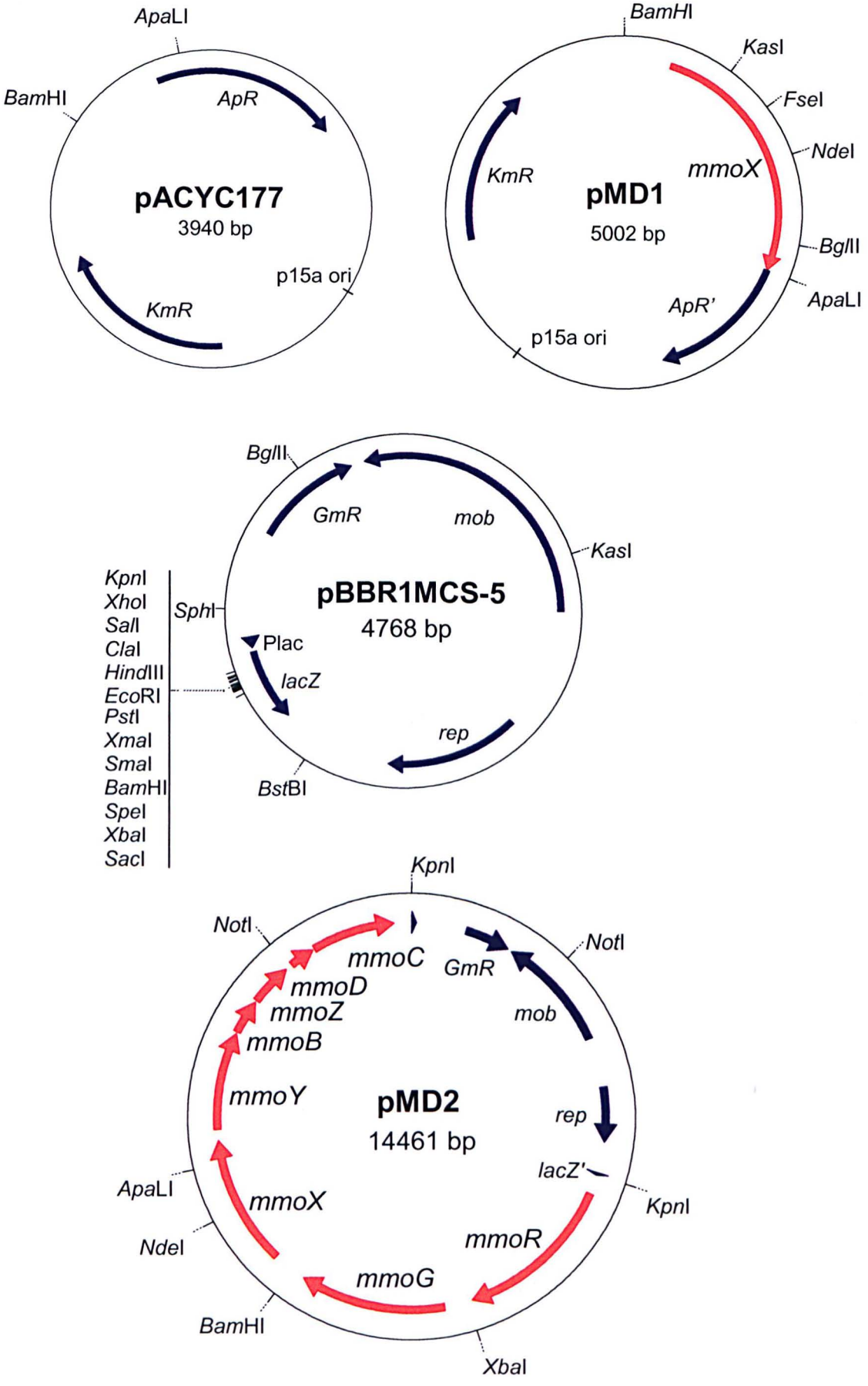
3.2 sMMO mutagenesis and recombinant expression system

The pNB101 and pTJS176 plasmids used in the previous *mmoX* mutagenesis study allowed the mutagenesis of residues 5' of an *Nde*I restriction site in the *mmoX* gene (Smith *et al.*, 2002). The region between 291 bp and 771 bp of the *mmoX* gene (relative to the start codon) encodes the active site of the enzyme (Elango *et al.*, 1997), and the cutting position of *Nde*I is at position 736 bp. A second cavity is adjacent to the active site pocket may also be involved in substrate binding (see Section 1.5.1.1). Therefore, mutagenesis targets using the pNB101 and pTJS176 plasmid system is limited to a segment of the core site. A system which enabled the substitution of residues over the entire core region of the enzyme was desired for this experiment, and therefore new plasmids were constructed.

An *Apa*LI site was present in *mmoX* and cut at position 1,312 bp, which was situated a suitable distance 3' of the region encoding the active site. A 1,559 bp *Bam*HI/*Apa*LI segment of pTJS176 was subcloned into pACYC177, and the plasmid was designated pMD1 (Figure 3.1). Within the *mmoX* fragment on pMD1, the restriction enzyme sites that occurred once in the plasmid molecule included the recognition sequences for *Kas*I, *Fse*I, *Nde*I and *Bgl*II. These enzymes generated cohesive ends, and the mutagenesis could be performed by cutting with two of the enzymes and replacing the gene segment between the sites with the corresponding *amoC* gene sequence, as described in the following section.

Following the modification of the *mmoX* sequence in pMD1, an additional plasmid was required for reincorporating the *Bam*HI/*Apa*LI segment into the *mmo* cluster. The requirements of the plasmid were that the vector backbone possess no *Bam*HI or *Apa*LI restriction sites. The pBBR1MCS5 plasmid was chosen. Since pBBR1MCS5 is a mobilizable broad-host-range plasmid, it should be suitable for introduction and maintenance in *Ms. trichosporium* Mutant F, therefore eliminating the need to subsequently subclone the mutagenized fragment into pTJS140 (or another suitable mobilizable and broad-host-range plasmid). The pBBRIMCS5 plasmid contained no *Apa*LI restriction sites, and the single *Bam*HI site was removed by digesting the plasmid with *Cla*I and *Bst*BI, which possess compatible cohesive ends, and re-circularizing the plasmid (Figure 3.1). The fragment from pTJS176 was transferred to the modified pBBRIMCS5 plasmid as a *Kpn*I fragment; this plasmid was designated pMD2 (Figure 3.1).

Figure 3.1 Plasmids used for the mutagenesis experiments. The *mmo* genes are shown in red. The mutagenesis could be performed on pMD1, and the mutagenized segment re-incorporated into the *mmo* cluster on pMD2. The pMD2 plasmid could be conjugated and maintained in *Ms. trichosporium* Mutant F.


ApaLI
BamHI
ApR
pACYC177
3940 bp
p15a ori
KmR
BamHI
KasI
FseI
NdeI
mmoX
KmR
pMD1
5002 bp
BglII
ApaLI
ApR'
p15a ori
BglII
GmR
mob
KasI
KpnI
XhoI
SalI
ClaI
HindIII
EcoRI
PstI
XmaI
SmaI
BamHI
SpeI
XbaI
SacI
SphI
pBBR1MCS-5
4768 bp
Plac
lacZ
rep
BstBI
KpnI
NotI
mmoC
GmR
NotI
mmoD
mmoZ
mmoB
mob
mmoY
pMD2
14461 bp
rep
lacZ'
ApaLI
mmoX
KpnI
NdeI
mmoR
mmoG
BamHI
XbaI

3.3 Construction of chimaeric *mmoX-amoC smmo* mutants

Mutants were created by replacing *mmoX* sequence between restriction enzyme sites with the corresponding sequence of the *amoC* gene. The restriction enzymes sites used were *Kas*I, *Fse*I, *Nde*I and *Apa*LI, and the approximate position the sites compared with the aminoacyl sequences of the MmoX and AmoC sequences is shown in Figure 3.2. The construction of mutants was performed by PCR amplifying *amoC* gene sequence corresponding to the sequence of the *mmoX* gene being replaced. The 3' –end of the PCR primers corresponded to the *amoC* gene, and the 5' –end, including the restriction enzyme recognition sequence, corresponded to the *mmoX* sequence (Table 3.1). The reading frame was maintained so that translation of the chimaeric gene would result in a chimaeric protein (Figure 3.3). The PCR was performed with *Pfu*Turbo proof-reading polymerase to minimize the possibility of incorporating PCR errors into the sequence. The *amoC* on pTJS104 was used as a template (Smith *et al.*, 1999). The PCR products were cloned into pCR2.1 (Section 2.8.7) for sequencing using the M13F and M13R sequencing primers. The fragments were released from the pCR2.1 vector by digestion with the restriction endonucleases at the temperature appropriate for the enzyme. The pMD1 plasmid was digested with the same enzymes. The DNA fragments were gel purified (Section 2.8.9), ligated (Section 2.8.6) and transformed into chemically competent *E. coli* (Section 2.7.2). Plasmids were screened by restriction enzyme digestion and RFLP analysis. The proper ligation of several fragments was achieved, and the plasmids were designated pMD13, pMD14, pMD17 and pMD19 (Figure 3.4). The mutagenized fragments were united with the operon by releasing the fragment with *Bam*HI/*Apa*LI and ligating into gel purified pMD2 that had been cut with *Bam*HI/*Apa*LI to remove the sequence between the enzyme sites. The resulting plasmids were designated pMD13-2, pMD14-2, pMD17-2 and pMD19-2 (Figure 3.4).

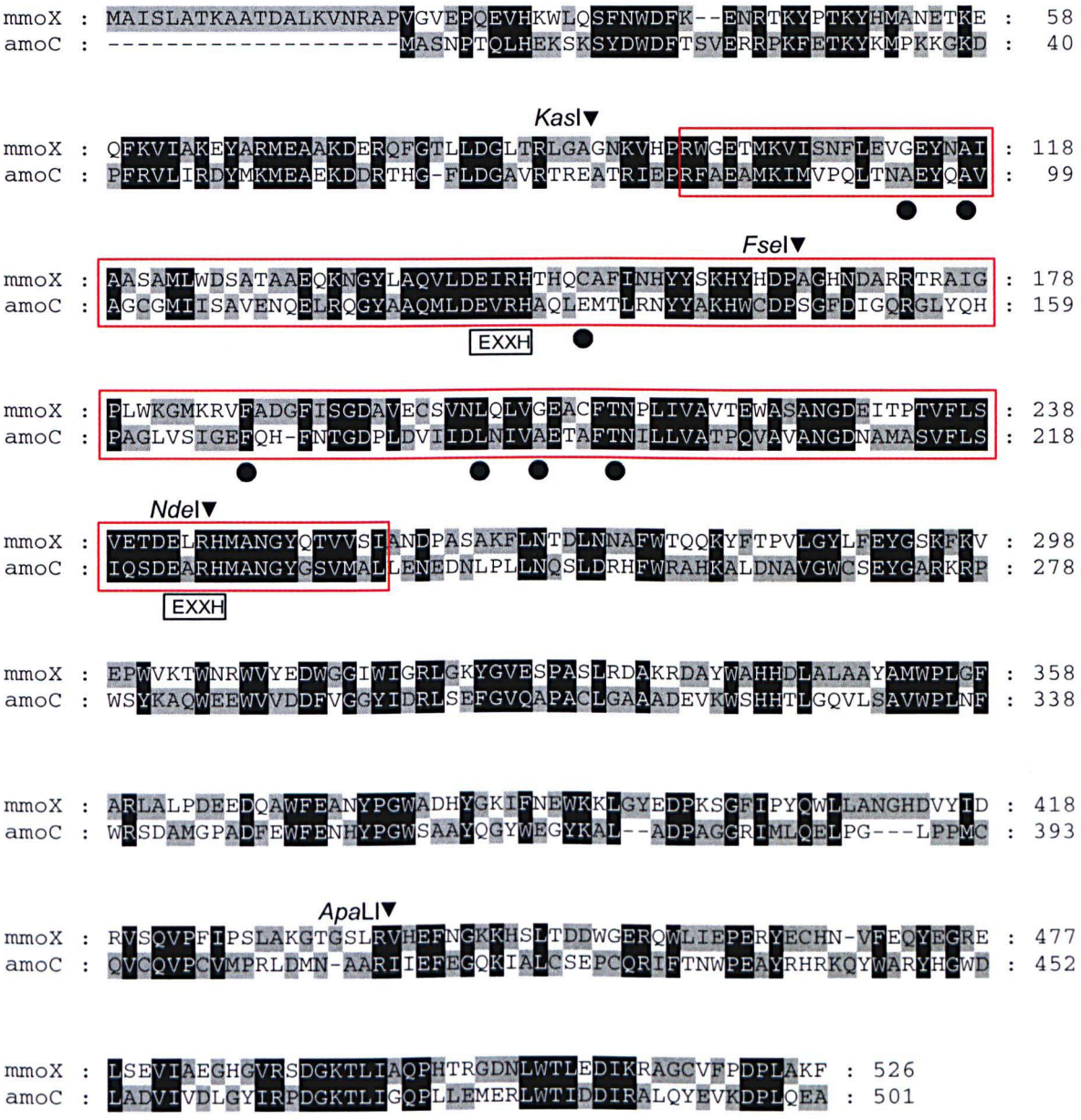


Figure 3.2 Alignment of the derived aminoacyl sequences of the *mmoX* and *amoC* genes from *Ms. trichosporium* OB3b and *Rhodococcus corallinus* B-276 (accession numbers CAA39068 and BAA07114, respectively). The putative active site region of the sMMO α-subunit is boxed in red. The position in the protein sequence corresponding to the cutting sites of the *Kas*I, *Fse*I, *Nde*I and *Apa*LI restriction enzymes is indicated. Dots are positioned underneath the alanine91$_{AmoC}$, alanine95$_{AmoC}$, cysteine151$_{MmoX}$, phenylalanine169$_{AmoC}$, leucine181$_{AmoC}$, alanine185$_{AmoC}$, threonine213$_{MmoX}$; the relevance of these residues are discussed in the text. The alignment is shaded according to conservation of property (GeneDoc, http://www.psc.edu/biomed/genedoc/).

Table 3.1 PCR primers used to amplify the *amoC* gene. The primers contained restriction enzyme sites (identified in the primer name, and underlined in the primer sequence), which were used to introduce the PCR fragments into the *mmoX* gene on pMD1. A suffix of 'F' or 'R' indicate that the sequence primes in a forward or reverse orientation to the *amoC* gene.

Primer name	Primer sequence (5' – 3')
amoC-KasI-F	CGG ACG GGC GCC GCC ACC AGG ATT GAG CCG
amoC-BglII-F	ACC AGG AGA TCT GGG AGG GCT ACA AGG CGC TCG
amoC-BglIIF-R	CCT CCC AGA TCT CCT GGT AGG CCG CGC TCC AG
amoC-FseI-F	CGA TCC GGC CGG CCT CGA CAT CGG TCA GC
amoC-FseI-R	TGT CGA GGC CGG CCG GAT CGC ACC AGT GCT TC
amoC-NdeI-F	GCC AGG CAT ATG GCC AAC GGG TAC GGC TC
amoC-ApaLI-R	GAA CTC GTG CAC CCG CGC GGC GTT CAT ATC

```
MmoX     : VIAKEYARMEAAKDERQFGTLLDGLTRLGAGNKVHPRWGETMKVISNFLEVGEYNAIAASAM
AmoC     : VLIRDYMKMEAEKDDRTHG-FLDGAVRTREATRIEPRFAEAMKIMVPQLTNAEYQAVAGCGM
primers  : ---------------------------LGAATRIEP-------------------------
ligated  : VIAKEYARMEAAKDERQFGTLLDGLTRLGAATRIEPRFAEAMKIMVPQLTNAEYQAVAGCGM



MmoX     : LWDSATAAEQKNGYLAQVLDEIRHTHQCAFINHYYSKHYHDPAGHNDARRTRAIGPLWKGMK
AmoC     : IISAVENQELRQGYAAQMLDEVRHAQLEMTLRNYYAKHWCDPSGFDIGQRGLYQHPAGLVSI
primers  : ----------------------------------KHWCDPAG-----------------
ligated  : IISAVENQELRQGYAAQMLDEVRHAQLEMTLRNYYAKHWCDPAGHNDARRTRAIGPLWKGMK
```

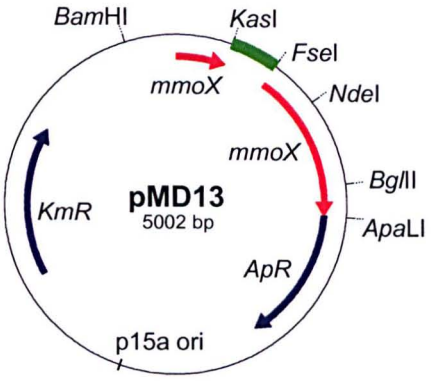
Figure 3.3 Example of the design of primers for a mutagenesis experiment. The example shown is for amoC-Kas1-F and amoC-FseI-R primers in the construction of 'Mutant 13'. An alignment of the derived partial MmoX and AmoC sequences is shown. The derived aminoacyl sequence of the primers is shown for illustrative purposes. The 'ligated' sequence corresponds to the derived aminoacyl sequence of the chimaeric *mmoX-amoC* gene of Mutant 13.

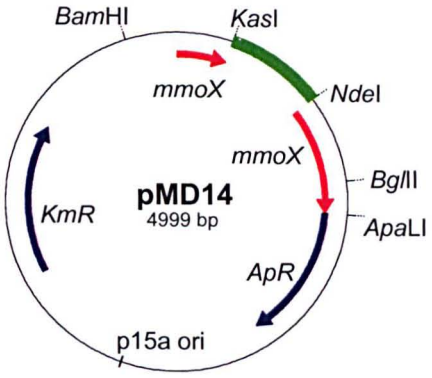
Table 3.2 sMMO mutants.

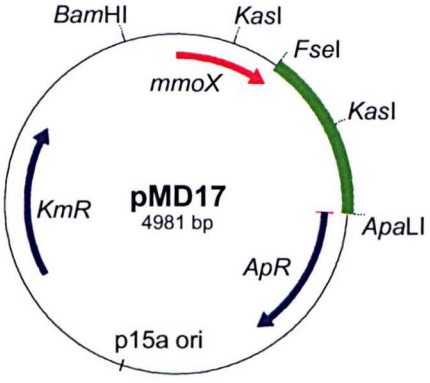
sMMO Mutant	Enzymes used	amino acids removed/substituted	no. amino acid changes[1]
Mutant 13	*Kas*I-*Fse*I	76/76	47
Mutant 14	*Kas*I-*Nde*I	155/154	97
Mutant 17	*Fse*I-*Apa*LI	272/265	173
Mutant 19	*Nde*I-*Apa*LI	192/186	123

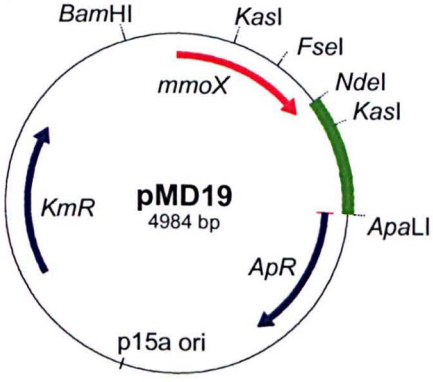
1. does not include gaps

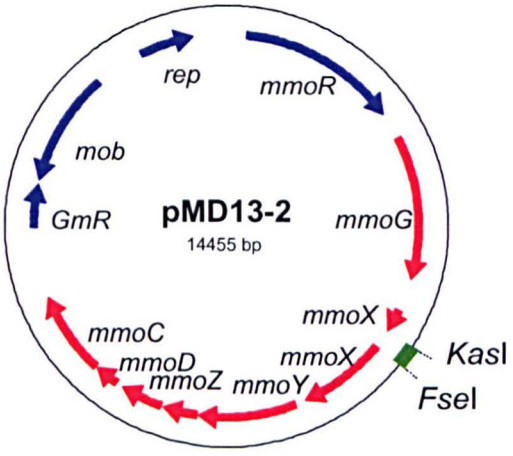
Figure 3.4 Plasmid maps of constructs created. The restriction enzyme used to incorporate *amoC* sequence (green) into the *mmoX* gene (red) is indicated. The pMD13, pMD14, pMD17 and pMD19 plasmids are the pMD1 plasmid containing *amoC* gene sequence. Plasmids pMD13-2, pMD14-2, pMD17-2 and pMD19-2 are pMD2 with the sequence between *Bam*HI/*Apa*LI replaced with the chimaeric sequence from the corresponding mutagenized pMD1 plasmid.

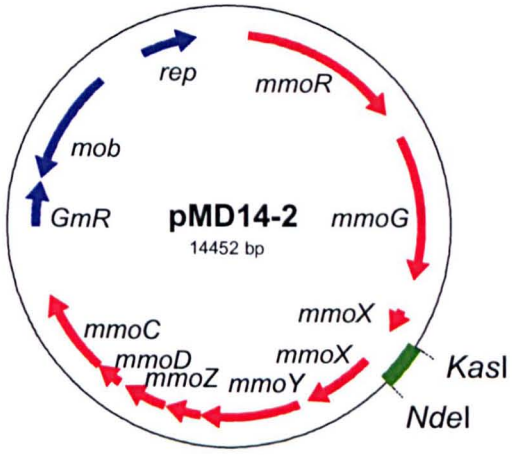

BamHI
KasI
FseI
NdeI
mmoX
mmoX
BglII
ApaLI
KmR
pMD13
5002 bp
ApR
p15a ori

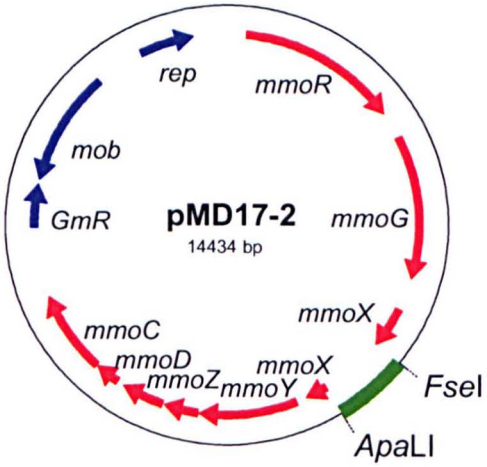

BamHI
KasI
mmoX
NdeI
mmoX
BglII
ApaLI
KmR
pMD14
4999 bp
ApR
p15a ori

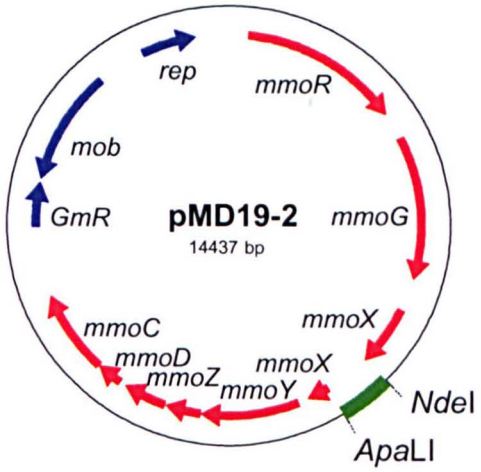

BamHI
KasI
FseI
mmoX
KasI
KmR
pMD17
4981 bp
ApaLI
ApR
p15a ori


BamHI
KasI
FseI
mmoX
NdeI
KasI
KmR
pMD19
4984 bp
ApaLI
ApR
p15a ori


rep
mmoR
mob
GmR
pMD13-2
14455 bp
mmoG
mmoX
mmoC
mmoD
mmoX
KasI
mmoZ
mmoY
FseI


rep
mmoR
mob
GmR
pMD14-2
14452 bp
mmoG
mmoX
mmoC
mmoD
mmoX
KasI
mmoZ
mmoY
NdeI


rep
mmoR
mob
GmR
pMD17-2
14434 bp
mmoG
mmoX
mmoC
mmoD
mmoX
FseI
mmoZ
mmoY
ApaLI


rep
mmoR
mob
GmR
pMD19-2
14437 bp
mmoG
mmoX
mmoC
mmoD
mmoX
NdeI
mmoZ
mmoY
ApaLI

3.4 Expression of mutant sMMO in *Ms. trichosporium* Mutant F

Ms. trichosporium Mutant F was transformed with pMD2, pMD13-2, pMD14-2, pMD17-2 and pMD19-2 plasmids by conjugation as described in Section 2.7.1. Transconjugants were selected on NMS agar containing 5 μg ml^{-1} gentamycin and 12.5 μg ml^{-1} kanamycin, and colonies appeared after two weeks incubation with methane in air (Section 2.3.2). Five colonies for each conjugation were subcultured onto NMS agar containing antibiotic selection and grown for three weeks on methane. For a sMMO naphthalene assay, the colonies were patched onto NMS agar plates containing copper and plates without added copper. Gentamycin and kanamycin selection was included and the plates incubated with methane for four weeks. The naphthalene oxidation sMMO assay was performed as described in section 2.16. *Ms. trichosporium* Mutant F carrying pMD2 gave a strong positive reaction. There was no colour change for *Ms. trichosporium* Mutant F carrying pMD13-2, pMD14-2, pMD17-2 and pMD19-2 (results not shown).

SDS-PAGE analysis was performed by removing biomass from a NMS plate containing confluent growth of the transconjugants of *Ms. trichosporium* Mutant F. Untransformed *Ms. trichosporium* Mutant F was included as a negative control and was grown as for the transconjugants, except gentamycin was not included in the growth medium. Sample preparation and SDS-PAGE was performed as described in Section 2.17. The sMMO hydroxylase subunits were visible from *Ms. trichosporium* Mutant F carrying pMD2. No sMMO proteins were visible in *Ms. trichosporium* Mutant F, or transconjugants of the pMD13-2, pMD14-2, pMD17-2 and pMD19-2 plasmids (Figure 3.5). These results suggested that the mutant enzymes were unstable (Smith *et al.*, 2002).

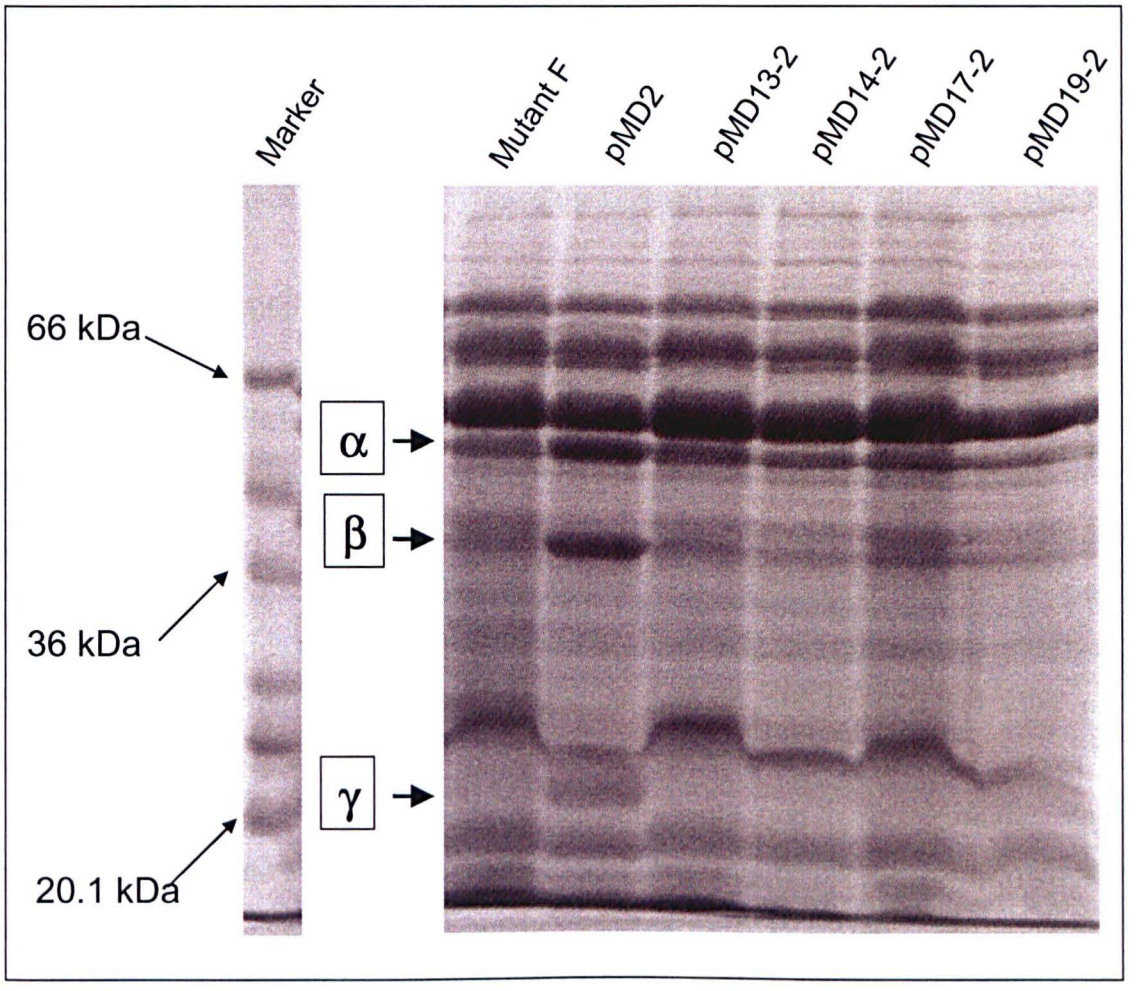


Figure 3.5 SDS-PAGE analysis of *Ms. trichosporium* Mutant F transconjugants grown under sMMO expressing conditions. 'Mutant F' is untransformed *Ms. trichosporium* Mutant F; other lanes are labelled according to the plasmid carried by the *Ms. trichosporium* Mutant F. The positions of the α–, β– and γ–subunits of the sMMO hydroxylase are indicated.

3.5 Discussion

This study describes the first attempt at combinatorial mutagenesis of the sMMO. The gene encoding the α-subunit of the AMO enzyme was used as parent sequence since, at the time, the AMO was the closest known relative of the sMMO and has the desirable characteristic of performing stereospecific epoxidations (Gallagher *et al.*, 1997). Therefore, an sMMO enzyme containing a hybrid of the AMO and sMMO core regions may have unique and potentially industrially relevant catalytic properties.

A major problem with the current *Ms. trichosporium* Mutant F expression system is that mutagenized sMMOs are frequently unstable. Smith and co-workers constructed four sMMO mutants (as discussed in Chapter 1) and only the Thr-213-Ser mutant sMMO enzyme, which is a conservative mutation, was sufficiently stable to be purified. The other mutagenized enzymes were not detectable by SDS-PAGE and could not be purified. The C151E mutant was particularly interesting since naphthalene oxidation activity of whole cells expressing the mutant was observed; however the enzyme was highly unstable and could not be detected after cell disruption.

A similar problem with enzyme stability/expression occurred in this study since no naphthalene activity with whole cells expressing the chimaeric enzymes was observed, and the enzymes could not be detected by SDS-PAGE. The possibility exists that there was low-level expression of the chimaeras comparable to that of the C151E mutant, but naphthalene oxidation was absent because the chimaeric enzymes had lost the ability to oxidise naphthalene.

The fact that the enzymes were not visible by SDS-PAGE suggests that they were unstable. It is most likely that the chimaeric α–subunit does not fold properly, although it is also possible that it no longer binds the active site iron atoms which could render the enzyme unstable (Pfeil *et al.*, 1993). Interactions between the other hydroxylase subunits (Rosenzweig *et al.*, 1993), or even between the coupling protein or reductase (Fox *et al.*, 1991), may have been disrupted. Another possible problem may be the presence of codons introduced from the *amoC* gene that were not recognized in *Ms. trichosporium*. The *amoC* codons introduced into the chimaeras were analysed and examples of all could be found in at least one *Ms. trichosporium* orf sequence. It should be noted that the CTA and CCA codons (leucine and proline,

respectively) were introduced into the mutants, and these codons have so far only been found in the *hppk* gene (present upstream of copy 2 of *pmoA*, unpublished) and the *clpX* gene (this study, Chapter 6), and there is no direct evidence that these genes are actively transcribed in *Ms. trichosporium* OB3b.[3]

The extent of the mutagenesis performed here was more drastic than the site-directed mutagenesis performed by Smith and co-workers, and therefore it is not necessarily surprising that the chimaeric mutants were unstable. Like the unstable C151E mutagenized enzyme, mutants "13" and "14" included the C151E substitution in addition to a further 46 and 96 changes, respectively. The more extensive mutagenesis was performed based on the hypothesis that the substitution of the cysteine-151 to glutamate requires compensatory changes elsewhere in the protein molecule. Therefore, more extensive mutagenesis may lead to better folding than a single residue change. Perhaps the best experiment to test this hypothesis would be to replace the entire *mmoX* gene with the *amoC* gene; however the cloning reactions that would be required to perform this experiment in this sMMO expression system are numerous and awkward, and it was not performed. The potential loss of proper intermolecular interaction between the α-subunit of the epoxygenase and the β– and γ–subunits of the hydroxylase would still exist.

The idea of preserving intra-molecular interactions when performing combinatorial mutagenesis is the basis of the recently developed 'SCHEMA' analysis (Voigt *et al.*, 2002). SCHEMA is a computational algorithm that determines the interactions between residues in a folded protein and then calculates the amount of disruption in the creation of a chimaera. This information predicts that properly folded chimaeras result when protein blocks containing a high degree of interaction are left undisturbed. The authors applied the algorithm to several proteins that had previously been the subject of random chimaeragenesis studies, and found that the active mutants nearly always contained crossovers in regions of low interaction. This information can be used to direct the location of crossovers between parent sequences and reduce the large number of inactive mutants normally generated by random shuffling (Bernhardt, 2004). This approach was used to generate chimaeras of two bacterial cytochrome P450 enzymes that shared 63 % amino acid identity (Otey *et al.*, 2004). 17 mutant enzymes were generated, 14 of which were properly folded and 13

[3] One of the 26 leucines in *mmoY* is also encoded by CTA, but it should be noted that there are apparent sequencing errors in the *mmoY* gene (Elango *et al.*, 1997).

contained the haem correctly bound. The 13 functional enzymes demonstrated functional diversity and three of them could oxidise a substrate not oxidised by either parent. This approach requires X-ray crystal structure data for the enzyme, which is available for the *Ms. trichosporium* OB3b sMMO hydroxylase (Elango *et al.*, 1997), and could be applied for future sMMO mutagenesis experiments.

More recently, a butane monooxygenase (BMO) from "*Pseudomonas butanovora*" has been characterized that is far more similar to the sMMO than is AMO (Sluis *et al.*, 2002). The derived sequences of the α-subunits of sMMO and BMO share 65 % amino acid identity, and the gene arrangement of the operon (*bmoXYBZDC*) encoding BMO is identical to that of *mmo*. The BMO enzyme does not oxidise methane and has less broad substrate specificity than sMMO. Since the sMMO is more closely related to the BMO than the AMO, the mutants generated would have a greater chance of being active than those created in this study. Combinatorial mutagenesis between the sMMO and BMO genes applied in a manner similar to that used by Otey and colleagues (2004) for the cytochromes P450 may result in active mutants that could provide valuable information about the factors affecting the substrate specificities of these enzymes. Also, like the cytochrome P450 mutants, unexpected characteristics may also be generated.

Additional perspectives

The characteristics of the C151E mutant described by Smith and co-workers (2002) are intriguing. The fact that the whole cell naphthalene assay was positive indicates that the enzyme was active; however the enzyme was present at levels too low to be purified. The low levels of the enzyme could be explained by either attenuation of transcription or proteolysis. The possibility that the mRNA transcript was less stable is also possible. Benzamidine is routinely added to cell free extract preparations for sMMO analyses, and the authors attempted to further reduce proteolysis of the C151E mutant by adding the protease inhibitor PMSF; however there was no noticeable improvement.

It seems most likely that enhanced susceptibility to proteolysis is responsible for the low level of expression of the mutagenized enzymes. Smith and co-workers apparently did not attempt the expression of mutant sMMOs in *Ms. trichosporium* mutant F at lower growth temperatures, which is a strategy that often improves

expression of recombinant proteins in *E. coli* (Enfors, 1992). Expression in a protease-deficient host can improve expression of some unstable proteins and there is the possibility that improved expression of mutagenized sMMO could be obtained in a protease-deficient strain of *Ms. trichosporium* mutant F. For this reason, the disruption of the gene encoding the Lon protease of *Ms. trichosporium* Mutant F was attempted in this study (Chapter 6).

Chapter 4

PCR Amplification of *mmoX* Genes from Environmental Samples

4.1 Introduction

Unlike the pMMO which has been detected in nearly all methanotroph isolates, the presence of sMMO amongst methanotrophs varies between strains. Until recently, the sMMO was believed to be mainly found in α–*Proteobacteria* and *Methylococcus capsulatus*. However, recently it has been discovered that some *Methylomonas* spp. possess an sMMO (Auman *et al.*, 2000; Shigematsu *et al.*, 1999) and that the presence of this enzyme is probably no more a characteristic of α–*Proteobacteria* methanotrophs than it is of γ–*Proteobacteria* methanotrophs.

Because the presence of sMMO is less common amongst methanotrophs than the pMMO, the *mmoX* gene is less useful than the *pmoA* gene as a marker for methanotrophs. Few extensive PCR surveys of *mmoX* genes in environmental samples have been performed. The limited studies include the detection of *mmoX* genes in a soil and groundwater sample (Miguez *et al.*, 1997), and small clone libraries from Sphagnum peat (Dedysh *et al.*, 1998), a moorland peat soil (McDonald *et al.*, 1995; Morris *et al.*, 2002) and the roots of rice plants (Horz *et al.*, 2001). More extensive surveys include an analysis of the *mmoX* genes from Lake Washington sediment (Auman & Lidstrom, 2002), from an aquifer (Newby *et al.*, 2004), and from a cave water sample (this study) (Hutchens *et al.*, 2004). Many methanotroph isolates have been screened by PCR for *mmoX* and some of these studies have been published (Auman *et al.*, 2000; Heyer *et al.*, 2002).

The objective of this work was to try to amplify novel *mmoX* genes from environmental samples. The ultimate goal was to obtain an *mmoX* gene which encoded an MmoX with a significantly different amino acid sequence to those of *Methylosinus trichosporium* OB3b and *Methylococcus capsulatus* (Bath), and particularly with differences in aminoacyl residues that might affect the catalytic properties of the enzyme. This "unique environmental *mmoX* gene" could then be "stitched" into the *Methylosinus trichosporium* OB3b *mmoXYBZDC* gene cluster as described in Chapter 3.

4.2 Design of *mmoX* PCR primers

For this study it was necessary to use *mmoX* PCR primers that amplify the complete active site region of the sMMO. At the beginning of this study, three *mmoX* PCR primer pairs were published: one that amplifies a 350 bp segment of the gene (Fuse *et al.*, 1998), another amplifying a 369 bp segment (Miguez *et al.*, 1997) and a third that amplifies a 521 bp fragment near the 3' end of the gene (McDonald *et al.*, 1995). None of these primer sets were suitable and therefore new primers were designed in this study.

The PCR primers were designed by aligning all the *mmoX* sequences available at the time (Table 4.1) and searching for regions that had conserved amino acid sequences and showed a high level of conservation in the nucleotide sequence. Primers were chosen that hybridized outside the segment of the gene that encodes the proposed core region of MmoX (the sMMO α-subunit) (Figure 3.2) such that *mmoX* gene amplified by PCR using the primers could potentially be used to make a chimaeric sMMO (Chapter 3).

Two forward and three reverse primers were designed (Table 4.2). The six combinations of primer pairs were tested in the PCR using genomic DNA from *Methylococcus capsulatus* (Bath), *Methylosinus trichosporium* OB3b, *Methylocystis* sp. strain M, and "*Methylothermus*" strain HB as a negative control. The optimal concentration of DNA template was determined by comparing the PCR amplification of the 16S rRNA gene from various amounts of template DNA using the 27F/ 1492R primers (Lane, 1991). All the *mmoX* targeted reactions were performed using 1 x PCR buffer (Invitrogen), 200 μM dNTPs, 3 mM $MgCl_2$ in a volume of 50 μl, and cycled 30 times using a denaturing temperature of 92 °C (Section 2.9). The primer pairs were tested using 2 pmole (40 nM) and 10 pmole (200 nM) of primer per PCR. Annealing temperatures of 50 °C, 55 °C, 60 °C and in some instances 62 °C were tested.

The primer pair consisting of the oligonucleotides annealing to positions 187-206 bp (termed 'mmoX206F') and 886-905 bp (termed 'mmoX886R') (Table 4.2) resulted in good amplification of the correct size product (719 bp) from the test DNA. 10 pmole of each primer and an annealing temperature of 60 °C were found to be optimal.

Table 4.1. *mmoX* sequences in Genbank (November 2000). These sequences were used in this study to design the *mmoX* PCR primers.

Organism	Genbank accession	Size (bp)
"Uncultured peat bog bacterium 24"	AF004555	524
"Uncultured groundwater bacterium 1"	AF155586	451
"Uncultured groundwater bacterium 2"	AF155587	451
Methylococcus capsulatus (Bath)	M90050	7188
"Methanotroph FL20"	AF183838	456
"Methanotroph FL-DIKO"	AF183839	447
Methylomonas sp. KSPIII	AB025021	7531
Methylomonas sp. KSWIII	AB025022	8107
"Unidentified alpha proteobacterium LR1"	UPR18440	699
Methylosinus sp. LW3	AY007287	1235
Methylosinus sp. LW4	AY007288	1235
Methylosinus sp. LW8	AY007289	1235
Methylomonas sp. LW13	AY007290	1235
Methylomonas sp. LW15	AY007291	1235
"Methanotroph M5"	AF183844	462
"Methanotroph M8"	AF183845	456
"Methanotroph M11"	AF183846	474
"Methanotroph M13"	AF183847	474
Methylosinus trichosporium OB3b	X55394	6042
Methylocella palustris	AF004554	524
Methylosinus sp. PW1	AY007292	1235
Methylocystis sp. M	MSU81594	5873
"Methanotroph WI-11"	AF183840	456
"Methanotroph WI-13"	AF183841	444
"Methanotroph WI-14"	AF183842	453
"Methanotroph WI-15"	AF183843	456
Methylocystis sp. WI14	AF153282	5253

Table 4.2. Degenerate PCR primers designed to anneal to all known *mmoX* genes. The IUPAC symbols used are A = adenine; C = cytosine; G = guanine; T = thymine; R = A or G; Y = C or T; S = C or G; B = C or G or T; M = A or C; N = A or C or G or T. The numbering for the primer annealing position corresponds to the *mmoX* gene of *Ms. trichosporium* OB3b.

Position	orientation	sequence 5' – 3'
187-206	forward	ATC GCB AAR GAA TAY GCS CG
219-238	forward	AAG GAC GAR MGY CAR TTC GG
815-834	reverse	TGS GTC CAG AAR GCG TTG TT
886-905	reverse	ACC CAN GGC TCG ACY TTG AA
1382-1401	reverse	CAC TCR TAR CGC TCS GGY TC

Subsequent to the development of these primers, a study was published in which new *mmoX* PCR primers had been designed (Auman *et al.*, 2000). These primers, mmoXA (ACCAAGGARCARTTCAAG) and mmoXB (TGGCACTCRTARCGCTC) amplify a 1,230 bp region of the *mmoX* gene. These primers were used in this study alongside the mmoX206F/ mmoX886R primers.

An attempt was made to design primers that annealed to *mmoX* and the corresponding genes of related monooxygenases (such as alkene monooxygenase of *Rhodococcus corallinus* (Saeki & Furuhashi, 1994) and toluene-*o*-monooxygenase from *Burkholderia cepacia* G4), but the genes were too divergent to align the nucleotide sequences. A set of primers was designed that was specific for both the *amoC* gene of the alkene monooxygenase (Saeki & Furuhashi, 1994) and *mmoX*. The primers 'RHMANGY-forward' (5' – GY CAY ATG GCB AAY GGB TA – 3') and 'DGKTLI-reverse' (5' – GAT SAR SGT YTT TCC RTC – 3') were tested under various PCR conditions with *Methylococcus capsulatus* (Bath) and *Methylosinus trichosporium* OB3b genomic DNA, but a PCR product could not be obtained.

4.3 Amplification of *mmoX* from environmental samples with mmoXA/ mmoXB PCR primers

4.3.1 Bath Spring

An attempt was made to amplify the *mmoX* gene using DNA isolated from the Roman Bath (Bath, England). The DNA was isolated by Stefan Radajewski (unpublished data) using the bead-beating extraction method (Morris *et al.*, 2002). The PCR of the *mmoX* gene using the primers mmoXA/ mmoXB was performed as described in the original paper (Auman *et al.*, 2000). The PCR was analysed by agarose gel electrophoresis, and a weak product of the correct size was visible. The fragment was gel purified (Section 2.8.9), ligated into the pCR2.1 TA-cloning vector (Section 2.8.7) and an *mmoX* gene library was constructed in *E. coli* TOP10 (Section 2.7.2).

A library of 100 clones was made and analysed by RFLP by digestion with *Eco*RI and double digestion with *Eco*RI/*Ava*II. The pCR2.1 plasmid contains *Eco*RI enzyme cutting sites on either side of the cloning site and therefore *Eco*RI digests result in the release of the cloned DNA fragment from the vector. Analysis by agarose gel electrophoresis showed that many of the cloned fragments were of slightly different sizes. There was a possibility that these were *mmoX* genes with insertions or deletions in the gene, or they could be spurious PCR products that had been cloned. Ten clones were sequenced using the M13R sequencing primer and the results are summarized in Table 4.3A. Only two of the ten clones contained an *mmoX* gene. The low quality of the library is probably a reflection of the poor amplification of the *mmoX* in the PCR. Further analysis of the Roman Bath sample was discontinued.

4.3.2 Moorhouse Peat

The *mmoX* gene was amplified from DNA extracted during a previous study of the Moorhouse peat in the north of England (McDonald *et al.*, 1999) and had been stored at - 80 °C. The stock DNA from the 11 cm depth sample was diluted tenfold with water and 1 µl was used as template in PCR with the mmoXA/ mmoXB primers (Auman *et al.*, 2000). A library of 24 clones in *E. coli* TOP10 was made. The plasmids were analysed by digestion with *Eco*RI and the presence of an insert of the correct size was verified by agarose gel electrophoresis. The plasmids were analysed

by double digestion with *Eco*RI/*Ava*II and *Eco*RI/*Xho*I. Four OTUs were identified and the inserts were sequenced with the M13R primer.

An analysis of the sequences indicated that the genes were very similar to the *mmoX* gene of representative species of *Methylosinus* (Table 4.3B). This is consistent with the findings of previous studies that indicate that *Methylosinus*-like organisms are abundant in this blanket bog peat (McDonald & Murrell, 1997b; McDonald *et al.*, 1999). Further analysis of *mmoX* gene diversity in Moorhouse peat DNA was not undertaken since it was considered less likely to yield novel *mmoX* genes than other environmental samples.

4.3.3 Movile Cave

The diversity of methanotrophs in samples taken from the Movile Cave system (described in Section 4.5; see also Hutchens *et al.* (2004) in appendix) was being investigated by Elena Hutchens and Stefan Radajewski. A SIP experiment with $^{13}CH_4$ had been performed and the analysis of SSU rRNA, *pmoA* and *mxaF* genes indicated that unique and as yet uncultivated methanotrophs were present in this unusual ecosystem. An analysis of these genes indicated that methanotrophs of the γ-*Proteobacteria* were dominant in the sample, but there was some evidence of β-*Proteobacteria* incorporating the $^{13}CH_4$. There were also several unique *pmoA* and *mxaF* genes detected in the PCR libraries generated from the ^{13}C-DNA. Therefore, an *mmoX* library was made using the ^{13}C-DNA in order to potentially detect novel *mmoX* genes.

A library of 67 *mmoX* clones was generated using the mmoXA/ mmoXB primers, essentially as described in Sections 4.3.1 and 4.3.2. Subsequently, a library was made using the mmoX206F/ mmoX886R primers, and it was these data that were published with the other data from the Movile Cave SIP experiment (Hutchens *et al.*, 2004); the results using the mmoXA/mmoXB primers and the mmoX206F/mmoX886R primers were similar, and a more in depth analysis is provided for the latter (Section 4.4).

Table 4.3 Result of BLASTX searches (Altschul *et al.*, 1997) of *mmoX* clones (mmoXA/mmoXB PCR primers).

A. Bath spring

Clone	Protein Name closest BLASTX match	Accession no.	Source	% Identity	BLAST Score/Expect
Bs1	Amide hydrolase	AAL41943	*Agrobacterium tumefaciens*	63 %	150/ $5e^{-35}$
Bs3	hypothetical protein	NP_415342	*Escherichia coli* K12	98 %	360/ $1e^{-99}$
Bs4	MmoX subunit, sMMO	AAF01268	*Methylosinus* sp. WI14	100 %	246/ $8e^{-65}$
Bs5	DNA helicase	NP_662505	*Chlorobium tepidum*	46 %	164/ $4e^{-39}$
Bs6	Hypothetical protein	NP_960489	*Mycobacterium avium*	46 %	40.4/ 0.0054
Bs7	Aspartate aminotransferase	NP_987511	*Methanococcus maripaludis*	45 %	49.3/ $1e^{-04}$
Bs8	MmoX subunit, sMMO	AAF01268	*Methylosinus* sp. WI14	100 %	246/ $8e^{-65}$
Bs9	Penicillin binding protein	ZP_00090564	*Azotobacter vinelandii*	43 %	141/ $3e^{-32}$
Bs10	Transcriptional regulator	ZP_00086051	*Pseudomonas fluorescens*	85 %	193/ $1e^{-64}$

B. Moorhouse peat

Clone	Protein Name closest BLASTX match	Accession no.	Source	% Identity	BLAST Score/Expect
Mhp1	MmoX subunit, sMMO	CAA39068	*Methylosinus trichosporium*	94 %	205/ $2e^{-55}$
Mhp2	MmoX subunit, sMMO	AAG31818	*Methylosinus* sp. LW8	94 %	233/ $1e^{-60}$
Mhp3	MmoX subunit, sMMO	AAG31817	*Methylosinus* sp. LW4	90 %	231/ $5e^{-60}$
Mhp4	MmoX subunit, sMMO	CAA39068	*Methylosinus trichosporium*	96 %	231/ $4e^{-60}$

C. Movile Cave, ^{13}C-DNA:

Clone	Protein Name closest BLASTX match	Accession no.	Source	% Identity	BLAST Score/Expect
Mc1	MmoX subunit, sMMO	AAG31820	*Methylomonas* sp. LW15	96 %	447/ $1e^{-124}$
Mc3	MmoX subunit, sMMO	AAG31820	*Methylomonas* sp. LW15	96 %	451/ $1e^{-126}$
Mc6	MmoX subunit, sMMO	AAG31820	*Methylomonas* sp. LW15	91 %	438/ $1e^{-122}$
Mc7	MmoX subunit, sMMO	AAG31820	*Methylomonas* sp. LW15	93 %	449/ $1e^{-125}$
Mc11	MmoX subunit, sMMO	AAG31816	*Methylosinus* sp. LW3	96 %	455/ $1e^{-127}$
Mc16	MmoX subunit, sMMO	AAG31820	*Methylomonas* sp. LW15	93 %	449/ $1e^{-125}$
Mc17	MmoX subunit, sMMO	AAG31820	*Methylomonas* sp. LW15	94 %	452/ $1e^{-126}$
Mc22	MmoX subunit, sMMO	AAG31820	*Methylomonas* sp. LW15	93 %	440/ $1e^{-125}$
Mc40	MmoX subunit, sMMO	AAG31820	*Methylomonas* sp. LW15	99 %	468/ $1e^{-131}$
Mc65	MmoX subunit, sMMO	AAG31820	*Methylomonas* sp. LW15	91 %	438/ $1e^{-122}$

The 67 *mmoX* clones were analysed by RFLP with *Eco*RI and *Eco*RI/ *Rsa*I. The clones were clustered into OTUs according to their banding patterns with *Eco*RI/ *Rsa*I, and representatives were sequenced with the M13F and M13R sequencing primers. The sequencing reactions with M13F and M13R did not generate overlapping sequence data of the 1,230 bp *mmoX* fragments. Therefore the complete sequence of these clones was not obtained. The results of BLASTX searches with the sequence data corresponding to the 5' end of the *mmoX* gene fragment are presented in Table 4.3C.

70 % of the *mmoX* clones were very similar to the *mmoX* of *Methylomonas* species. There were three *mmoX* clones, including mc11 (Table 4.3C), that had sequences which were very similar to the *mmoX* from representative species of *Methylosinus*. Four of the clones, including mc6 & mc65, were unusual and their sequences did not group closely with the *mmoX* sequences of any characterized methanotrophs; analysis of these *mmoX* genes showed that they clustered with the mvs13.14 sequence (Genbank accession AAR04306) from the mmoX206F/ mmoX886R amplified *mmoX* PCR library (Section 4.4).

4.4 PCR amplification of *mmoX* from Movile Cave DNA with mmoX206F and mmoX886R PCR primers

An *mmoX* PCR library was generated with primers mmoX206F/ mmoX886R using the ^{13}C-DNA from the Movile Cave SIP experiment as template (Hutchens *et al.*, 2004). This library was constructed to compare the types of *mmoX* genes retrieved with those obtained using the mmoXA/ mmoXB primers (Section 4.3.3). The analysis indicated that many of the same sequence types were obtained, and the proportion of clones corresponding to each OTU was similar in each library. Unlike the 1,230 bp *mmoX* PCR fragment obtained with the mmoXA/ mmoXB primers, which required a third sequencing reaction with an internal primer (Section 4.3.3), the cloned 719 bp *mmoX* fragments obtained with the mmoX206F/ mmoX886R primers were completely sequenced using the M13F and M13R primers.

A library of 50 *mmoX* clones was analysed. OTUs were assigned on the basis of an RFLP of the clones with *Eco*RI/ *Ava*II and *Eco*RI/ *Rsa*I. Fifteen OTUs were identified and representatives were sequenced. Phylogenetic analysis was performed using the corresponding fragment of the BmoX sequence of the butane

monooxygenase of *'Pseudomonas butanovora'* (Sluis *et al.*, 2002) as the outgroup (Figure 4.1). The results indicated that MmoX sequences from ten of the OTUs, representing 39 clones, grouped near to the MmoX sequences of *Methylomonas* species. Two OTUs (mvs13.13 & mvs13.4), corresponding to seven clones, had derived MmoX sequences that grouped with those of *Methylosinus* spp. and *Methylocystis* spp. The MmoX sequence of mvs13.15 grouped with that of MmoX of *Methylococcus capsulatus*. The derived MmoX sequences of two OTUs (mvs13.7 & mvs13.14) did not group with the MmoX of any characterized methanotrophs. An alignment of the derived MmoX sequences of mvs13.7 and mvs13.14 with those of *Methylosinus trichosporium* OB3b and *Methylococcus capsulatus* (Bath) is shown in Figure 4.2. The derived mvs13.7 MmoX sequence contains 21 aminoacyl residue differences from those of the *Methylosinus trichosporium* OB3b and *Methylococcus capsulatus* (Bath) MmoX sequences; however many of the differences are conservative changes and none are residues predicted to be critical for catalysis (Section 1.5.1).

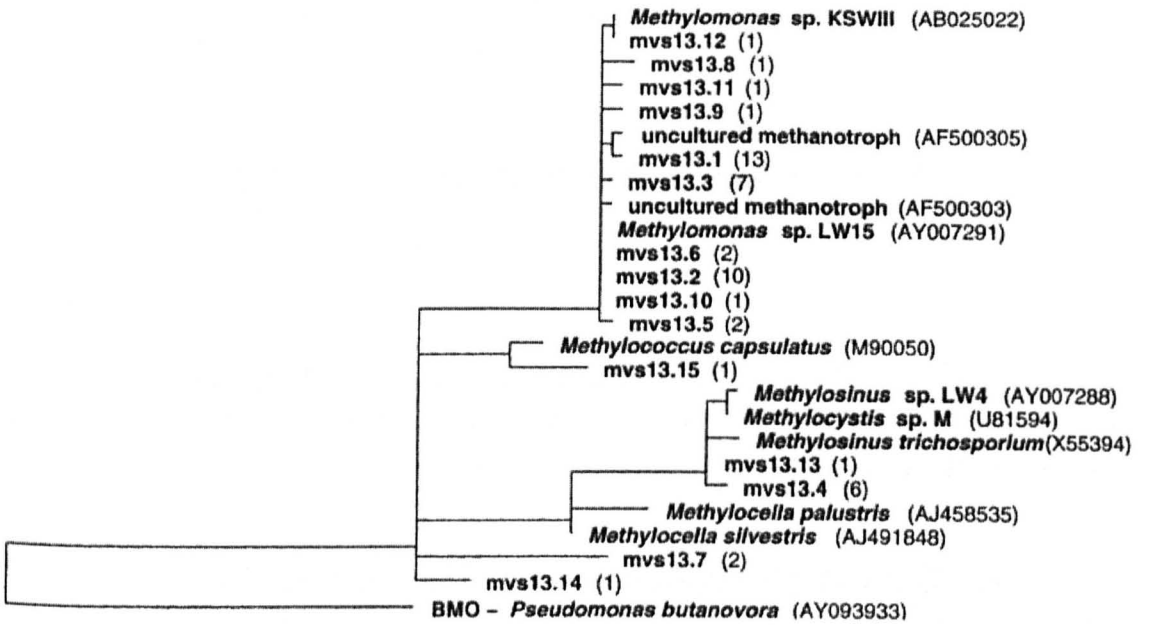


Figure 4.1 A maximum likelihood tree of derived MmoX sequences detected in the ^{13}C-DNA fraction. The tree was constructed using the derived amino acid sequence of the butane monooxygenase (BMO) of '*Pseudomonas butanovora*' as an outgroup and a filter that excluded the primer sequences. The number of clones assigned to each OTU by RFLP analysis is shown in parentheses. Taken from Hutchens *et al.* (2004).

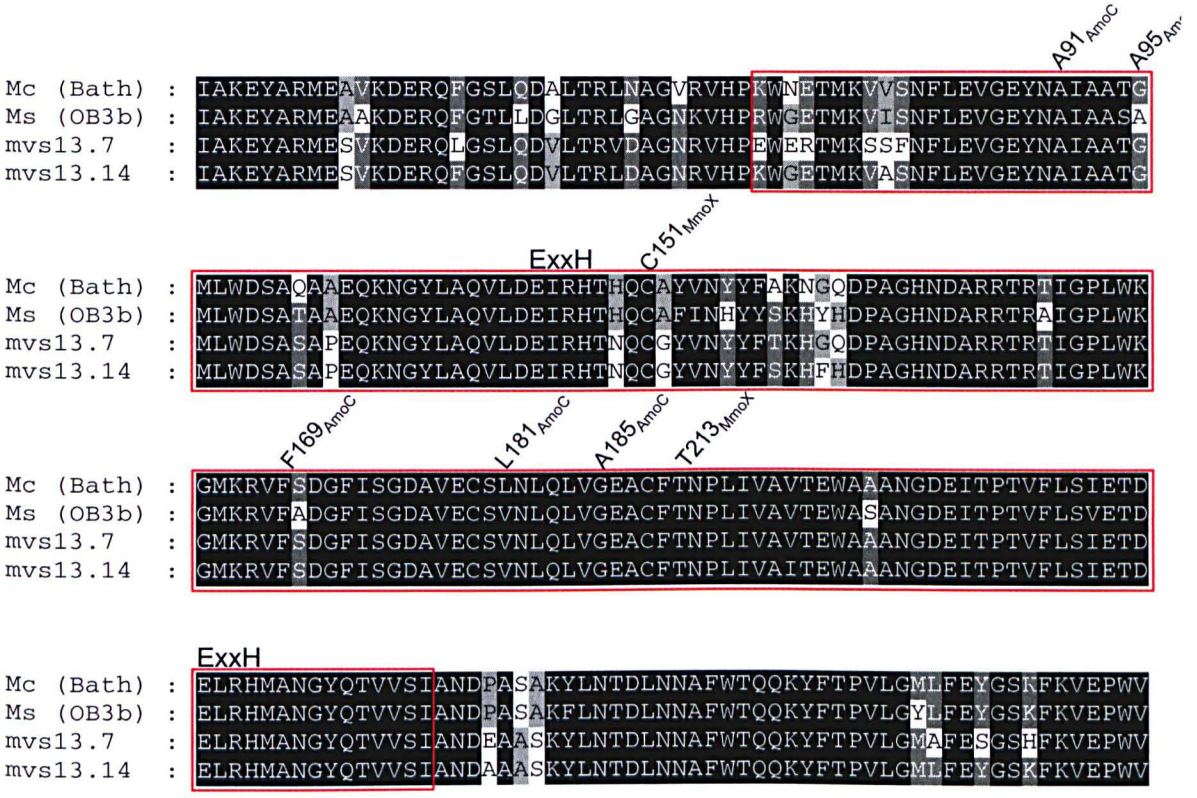


Figure 4.2 Alignment of the derived MmoX sequences of clones mvs13.7 and mvs13.14 with the sequences of *Methylococcus capsulatus* (Bath) ['Mc (Bath)'] and *Methylosinus trichosporium* OB3b ['Ms (OB3b)']. The putative active site region is boxed in red (Elango *et al.*, 1997). The positions of the iron-ligating glutamate and histidine residues (ExxH) are indicated. Aminoacyl residues of interest with respect to catalysis and stereoselectivity of the enzyme are indicated (refer to Chapter 3 for further information).

4.5 Discussion

Like the *pmoA* gene of the particulate methane monooxygenase (pMMO), the *mmoX* is an important functional gene marker for methane oxidisers (McDonald *et al.*, 1995). Although the pMMO is common to nearly all characterized methanotrophs and the sMMO is relatively poorly distributed, the diversity of *mmoX* sequences is valuable when studying the ecology of methanotrophs since it encodes a polypeptide of a key methanotroph enzyme and reflects the catalytic diversity of the methane-oxidizing population. Moreover, the recent sequencing of *mmoX* genes from numerous methanotroph isolates has increased the value of this gene as a phylogenetic marker (Auman *et al.*, 2000; Heyer *et al.*, 2002).

In this study, two sets of *mmoX* PCR primers were tested on DNA from environmental samples. One primer pair, mmoX206F & mmoX886R, was designed in this study and amplified a 719 bp product that spans the region encoding the active site of the sMMO enzyme (Rosenzweig *et al.*, 1993). The other pair had been used previously to amplify *mmoX* genes from Lake Washington sediment (Auman & Lidstrom, 2002) and numerous methanotroph isolates (Auman *et al.*, 2000; Heyer *et al.*, 2002). In this study, the primers were used to amplify *mmoX* from DNA isolated from peat and from ^{13}C-DNA from a SIP experiment with cave water.

The entrance to Movile Cave, situated in Romania, is an artificial shaft that was dug for geological purposes in 1986 and uncovered a previously sealed groundwater ecosystem (Sarbu *et al.*, 1996). The ecosystem of the cave is fuelled by *in situ* chemoautotrophic metabolism that sustains a unique community of unusual organisms, including 33 endemic species of terrestrial and aquatic invertebrates (Sarbu & Kane, 1995; Sarbu *et al.*, 1996). The microbial ecology of Movile Cave is relatively poorly characterized, but autotrophic sulfur oxidising *Thiobacillus* and *Beggiatoa* species have been isolated (Sarbu *et al.*, 1994; Vlasceanu *et al.*, 1997) and the activity and diversity of methanotrophs has been determined (this study; Hutchens *et al.* (2004)), and both groups of organisms probably contribute to primary production in the system.

A DNA-SIP experiment using $^{13}CH_4$ was performed using Movile Cave water which identified bacteria related to the methanotroph genera *Methylomonas*, *Methylococcus* and *Methylosinus/Methylocystis* (Hutchens *et al.*, 2004). The *mmoX* library data obtained in this study complemented the DNA-SIP experiment in which

libraries of 16S rRNA, *pmoA* and *mxaF* genes were also analysed. RFLP and sequence analysis of *mmoX* clones indicated that *mmoX* genes related to those of *Methylomonas* spp. were the most abundant in the library. 16S rRNA, *pmoA* and *mxaF* genes most closely resembling those of *Methylomonas* spp. were also the most abundant clones in these libraries. All four clone libraries indicated the presence of organisms related to *Methylosinus/Methylocystis* spp. Only the *pmoA* and *mmoX* libraries detected genes most resembling those from *Methylococcus capsulatus.* In addition to genes resembling those of characterized organisms, several genes were detected that did not display any affiliation with the corresponding genes from cultivated methylotrophs. These include the *mmoX* OTUs mvs13.7 and mvs13.14. Without a cultivated representative it is difficult to determine the organism to which these *mmoX* genes belong. These *mmoX* genes could be a divergent version of the gene in a methanotroph known to possess sMMO. Alternatively it could be from a methanotroph, such as a *Methylobacter,* that is not known to possess sMMO; the *pmoA* library included genes that group with those of characterized *Methylobacter* spp., suggesting that these organisms are present in Movile Cave. A third possibility is that the unusual *mmoX* genes belong to completely uncharacterized methanotroph genera, such as representatives of the β-*Proteobacteria* identified in the 16S rRNA library.

Although the analysis of *mmoX* genes from environmental samples performed in this study indicated that it can be a useful marker in ecology studies, the primary objective of the study was to amplify *mmoX* genes which were significantly different from those of characterized organisms. An analysis of the derived MmoX sequences indicated that the sequences were not significantly different from those of cultivated methanotrophs. A possible problem with the approach used is that the PCR primers were designed according to the sequences of known *mmoX* genes and are therefore likely to retrieve similar sequences. The alternative is to broaden the specificity of the primers to include the genes of related enzymes. This was attempted by using the sequence of the *amoC* gene of the alkene monooxygenase; however the primers would not amplify the *mmoX* from the genomic DNA of representative methanotrophs. The *amoC* and *mmoX* are probably too divergent to design effective primers. Recently, the sequence of a soluble butane monooxygenase has been made available (Sluis *et al.*, 2002) and the genes show relatively high similarity to those of

the sMMO. It may be possible to design primers that can amplify both the *mmoX* and *bmoX* genes and these primers may also retrieve novel genes from DNA isolated from environmental samples. Another approach to potentially isolate novel genes is to use a direct cloning method, which eliminates the PCR bias. The use of this approach to isolate methanotroph genes is described in the following chapter.

Chapter 5

Direct Cloning of $^{13}CH_4$-Labelled DNA from Gisburn Forest Soil

5.1 Introduction

Stable isotope probing (SIP) is a technique to study the active organisms in environmental samples (Radajewski *et al.*, 2003). The carbon in the biosphere consists predominantly of the ^{12}C isotope and the SIP technique is based on the principle that the biological molecules of the organisms that incorporate a ^{13}C substrate will become enriched with the heavy isotope. It is possible to purchase substrates that have been synthesized which contain > 99 % of the stable ^{13}C isotope. In DNA-SIP, the buoyant density of ^{13}C-DNA is greater than the community ^{12}C-DNA and can be separated by CsCl gradient centrifugation (Radajewski & Murrell, 2002; Radajewski *et al.*, 2003). The DNA-SIP technique was originally developed using $^{13}CH_4$ and $^{13}CH_3OH$ substrates, but has since been used with $^{13}CO_2$ (Whitby *et al.*, 2001) and ^{13}C-naphthalene (Jeon *et al.*, 2003).

In previous DNA-SIP studies, the microbial population that incorporated the isotopically heavy substrate has been characterized by PCR amplifying the small subunit rRNA (SSU rRNA) genes (i.e. 16S and 18S rRNA) and functional genes (ie. genes encoding methane monooxygenase and methanol dehydrogenase) using the ^{13}C-DNA as template (Hutchens *et al.*, 2004; Jeon *et al.*, 2003; Lueders *et al.*, 2004b; Morris *et al.*, 2002; Radajewski *et al.*, 2000; Radajewski *et al.*, 2002; Whitby *et al.*, 2001). In experiments using the C-1 substrates $^{13}CH_4$ and $^{13}CH_3OH$, the majority of SSU rRNA clones in a library usually belonged to known methylotrophic genera, but these 'usual suspects' are often accompanied by sequences pointing to unexpected organisms. One example is the discovery of many *Acidobacterium* sequences in PCR libraries generated from ^{13}C-DNA purified from methanol SIP experiments (Radajewski *et al.*, 2000; Radajewski *et al.*, 2002).

Metagenomics is a term for the study of the microbial community by the culture independent analysis of genomes. This has been reviewed recently (Schloss & Handelsman, 2003). A library of community genomic fragments is constructed by isolating DNA directly from environmental samples and cloning it into a suitable vector. The DNA can be randomly sequenced (Venter *et al.*, 2004), or clones containing genes of interest can be identified and sequenced (Beja *et al.*, 2000; Beja *et al.*, 2002; Liles *et al.*, 2003; Quaiser *et al.*, 2002).

There is considerable interest in the possibility of combining DNA-SIP with metagenomic studies (Schloss & Handelsman, 2003; Wellington *et al.*, 2003). A

difficulty studying a "metagenome" is the large complexity of the clone libraries that result from the huge microbial diversity in an environmental sample. Estimates of the microbial diversity in a forest soil using single-stranded DNA reassociation kinetics indicated between 12,000 to 18,000 different species (Torsvik *et al.*, 1990). In a metagenomic library constructed from ^{13}C-DNA from a DNA-SIP experiment, the cloned DNA is restricted to that from the genomes of organisms that have obtained the majority of their carbon, either directly or indirectly via cross-feeding, from the ^{13}C-substrate. This limits the library to genome fragments of organisms involved in the metabolic process of interest and increases the likelihood of obtaining and sequencing genes of relevance.

A potential problem with directly cloning ^{13}C-DNA into a plasmid vector, is the inevitable shearing of DNA molecules during the CsCl ultracentrifugation and purification steps (Schloss & Handelsman, 2003). The maximum size of DNA inserts that can be cloned is a fraction of the size of the original fragments, since the molecules must first be treated with a restriction enzyme to obtain cohesive ends that can be ligated with suitably high efficiency into a plasmid vector. Larger clones have the potential to provide more information and potentially to link a metabolic gene cluster to a phylogenetic gene, such as that of the SSU rRNA.

In this study, a metagenomic library was constructed using ^{13}C-DNA from a DNA-SIP experiment with $^{13}CH_4$. The objective was to determine the feasibility of cloning the ^{13}C-DNA and identifying a clone containing methanotroph genes.

5.2 Cloning Gisburn soil DNA

5.2.1 Stable isotope probing of Gisburn soil with $^{13}CH_4$

The SIP experiments were performed in a manner similar to that described previously (Radajewski *et al.*, 2002). Soil was collected from an oak forest near Gisburn (UK) in October 2002. 5 g samples of soil were placed in 125 ml serum vials and a soil slurry was established by the addition of 10 ml of NMS. In the first experiments, the soil was amended with 0.1 x NMS medium (pH 3.5 phosphate buffer) and incubated with ^{13}C-DNA until a total of 50 ml had been consumed. The $^{13}CH_4$ was introduced by five injections of 10 ml aliquots, and the headspace was flushed with air between injections. Although a heavy DNA band was visible in the gradient (results not shown), agarose gel electrophoresis of the purified DNA revealed that relatively little ^{13}C-DNA was present (~ 2 μg) when considering the amount of carbon consumed (~ 2 mmoles). Although this amount of ^{13}C-DNA would be sufficient to create PCR libraries, it was insufficient to optimize the cloning conditions. In the original SIP experiment with Gisburn soil (Radajewski *et al.*, 2002), dried soil without added medium that consumed 50 ml of did not generate a visible ^{13}C-DNA band, but soil slurries with 0.1 x NMS did result in a visible ^{13}C-DNA band in the CsCl gradient. The authors speculated that nutrient limitation was causing uncoupling of the methane oxidation to growth, resulting in the complete oxidation of $^{13}CH_4$ to CO_2 or the incorporation of carbon into poly-β-hydroxybutyrate. Therefore, to generate enough ^{13}C-DNA to optimize and obtain a clone library, the SIP experiment was repeated with soil slurried in 1.0 x NMS medium. Approximately 100 μg of ^{13}C-DNA was subsequently purified from the 5 g of soil that had consumed 50 ml of $^{13}CH_4$ in this SIP experiment

Optimizing the cloning of this ^{13}C-DNA took a considerable amount of time and required large amounts ^{13}C-DNA. When a successful cloning method was finally established, the only ^{13}C-DNA available was from a SIP experiment that had not oxidised the $^{13}CH_4$ during the first four weeks of incubation. In order to initiate oxidation, the microcosm had been spiked with 100 μl of slurry from a replicate SIP microcosm that had already consumed 50 ml of $^{13}CH_4$. The spiked slurry began oxidizing immediately and consumed a total of 50 ml of $^{13}CH_4$ during three weeks of incubation.

The DNA was extracted from the $^{13}CH_4$ labelled soil using a gentle SDS/CTAB extraction protocol as described in Section 2.13.1. After centrifugation (Section 2.13.2), the DNA was visible in ambient light and the ^{13}C-DNA band could be collected without exposure to DNA-damaging UV irradiation (Figure 5.1). Little precaution towards UV-induced DNA damage during DNA manipulations was taken in the original attempts to clone ^{13}C-DNA, and may have been a contributing factor to an initial inability to clone soil DNA, since even short exposures can result in severe decreases in the ability of plasmid DNA to transform *E. coli*. As an alternative to obtaining a gradient where the bands are visible in ambient light, it may be possible to fractionate the gradients and subsequently screen the fractions for DNA (Lueders *et al.*, 2004a). The DNA from the CsCl gradient was purified as described in Section 2.13.3. The ^{13}C-DNA prepared by this method was analysed by PFGE to estimate the size of the DNA fragments obtained, and the majority of the fragments were found to be between ~ 20 kb and ~ 100 kb (Figure 5.2).

A small 16S rRNA PCR library was generated from the ^{13}C-DNA. Six clones were analysed and sequenced. Two of the clones were related to methanotroph 16S rRNA genes and four were most closely related to the genes of *Frateuria* spp. (results not shown). It is not clear why *Frateuria* spp. incorporated the ^{13}C, but is likely a result of cross-feeding from the methanotroph population.

5.2.2 Preparation of ^{13}C-DNA for cloning and ligation into pCC1BAC™

It is known that restriction enzyme activity is very sensitive to inhibition by soil organics which are often co-extracted with DNA (Zhou *et al.*, 1996). Although the DNA had been purified by several rounds of CsCl gradient centrifugation, the DNA was not sufficiently pure to be digested with the *Bam*HI restriction enzyme (results not shown). Therefore, further purification was achieved by electrophoresis of the ^{13}C-DNA through a 1 % (w/v) agarose gel, as described in Section 2.14.1. The electrophoresis through a 1 % agarose gel had two convenient purposes: first, it eliminated the contamination of the DNA such that it could be digested with restriction enzymes, and secondly the DNA formed a tight band that could be excised in an agarose plug (Figure 5.3A); the agarose plug containing the ^{13}C-DNA was convenient since the restriction enzyme digestion of the DNA could be performed in

the plug. This minimized additional DNA damage by shearing that could potentially occur when pipetting the DNA solution.

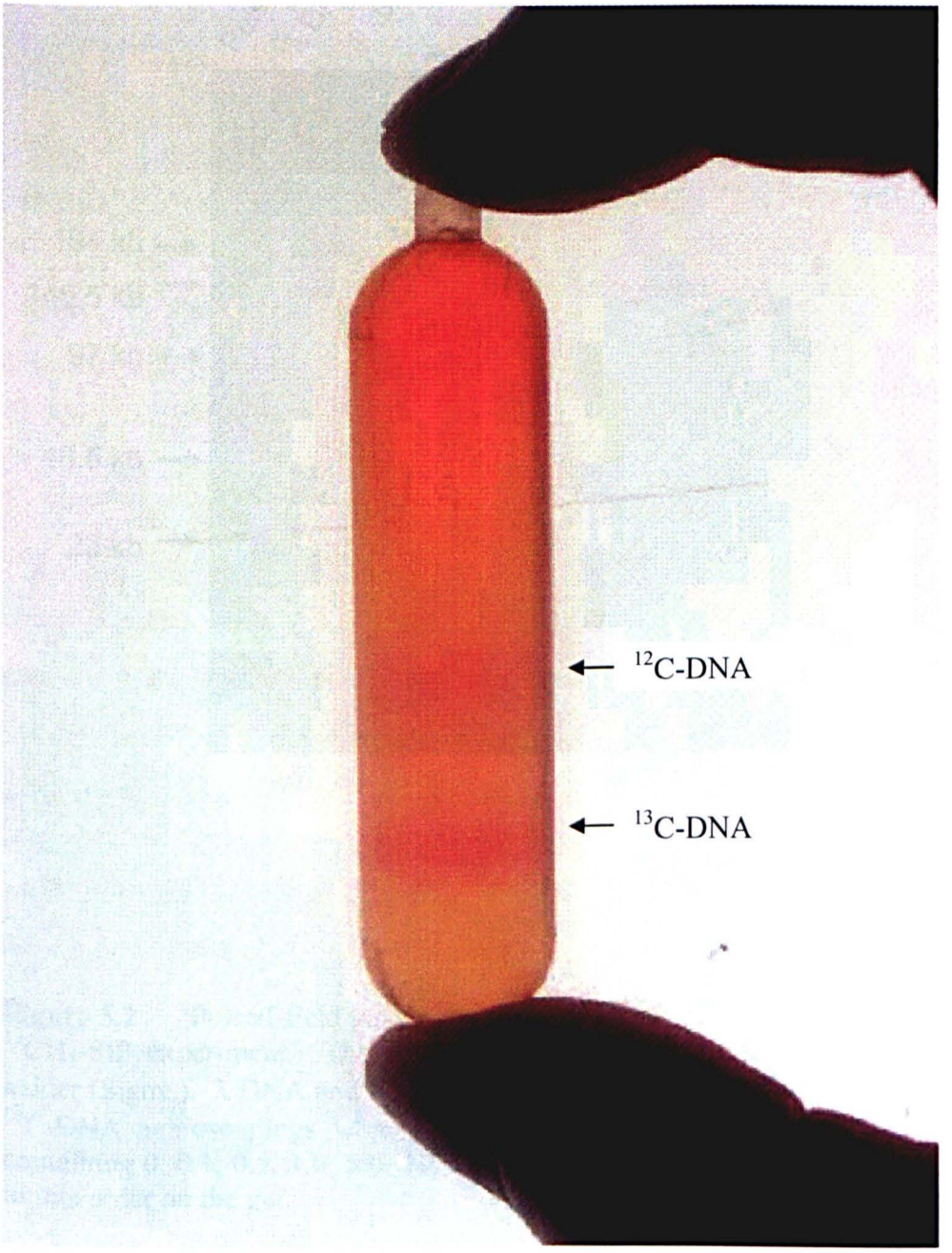


Figure 5.1 CsCl gradient of $^{13}CH_4$ SIP – labelled DNA from Gisburn soil. The 'heavy' ^{13}C-DNA and 'light' ^{12}C-DNA bands were clearly visible in ambient light.

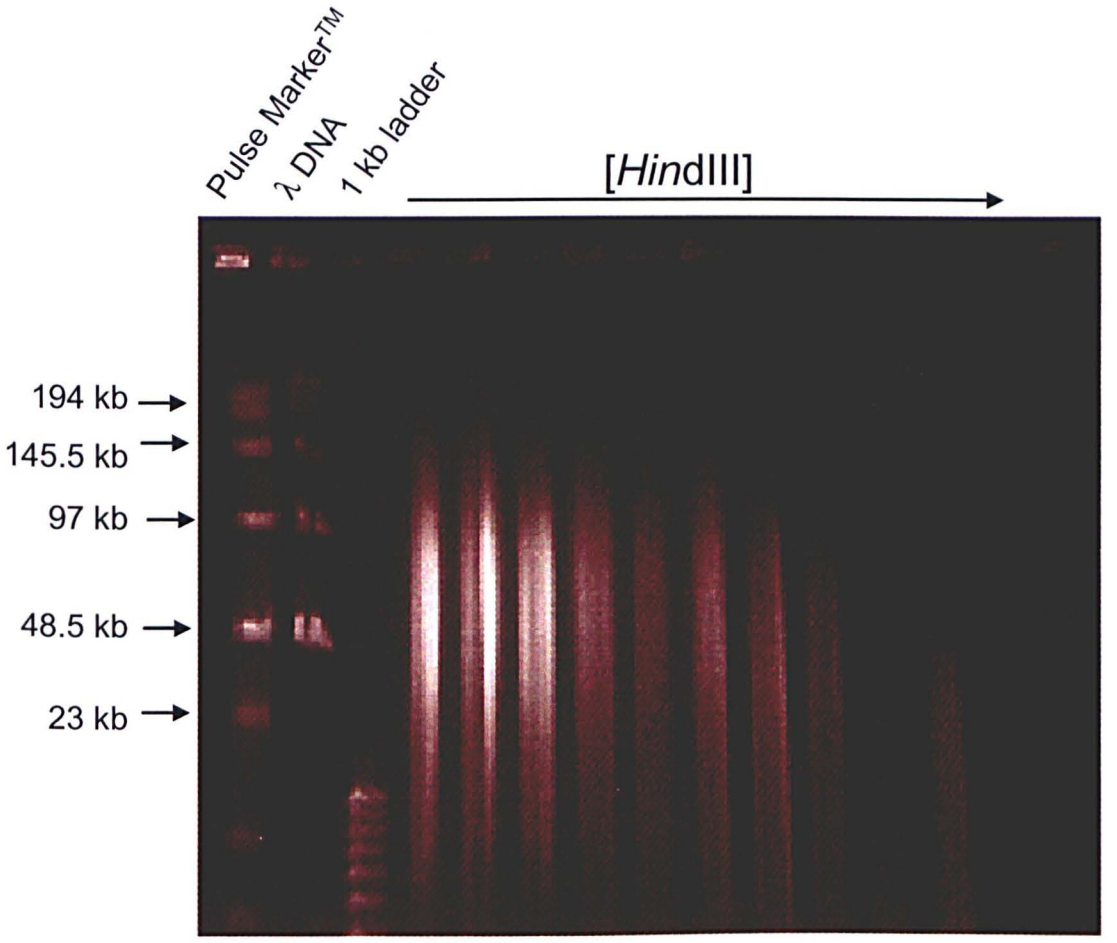
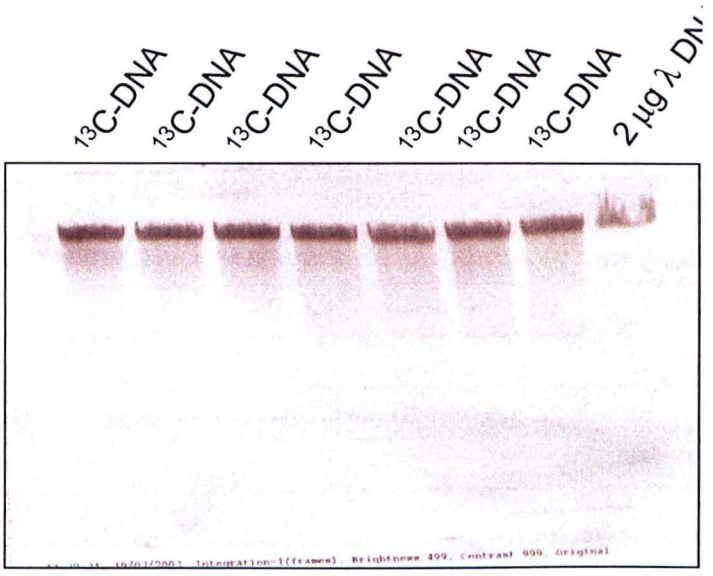


Figure 5.2 Pulsed-field gel electrophoresis of ^{13}C-DNA from a Gisburn forest soil $^{13}CH_4$-SIP experiment. The first and last lanes contain the Pulse Marker™ DNA ladder (Sigma). λ DNA and 1 kb ladder (Invitrogen) were also included, as indicated. ^{13}C-DNA agarose plugs were incubated in 500 µl 1 x restriction enzyme buffer containing 0, 0.1, 0.5, 1.0, 5.0, 10, 25, 50, 100, 250, 500 units of *Hin*dIII, and loaded in this order on the gel.

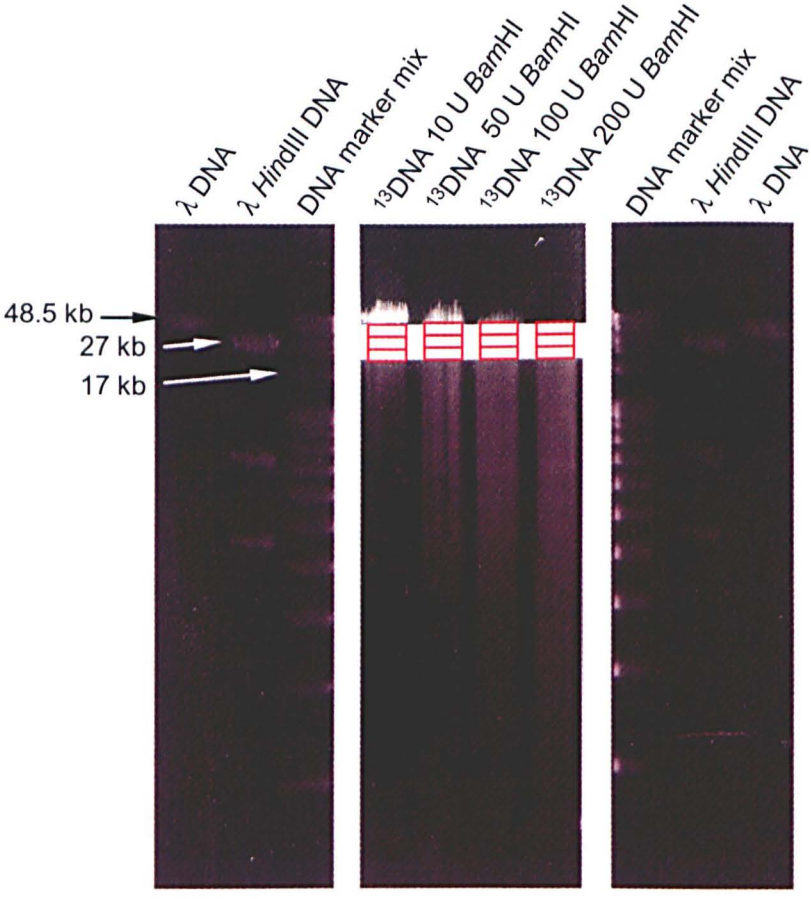
Partial digestion of ^{13}C-DNA in the plugs was performed as described in Section 2.14.2. The amount of DNA was estimated to be approximately 4 μg per agarose plug and half a plug was used for each digestion. The amount of enzyme required to get partial cutting of the DNA was determined empirically for each experiment by incubating with a range of enzyme concentrations. Partial cutting of the ^{13}C-DNA in this experiment was found to occur with between 50 and 100 units of *Bam*HI in 500 μl of 1 x restriction enzyme buffer. 75 units were used for the partial digestion of ^{13}C-DNA that was cloned into pCC1BAC™ to generate the library screened in this study. The partially digested ^{13}C-DNA was resolved on a standard agarose gel (20 cm x 20 cm) rather than a PFGE system, since standard electrophoresis resulted in a more tight distribution of the desired size fragments within the gel and easier elution from the agarose plug (Section 2.8.9). Locating the segment of the gel without exposure of the ^{13}C-DNA to UV irradiation was performed as described in Section 2.14.2. In summary, the DNA marker lanes were stained and visualized, and the positions of the bands marked on the gel. By aligning the marker lanes with the rest of the gel, the regions corresponding to approximately 20 – 25 kb, 25 – 30 kb and 30 – 50 kb in the ^{13}C-DNA lanes were identified, and small slabs of agarose in these regions were removed. The remaining gel was stained with ethidium bromide and photographed under UV (Figure 5.3B). The partially *Bam*HI-digested DNA was electro-eluted from the gel slices as described in Section 2.8.9. The partially *Bam*HI-digested ^{13}C-DNA was quantified by agarose gel electrophoresis by comparison with known quantities of DNA marker. Approximately 100 ng of 30 - 50 kb size-selected ^{13}C-DNA was ligated into pCC1BAC™ (Epicentre) according to the manufacturer's instructions (Section 2.14.3).

Figure 5.3 A. photograph of ^{13}C-DNA electrophoresed through an agarose gel and stained with 15 μg ml^{-1} Nile blue. Each lane contained ~ 60 μl of dialysed ^{13}C-DNA from the Gisburn SIP experiment. B. Photographic reconstruction of an agarose gel where size-selected DNA fragments had been excised. The DNA size markers used were λ DNA, *Hin*dIII digested λ DNA, and a mixture of 1 kb DNA ladder (Invitrogen) with uncut λ DNA (48.5 kb), λ DNA cut with *Kpn*I (31 kb, 17 kb fragments) and *Xba*I (24.0 kb and 24.5 kb fragments), as indicated. The ^{13}C-DNA was cut with various amounts of *Bam*HI enzyme, as indicated. The red blocks indicate the approximate position of where the agarose gel slabs excised had been positioned.

A


^{13}C-DNA
^{13}C-DNA
^{13}C-DNA
^{13}C-DNA
^{13}C-DNA
^{13}C-DNA
^{13}C-DNA
2 µg λ DN

B


λ DNA
λ HindIII DNA
DNA marker mix
^{13}DNA 10 U BamHI
^{13}DNA 50 U BamHI
^{13}DNA 100 U BamHI
^{13}DNA 200 U BamHI
DNA marker mix
λ HindIII DNA
λ DNA
48.5 kb
27 kb
17 kb

5.2.3 *E. coli* Transformax EC300 library of Gisburn soil DNA clones

The transformation of *E. coli* TransforMax EC300™ (Epicentre) was performed as described in Section 2.14.4. The voltage and resistance settings were optimized by testing a variety of settings with control BAC plasmid DNA (results not shown). Transformation of *E. coli* TransforMax EC300™ with 2 µl of the ligation mix resulted in approximately 1000 clones. A clone library was generated by filling a total of 24 96-well plates as described in Section 2.14.5. More than 99 % of the clones grew in the medium and therefore the library consisted of nearly 2300 clones.

Plasmids were isolated from 48 clones (Section 2.6.2) chosen at random and analysed by digestion with *Bam*HI and agarose gel electrophoresis. Most contained inserts between 20 and 30 kb in size (Figure 5.4) and two of the 48 plasmids contained no visible insert DNA. The sequencing primers T7 (3'-TAA TAC GAC TCA CTA TAG GG -5') and RP2 (3'- TAC GCC AAG CTA TTT AGG TGA GA - 5') anneal adjacent to the cloning site of pCC1BAC™ and were used to sequence the ends of several of the cloned inserts. The sequence data were compared to the database Genbank database by BLASTX analysis (Altschul *et al.*, 1997) and the results are shown in Table 5.1. All of the end-sequenced clones were different to each other with the exception of GSC1 and GSC10, which were identical.

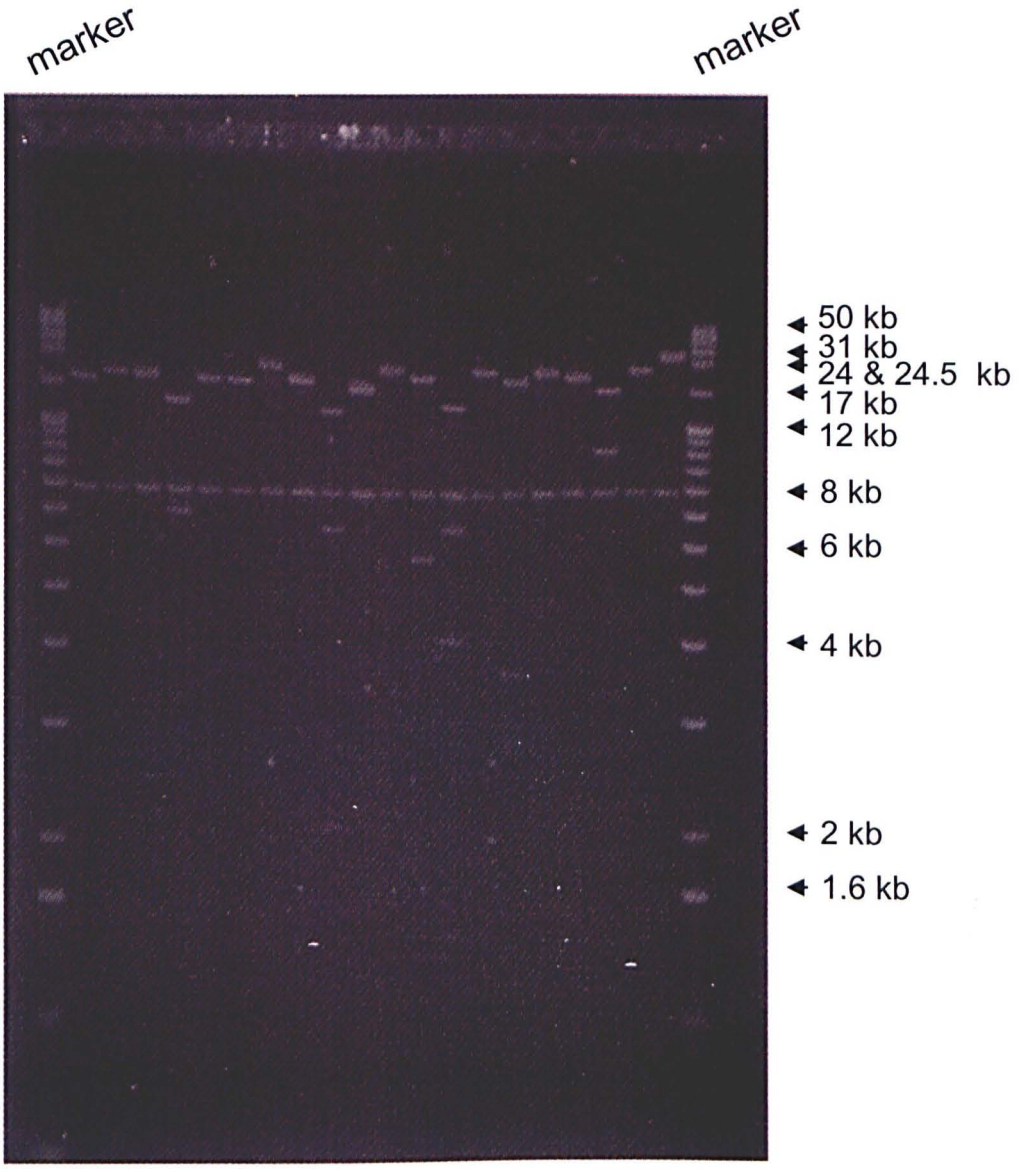


Figure 5.4 RFLP analysis of 20 Gisburn soil clones (GSCs). The DNA size marker consists of 1 kb ladder (Invitrogen) and fragments of λ DNA (as in Figure 5.3B). The gel contained 0.8 % agarose.

Table 5.1 Result of BLASTX searches (Altschul *et al.*, 1997) of end-sequenced clones.

Clone	Protein Name Closest BLASTX Match	Accession no.	Source	% Identity	BLAST Score/Expect
GSC1	Cytochrome d ubiquinol oxidase, subunit II	NP_761021	*Vibrio vulnificus* CMCP6	35 %	169/ $7e^{-41}$
	probable signal peptide protein	NP_523060	*Ralstonia solanacearum*	35 %	72/ $2e^{-11}$
GSC2	TPR domain protein	NP_421922	*Caulobacter crescentus*	42 %	199/ $7e^{-50}$
	ABC transporter permease	NP_643948	*Xanthomonas axonopodis*	34 %	117/ $2e^{-25}$
GSC3	Oxoglutarate dehydrogenase	NP_636859	*Xanthomonas campestris*	65 %	265/ $5e^{-86}$
	Predicted transcriptional regulator (COG1386)	ZP_00127917	*Pseudomonas syringae*	62 %	94.7/ $2e^{-18}$
	Pseudouridylate synthase	NP_779666	*Xylella fastidiosa*	60 %	75.1/ $2e^{-12}$
GSC4	Hypothetical protein	ZP_00091490	*Azotobacter vinelandii*	29 %	58.9/ $1e^{-7}$
	Conserved hypothetical protein	NP_636148	*Xanthomonas campestris*	57 %	271/ $1e^{-71}$
GSC5	uncharacterized (COG4625)	ZP_00084992	*Pseudomonas fluorescens*	29 %	99.8/ 1e-^{20}
	Cation/multidrug efflux pump (COG0841)	ZP_00127150	*Pseudomonas syringae*	67 %	312/ $6e^{-85}$
GSC6	YapH protein	NP_637389	*Xanthomonas campestris*	26 %	37/ 0.59
GSC7	Thymidylate synthase (COG0207)	ZP_00149890	*Dechloromonas aromatica*	66 %	345/ $6e^{-94}$
	Conserved hypothetical protein	NP_882629	*Bordetella parapertussis*	36 %	82.4/ $3e^{-17}$
GSC8	ATP-dependent Clp protease subunit	NP_638417	*Xanthomonas campestris*	69 %	216/ $7e^{-55}$
	Uncharacterized	AAD32752	*Streptomyces lavendulae*	37 %	109/ $9e^{-23}$
GSC9	ABC transporter ATP-binding protein	NP_825718	*Streptomyces avermitilis*	50 %	64/ $4e^{-9}$
	Outer membrane protein (COG4775)	ZP_00041072	*Xylella fastidiosa*	40 %	186/ $5e^{-46}$
GSC10	Cytochrome d ubiquinol oxidase, subunit II	NP_761021	*Vibrio vulnificus* CMCP6	39 %	207/ $3e^{-52}$
	Probable signal peptide protein	NP_523060	*Ralstonia solanacearum*	40 %	126/ $6e^{-28}$
GSC11	Biotin synthase	NP_640744	*Xanthomonas axonopodis*	71 %	131/ $2e^{-29}$
	7-Keto-8-aminopelargonate synthetase (COG0156)	ZP_00038860	*Xylella fastidiosa*	54 %	81/ $8e^{-21}$
	Transcription-related protein	NP_642379	*Xanthomonas axonopodis*	70 %	350/ $2e^{-98}$
GSC12	Restriction enzyme A (COG1403)	YP_005265	*Thermus thermophilus*	63 %	103/ $7e^{-21}$
GSC13	UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine	NP_641746	*Xanthomonas axonopodis*	45 %	156/ $7e^{-37}$
GSC14	Uncharacterized conserved protein (COG2930)	ZP_00039701	*Xylella fastidiosa*	56 %	224/ $2e^{-57}$

	TonB-dependent receptor	NP_642846	*Xanthomonas axonopodis*	37 %	186/ $9e^{-46}$
GSC15	Heavy-metal transporting P-type ATPase	NP_767340	*Bradyrhizobium* USDA 110	54 %	233/ $4e^{-60}$
	Diadenosine tetraphosphate hydrolase (COG0537)	ZP_00102122	*Desulfitobacterium hafniense*	66 %	145/ $1e^{-33}$
GSC2794	2-Oxoglutarate dehydrogenase E1 component	YP_008089	*Parachlamydia* sp. UWE25	40 %	162/ $6e^{-40}$
GSC1953	50S Ribosomal protein L16	NP_641326	*Xanthomonas axonopodis*	81 %	117/ $3e^{-25}$
	30S Ribosomal protein S17	NP_641328	*Xanthomonas axonopodis*	63 %	83.6/ $2e^{-15}$
	Ribonuclease G (COG1530)	NP_637957	*Xanthomonas campestris*	75 %	138/ $2e^{-31}$
GSC7794	Hypothetical protein (COG2331)	NP_824855	*Streptomyces avermitilis*	62 %	92/ $2e^{-17}$
	Predicted Zn-dependent protease (COG0312)	ZP_00172835	*Methylobacillus flagellatus*	39 %	36/ 1.5
GSC7806	ABC-type phosphate transport system (COG0581)	ZP_00123767	*Pseudomonas syringae*	56 %	156/ $1e^{-40}$

5.2.4 Probing Gisburn soil DNA clone library with methylotroph functional genes

The library was screened for *mxaF*, *pmoA* and *mmoX* genes by colony hybridization (Section 2.10 & 2.14.5). The *mxaF* primers f1003 & r1561 (McDonald & Murrell, 1997a), *pmoA* primers A189 & A682 (Holmes *et al.*, 1995) and *mmoX* primers mmoX206F & mmoX886R (Chapter 4) were used to PCR amplify the probe DNA using the ^{13}C-DNA as template. Since the probes were amplified from ^{13}C-DNA, there was a good chance that the homologous probe was present for the gene of any potential clone. The three gene hybridizations were performed separately and a total of four membranes with 6 x 96 clones were hybridized. The hybridized membranes were washed at a moderate stringency of 55 °C with 2 x SSC.

Clones that showed weak hybridization with a probe were screened by PCR using the original gene primers and also end-sequenced with the T7 and RP2 sequencing primers. None of the clones demonstrating weak hybridization signals appeared to contain the target gene. No clones were identified that carried *mxaF* or *mmoX*. Two clones, GSC357 (Gisburn Soil Clone) and GSC1346 (not shown) hybridized strongly to the *pmoA* probe (Figure 5.5). The presence of *pmoA* on these clones was first confirmed by PCR amplification using the A189/ A682 PCR primers, and the GSC357 insert was fully sequenced (Section 5.2.5). The RFLP patterns of GSC357 and GSC1346 digested with *Not*I were identical (results not shown).

Difficulties have been observed in trying to clone complete *pmoCAB* operons into *E. coli* using high-copy cloning vectors (Gilbert *et al.*, 2000; Semrau *et al.*, 1995). Presumably there is some expression of the pMMO proteins which exhibits a "toxic effect", possibly by disrupting the *E. coli* cell membranes. This study and the recent description of the cloning *pmoCAB* operons from *Methylocystis* sp. SC2 into a BAC vector (Ricke *et al.*, 2004) indicate that this "toxicity" is not lethal in single-copy vectors. The copy number of the pCC1BAC™ vector in *E. coli* TransforMax EC300™ can be up-regulated approximately 15-fold (Epicentre, product literature); however it was found that the copy-number of GSC357 and GSC1346 could not be increased, whereas it could be for other GSC clones (results not shown).

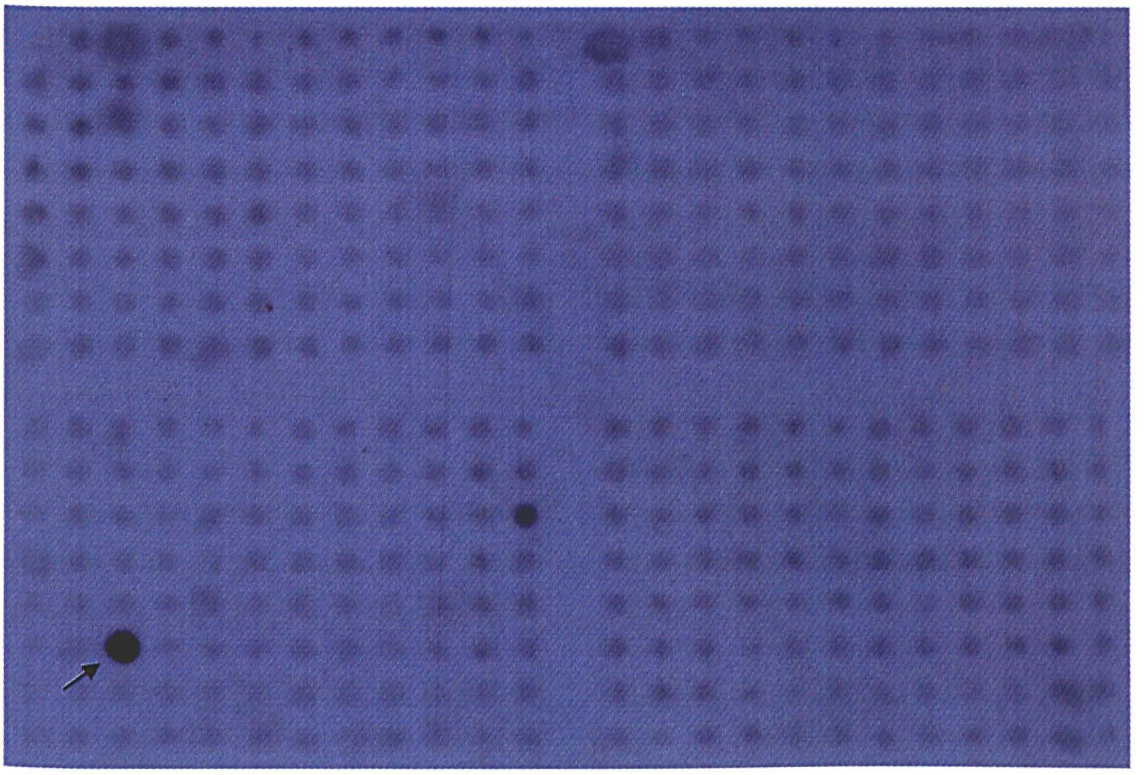


Figure 5.5 Section of X-ray film showing the hybridization of the *pmoA* probe to GSC357 (indicated by arrow).

5.2.5 Characterization and sequencing of the Gisburn soil *pmoA* clone

GSC357 was completely sequenced using a combination of shotgun sequencing and direct sequencing. The direct sequencing was performed to join contiguous sequence data and to improve ambiguous sequence. The complete sequence was found to be 15,253 bp. A summary of the genes and their arrangement is given in Figure 5.6 and Table 5.2. The ORFs were identified using NCBI's Open-Reading-Frame-Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and the sequences were analysed and annotated using BLASTX (Altschul *et al.*, 1997). GSC1346, which also hybridized with the probe and had similar RFLP patterns to GSC357, was partially sequenced and the genes found to possess > 95 % sequence identity with those on GSC357. It is likely that GSC357 and GSC1346 are from different strains of the same species of methanotroph.

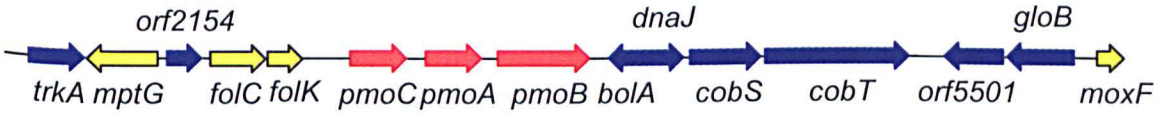


Figure 5.6 (A) Gene map of GSC357 (15,253 bp). The *pmoCAB* operon is shown in red. Genes with (possible) roles in methylotrophy are in yellow.

Table 5.2 Result of BLASTX searches (Altschul *et al.*, 1997) of ORFs identified on GSC357.

Clone	Protein Name of Closest BLASTX Match	Accession no.	Source	Identity	BLAST Score/Expect
trkA	component of K^+ transport system	NP_422290	*Caulobacter crescentus*	50 %	$149/5e^{-35}$
mptG	β-ribofuranosylaminobenzene 5'-phosphate synthetase	AAS86339	*Hyphomicrobium zavarzinii*	42 %	$212/7e^{-54}$
orf2154	pterin-4α-carbinolamine dehydratase	ZP_00034540	*Burkholderia fungorum*	63 %	$170/9e^{-42}$
folP	dihydropteroate synthase	ZP_00050905	*Magnetospirillum magnetotacticum*	53 %	$238/ 6e^{-62}$
folK	2-amino-4-hydroxy-hydroxymethyldihydopteridine pyrophosphokinase (Hppk)	NP_102516	*Mesorhizobium loti*	46 %	$324/2e^{-29}$
pmoC	**pMMO subunit C**	AAF37896	*Methylocystis* sp. M	95 %	$400/ 1e^{-110}$
pmoA	**pMMO subunit A**	CAE47800	*Methylocystis* sp. SC2	94 %	$417/ 1e^{-116}$
pmoB	**pMMO subunit B**	CAE47801	*Methylocystis* sp. SC2	92 %	706/ 0.0
bolA	BolA-like protein	NP_945854	*Rhodopseudomonas palustris*	40 %	$55.8/2e^{-07}$
dnaJ	DnaJ-family molecular chaperone	ZP_00052274	*Magnetospirillum magnetotacticum*	48 %	$177/8e^{-44}$
cobS	cobalt insertion protein	NP_766820	*Bradyrhizobium japonicum*	84 %	$572/1e^{-162}$
cobT	cobyrinic acid synthase	NP_945849	*Rhodopseudomonas palustris*	61 %	702/0.0
orf5501	predicted secreted protein	NP_772837	*Bradyrhizobium japonicum*	55 %	$270/2e^{-71}$
gloB	Zn-dependent hydrolase (GloB)	NP_772836	*Bradyrhizobium japonicum*	54 %	$317/2e^{-25}$
moxF	methanol dehydrogenase large subunit-like protein	NP_772853	*Bradyrhizobium japonicum*	63%	$166/1e^{-40}$

GSC357 contains a complete *pmoCAB* operon. The *pmoC* gene is 765 bp, *pmoA* is 759 bp and *pmoB* is 1257 bp. The intergenic region between *pmoC* and *pmoA* is 269 bp and between *pmoA* and *pmoB* is 227 bp. The sequences of *pmoCAB* operons and derived aminoacyl sequences demonstrate considerable similarity to those of *Methylocystis* sp. strain M (Gilbert *et al.*, 2000), *Methylosinus trichosporium* OB3b (Gilbert *et al.*, 2000), *Methylocystis* sp. strain SC2 (Ricke *et al.*, 2004) and *Methylococcus capsulatus* (Bath) (Semrau *et al.*, 1995; Stolyar *et al.*, 1999) (Table 5.3). Putative ribosome binding sites are present 5' of all three *pmo* genes. The *pmoCAB* operon in *Methylocystis* strains has been shown to be expressed from a σ^{70} promoter 5' of the *pmoC* and the transcriptional start has been mapped by primer extension for copies of *pmoCAB* (Gilbert *et al.*, 2000; Ricke *et al.*, 2004). A putative -35, -10 σ^{70} consensus sequence is present 5' of the *pmoC* gene on GSC357 (Figure 5.7).

In the original SIP experiment with the Gisburn forest soil (Radajewski *et al.*, 2002) partial *pmoA* genes were PCR amplified from ^{12}C-DNA and ^{13}C-DNA. Two clone types, P12.9 (Genbank accession AY080955) and P13.7 (Genbank accession AY080951) represented a significant proportion of the ^{12}C-DNA and ^{13}C-DNA *pmoA* libraries, respectively. The derived PmoA amino acid sequences of P12.9 and P13.7 are essentially identical to the derived PmoA of the gene on GSC357[4]. Therefore, the *pmoA* gene of GSC357 appears to be related to that detected by PCR in the previous study.

Sequencing studies have identified genes flanking the *pmoCAB2* operon in *Methylocystis* sp. strain SC2 (Ricke *et al.*, 2004), one of the *pmoCAB* operons of *Methylococcus capsulatus* (Bath) (Semrau *et al.*, 1995), and one of the operons of *Ms. trichosporium* OB3b (Gilbert *et al.*, 2000). None of these ORFs correspond to those found on GSC357. However, unpublished data (I.R. McDonald & J.C. Murrell) for the second copy of *pmoCAB* in *Ms. trichosporium* OB3b indicates that, like GSC357,

[4] There is a single difference in the region amplified by the A682 PCR primer (5' – GAA SGC NGA GAA GAA SGC – 3') used to generate the soil *amoA/pmoA* libraries, and is likely an artifact caused by the primer. This primer assumes an alanine is present at position 226 of the sequence, which is true for the AmoA proteins and the PmoA of *Methylococcus capsulatus* (Bath), whereas in *Methylosinus trichosporium* OB3b and *Methylocystis* strains M & SC2 it is a glycine. This primer has been shown to be biased towards *amoA* sequences when constructing libraries from environmental samples (Bourne *et al.*, 2001). An analysis of the full-length *amoA/pmoA* sequences currently in the database (results not shown) indicate that the 5' – GAA NSC SGA GAA GAA NGC – 3' sequence may be more suitable when targeting both *pmoA* and *amoA* from an environmental samples; regardless of this, the original primer is still effective.

the *folP* and *folK* genes are immediately 5' of *pmoC*. Therefore it is possible that these *pmoCAB* operons are situated within similar regions of the chromosome in both organisms.

Table 5.3 Identity of the derived pMMO protein sequences of the *pmoCAB* of GSC357 with those from other sequenced operons.

Organism	PmoC	PmoA	PmoB
Methylocystis strain M	93 %	92 %	88 %
Methylocystis strain SC2 (copy 1)	83 %	94 %	87 %
Methylocystis strain SC2 (copy 2)	65 %	64 %	58 %
Methylosinus trichosporium OB3b	85 %	88 %	77 %
Methylococcus capsulatus (Bath)	52 %	57 %	47 %

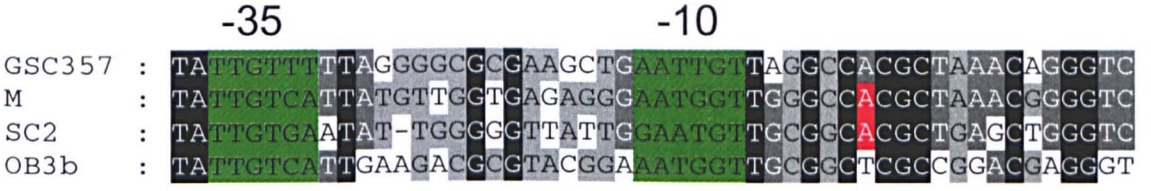


Figure 5.7 Sequence alignment of the promoter regions upstream (5') of the *pmoC* genes of GSC357, *Methylocystis* sp. strain M (M), *Methylocystis* sp. strain SC2 (SC2) and *Methylosinus trichosporium* OB3b (OB3b). The *pmoCAB1* sequence of *Methylocystis* sp. strain SC2 was used. The putative -35, -10 consensus sequences are shaded in green (based on the *E. coli* σ^{70} promoter consensus TTGACA-N17-TATAAT). Experimentally determined transcriptional start residues are highlighted in red. The position of the *Ms. trichosporium* OB3b promoter was determined by RT-PCR (Stafford 2002).

Analysis of the genome sequence of *Methylobacterium extorquens* AM1 indicated that many methylotrophy genes are linked in clusters on the chromosome (Chistoserdova *et al.*, 2003). With the possible exception of *moxF*, none of the genes situated downstream (3') of the *pmoCAB* operon on GSC357 appear to be specifically related to growth on methane. The *moxF* gene is an apparent homologue of the *mxaF*, which encodes the large subunit of the methanol dehydrogenase. Similar *moxF* genes are present in numerous organisms and its function has not been determined.

Three genes, *mptG*, *folP* and *folK* are present upstream (5') of *pmoC* and have proposed roles in C-1 metabolism (Chistoserdova *et al.*, 2003). Dihydropteroate synthase (FolP) and 2-amino-4-hydroxy-6-hydroxymethyl-7,8-dihydopteridine pyrophosphokinase (FolK) are involved in folate biosynthesis. In methylotrophs that use the serine pathway for formaldehyde fixation, a tetrahydrofolate cofactor is involved in transferring the methyl group (Chistoserdova *et al.*, 2003), therefore the folate biosynthesis is linked to growth on C-1 compounds. The *mptG* gene encodes β-ribofuranosylaminobenzene 5'-phosphate synthetase (β-RFAP synthetase), which is involved in methanopterin biosynthesis (Scott & Rasche, 2002). Tetrahydromethanopterin is an "archaeal" cofactor that is also present in methylotrophic bacteria and involved in the transfer of the methyl group during oxidation or assimilation of formaldehyde (Chistoserdova *et al.*, 1998; Chistoserdova *et al.*, 2000). A phylogenetic tree of the β-RFAP synthetase from GSC357 and other β-RFAP synthetases from the database is shown in Figure 5.8. The sequence appears to group with that of the corresponding gene from *Hyphomicrobium zavarzinii*, which is a α-*Proteobacteria* methylotroph. The fact the sequence is most similar to that of a α-*Proteobacteria* representative is consistent with the hypothesis that GSC357 is a chromosome fragment from a *Methylocystis* species.

The ORF between *mptG* and *folP* encodes pterin-4α-carbinolamine dehydratase, which is involved in the biosynthesis of the tetrahydrobiopterin (BH_4) cofactor (Thöny *et al.*, 2000). Known BH_4-dependent enzymes include the phenylalanine-, tyrosine- and tryptophan-hydroxylases, NO-synthases and the glyceryl-ether monooxygenase. The position of the gene within a methylotrophy island containing *pmoCAB* may be a coincidence.

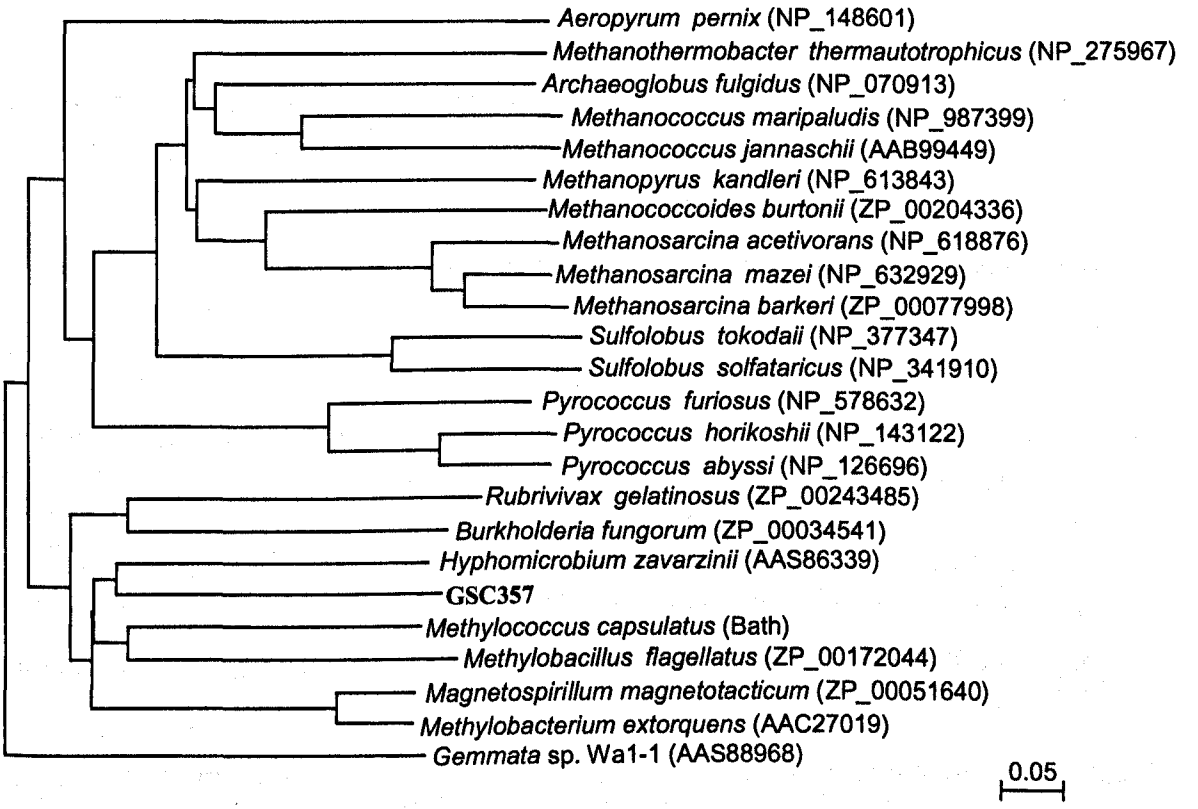


Figure 5.8 Phylogenetic tree of β-RFAP synthetase sequences. The sequence for *Methylococcus capsulatus* (Bath) was identified by BLAST of the unfinished genome (http://tigrblast.tigr.org/ufmg/index.cgi?database=m_capsulatus|seq).

5.3 Discussion

It is estimated that less than 1 % of bacteria in soil have been isolated in pure culture (Torsvik & Ovreas, 2002). Growth of bacteria in laboratory cultures is often biased and the organisms that dominate enrichment cultures are not necessarily those that are most active in the environment. Stable isotope probing (SIP) is a method of identifying microorganisms in an environment, or an environmental sample, that incorporate a particular ^{13}C-substrate into cellular biomass. The technique is powerful since several studies have shown that SIP often identifies organisms not previously known to be involved in the process. Traditionally, DNA-SIP has been used in conjunction with PCR of SSU rRNA and functional genes to identify the organisms that have incorporated the heavy substrate. Although a small number of genes are usually targeted by PCR and analysed, DNA-SIP results in the isolation of the complete genomic DNA complement of the labelled population. This is the first study where the ^{13}C-DNA from a DNA-SIP experiment has been cloned into a vector in order to directly sequence the genomic fragments of the active population, without needing to first isolate the genes by PCR.

There are several limitations to SIP including the potential that organisms that had not been active prior to substrate addition have been stimulated or enriched (Torsvik & Ovreas, 2002) and the potential for cross-feeding that can occur during the long incubations that are needed to obtain an adequate amount of ^{13}C-DNA (Radajewski *et al.*, 2003). Indeed, there has been direct evidence of cross-feeding in SIP experiments after incubation with ^{13}C-methane and methanol, such as the appearance of a chrysophyte algae (Hutchens *et al.*, 2004) and the bacteriovorous *Bdellovibrio* species (Hutchens *et al.*, 2004; Morris *et al.*, 2002). In this study there was evidence that a considerable amount of the ^{13}C-methane was used by species belonging to the genus *Frateuria*. This particular problem was not the focus of this study; other work in the JCM laboratory is addressing the issue of ^{13}C recycling in DNA-SIP experiments and determining the minimum incubation times with the ^{13}C-substrate (J. Vohra, unpublished).

In this study, a library of 2300 clones was constructed from ^{13}C-DNA obtained from a DNA-SIP experiment with $^{13}CH_4$. A clone was identified by gene probing that contained a *pmoCAB* operon, which encodes the particulate methane monooxygenase enzyme (pMMO) catalysing the first step in the metabolism of methane by

methanotrophic bacteria (Hanson & Hanson, 1996). Sequence analysis of the *pmoA* gene, for which a substantial amount of phylogenetic data is available (McDonald & Murrell, 1997b), indicates that clone was most probably a genomic fragment from an uncultivated *Methylocystis* strain. The complete sequence of the clone was determined and additional genes encoding enzymes which may also be involved in biochemical pathways required for growth on methane were identified.

One example was the gene for β-ribofuranosylaminobenzene 5'-phosphate (β-RFAP) synthetase, which is involved in the synthesis of the "archaeal" tetrahydromethanopterin cofactor (Scott & Rasche, 2002). This enzyme and a pathway for the assimilation of formaldehyde using a methenyl-tetrahydromethanopterin-linked pathway has been reported for the methylotrophic organisms *Methylobacterium extorquens* AM1 (Chistoserdova *et al.*, 2003) and *Methylobacillus flagellatus* (Chistoserdova *et al.*, 2000). The sequencing results here suggest that the gene for β-RFAP synthetase is situated close to the *pmoCAB* operon in the genome of a methanotroph.

The clone library screened in this study was relatively small and represents less than ten genome equivalents (assuming the average insert is 20 kb and the average genome is 5 Mb: 20 kb x 2,300 clones ÷ 5,000 kb = 9.2). Methanotrophs probably constitute a relatively small proportion of the total bacterial community in this soil, and therefore to identify a methanotroph gene in a library constructed from total community DNA would require screening many thousands of genome equivalents. Most SIP experiments implicate organisms not predicted to be involved in the process being studied. These include possible methylotrophic *Acidobacterium* representatives (Radajewski *et al.*, 2003) and a phenol-degrading *Thauera* species (Manefield *et al.*, 2002a). In *Methylobacterium extorquens* AMI, many of the genes involved in growth on one-carbon compounds are clustered together on the chromosome, in what have been termed methylotrophy islands (Chistoserdova *et al.*, 2003). Therefore the sequence of cloned genomic fragments from an as yet uncultivated, and perhaps unculturable, methylotroph could provide insights into the biochemistry of the organism.

In this study the majority of the ^{13}C-DNA purified was between 15 kb and 100 kb in length, and the maximum insert size observed in the clone library was approximately 30 kb. It may be possible to increase the size of the inserts by size-

selecting the partially digested ^{13}C-DNA by PFGE, and using a restriction enzyme, such as *Sau*3A1, that recognizes a frequently occurring four-base DNA sequence. The cloning of ^{13}C-DNA partially cut with *Sau*3A1 was attempted during this work and found to result in a high percent of clones lacking DNA inserts (results not shown); however it should be possible to overcome this technical difficulty. Previous studies have shown that with extensive optimization it is possible to obtain clones from soil with inserts as large as 80 kb (Rondon *et al.*, 2000). The PFGE analysis of the ^{13}C-DNA obtained in this study indicates that DNA fragments larger that 100 kb were present, and therefore it should be possible to obtain ^{13}C-DNA clones significantly larger than 30 kb.

The clone library in this study was generated by partial digestion with *Bam*HI, which recognizes a 5' – GGATTC – 3' sequence. The library may have been biased towards *Bam*HI fragments that fall within the 10-30 kb size range of the library. If two *Bam*HI sites were not within 30 kb of each other in a genome, it is unlikely that this region of DNA would be represented in the library. Nearly half of the 48 clones analysed had additional *Bam*HI sites within the insert fragment as determined by RFLP with *Bam*HI, which indicates that the cloned fragments were in fact incompletely digested, or alternatively had been partially protected from cutting by methylation. Two end-sequenced clones, GSC1 and GSC10, appeared to be identical. Also, GSC357 and GSC1346 that hybridized with the *pmoA* appeared to be the same fragment from similar strains of a bacterial species, which suggests that the *Bam*HI library was somewhat biased towards this DNA fragment. The bias could be decreased by increasing or decreasing the size of the ligated ^{13}C-DNA fragments, using an enzyme, such as *Sau*3A1, that cuts the average DNA more frequently than *Bam*HI, or finally by making several libraries, each using a different restriction enzyme.

In summary, using SIP-generated ^{13}C-DNA as a source of genomic DNA from environmental samples provides access to the genomes of bacteria in the environment that are involved in specific metabolic processes. With continued advances in genomics and DNA sequence analysis it may soon be possible to reconstruct the complete genomes of microbial populations and consortia directly from the environment (Venter *et al.*, 2004) thereby overcoming the potential problems of "unculturability" of certain microbes. DNA-SIP offers a way of targeting specific metabolic processes in the environment and generating DNA sequences highly

enriched in certain physiological groups of bacteria which contain sets of genes involved in that metabolic process.

Chapter 6

Cloning, Sequence Analysis and Attempted Marker-Exchange Disruption of the Lon Protease in *Ms. trichosporium* OB3b

6.1 Introduction

A recombinant expression system for sMMO is essential for site-directed mutagenesis studies that can be used to ascertain the precise mechanism of sMMO catalysis. This will be invaluable for directed evolution experiments aimed at tailoring the enzyme for the industrial synthesis of fine chemicals (see Section 1.7). Although there have been considerable efforts to express active sMMO in heterologous hosts (Murrell *et al.*, 2000), the homologous *Ms. trichosporium* Mutant F sMMO expression system remains the only high-level and stable system (Lloyd *et al.*, 1999b; Murrell *et al.*, 2000). A site-directed mutagenesis study of two residues in the active site of the enzyme was performed using the Mutant F system (Smith *et al.*, 2002) which illustrated its value for studying the roles of individual aminoacyl residues, and also demonstrated its potential as an expression system for directed evolution experiments.

In the mutagenesis study by Smith and co-workers (2002), the T213S mutant was purified to homogeneity and characterized *in vitro*. The other three mutant enzymes (C151E, C151Y and T213A) were present at levels too low to be detected by SDS-PAGE after cell breakage, even though both the C151E and T213A exhibited activity in whole cell sMMO assays. Apparently, expression of the mutant enzymes was unstable, or the enzymes were highly susceptible to proteolytic degradation upon cell breakage. In this study (Chapter 3), four new sMMO mutants were created by making chimeras of the sMMO α-subunit with that of the alkene monooxygenase sequence, and none of the modified enzymes could be detected by SDS-PAGE after expression in the Mutant F system. Like the site-directed C151 mutants, these mutants appeared to be either more susceptible to proteolysis, or transcription from the *mmoX* promoter was attenuated.

This study was intended to investigate the effect of expressing recombinant mutagenized sMMO in a protease-deficient version of the homologous *Ms. trichosporium* expression system (Lloyd *et al.*, 1999b; Smith *et al.*, 2002). This would be done by disrupting the *lon* protease gene of *Ms. trichosporium* Mutant F. The Lon protease is a member of the AAA^+ superfamily (ATPases associated with diverse cellular activities) of ATPases and homologues have been identified in all organisms from all domains of life (Latterich & Patel, 1998). Lon is believed to be responsible for half the proteolysis in the *E. coli* cytoplasm (Maurizi, 1992) and *E.*

coli expression strains deficient in *lon* are believed to exhibit increased yields of some recombinant proteins (Gottesman, 1990; Murby *et al.*, 1996).

Lon mutants were first described in *E. coli* and were found to stabilize the expression of nonsense and mis-sense β-galactosidase (Bukhari & Zipser, 1973; Gottesman & Zipser, 1978). The *lon* gene is expressed from a heatshock promoter and Lon degrades misfolded cytoplasmic proteins (Gottesman, 1996; Rosen *et al.*, 2002). The Lon protease not only degrades mis-folded proteins, and over-expression of plasmid encoded *lon* compromises growth of *E. coli* by increasing the degradation of normal proteins (Goff & Goldberg, 1987). Lon is also involved in regulating gene expression in *E. coli* through the degradation of regulatory proteins. The best studied example is the cell division inhibitor SulA, which is expressed in response to DNA damage (Mizusawa & Gottesman, 1983; Nishii *et al.*, 2002). *E. coli lon* mutants are therefore more sensitive to DNA damaging agents. Lon also degrades the RcsA protein, responsible for activating the capsule biosynthesis genes (Torres-Cabassa & Gottesman, 1987).

Several *lon* mutants of other bacterial species have been isolated or created. Like *E. coli lon* mutants, some of these strains are also more susceptible to DNA damage indicating a conserved role of Lon in regulating the SOS response (Bretz *et al.*, 2002; Stewart *et al.*, 1997; Whistler *et al.*, 2000). In all cases where characterized, Lon appears to be a heatshock protein and therefore is likely to have a universal role in degrading misfolded proteins. Other roles of Lon include cell differentiation and flagellar synthesis in *Vibrio parahaemolyticus* (Stewart *et al.*, 1997), activation of type III secretion in *Pseudomonas syringae* (Bretz *et al.*, 2002), antibiotic production in *Pseudomonas fluorescens* Pf-5 (Whistler *et al.*, 2000), exopolysaccharide production and nodulation of alfalfa by *Sinorhizobium meliloti* (Summers *et al.*, 2000), carbon monoxide utilization by *Oligotropha carboxidovorans* (Santiago *et al.*, 1999), expression of invasion genes in *Salmonella enterica* (Takaya *et al.*, 2002; Takaya *et al.*, 2003), iron uptake in *Azospirillum brasilense* (Mori *et al.*, 1996), cell cycle progression of *Caulobacter crescentus* (Wright *et al.*, 1996) and sporulation of *Bacillus subtilis* (Schmidt *et al.*, 1994). The only report of the Lon protease being indispensable for vegetative growth is in *Myxococcus xanthus*, which has two copies of Lon; one of which one is necessary for morphological development

(Gill *et al.*, 1993; Tojo *et al.*, 1993a) and the other for vegetative growth (Tojo *et al.*, 1993b). The mitochondrial Lon protease is also essential (Suzuki *et al.*, 1994).

The aim of this study was to attempt to create a Lon-protease-deficient mutant of *Ms. trichosporium* for expression of recombinant sMMO.

6.2 Cloning the *Ms. trichosporium* OB3b *lon* gene

The *Ms. trichosporium* OB3b *lon* gene was identified by Southern blotting using a gene probe from *Sinorhizobium meliloti* 1021. A 1.7 kb segment of the *S. meliloti lon* was amplified by PCR using primers 604F (5'-TTCATGGAAGGCGAGATTTCG-3') and 2261R (5'-TTGACGTTGTCAGGGATGTCG-3') designed using the genome sequence data from http://sequence.toulouse.inra.fr/meliloti.html (Galibert *et al.*, 2001), and labelled by random priming (Section 2.10.3). The *S. meliloti lon* gene was chosen as a probe since *S. meliloti* and *Ms. trichosporium* OB3b are both in the α-*Proteobacteria* and an *S. meliloti* probe was successful in identifying the *rpoN* gene in *Ms. trichosporium* OB3b (Stafford *et al.*, 2003). *Ms. trichosporium* OB3b genomic DNA was digested with several restriction enzymes (Figure 6.1) and hybridized with the *S. meliloti lon* gene probe by Southern blotting (Figure 6.2). Signals corresponding to a 2.5 kb *Bgl*II fragment and a 3.5 kb *Sst*I fragment were identified as good candidates for cloning. *Ms. trichosporium* OB3b DNA was digested with these enzymes, size fractionated by agarose gel electrophoresis, and the gel-purified fragments were cloned into the pUC19 multiple cloning site; the *Bgl*II fragments were ligated into the *Bam*HI site of the vector. A library of 223 *Bgl*II and 400 *Sst*I *E. coli* TOP10^{TM} clones were screened for *lon* by colony hybridization (Section 2.10.2) using the *S. meliloti* probe (Figure 6.3). Five *Bgl*II clones (numbers 23, 28, 61, 99 and 193) hybridized strongly and a restriction map indicated that there were two different cloned fragments (Figure 6.4). Clones 28, 61 and 193 are identical (restriction pattern not shown for clone 61), and clones 23 and 99 are the same. Analysis of the restriction map and comparison with the Southern blot suggested that the *Bgl*II enzyme cut within the *lon* gene generating two fragments of approximately equal size; this was confirmed by sequencing, as described in the following section.

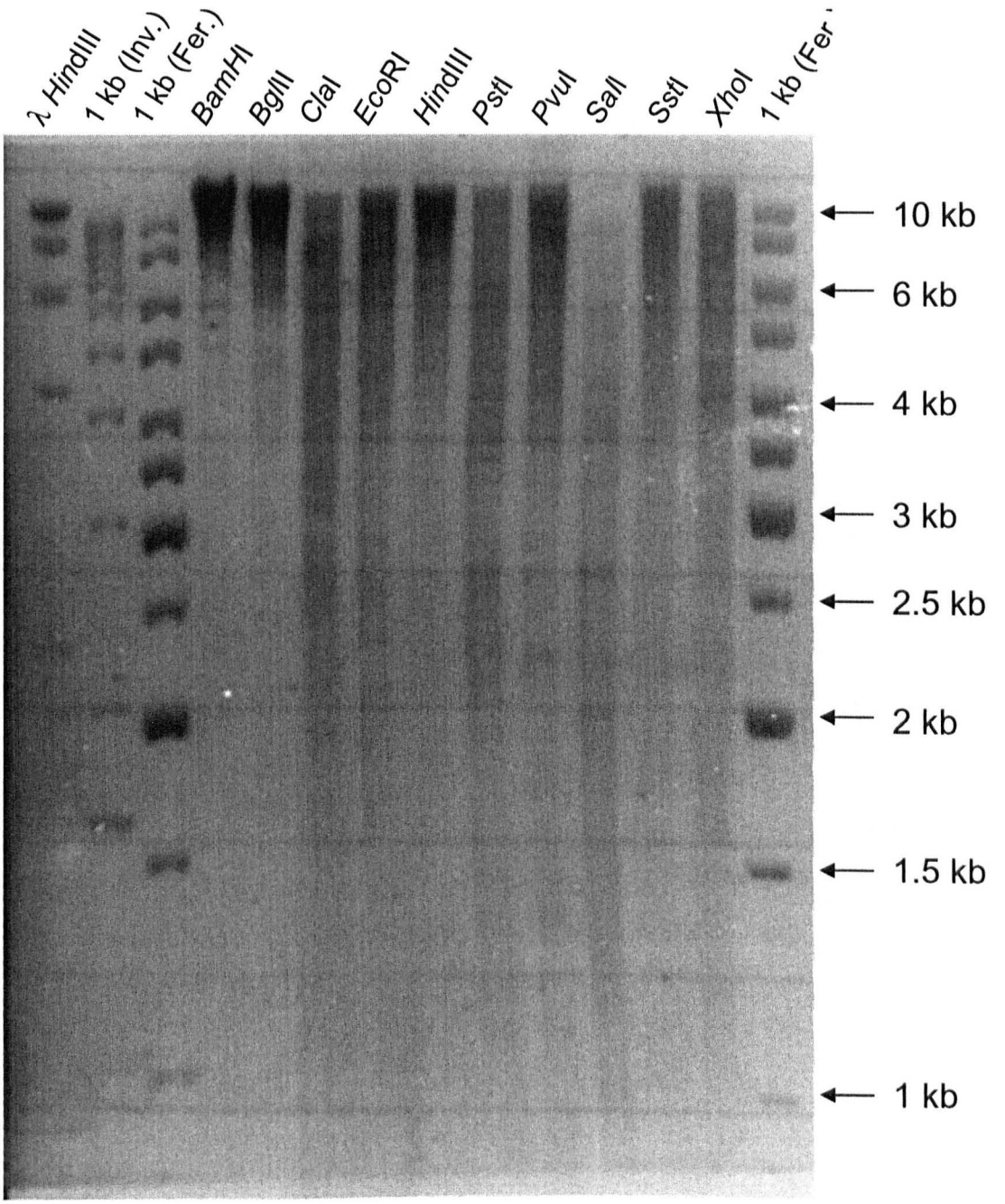


Figure 6.1 Agarose gel of *Ms. trichosporium* OB3b genomic DNA digested with various restriction enzymes. Lane 1, *Hin*dIII digested λ DNA; lanes 2 & 14, 1 kb ladder (Invitrogen); lane 3, 1 kb ladder (Fermentas); lanes 4 to 13, *Ms. trichosporium* OB3b DNA digests. The DNA was blotted and probed with the *lon* gene of *S. meliloti* strain 1021 (Figure 6.2).

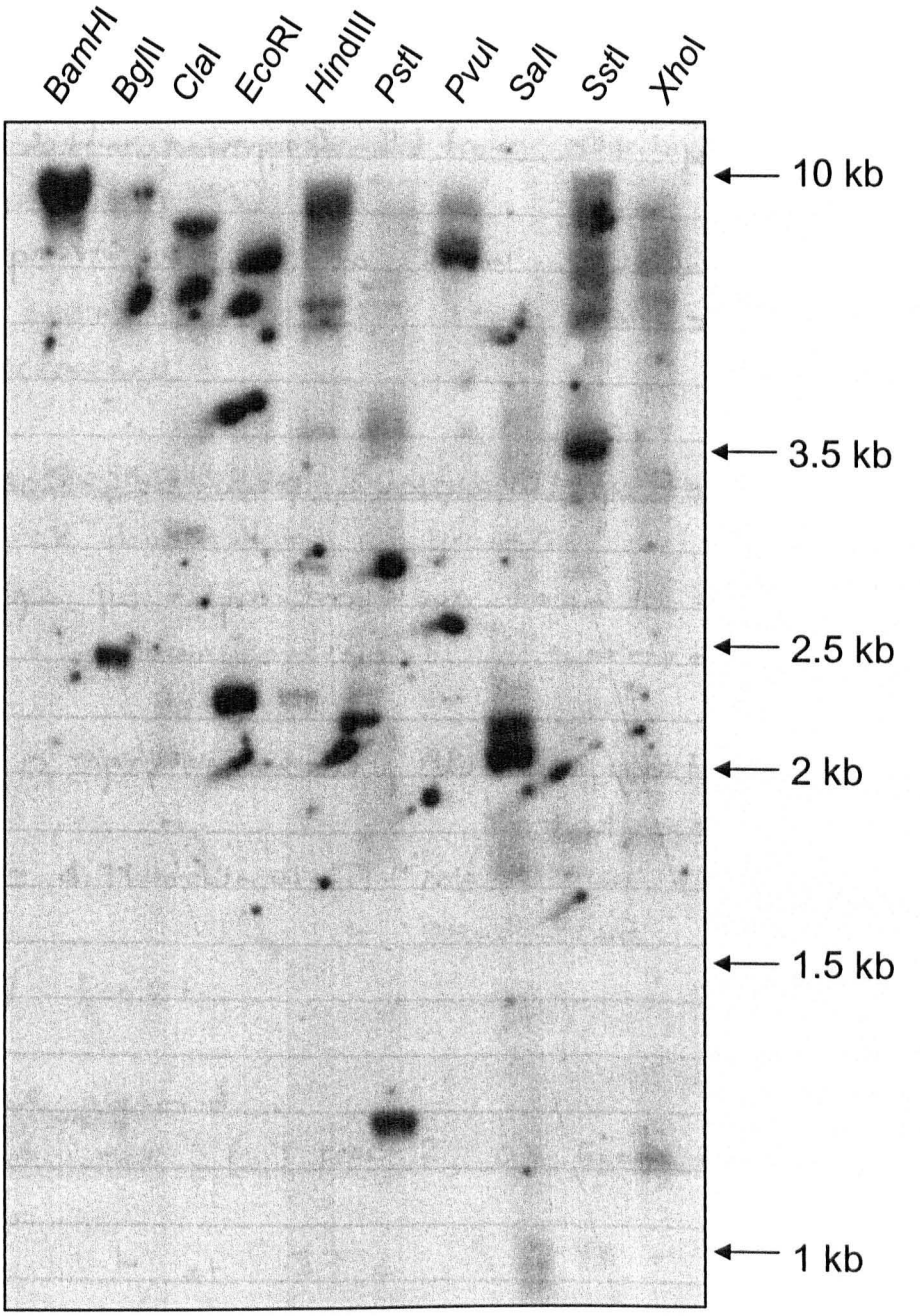


Figure 6.2 Southern blot of *Ms. trichosporium* OB3b DNA (Figure 6.1) hybridized with a *lon* gene fragment of *Sinorhizobium meliloti* 1021 (see text for details).

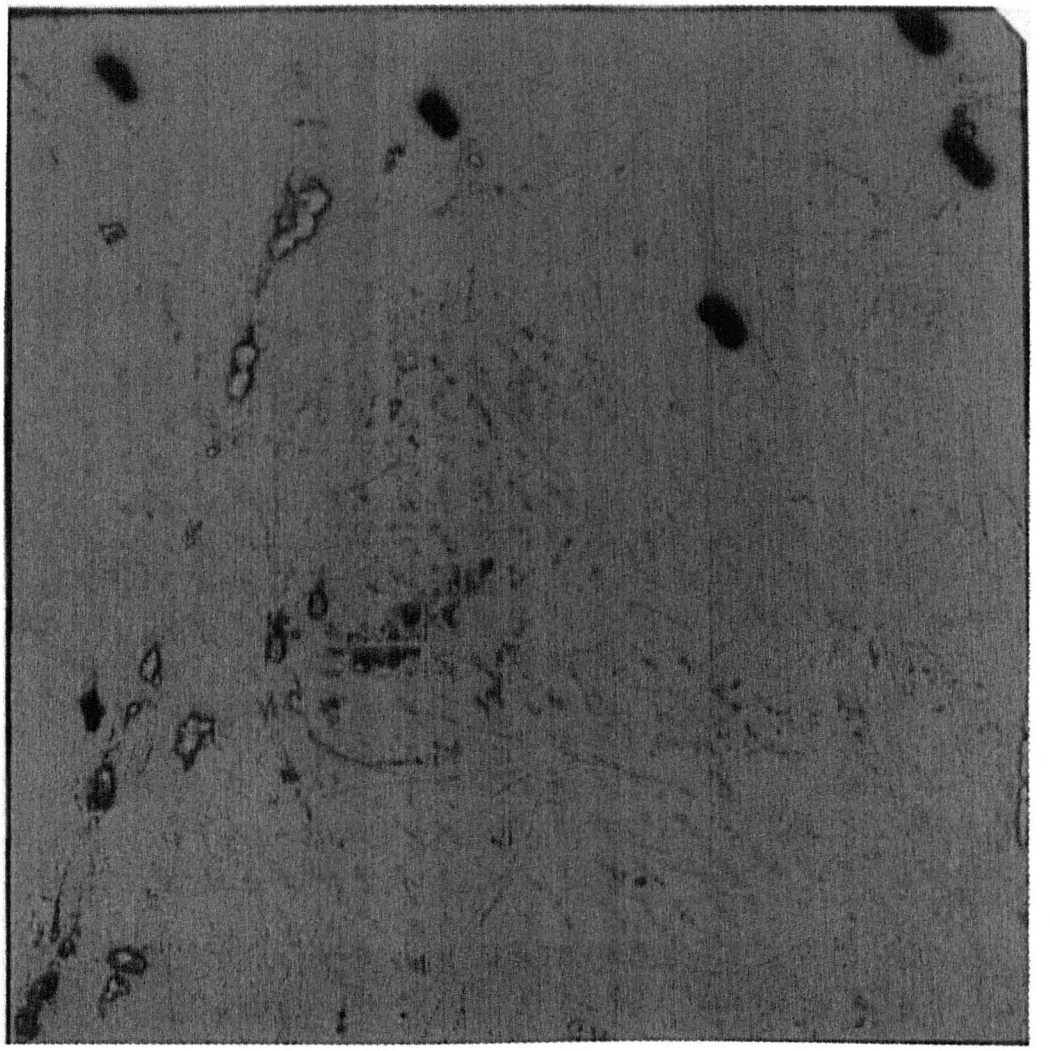


Figure 6.3 Colony blot of 223 *Bgl*II and 400 *Sst*I clones containing *Ms. trichosporium* OB3b genomic fragments, hybridized with a *Sinorhizobium meliloti* 1021 *lon* gene fragment. Five *Bgl*II fragments hybridized strongly with the probe and these were shown by DNA sequencing to contain the *Ms. trichosporium* OB3b *lon* gene.

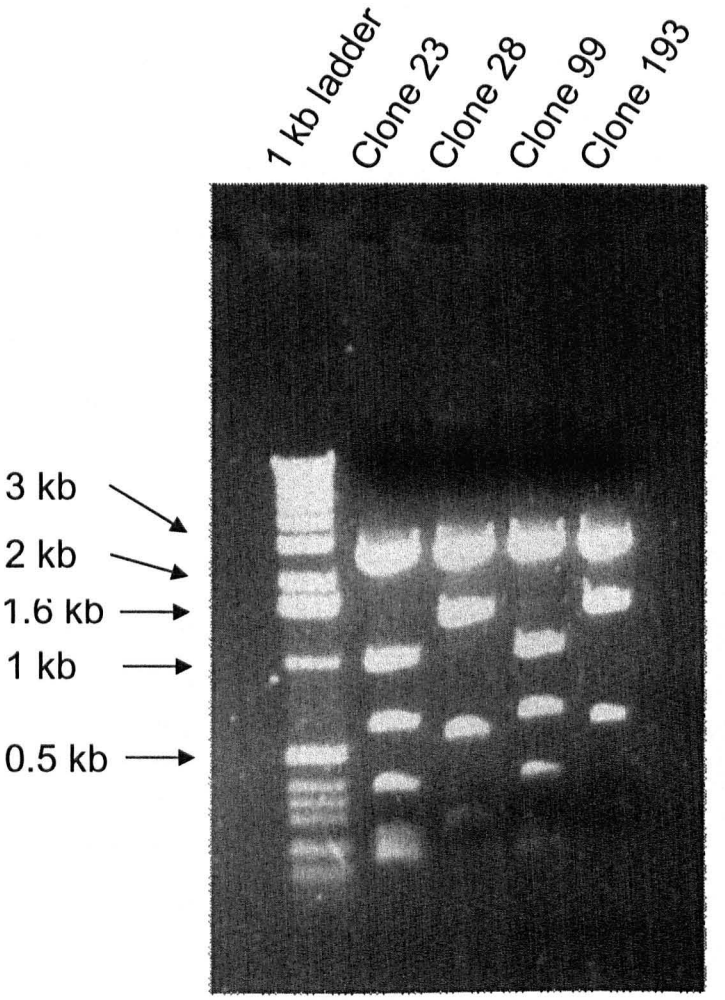


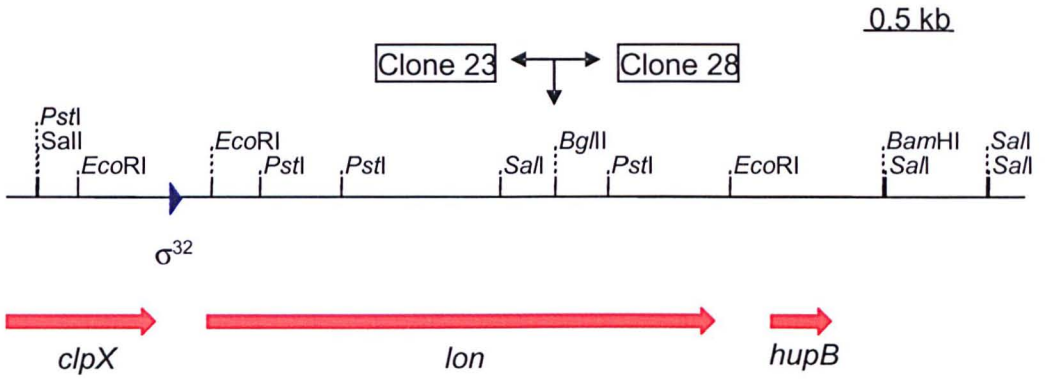
Figure 6.4 *Pst*I and *Eco*RI double digest of the *Ms. trichosporium* OB3b *lon* clones. The restriction enzymes were chosen since they cut on either side of the *Bam*HI site of the vector. The banding patterns resulting from *Pst*I and *Eco*RI sites within the inserts suggested that there were two different inserts.

6.3 *Ms. trichosporium* OB3b *lon*, nucleotide sequence analysis

Clones 23 and 28 were sequenced as described in Section 2.15. The sequencing indicated that a *Bgl*II site was present within the *Ms. trichosporium* OB3b *lon* gene and that clone 23 contained the 5'–end, and clone 28 the 3'–end, of the gene (Figure 6.5). Clone 23 was 2.6 kb and clone 28 was 2.2 kb in length. It is not clear why only one *Bgl*II band was visible in the Southern blot (Figure 6.2). It is possible that the 2.2 kb fragment (clone 28), which overlaps less with the probe than the 2.6 kb fragment (clone 23), did not hybridize sufficient probe to produce a signal in the Southern hybridization. Perhaps it was possible to detect the 2.2 kb fragment (clones 28, 61 and 193) in the colony blot due to the high plasmid copy number.

The derived aminoacyl sequence of the *Ms. trichosporium* OB3b Lon is 801 residues in length and shares 75 % identity with the *S. meliloti* 1021 Lon and 60 % identity with the *E. coli* Lon, which are shown as an alignment in Figure 6.6. Currently, the closest Genbank match of *Ms. trichosporium* OB3b Lon is the sequence from *Rhodopseudomonas palustris* CGA009, also a member of the α*–Proteobacteria.* The Walker A and B motifs associated with ATP binding and hydrolysis (Walker *et al.*, 1982) are present, as is the proteolytically active serine residue (Botos *et al.*, 2004), as indicated in Figure 6.6.

The *clpX* and *hupB* genes flank *lon* in the *Ms. trichosporium* OB3b genome in the same arrangement as in the *E. coli* genome, which is additional proof that this is indeed the *Ms. trichosporium* OB3b *lon* gene. In *E. coli* and other bacterial genomes, *clpP* lies upstream of *clpX*, and the ClpXP is a serine protease, not unlike Lon in both structure and function (Gottesman, 1996; Wickner & Maurizi, 1999). 706 bp of the *Ms. trichosporium* OB3b *clpX* gene was present on the clone, and the deduced amino acid sequence (234 residues) shares 68 % identity with carboxy terminus of the *E. coli* ClpX protein. The *hupB* gene encodes HU-1, a small histone-like DNA binding protein (Goshima *et al.*, 1990).



▸ σ^{32}: RpoH recognition sequence of a putative heatshock promoter:

Sequence:	CAGCGTTCTTGAAGTGCGGCGCCGCCAGGCCCATTTG
E. coli consensus (-35, -10):	TNTCNCCCTTGAA CCCCATTTA

Figure 6.5 The genetic map of the *lon* clone (4835 bp) of *Ms. trichosporium* OB3b. The contiguous sequence was formed by joining Clone 23 and Clone 28 at the *Bgl*II site, as indicated. 706 bp of the *clpX* gene is present. Between the *clpX* and the *lon* gene is a putative σ^{32} consensus sequence, shown in the bottom panel, which potentially controls the expression of the *Ms. trichosporium lon*. Downstream (3') of *lon* is the *hupB* gene.

Figure 6.6 Alignment of the derived *Ms. trichosporium* OB3b Lon sequence (MtLON) with the Lon sequences from *S. meliloti* 1021 (SmLON; Genbank accession O69177) and *E. coli* K12 (EcLON; Genbank accession NP_414973). The active site serine residue is indicated with an arrow and the Walker A an B regions are overlined.


Walker A
Walker B

6.4 Attempted complementation of *E. coli lon*⁻ strain with the *Ms. trichosporium* OB3b *lon* gene

In order to prove that the protein encoded by the *Ms. trichosporium* OB3b *lon* gene cloned and sequenced in this study was indeed homologous in function to the *E. coli* Lon protease, the complementation of a Lon deficient *E. coli* strain B (SaiSree *et al.*, 2001) was attempted. Unfortunately, the PCR amplification of the *lon* gene from the *Ms. trichosporium* OB3b genome was not successful. Primers 23F1: 5'- AGT TTC TGC AGG TCG ACA CG -3' and 28R1: 5'- CAT GCC GAG CTT GCG CAG C -3' were used, and the PCR was performed using *Pfu*Turbo polymerase since it is has greater fidelity than *Taq* polymerase, and can generate products as large as 10 kb. The PCR conditions used were 95 °C for 1 minute, 50 °C for 1 minute and 72 °C for 5 minutes, over 30 cycles. A product of the correct size was not obtained (results not shown). Amplification was attempted using a touchdown PCR programme (Don *et al.*, 1991), but also failed to generate the desired product. More optimization was required, or perhaps different PCR primers would have been effective. The experiment was abandoned due to a lack of time.

If the PCR of the *lon* gene had been successful, the *Ms. trichosporium* OB3b *lon* gene would have been expressed in *E. coli* BLR, either from its native promoter, or using an *E. coli* expression vector. A UV sensitivity test would have been performed to test if *Ms. trichosporium* OB3b Lon restored the wild-type *E. coli* phenotype.

6.5 Analysis of the putative *Ms. trichosporium* OB3b *lon* promoter

The presence of a putative σ^{32} consensus sequence approximately 150 bp upstream of the *lon* start codon (Figure 6.5) suggests that the *Ms. trichosporium* OB3b Lon is a heatshock protein, as is the *E. coli* Lon (Gottesman, 1996; Phillips *et al.*, 1984). The putative σ^{32} (RpoH) consensus sequence is a relatively strong match with that from *E. coli*, with the -10 and -35 having the correct spacing, and differing in only 6 of 23 nucleotide residues (Cowing *et al.*, 1985).

A test to indicate if the *Ms. trichosporium* OB3b Lon is up-regulated in response to heatshock would be to express GFP under the control of the *lon* promoter, using the broad host range promoter-trap plasmid pBBR1-GFP (Ouahrani-Bettache *et al.*, 1999). The region upstream of *Ms. trichosporium* OB3b *lon* containing the putative σ^{32} consensus sequence was amplified by PCR using primers s32F-*Bgl*II (5'- ATA ATA AGA TCT TGA CGC GAA TTT CGG CAG G -3') and s32R-*Bam*HI (5'- ATA ATA GGA TCC CAT TTC GTG TAT CCT GTT CC -3'). This experiment was not completed. Depending on the strength and the level of inducibility of this promoter, it could be useful in creating an expression vector for *Ms. trichosporium* OB3b. An expression vector for *Ms. trichosporium* OB3b does not currently exist and is much needed for ongoing experiments in the JCM laboratory.

6.6 Attempted marker-exchange disruption of the *Ms. trichosporium* OB3b *lon* gene

6.6.1 Plasmid constructs

Marker-exchange mutagenesis of the *Ms. trichosporium* OB3b *lon* gene was attempted using the pK18 system (Schäfer *et al.*, 1994). A *Cla*I site was identified in the *lon* gene segment in Clone 23, which appeared to be at a suitable site to introduce an antibiotic resistance cassette. A second restriction enzyme, that would have enabled the deletion of a portion of the gene, was not readily available. However, the interruption of the gene, especially if the antibiotic cassette was introduced in the opposite orientation to *lon*, should be sufficient for disruption. Clone 23 contained approximately 1 kb and 1.5 kb of genomic sequence flanking the *Cla*I site, which is generally sufficient to obtain a homologous recombination event.

The pLONDEL-mob plasmid, designed to disrupt the chromosomal *lon* gene, was constructed using Clone 23, pK18mob and p34S-Gm (Figure 6.7). Clone 23 was digested with *Kpn*I, *Sph*I, *Cla*I, and the two insert fragments were purified. The gentamycin resistance gene was released from p34S-Gm with *Acc*I, which has cohesive ends compatible with *Cla*I. pK18mob was linearized with *Kpn*I, *Sph*I and agarose gel purified. To prevent ligation of the *Cla*I overhangs, the Clone 23 DNA fragments were dephosphorylated; the other molecules were not dephosphorylated, which would have prevented ligation of the molecules. The four DNA fragments were ligated to form pLONDEL-mob, and the proper assembly was confirmed by restriction mapping (Figure 6.8). A version of the construct was also made using the pK18mobsacB backbone by transferring the insert as an *Apa*LI-*Hin*dIII fragment (results not shown). This plasmid, pLONDEL-mobsacB, contained the *sacB* gene, which confers sucrose sensitivity, and therefore allows selection for double recombinants on growth medium containing sucrose (Schäfer *et al.*, 1994).

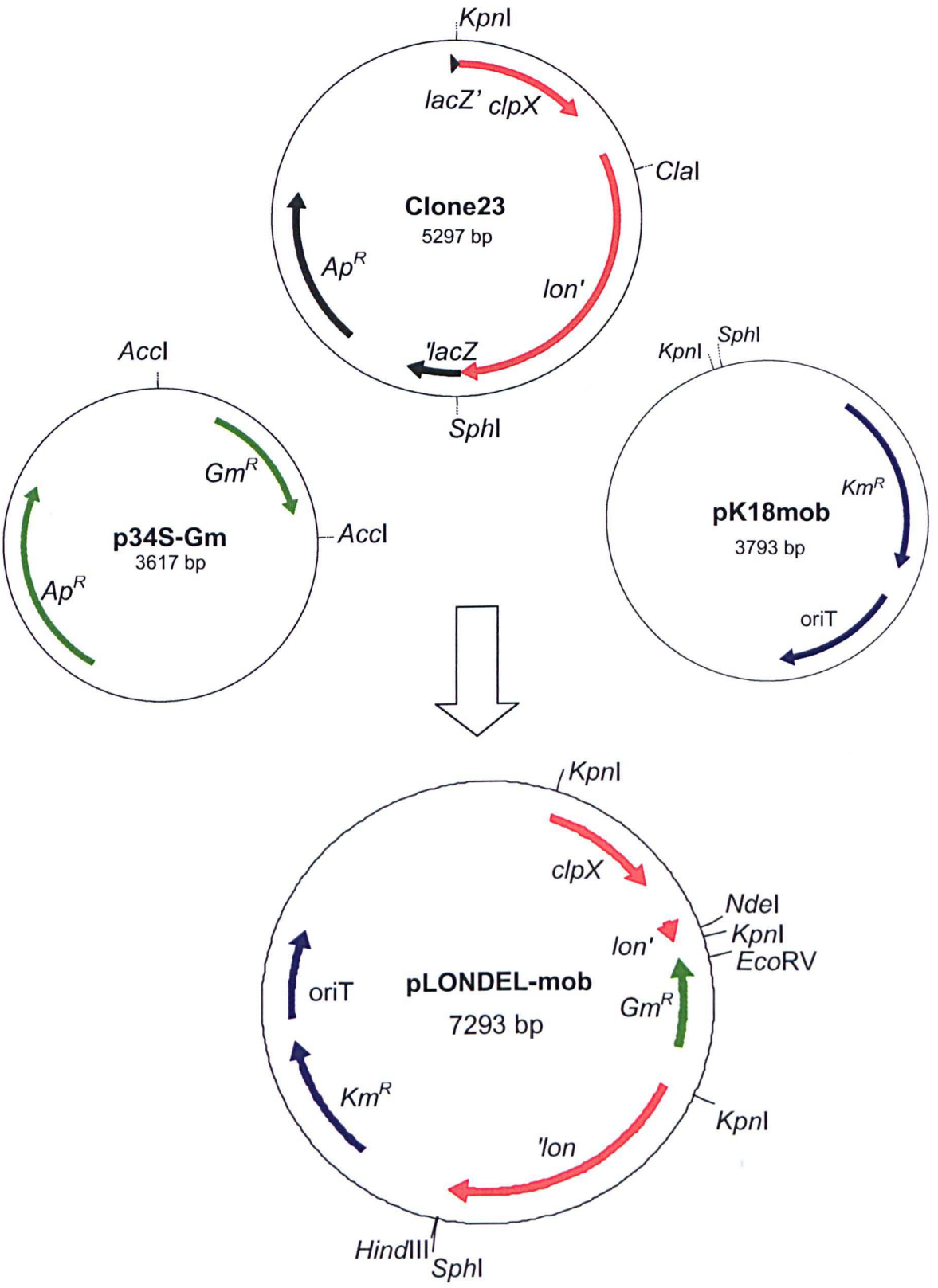


Figure 6.7 Construction of pLONDEL-mob, a suicide vector for disruption of the chromosomal *lon* gene with a gentamycin resistance cassette. Refer to the text for details. A second version of this plasmid (pLONDEL-mobsacB) containing *sacB* on the vector backbone (Schäfer *et al.*, 1994), was also created.

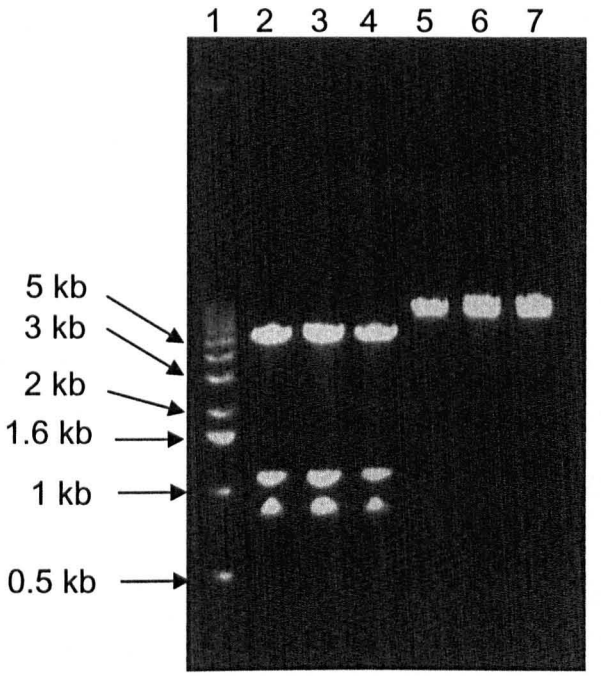


Figure 6.8 Restriction map of three pLONDEL-mob clones. The clones were analysed to confirm the correctness of the ligation and to determine the orientation of the gentamycin resistance gene cassette. Lane 1, 1 kb DNA marker (Invitrogen); lanes 2 – 4 are the clones cut with *Kpn*I; lanes 5 – 7 are the clones cut with *Nde*I & *Eco*RV. The apparent molecule sizes of the *Kpn*I fragments correspond well with the predicted sizes of 847 bp, 1,110 bp & 5,337 bp. Had the orientation of the of the gentamycin resistance gene been the reverse of that shown in Figure 6.5, a band of 796 bp would be seen in the *Nde*I/ *Eco*RV digest. The orientation of the gentamycin resistance gene was also confirmed by *Hin*dIII/ *Eco*RV mapping (not shown).

6.6.2 Conjugation and selection of recombinants

The pLONDEL-mob construct was conjugated into *Ms. trichosporium* OB3b as described in Section 2.7.1, and transconjugants, either single or double recombinants, were isolated on NMS agar plates containing gentamycin. 200 colonies were subcultured by patching onto NMS agar plates containing 5 $\mu g\ ml^{-1}$ gentamycin. After subculturing, the transconjugants were tested for sensitivity to 12.5 $\mu g\ ml^{-1}$ kanamycin, which would suggest a loss of the plasmid backbone, by patching onto NMS agar plates containing both gentamycin and kanamycin selection. All the clones grew in the presence of kanamycin.

These results indicated that a double recombination event was relatively rare, if indeed the *lon* gene could be disrupted. To increase the ease of isolating a double recombinant (*lon*-) strain, the pLONDEL-mobsacB construct was used. This plasmid confers sensitivity to sucrose and therefore a double recombinant can be selected with gentamycin and sucrose. The plasmid was conjugated into *Ms. trichosporium* Mutant F and transconjugants were isolated on NMS agar containing gentamycin and kanamycin. A transconjugant was grown in NMS batch culture (containing gentamycin/ kanamycin) to an OD_{540} of 0.7 (approximately $2 \times 10^8\ ml^{-1}$). A dilution series of the culture was plated onto NMS agar containing 10 % sucrose, and incubated with methane until colonies appeared. A very low frequency for loss of sucrose sensitivity was observed, with 45 colonies appearing on the NMS plate spread with 100 µl of undiluted culture, 3 colonies on the plate spread with 100 µl of a 10^{-1} dilution, and no colonies appearing on plates spread with more dilute cell suspensions. The colonies were subcultured onto NMS gentamycin/ kanamycin agar plates[5], with, and without, the addition of 10 % sucrose. The cells grew less quickly on sucrose-containing medium and microscopically the cells were pleiomorphic and generally elongated, but exhibited wild-type cell morphology when grown in the absence of sucrose. As a control, *Ms. trichosporium* OB3b was grown on NMS in the presence of 10 % sucrose, and when observed under phase-contrast microscopy, the cells did not have the elongated cell shape observed with pLONDEL-mobsacB transconjugants. These results suggest that the sucrose-resistant colonies were a result of a partial loss in sucrose-sensitivity and not from a loss of the entire vector

[5] *Ms. trichosporium* mutant F contains resistance to kanamycin, and therefore resistance to kanamycin in this situation does not indicate the presence of the pK18 vector within the genome

sequence. To confirm if the pK18mobsacB backbone was still present in the chromosome of the clones, they were screened by colony hybridization by probing with pK18mobsacB labelled DNA. *Ms. trichosporium* Mutant F was included as a control. All the colonies hybridized with the probe, but the results were inconclusive because the *Ms. trichosporium* Mutant F control also hybridized (results not shown). The hybridization of *Ms. trichosporium* Mutant F probably resulted from the kanamycin resistance cassette in the *mmoX* gene. The experiment would need to be repeated by probing with a segment of the pK18mobsacB vector not containing kanamycin resistance gene sequence.

6.7 Discussion

The homologous expression of recombinant sMMO in *Ms. trichosporium* OB3b (Mutant F) (Lloyd *et al.*, 1999b) remains the only high-level sMMO expression system. This is true despite considerable effort to obtain sMMO expression in heterologous hosts, which seem incapable of folding or assembling the hydroxylase component (Murrell *et al.*, 2000). The Mutant F system has been used to express cysteine 151 and threonine 213 α-subunit mutants, created by site-directed mutagenesis (Smith *et al.*, 2002). Several of the mutants were unstable and the authors were unable to improve the expression by adding protease inhibitors. The instability of some sMMO mutants was also found in this study (discussed in Chapter 3) and has been observed in other experiments in the JCM lab (L. De Nagornoff, H. Dalton, J.C. Murrell, unpublished).

E. coli protein expression systems are by far the most common and best characterized, and there are many different strains and vectors available which can be evaluated for each expression experiment (Makrides, 1996). The natural Lon-deficient *E. coli* B strain and its derivatives (SaiSree *et al.*, 2001), including strains BL21 and BLR, are widely used *E. coli* expression hosts. The contribution of the Lon-minus phenotype to the effectiveness of these strains as expression hosts for heterologous proteins is unclear. It has been shown that *E. coli lon* mutants stabilize some temperature sensitive mutants, indicating that Lon is degrading active protein and that this activity is partly responsible for the phenotype (Gottesman, 1996; Grossman *et al.*, 1983) Some authors claim that *E. coli lon* mutants can improve yields of unstable recombinant proteins (Enfors, 1992; Meerman & Georgiou, 1994; Murby *et al.*, 1996), whereas others claim the improvements are minimal (Makrides, 1996). To test if the Lon protease was partly responsible for the degradation of unstable mutagenized-sMMO in *Ms. trichosporium* Mutant F, the disruption of the *lon* gene of *Ms. trichosporium* Mutant F was attempted in this study.

The *lon* gene in *Ms. trichosporium* OB3b was identified and found it to share a high degree of identity with those from other bacteria. The banding pattern in the hybridization experiment suggests that there is only one copy of the gene within the genome. The arrangement of genes next to *lon* on the chromosome is identical to that in many bacteria, including *E. coli* (Whistler *et al.*, 2000). In addition, a putative σ^{32} consensus sequence was identified upstream of *lon*, indicating that, like *E. coli*, the

Lon protease is probably a heatshock protein in *Ms. trichosporium* OB3b (Eastgate *et al.*, 1995; Phillips *et al.*, 1984).

Attempts were made to disrupt *lon* by insertion of an antibiotic resistance marker into the gene on the chromosome. The pK18 system (Schäfer *et al.*, 1994) was used since it has been utilized for allelic exchange in this *Ms. trichosporium* OB3b previously (Stafford *et al.*, 2003) and also in this study (Chapter 8). The pK18-mobsacB plasmid contains the *sacB* gene from *Bacillus subtilis*, which confers sucrose sensitivity which can be used as positive selection for double recombination events. In other marker-exchange mutagenesis studies with *Ms. trichosporium* OB3b, it has not been necessary to select with sucrose since double recombinants arose spontaneously and at high frequency. Sucrose resistant colonies were obtained using the pK18mobsacB-derivative; however they appeared to be SacB mutants, indicating that the frequency of SacB-minus emergence was greater than the frequency of the double-crossover event. This result suggests that the Lon protease may be essential for vegetative growth of *Ms. trichosporium* OB3b under the growth conditions used, which is only the second report of an inability to isolate a bacterial *lon* null-mutant after that of *Myxococcus xanthus* (Tojo *et al.*, 1993b).

The Lon protease has numerous functions in biological systems. Apart from the degradation of abnormal proteins, there is evidence that Lon specifically cleaves numerous transcriptional regulators in bacteria and therefore indirectly regulates gene transcription. The best studied example is the cell division inhibitor SulA from *E. coli* (Canceill *et al.*, 1990). The wide range of Lon protease functions in biological systems makes it difficult to speculate why Lon may be indispensable in *Ms. trichosporium* OB3b. The possibility that *Ms. trichosporium* OB3b contains a SulA protein that is expressed under the growth conditions used in this experiment has been considered. SulA is expressed in *E. coli* and some other bacteria in response to DNA damage. Methanotrophs may be exposed to high levels of mutagens, including formaldehyde from the oxidation of methanol (Anthony, 1986) and hydroxylamine from the co-oxidation of ammonia by the MMO enzyme (Bédard & Knowles, 1989). However, the intracellular levels of these metabolites are likely kept very low by the activity of the formaldehyde metabolism pathways (Lidstrom, 2001) and the hydroxylamine oxidoreductase enzyme (Zahn *et al.*, 1994). In addition, there is no evidence that *Ms. trichosporium* OB3b possesses a SulA protein, which is common mainly to members of the *Enterobacteriaceae*, and is not present in the genome of the

α-*Proteobacterium S. meliloti* 1021 (Galibert *et al.*, 2001; Summers *et al.*, 2000) nor was it found in a search through the genome sequence of the methylotroph *Methylobacterium extorquens* AM1 (http://pedant.gsf.de/cgi-bin/wwwfly.pl?Set=Mextorquens&Page=index).

It is possible that the inability to obtain a *lon*⁻ mutant of *Ms. trichosporium* OB3b was specific to the growth conditions used in this study. However, due to the restricted growth requirements of methanotrophs, there are relatively few growth variables that can be modified to test this possibility. Theoretically, if the Lon protease is necessary for expression of the pMMO enzyme, it could be possible to obtain a *lon*⁻ mutant under low copper: biomass growth conditions, which favour sMMO expression, or during growth on methanol. Other variables that could be changed include nitrogen source and growth temperature (Bowman *et al.*, 1993). It would be interesting to test if the Lon indispensability trait is common to other methylotrophs, or if it is limited to *Ms. trichosporium* OB3b. If the Lon protease is required for the regulation of genes involved in methylotrophy, it may be possible to isolate a *lon*⁻ mutant of the facultative methylotroph *Methylobacterium extorquens* AM1 during growth on glucose.

Chapter 7

Investigation into the Role of *mmoG* in sMMO Expression

7.1 Introduction

Transcription of the sMMO structural genes in *Methylosinus trichosporium* OB3b is RpoN mediated from a σ^{54} promoter upstream of the *mmoX* gene (Nielsen *et al.*, 1997; Stafford *et al.*, 2003). Transcription from the promoter only occurs at low copper-to-biomass growth conditions (Murrell *et al.*, 2002).

The sequence of the σ^{54}-dependent transcriptional regulator responsible for activating transcription from the *mmoX* σ^{54} promoter has recently been described (Stafford *et al.*, 2003). This gene, termed *mmoR*, was disrupted by marker-exchange mutagenesis and shown to be essential for expression of sMMO. A gene designated *mmoG* is situated immediately downstream (3') of *mmoR* and was also shown to be essential for sMMO expression. Analysis of the MmoG sequence indicates that it is a member of the GroEL family of chaperonins. Homologues of *mmoR* and *mmoG* in *Methylococcus capsulatus* (Bath) were analysed and also found to be essential for sMMO expression in this organism (Csáki *et al.*, 2003).

The potential mechanism of MmoG involvement in sMMO expression is intriguing. The initial results suggest that MmoG is necessary for transcription of sMMO structural genes from the *mmoX* σ^{54} promoter and therefore may be involved in assembling a component of the transcription machinery (Csáki *et al.*, 2003; Stafford *et al.*, 2003). To our knowledge this is the only report of a chaperonin functioning in gene regulation.

Since the α–, β– and γ–subunits of the sMMO hydroxylase are not assembled in heterologous hosts (Section 1.8), it is suspected that an sMMO-specific chaperone is present in methanotrophs and assembles the hydroxylase component (Murrell *et al.*, 2000). Although the evidence available to date indicates that the MmoG is involved in the transcription of sMMO genes, it was initially hypothesized to be the hydroxylase assembling chaperone. The purpose of this study was to determine if indeed the MmoG is required for the assembly or maturation of the sMMO enzyme of *Methylosinus trichosporium* OB3b.

7.2 Sequence analysis

The *mmoG* gene of *Ms. trichosporium* OB3b is situated immediately upstream (5') of the *mmoX* gene (Figure 1.8), and preliminary sequence analysis has been performed (Stafford *et al.*, 2003). The deduced MmoG peptide sequence indicated that it is 581 amino acids in length, which is considerably longer than most GroEL proteins and longer than the inferred length of the MmoG of *Methylococcus capsulatus* (Bath), which is 559 residues (Csáki *et al.*, 2003), and the MmoG of *Methylocella silvestris* BL2^T (unpublished data; see Chapter 9), which was found to be 558 residues. The additional peptide sequence in the *Ms. trichosporium* OB3b MmoG occurs at the amino terminus of the protein, corresponding to the 5' –end of the gene.

Several of the sMMO genes of *Methylosinus sporium*, including *mmoG*, have recently been sequenced in our laboratory by Hanif Ali (unpublished data). Since *Ms. sporium* is closely related to *Ms. trichosporium* OB3b, the *mmoG* nucleotide sequences were compared to determine if the *mmoG* ribosome binding site and the start codon were conserved. The proposed *mmoG* ribosome binding site in *Ms. trichosporium* OB3b is 19 bp upstream (5') of the ATG codon (Stafford *et al.*, 2003), which is quite distant considering most are situated 5 to 13 bp upstream (Ma *et al.*, 2002). The alignment showed that the proposed ATG start codon for the *mmoG* of *Ms. trichosporium* OB3b is not present in the *Ms. sporium* sequence (Figure 7.1). The only potential start codon present in both sequences is a GTG, which is also preceded by a putative ribosome binding site, six bp upstream in *Ms. trichosporium* OB3b and ten bp upstream in *Ms. sporium*. The derived aminoacyl sequence of MmoG using the GTG start codon is 560 residues in length for both the *Ms. trichosporium* OB3b and *Ms. sporium* proteins, which is closer to that of *Methylococcus capsulatus* (Bath) and *Methylocella silvestris* BL2^T. An alignment of the MmoG protein sequences indicates that the position of the GTG start codon corresponds with the ATG start codon of the *Methylococcus capsulatus* (Bath) and the *Methylocella silvestris* BL2^T MmoG proteins (Figure 7.2). There is a possibility that the GTG codons are coincident sequencing errors in both the *Ms. trichosporium* OB3b and *Ms. sporium* clones. If these are authentic translation start codons, this is the first evidence that translation initiation can occur at GTG in *Methylosinus*.

Figure 7.1 Alignment of the *mmoG* DNA sequences of *Ms. trichosporium* OB3b (OB3b) and *Ms. sporium* (sporium). Nucleotide sequence upstream and downstream of the *mmoG* open reading frames are included, and key the features of neighbouring genes are indicated. The green arrows indicate translational start codons and the red arrows indicate stop codons. The black dot indicates the position of the codon originally believed to be the start of the *Ms. trichosporium* OB3b *mmoG* gene. *orfR2* is a putative gene between *mmoR* and *mmoG*, which is discussed in Section 1.5.3.1. RBS, putative ribosome binding site; the -24, -12 consensus of the *mmoX* σ^{54} promoter is shown.

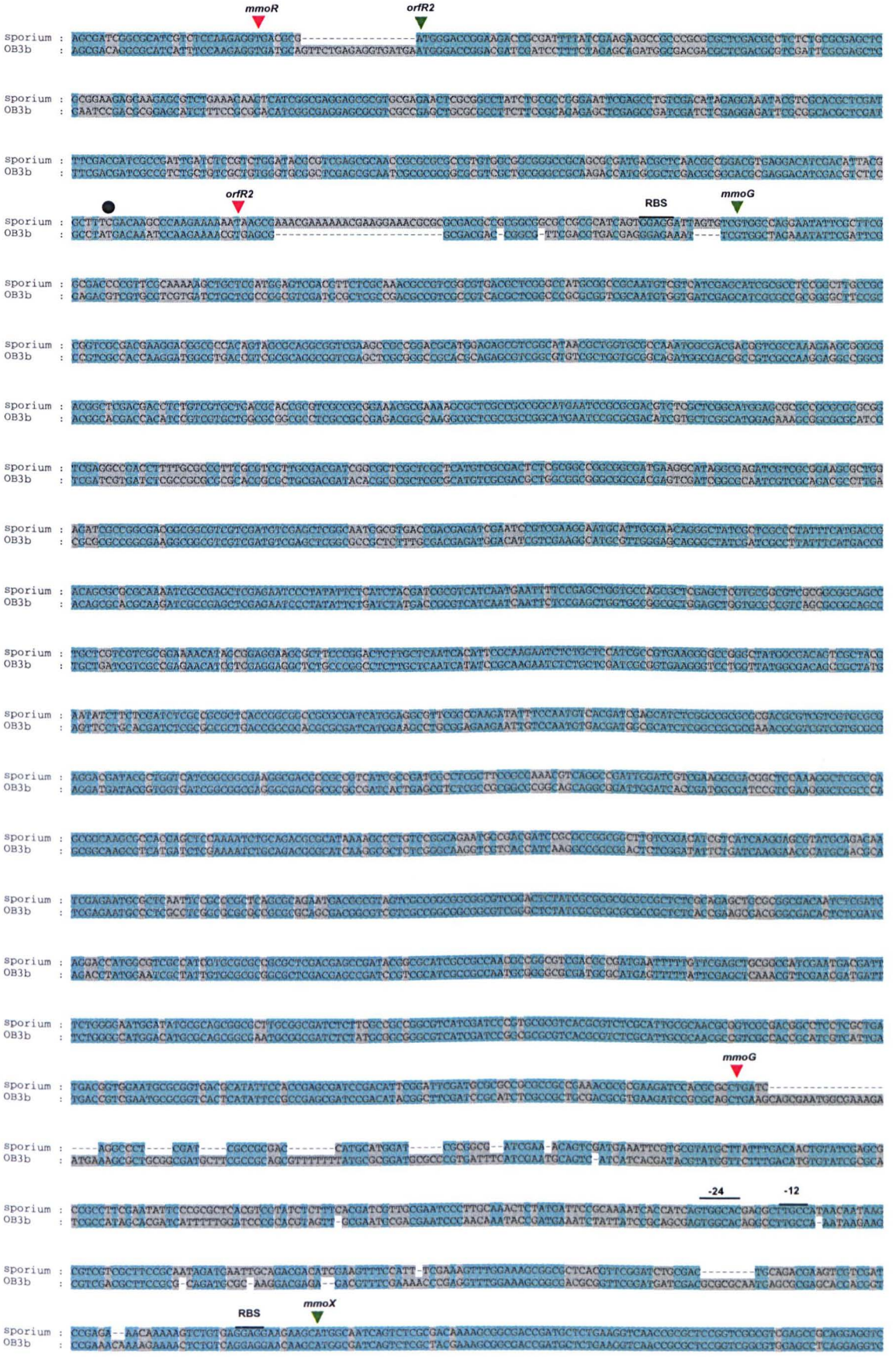
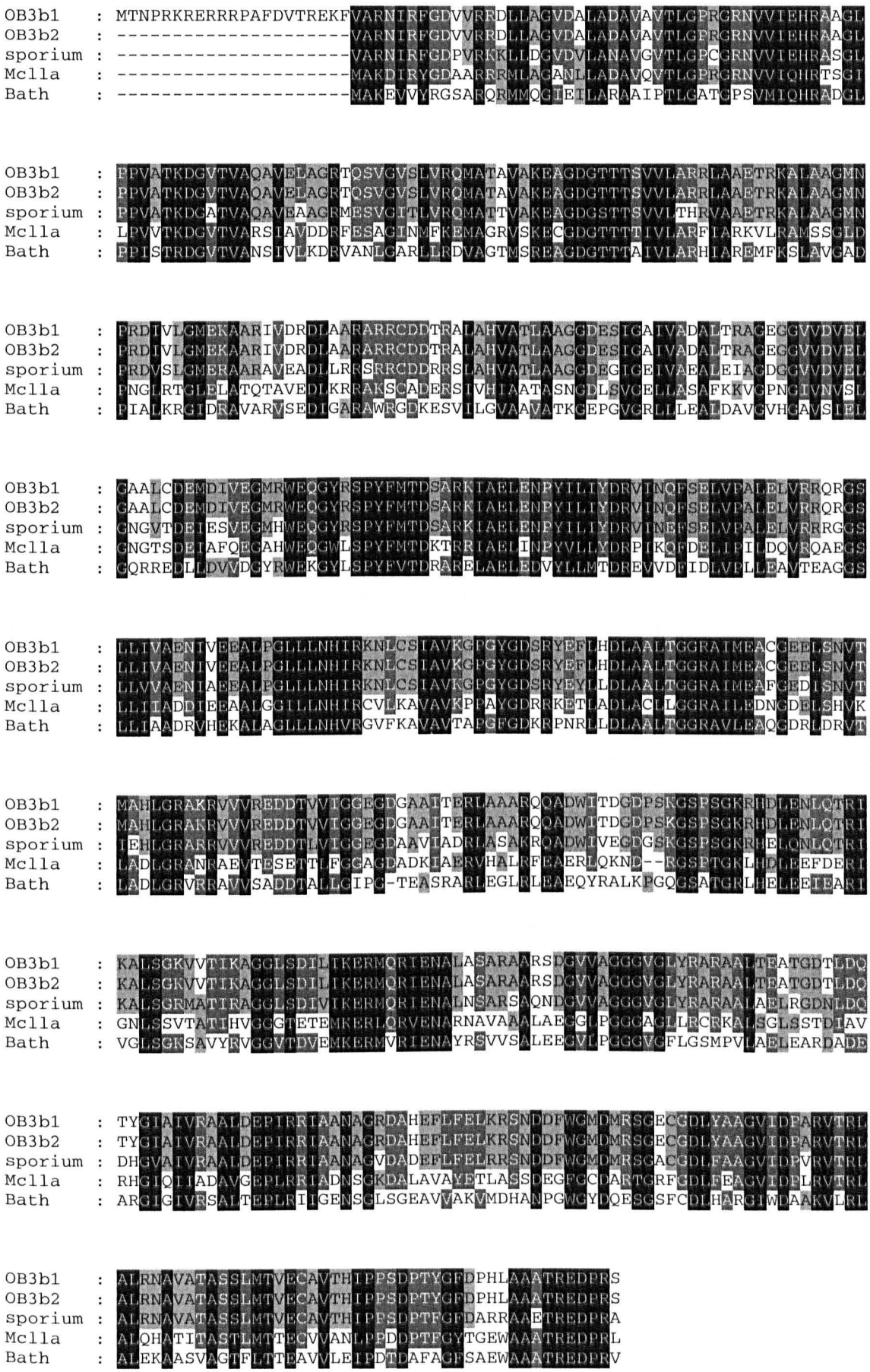

mmoR
orfR2
Sporium :
OB3b :
orfR2
RBS
mmoG
mmoG
-24
-12
RBS
mmoX

Figure 7.2 Alignment of the derived aminoacyl sequence of the available MmoG proteins. 'OB3b1', *Ms. trichosporium* OB3b MmoG using the previously assigned ATG start codon (Genbank accession number CAD61956); 'OB3b2', *Ms. trichosporium* OB3b derived MmoG using the GTG start codon; 'sporium', *Ms. sporium* derived MmoG; 'Mclla', *Methylocella silvestris* $BL2^T$ derived MmoG; 'Bath', *Methylococcus capsulatus* (Bath) derived MmoG. GTG encodes a valine residue. The alignment is shaded according to conservation of property (GeneDoc, http://www.psc.edu/biomed/genedoc/).


OB3b1 : MTNPRKRERRRPAFDVTREKFVARNIRFGDVVRRDLLAGVDALADAVAVTLGPRGRNVVIEHRAAGL
OB3b2 : ----------VARNIRFGDVVRRDLLAGVDALADAVAVTLGPRGRNVVIEHRAAGL
sporium : ----------VARNIRFGDPVRKKLLDGVDVLANAVGVTLGPCGRNVVIEHRASGL
Mclla : ----------MAKDIRYGDAARRRMLAGANLLADAVQVTLGPRGRNVVIQHRTSGI
Bath : ----------MAKEVVYRGSARQRMMQGIETLARAAIPTLGATGPSVMIQHRADGL
OB3b1 : PPVATKDGVTVAQAVELAGRTQSVGVSLVRQMATAVAKEAGDGTTTSVVLARRLAAETRKALAAGMN
OB3b2 : PPVATKDGVTVAQAVELAGRTQSVGVSLVRQMATAVAKEAGDGTTTSVVLARRLAAETRKALAAGMN
sporium : PPVATKDGATVAQAVEAAGRMESVGITLVRQMATTVAKEAGDGSTTSVVLTHRVAAETRKALAAGMN
Mclla : LPVVTKDGVTVARSIAVDDRFESAGINMFKEMAGRVSKECGDGTTTTIVLARFIARKVLRAMSSGLD
Bath : PPISTRDGVTVANSIVLKDRVANLGARLLRDVAGTMSREAGDGTTTAIVLARHIAREMFKSLAVGAD
OB3b1 : PRDIVLGMEKAARIVDRDLAARARRCDDTRALAHVATLAAGGDESIGAIVADALTRAGEGGVVDVEL
OB3b2 : PRDIVLGMEKAARIVDRDLAARARRCDDTRALAHVATLAAGGDESIGAIVADALTRAGEGGVVDVEL
sporium : PRDVSLGMERAARAVEADLLRRSRRCDDRRSLAHVATLAAGGDEGIGEIVAEALEIAGDGGVVDVEL
Mclla : PNGLRTGLELATQTAVEDLKRRAKSCADERSIVHIAATASNGDLSVGELLASAFKKVGPNGIVNVSL
Bath : PIALKRGIDRAVARVSEDIGARAWRGDKESVILGVAAVATKGEPGVGRLLLEALDAVGVHGAVSIEL
OB3b1 : GAALCDEMDIVEGMRWEQGYRSPYFMTDSARKIAELENPYILIYDRVINQFSELVPALELVRRQRGS
OB3b2 : GAALCDEMDIVEGMRWEQGYRSPYFMTDSARKIAELENPYILIYDRVINQFSELVPALELVRRQRGS
sporium : GNGVTDEIESVEGMHWEQGYRSPYFMTDSARKIAELENPYILIYDRVINEFSEIVPALELVRRRGGS
Mclla : GNGTSDEIAFQEGAHWEQGWLSPYFMTDKTRRIAELINPYVLLYDRPIKQFDELIPILDQVRQAEGS
Bath : GQRREDLLDVVDGYRWEKGYLSPYFVTDRARELAELEDVYLLMTDREVVDFIDLVPLLEAVTEAGGS
OB3b1 : LLIVAENIVEEALPGLLLNHIRKNLCSIAVKGPGYGDSRYEFLHDLAALTGGRAIMEACGEELSNVT
OB3b2 : LLIVAENIVEEALPGLLLNHIRKNLCSIAVKGPGYGDSRYEFLHDLAALTGGRAIMEACGEELSNVT
sporium : LLVVAENIAEEALPGLLLNHIRKNLCSIAVKGPGYGDSRYEYLLDLAALTGGRAIMEAFGEDISNVT
Mclla : LLIIADDIEEAALGGILLNHIRCVLKAVAVKPPAYGDRRKETLADLACLLGGRAILEDNGDELSHVK
Bath : LLIAADRVHEKALAGLLLNHVRGVFKAVAVTAPGFGDKRPNRLLDLAALTGGRAVLEAQGDRLDRVT
OB3b1 : MAHLGRAKRVVVREDDTVVIGGEGDGAAITERLAAARQQADWITDGDPSKGSPSGKRHDLENLQTRI
OB3b2 : MAHLGRAKRVVVREDDTVVIGGEGDGAAITERLAAARQQADWITDGDPSKGSPSGKRHDLENLQTRI
sporium : IEHLGRARRVVVREDDTLVIGGEGDAAVIADRLASAKRQADWIVEGDGSKGSPSGKRHELQNLQTRI
Mclla : LADLGRANRAEVTESETTLFGGAGDADKIAERVHALRFEAERLQKND--RGSPTGKLHDLEEFDERI
Bath : LADLGRVRRAVVSADDTALLGIPG-TEASRARLEGLRLEAEQYRALKPGQGSATGRLHELEEIEARI
OB3b1 : KALSGKVVTIKAGGLSDILIKERMQRIENALASARAARSDGVVAGGGVGLYRARAALTEATGDTLDQ
OB3b2 : KALSGKVVTIKAGGLSDILIKERMQRIENALASARAARSDGVVAGGGVGLYRARAALTEATGDTLDQ
sporium : KALSGRMATIRAGGLSDIVIKERMQRIENALNSARSAQNDGVVAGGGVGLYRARAALAELRGDNLDQ
Mclla : GNLSSVTATIHVGGGTETEMKERLQRVENARNAVAAALAEGGLPGGGAGLLRCRKALSGLSSTDIAV
Bath : VGLSGKSAVYRVGGVTDVEMKERMVRIENAYRSVVSALEEGVLPGGGVGFLGSMPVLAELEARDADE
OB3b1 : TYGIAIVRAALDEPIRRIAANAGRDAHEFLFELKRSNDDFWGMDMRSGECGDLYAAGVIDPARVTRL
OB3b2 : TYGIAIVRAALDEPIRRIAANAGRDAHEFLFELKRSNDDFWGMDMRSGECGDLYAAGVIDPARVTRL
sporium : DHGVAIVRAALDEPIRRIAANAGVDADEFLFELRRSNDDFWGMDMRSGACGDLFAAGVIDPVRVTRL
Mclla : RHGIQIIADAVGEPLRRIADNSGKDALAVAYETLASSDEGFGCDARTGRFGDLFEAGVIDPLRVTRL
Bath : ARGIGIVRSALTEPLRIIGENSGLSGEAVVAKVMDHANPGWGYDQESGSFCDLHARGIWDAAKVLRL
OB3b1 : ALRNAVATASSLMTVECAVTHIPPSDPTYGFDPHLAAATREDPRS
OB3b2 : ALRNAVATASSLMTVECAVTHIPPSDPTYGFDPHLAAATREDPRS
sporium : ALRNAVATASSLMTVECAVTHIPPSDPTFGFDARRAAETREDPRA
Mclla : ALQHATITASTLMTTECVVANLPPDDPTFGYTGEWAAATREDPRL
Bath : ALEKAASVAGTFLTTEAVVLEIPDTDAFAGFSAEWAAATREDPRV

7.3 Co-expression of the *Ms. trichosporium* OB3b *mmoG* and sMMO structural genes in *Escherichia coli* BLR

The *mmoG* gene is required for sMMO expression in *Ms. trichosporium* OB3b (Stafford *et al.*, 2003). It is possible that it is a specific chaperone for the assembly of the hydroxylase component. Previous attempts to express the sMMO in *E. coli* failed to yield active sMMO and evidence suggests that the hydroxylase is not properly assembled (Murrell *et al.*, 2000). Since the *mmoG* gene was not included in these failed *E. coli* sMMO expression experiments, the objective of this experiment was to co-express the *mmoG* with the sMMO structural genes in *E. coli* and determine if it resulted in the expression of active sMMO enzyme.

7.3.1 Plasmid constructs for co-expression

Two plasmids were constructed that could be simultaneously maintained in *E. coli*. One plasmid was based on the *E. coli* pET-3a expression vector, and the other on the broad host range pBBR1-MCS5 vector (Kovach *et al.*, 1994). The vectors can be maintained in the same cell since they are in different incompatibility groups and possess different antibiotic resistance markers.

7.3.1.1 pMD38: mmoG expression vector

The DNA fragment containing the *mmoG* gene was purified as an *Xba*I-*Bam*HI fragment from pMD2, and ligated into pET-3a which had been linearized using the same restriction enzymes. The resultant plasmid, pMD38, contained the *mmoG* gene downstream of the T7 promoter (Figure 7.3). The T7 promoter (PT7) is a strong bacteriophage promoter that is recognized by the T7 RNA polymerase enzyme. Expression from PT7 can be achieved in the T7 lysogen strain, *E. coli* BLR(DE3), which has the T7 RNA polymerase gene expression controlled by an IPTG inducible promoter (refer to the Novagen product literature for more information).

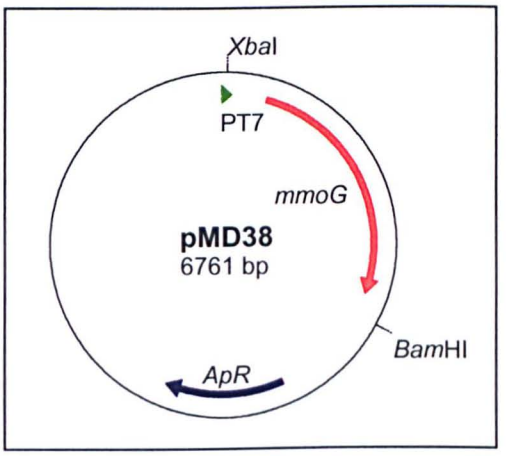


Figure 7.3 Schematic representation of the *mmoG* expression vector (pMD38).

7.3.1.2 pMD30: sMMO expression vector

The *mmoXYBZDC* genes were placed under the regulation of the neomycin phosphotransferase promoter (Pneo) from the RK2 vector pJB3Km1 (Blatny *et al.*, 1997), and placed on the pBBR1-MCS5 vector backbone. The Pneo promoter was chosen since it is a strong initiator of RNA transcription in *E. coli* (David Hodgson, personal communication), and may also be recognized in *Ms. trichosporium* OB3b (Section 7.4).

A series of cloning steps were used to make the pMD30 sMMO expression vector (Figure 7.4). The *mmoX* gene and upstream region was PCR amplified from pTJS176 (Smith *et al.*, 2002) using the primers 4266F (5' – GGG GTA CCA GCA CGC AAT GGC GAA AGA ATG -3') and 6255R (5' – GGG GTA CCT CAG AAT TTG GCG AGC GGA TC -3') and the product ends digested with *Kpn*I restriction enzyme (located at the GGTACC recognition sites in the PCR primers). The *mmoX* PCR fragment was ligated into pTJS140 linearized with *Kpn*I, to make the plasmid pMD3 (Figure 7.4A). The insert DNA was transferred to pJB3Km1 as a *Kpn*I fragment, resulting in the plasmid pMD3Km1 (Figure 7.4B). The Pneo promoter region from pJB3Km1 was amplified using the PCR primers PNEOF (5'- ATA GGA TCC TTA AGA CAG CAA GCG AAC CGG AAT TG -3') and PNEOR (5'- ATA ATA TTC GAA GAT CTT GAT CCC CCT GCG CCA T -3') and *Pfu*Turbo DNA polymerase. The *Bam*HI site in the PNEOF (GGATCC) and the *Bst*BI site in PNEOR (TTCGAA) were used to ligate the promoter into pMD3, resulting in the plasmid pMD3KmN (Figure 7.4C). The insert DNA was transferred to pBBR1-MCS5 as a *Bam*HI-*Kpn*I fragment, to yield pMD3N-BBR5 (Figure 7.4D). The rest of the *mmo* cluster was introduced as a *Apa*LI-*Kpn*I fragment from pMD2 into pMD3N-BBR5, resulting in the *mmo* expression plasmid pMD30 (Figure 7.4E).

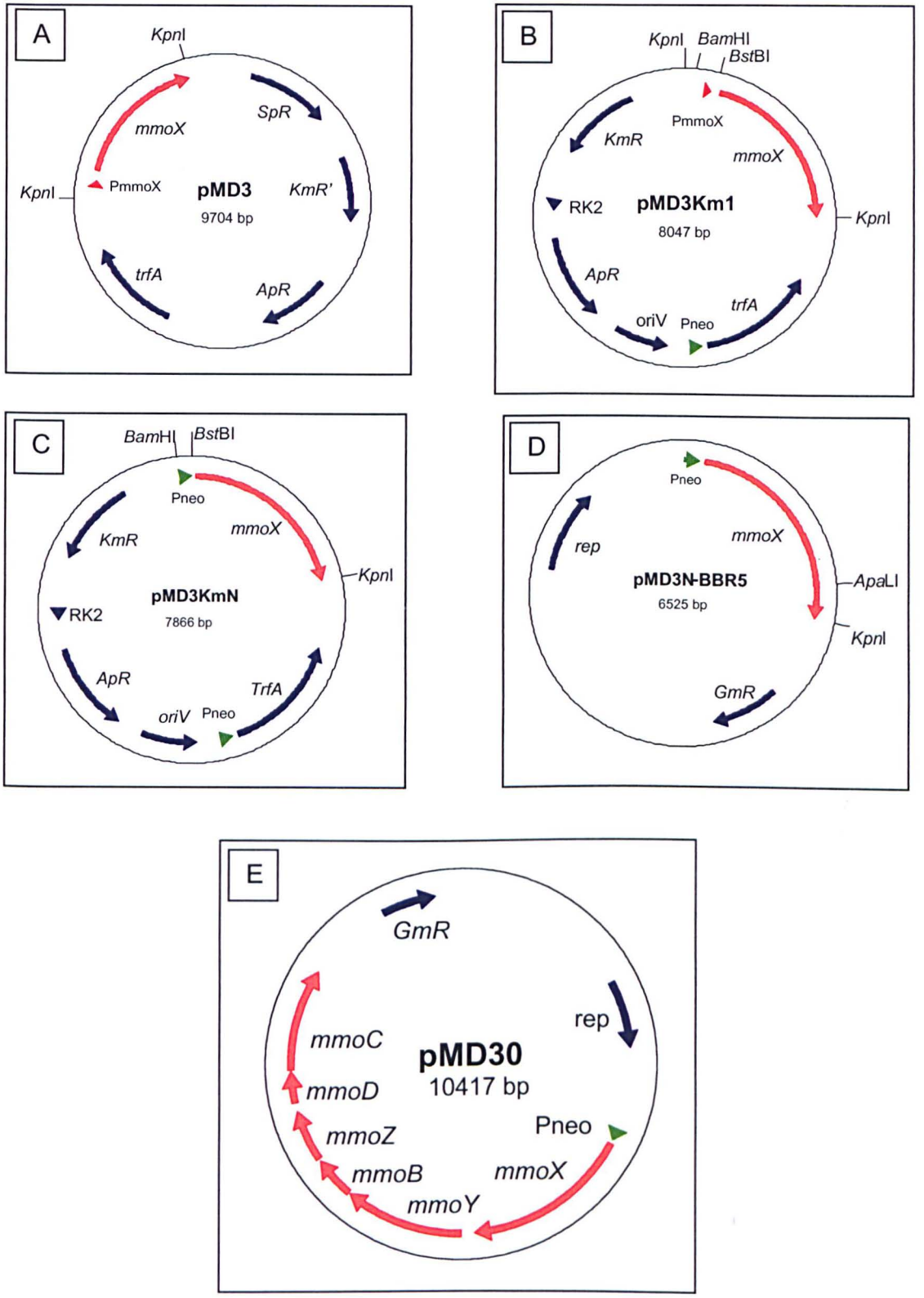


Figure 7.4 Schematic representation of plasmids pMD3, pMD3Km1, pMD3KmN and pMD3N-BBR5 and pMD30. The plasmids shown in panes B, C and D are intermediates in the construction of pMD30. Details of the plasmid constructions are given in the text.

7.3.2 Introduction of pMD30 and pMD38 into *Escherichia coli* BLR(DE3)

E. coli BLR(DE3) was made electrocompetent and transformed as described in Section 2.7.2. Transformants were selected on LB agar containing 50 μg ml^{-1} ampicillin (pET-3a, pMD38) or 5 μg ml^{-1} gentamycin (pBBR1-MCS5, pMD30). Double transformants (strains carrying two different plasmids) were obtained by making single transformant *E. coli* electrocompetent, and electroporating with the second plasmid; double transconjugants were maintained on medium containing both ampicillin and gentamycin antibiotics.

The stability of the plasmids in *E. coli* BLR(DE3) was verified by purifying the plasmids and analysing by agarose gel electrophoresis. No DNA was recovered when using the routine plasmid extraction method (Section 2.6.2), presumably since *E. coli* BLR possesses endonuclease activity and the extraction method did not adequately purify the DNA preparation. The plasmid DNA was sufficiently stable following purification with the QIAprep™ Miniprep Kit (Section 2.6.2). *E. coli* BLR(DE3) carrying pET-3a (4,640 bp), pBBR1-MCS5 (4,768 bp), pMD30 (10.4 kbp), pMD38 (6,761 bp), and both pMD30/pMD38 were analysed (Figure 7.5). The pET vector is a higher copy number plasmid than the pBBR1 vector, and the results suggest that pET-3a was present at approximately ten-fold higher number than the pBBR1-MCS5 plasmid (Figure 7.5, lanes 2, 3 & 4). Five double transformants were analysed and both plasmids could be clearly identified (Figure 7.5, lanes 7-11), indicating that the plasmids could be stably maintained in the cell simultaneously.

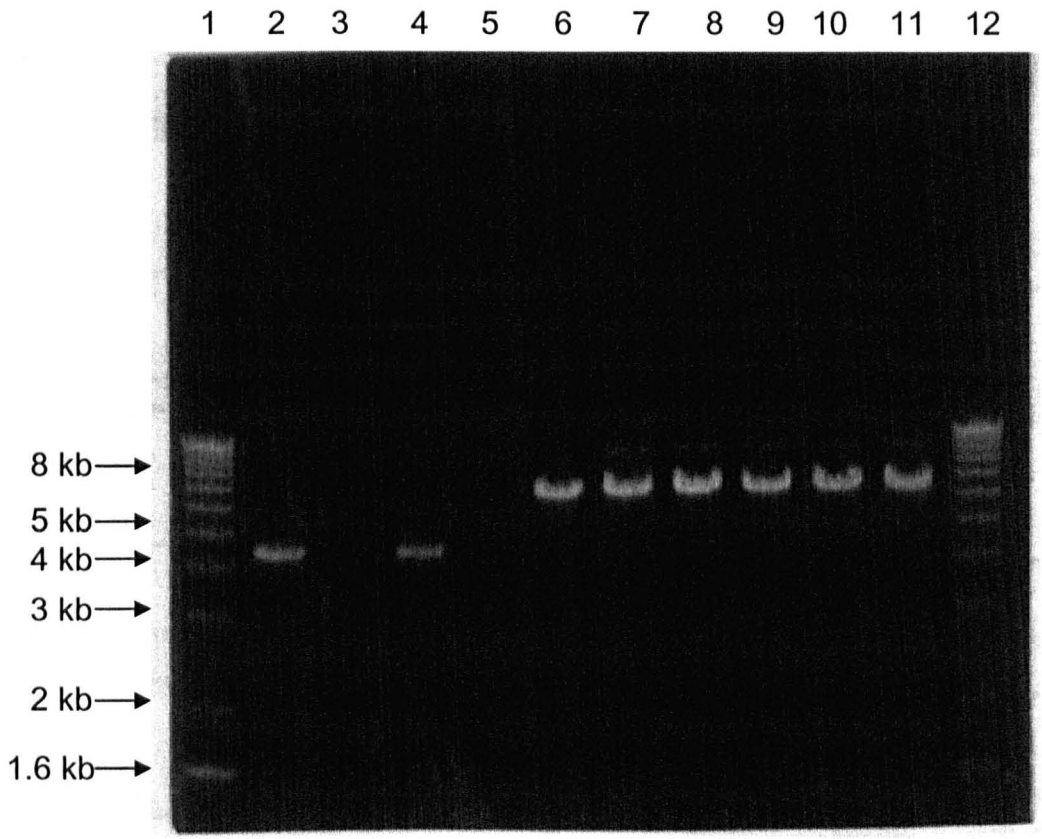


Figure 7.5 Agarose gel electrophoretic analysis of plasmids prepared from *E. coli* BLR(DE3) strains. Lane 1 & 12, 1 kb DNA ladder (Invitrogen); lane 2, 1 µl plasmid from *E. coli* BLR(DE3) [pET-3a] (4.6 kbp); lane 3, 1 µl plasmid from *E. coli* BLR(DE3)[pBBR1-MCS5] (4.8 kbp); lane 4, 10 µl plasmid from *E. coli* BLR(DE3)[pBBR1-MCS5]; lane 5, 1 µl plasmid from *E. coli* BLR(DE3)[pMD30] (10.4 kbp); lane 6, 1 µl plasmid from *E. coli* BLR(DE3)[pMD38] (6.8 kbp); lanes 7-11, 1 µl plasmid from five isolates of *E. coli* BLR(DE3)[pMD30/pMD38].

7.3.3 Expression experiments

The *E. coli* strains carrying the various plasmids were grown in 250 ml conical flasks containing 50 ml of LB medium. The cultures were incubated at 37 °C on a rotary shaker at 200 rpm. The LB medium contained appropriate antibiotic for plasmid selection (50 μg ml^{-1} ampicillin, 5 μg ml^{-1} gentamycin, or both) and were inoculated with 500 μl of overnight culture. Optical density readings were taken at intervals (540 nm) and 1 ml samples were removed for SDS-PAGE analysis (Section 2.17) and naphthalene oxidation sMMO assays (Section 2.16). Induction of the T7 DNA polymerase in *E. coli* BLR(DE3) cultures was performed by adding 50 μl of 0.8 M IPTG during the exponential growth phase (OD_{540} ~ 0.6). The growth of the cultures is represented in Figure 7.6.

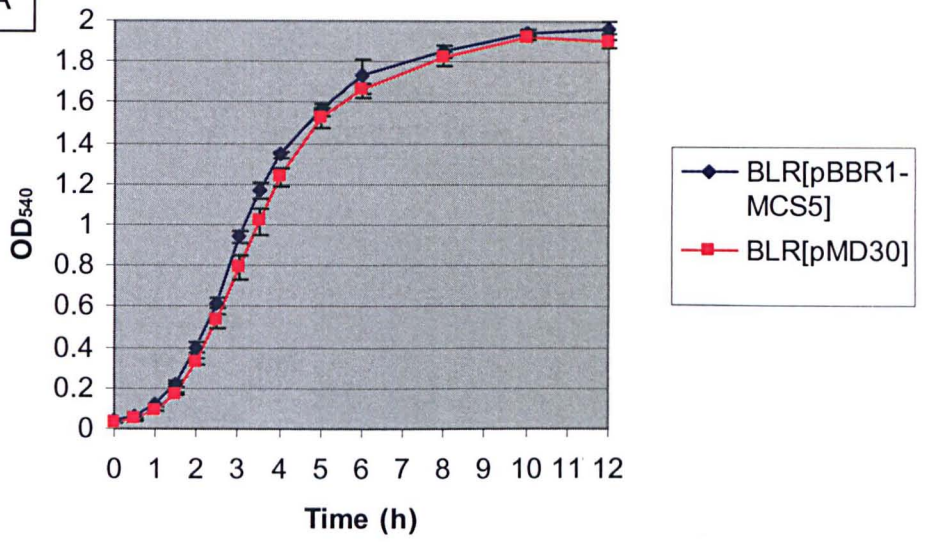
The cultures were monitored by phase contrast microscopy for inclusion body formation. Inclusion bodies were clearly visible in *E. coli* BLR(DE3)[pMD38] and *E. coli* BLR(DE3)[pMD38/pMD30] after the addition of IPTG, which suggested that at least some of the recombinant MmoG protein was packaged into inclusion bodies (results not shown). Total cell protein was analysed by SDS-PAGE (Section 2.17). *E. coli* BLR strains carrying the pET-3a or pBBR1-MCS5 vectors were used as negative controls to show the polypeptide profile of the expression host.

The result of SDS-PAGE analysis is shown in Figure 7.7, and the samples for this analysis were taken from the expression experiments represented in Figure 7.6. The pure sMMO hydroxylase in the second lane is from *Methylococcus capsulatus* (Bath), and was obtained from Sue Slade (University of Warwick). The β-subunit of the hydroxylase co-migrated with an *E. coli* protein in 12 % polyacrylamide gels (results not shown), but was sufficiently resolved in 10 % polyacrylamide gels, as shown (Figure 7.7). However, the γ-subunit of the hydroxylase migrated off the bottom of the 10 % polyacrylamide gel. Obtaining a gel that adequately resolved all the hydroxylase subunits could possibly be achieved by using a larger (20 x 20 cm) 12 % polyacrylamide gel, or a gel possessing a polyacrylamide concentration gradient.

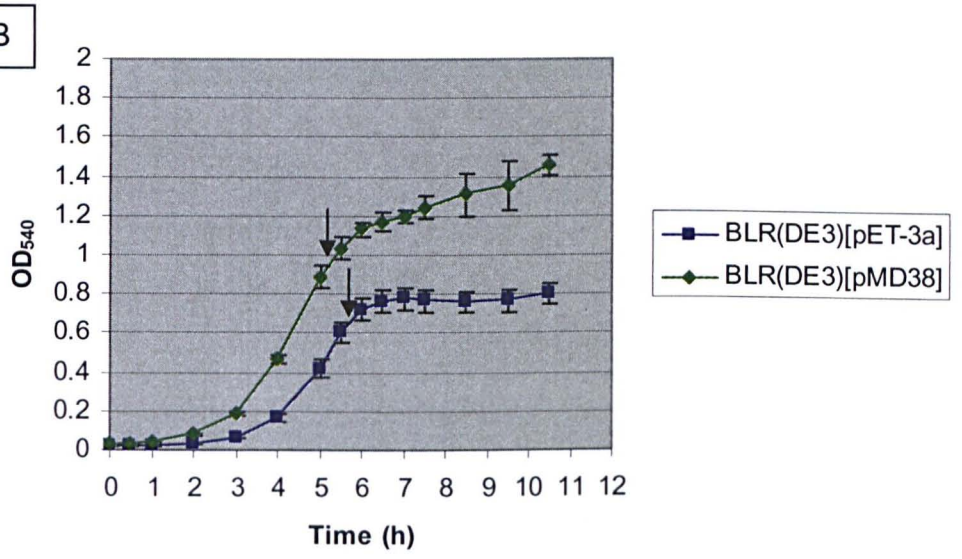
The presence of the pMD30 plasmid had no detectable effect on the growth of the *E. coli* strain (Figure 7.6A). Previous expression experiments with the *mmo* genes under the control of PT7 on a pET vector reduced the growth rate of the culture and prompted the formation of inclusion bodies after IPTG addition (Lloyd, 1997). The

Figure 7.6 Growth curves of *E. coli* strains carrying pET-3a and pBBR1MCS5 plasmid, or their derivatives. The experiments were performed in triplicate; the OD_{540} values are averages of the three readings and error bars correspond to ± standard deviation. The key next to the plots indicate the strain (*E. coli* BLR or *E. coli* BLR(DE3)) and the plasmid corresponding to the curve. The arrows in panels B and C show the point of IPTG addition to the cultures. Samples from these cultures were taken at various time points and analysed by SDS-PAGE (Figure 7.7).

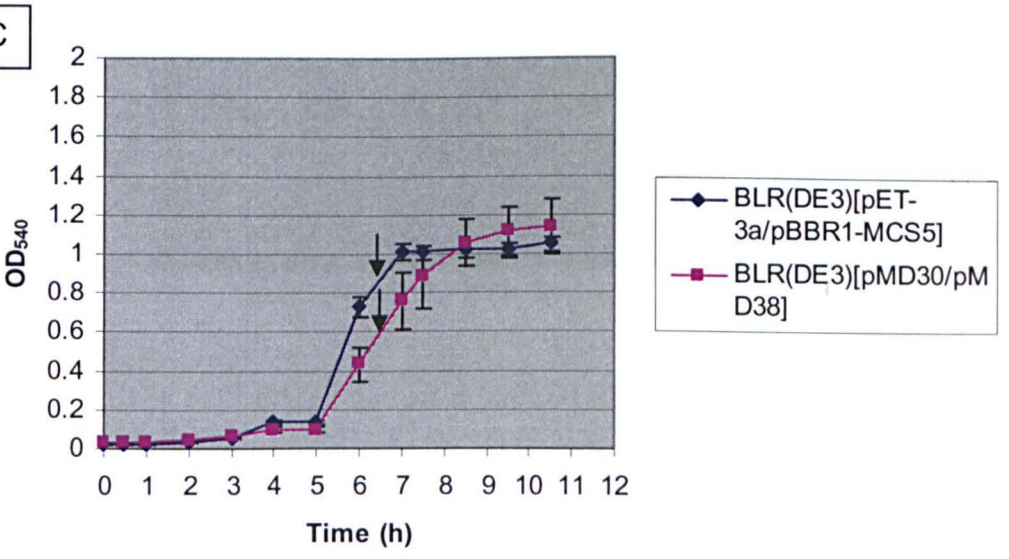
A


OD540
Time (h)
BLR[pBBR1-MCS5]
BLR[pMD30]

B


OD540
Time (h)
BLR(DE3)[pET-3a]
BLR(DE3)[pMD38]

C


OD540
Time (h)
BLR(DE3)[pET-3a/pBBR1-MCS5]
BLR(DE3)[pMD30/pMD38]

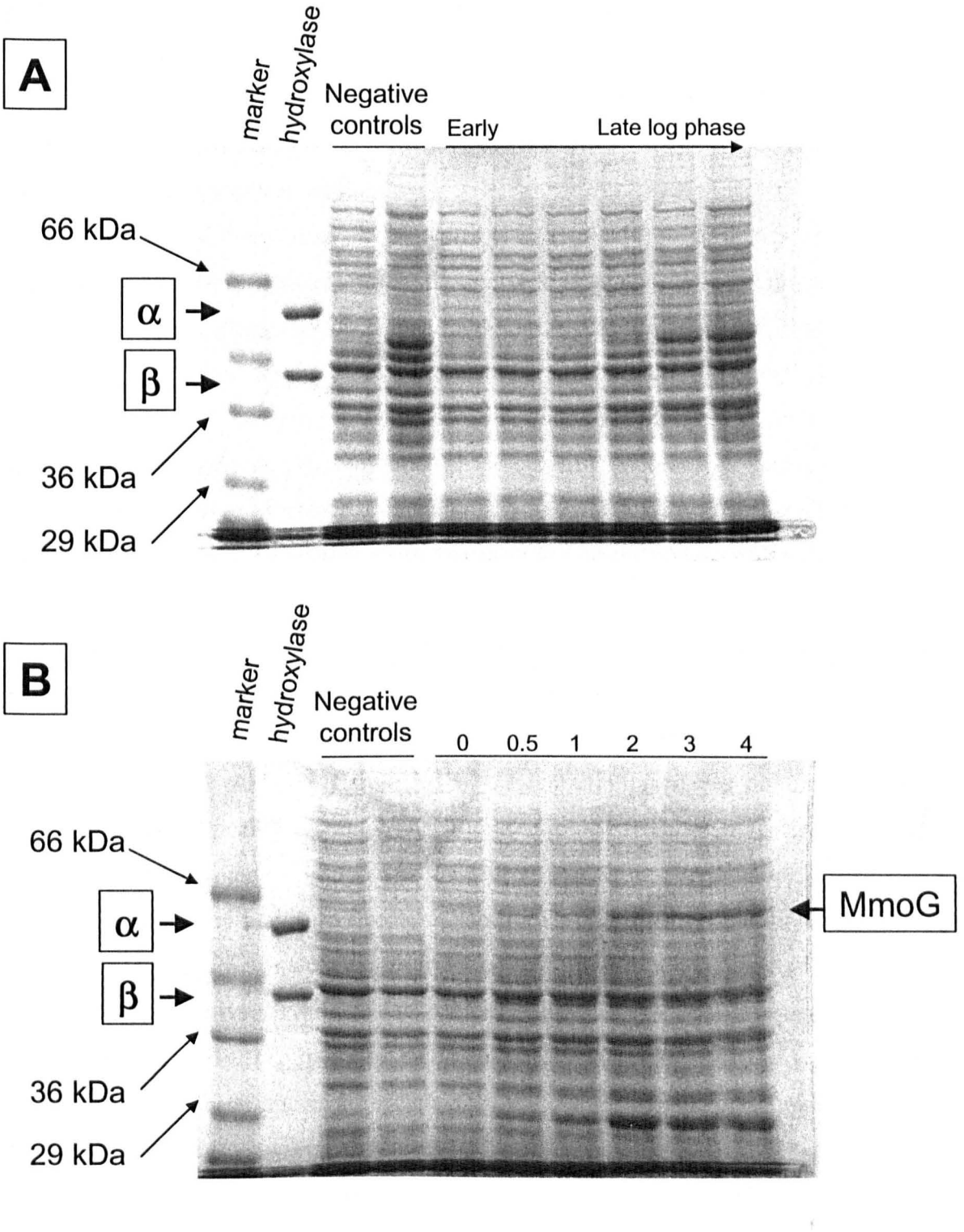


Figure 7.7 10 % SDS-PAGE gels of *E. coli* BLR[pMD30] expressing the sMMO (A) and *E. coli* BLR[pMD30/pMD38] co-expressing the sMMO and the MmoG (B). MmoG expression was induced (B) by the addition of 1 mM IPTG to the growth medium at an OD_{540} of 0.6 (indicated by an arrow in Figure 7.6), and samples taken for analysis at 0, 0.5, 1, 2, 3 and 4 h post-IPTG addition. In (A) the negative controls are *E. coli* BLR(DE3)[pBBR1-MCS5] at 1 h (left lane) and 12 h (right lane); and (B) *E. coli* BLR(DE3)[pET-3a/pBBR-MCS5] at 0 h (left lane) and 4 h (right lane) following IPTG addition.

difference is probably a result of the lower copy number of the pMD30 plasmid and the relatively weak Pneo promoter compared with the T7 promoter. The growth of *E. coli* BLR(DE3) carrying pET-3a plasmids appeared to have longer lag phases and lower maximum cell densities than *E. coli* BLR carrying pBBR1-MCS5 (Figure 7.6). The reason for this was not explored.

The α– and β–subunits of the hydroxylase expressed in *E. coli* BLR[pMD30] were faintly visible in the SDS-PAGE analysis (Figure 7.7A). The amount of expression did not appear to vary considerably during the growth phase of the culture, as was expected since the Pneo promoter is constitutive. Inclusion bodies formed in previous sMMO over-expression studies in *E. coli* (Lloyd, 1997). In this study, inclusion bodies were not visible under phase contrast microscopy, perhaps because they were too small to be visualized. Naphthalene sMMO assays (Section 2.16) were performed on samples taken at various stages in the growth phase of the culture and showed negative reactions.

The MmoG protein has a predicted molecular weight of 59.6 kDa. After IPTG induction, a protein of approximately this size was visible by SDS-PAGE in *E. coli* BLR(DE3)[pBBR1-MCS5/pMD38] (results not shown) and *E. coli* BLR(DE3)[pMD30/pMD38] (Figure 7.7B). Accumulation of this protein appeared to reach maximal levels two hours after the addition of IPTG. The cells developed inclusion bodies, which suggested that the MmoG protein was insoluble. In addition, the accumulation of additional protein bands of approximately 30 kDa was also visible. It is likely these peptides are fragments of the MmoG protein that have been degraded by proteolysis. Samples of proteins from the gel were analysed by mass spectroscopy, but the results were inconclusive. The stability of the recombinant protein may have been improved by lowering the culture incubation temperature, but this was not tested.

The quantity of recombinant α– and β–subunits of the sMMO hydroxylase appeared to be greater when co-expressed with the *mmoG* (Figure 7.7B). Most probably the over-expression of MmoG overwhelms the proteolysis machinery of the cell and therefore the α– and β–subunits are degraded less quickly. The possibility that the MmoG stabilized the ubunits by acting as a chaperone and assembling the sMMO hydroxylase was not tested. Simply examining if the subunits were in the soluble or insoluble fraction of the cell could provide an insight. Naphthalene sMMO

assays were performed on *E. coli* BLR(DE3)[pMD30/pMD38] 0.5, 1, 2, 3 and 4 hours following IPTG induction, and the assays showed no sMMO activity.

Evidence became available at this time that the *mmoG* of *Methylococcus capsulatus* (Bath) was necessary for transcription from the *mmoX* promoter (Rob Csáki, personal communication), and these data were recently published (Csáki *et al.*, 2003). Similar experiments in *Ms. trichosporium* OB3b also suggested that the *mmoG* gene product was involved in initiating transcription from the *mmoX* promoter (Stafford *et al.*, 2003). The co-expression experiments in *E. coli* were temporarily abandoned at this time, and a second approach, discussed in the following section, was taken to investigate whether the MmoG protein also had a role in assembling the sMMO enzyme.

7.4 Homologous expression of *mmoXYBZDC* genes in *Ms. trichosporium*

The results of Csáki *et al.* (2003) and Stafford *et al.* (2003) indicated that the *mmoG* of *Methylococcus capsulatus* (Bath) and *Ms. trichosporium* OB3b were necessary for transcription from the *mmoX* promoter. In both studies, RT-PCR experiments showed that *mmoX* transcripts were not present in *mmoG*$^-$ mutants under the low-copper growth conditions that favour sMMO expression. In an elegant experiment, Csáki and co-workers show that GFP expression from the *mmoX* promoter on a plasmid occurs in the wild-type *Methylococcus capsulatus* (Bath) but not in the Δ*mmoG* mutant, which clearly demonstrates that transcription from this promoter requires a functional *mmoG*. Although these results suggest that MmoG is involved in initiating transcription from the *mmoX* promoter, they do not show whether the MmoG has an additional role in assembling the sMMO hydroxylase.

An experiment aiming to test if the MmoG protein has a post-transcriptional role in sMMO activity was to express the sMMO in *Ms. trichosporium* JS2 (*mmoG*::Gm^R) (Stafford *et al.*, 2003) from an alternative promoter on a broad host range plasmid. If the plasmid is unable to complement the *mmoG*$^-$ mutant, then it would suggest that the MmoG also has a role in the assembly or maturation of the sMMO enzyme. The pMD30 plasmid (Section 7.3.1.2) was suitable for this experiment. The *mmoX* promoter on pMD30 has been replaced with the neomycin phosphotransferase promoter (Pneo), which should be recognized in *Ms.*

trichosporium OB3b since pJB3Km1 can be stably maintained (Smith *et al.*, 2002), and the Pneo controls the transcription of the *trfA* gene of pJB3Km1, which is essential for plasmid replication (Konieczny *et al.*, 1997). The control experiments would be to complement the *Ms. trichosporium* Mutant F (Table 2.1) and *Ms. trichosporium* JS1 (*mmoR*::Gm^R) with pMD30; if the pMD30 expression system was functional, then the plasmid should complement these mutant strains of *Ms. trichosporium.*

7.4.1 Conjugation of pMD30 into *Ms. trichosporium* mutant F

The pMD30 plasmid was conjugated into *Ms. trichosporium* mutant F as described in Section 2.7.1. Transconjugants were isolated on NMS (copper added) agar plates containing 12.5 μg ml^{-1} kanamycin and 5 μg ml^{-1} gentamycin antibiotics. Purified transconjugants were grown on both copper-added and low-copper NMS plates containing antibiotic selection. Naphthalene oxidation sMMO assays were performed. No activity was detected, and SDS-PAGE analysis did not reveal the presence of sMMO subunits (results not shown).

The results indicated that the sMMO was not expressed, or was very poorly expressed, from the pMD30 plasmid in *Ms. trichosporium* mutant F. It is possible that expression from the neomycin phosphotransferase promoter is very weak in *Ms. trichosporium* OB3b. Interestingly, the results were identical in a similar experiment attempted by Rob Csáki at the University of Szeged, except using a kanamycin resistance gene promoter and attempting expression in a *Methylococcus capsulatus* (Bath) mutant (R. Csáki, personal communication).

7.5 Discussion and Future Perspectives

These experiments were temporarily abandoned due to a lack of time. In particular, it will be important to re-examine plasmid expression systems in *Ms. trichosporium*. Obtaining a sMMO expression plasmid that is not controlled by the native *mmoX* promoter will have several uses in addition to the experiment described in Section 7.4.1. A constitutive recombinant sMMO expression system would be useful for expressing site-directed sMMO hydroxylase mutants (Smith *et al.*, 2002) that have lost the ability to oxidise methane. Since pMMO synthesis is reduced during sMMO expression conditions, methane-inactive sMMO mutants would be carbon–starved during sMMO expressing conditions. However, a constitutive sMMO expression system would enable the expression of the recombinant sMMO during pMMO expressing conditions, which would provide a supply of energy and reducing power for sMMO activity.

In addition to this experiment, there are other studies currently ongoing in the JCM laboratory that require expression from a plasmid that is mobilizable and can be maintained in *Ms. trichosporium* OB3b. One experiment will be to express the *mmoR* gene in the *mmoR*$^-$ mutant *Ms. trichosporium* JS1 (Stafford *et al.*, 2003). This would allow mutagenesis studies of MmoR, which could be used to discern the mechanism of MmoR regulation of sMMO gene transcription.

Several potential promoters can be tried. These include the *mxaF* gene promoter from *Ms. trichosporium* OB3b, but this promoter region has not yet been cloned and sequenced. The *mxaF* promoter from *Methylobacterium extorquens* AM1 has been used to create expression plasmids for *Methylobacterium extorquens* AM1 (Marx & Lidstrom, 2001), and the effectiveness of these expression plasmids could be tested in *Ms. trichosporium* OB3b. Another possible promoter is the λ phage promoter system. Expression from this promoter is temperature regulated and has been found to be highly effective in *Methylobacterium extorquens* AM1 (Carlos Miguez, personal communication). A mobilizable, broad host range GFP expression plasmid controlled by the λ phage promoter (Choi & Miguez, unpublished) is currently being tested in *Ms. trichosporium* OB3b (results not shown). Other options include the *pmoCAB* promoter and the *lon* gene promoter (Chapter 6) from *Ms. trichosporium* OB3b.

Chapter 8

Deletion of the sMMO Structural Genes from the Chromosome of *Ms. trichosporium* OB3b

8.1 Introduction

The development of the first sMMO marker-exchange mutant was described by Martin and Murrell (1995). This strain, called 'Mutant F', was created by disrupting the *mmoX* gene on the chromosome of *Ms. trichosporium* OB3b with a kanamycin resistance cassette.

The creation of Mutant F was critical to the development of a recombinant-sMMO expression system that functions by introducing the cloned *Ms. trichosporium* OB3b sMMO genes into sMMO$^-$ Mutant F (Lloyd *et al.*, 1999b). Efforts to obtain active recombinant sMMO expression in *E. coli* or other non-methanotrophs always resulted in inactive sMMO hydroxylase (Murrell *et al.*, 2000) or in highly unstable expression (Jahng & Wood, 1994). The level of recombinant sMMO expression in pMMO-only methanotrophs was also poor (Lloyd *et al.*, 1999a). However, the stability and expression level of the Mutant F system is comparable to that of the wild-type.

The creation of Mutant F and the development of a homologous high-level sMMO expression system enabled the first site-directed mutagenesis of the α-subunit of the sMMO hydroxylase (Smith *et al.*, 2002). The α-subunit of the sMMO contains the active site pocket of the enzyme (Elango *et al.*, 1997; Rosenzweig *et al.*, 1993) and is therefore the most attractive target for preliminary site-directed mutagenesis experiments. Mutagenesis experiments using the Mutant F system are limited to the α-subunit since the expression of the genes encoding the β– and γ–subunits of the hydroxylase appears to occur (Lloyd *et al.*, 1999b). The mutagenesis of component B can be performed in an *E. coli* expression system (Brazeau *et al.*, 2001b; Brazeau & Lipscomb, 2003; Lloyd *et al.*, 1997a), and recombinant expression of MmoD (Lloyd, 1997; Merkx & Lippard, 2002) and the sMMO reductase (Blazyk & Lippard, 2002; West *et al.*, 1992) in *E. coli* has been demonstrated. However, the development of a new sMMO mutant lacking *mmoY* and *mmoZ* would be required for the site-directed mutagenesis of the β– and γ–subunits of the hydroxylase or to express sMMO in the absence of *mmoD*.

This chapter describes the development of a *Ms. trichosporium* OB3b mutant with all the sMMO structural genes (*mmoX, mmoY, mmoB, mmoZ, mmoD* and *mmoC*) deleted or disrupted. This strain should prove a suitable expression host for sMMO

and would allow the mutagenesis of any sMMO component without the background transcription of wild-type genes from the chromosome. This mutant should also enable the study of the MmoD to an extent that was not previously possible (see Chapter 9). Finally, heterologous sMMO expression in this strain may be superior to that observed in pMMO-only methanotrophs, since other factors required for high level sMMO expression, such as regulatory proteins, assembly factors or chaperones, may be absent in these organisms (Lloyd *et al.*, 1999a).

8.2 Marker-exchange mutagenesis

8.2.1 Experimental design

The objective of the experiment was to delete the sMMO structural genes from the chromosome of *Ms. trichosporium* OB3b. This would be done by marker-exchange mutagenesis, essentially as described previously (Stafford *et al.*, 2003). First, the cloned *Ms. trichosporium* OB3b *mmo* gene cluster would be modified by replacing the structural genes with an antibiotic resistance gene. Second, the modified gene cluster would be introduced into *Ms. trichosporium* OB3b on a suicide vector and double recombinants selected (Schäfer *et al.*, 1994).

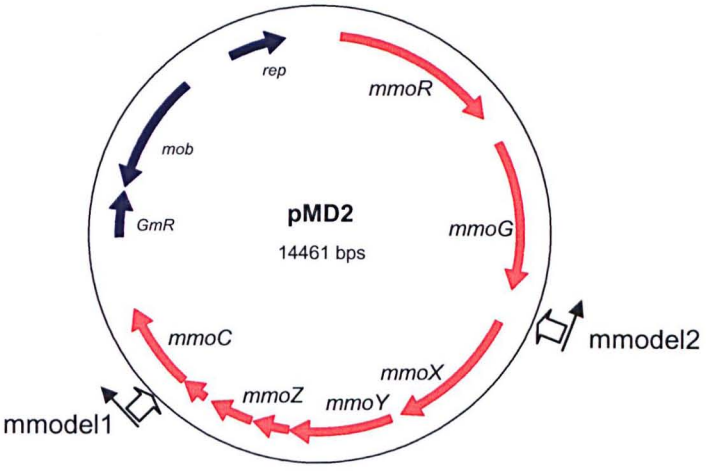
Two experimental strategies designed to replace the cloned sMMO structural genes with an antibiotic resistance gene were conceived:

Strategy 1 (Figure 8.1) was simple in that it involved few cloning steps, but it required PCR amplification of a large DNA fragment which can be difficult. The PCR was performed using *Pfu*Turbo™ polymerase (Section 2.9) which is suitable for amplifying targets up to 15 kb in length from plasmids (Stratagene product literature). An extension time of 15 min and an annealing temperature of 55 °C were used. The PCR did not work (results not shown) and the strategy was abandoned.

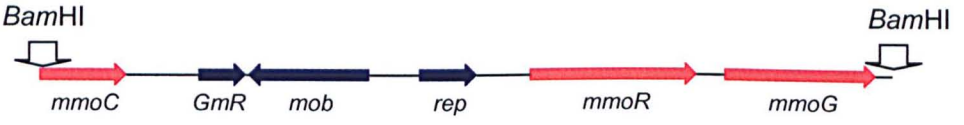
Strategy 2 was used successfully and is described in Section 8.2.2.

Figure 8.1 *Strategy 1*, designed to replace the sMMO structural genes with an antibiotic resistance marker. PCR primer mmodel1 (5'-ATA ATA GGA TCC ATC GTC ATC GAG ACC GAG GAC G-3') anneals to nucleotide positions 10 to 31 of the *mmoC* gene; this primer would be used to introduce a *Bam*HI site (underlined) into *pmoC*. PCR primer mmodel2 (5'-ATC GGT ATT TGT TGG GAT TCG TCG-3') anneals to the plasmid 201 to 224 bp upstream of the *mmoX* start. The arrows on the diagram indicate the direction of transcription from the primers (A). The 9,942 bp PCR product (B) would contain two *Bam*HI sites: one site is present 248 bp upstream of *mmoX* in pMD2 (Table 2.2) and the second would have been introduced by the mmodel1 primer. After cutting the product with *Bam*HI, a gentamycin cassette (GmR) released from p34S-Gm (Table 2.2) with *Bam*HI could be ligated, to produce the plasmid shown (C).

A



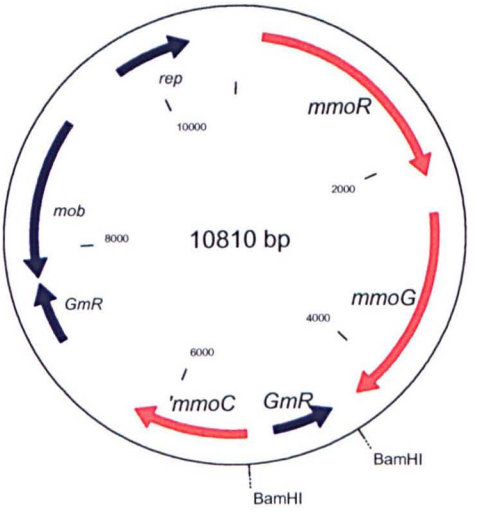
B



PCR product

9942 bp

C



8.2.2 Plasmid construction; *mmo* gene replacement with a gentamycin cassette

This section describes the cloning steps used to construct a plasmid where *mmoX* and its promoter, *mmoY*, *mmoB*, *mmoZ*, *mmoD* and the 5' – end of *mmoC* were replaced with a gentamycin gene cassette.

It was first necessary to introduce a *Bam*HI site into the *mmoC* gene, which was accomplished by PCR using the mmodel1 primer and an M13 primer that anneals to the vector as shown in Figure 8.2A. The 1,247 bp PCR product was amplified using *Taq* DNA polymerase (Section 2.9) using an annealing temperature of 55 °C. The PCR product was digested with *Bam*HI and *Kpn*I restriction enzymes. To generate the pMMODEL-pUC construct (Figure 8.2C), a three way ligation was performed using the digested PCR product, a fragment upstream of *mmoX* released from pMD2 with *Xba*I/ *Bam*HI restriction enzymes, and pUC19 plasmid digested with *Xba*I/ *Kpn*I (Figure 8.2B). The *Bam*HI/ *Xba*I fragment from pMD2 lies upstream (5') of the *mmoX* promoter, and therefore the subsequent marker-exchange mutagenesis will also result in the deletion of the promoter from the chromosome. *E. coli* was transformed with the ligation mix and the transformants screened by blue/white selection on ampicillin containing plates (Section 2.7.2). pMMODEL-pUC was identified by restriction mapping with *Bam*HI/ *Xba*I and *Bam*HI/ *Kpn*I (results not shown).

The next step was to insert a gentamycin gene cassette between the *mmoG* and *mmoC* sequence, which would be used as the selection marker for the mutagenesis. The gentamycin cassette was excised from p34S-Gm (Table 2.2) with *Bam*HI and ligated into the *Bam*HI site of pMMODEL-pUC. *E. coli* was transformed with the DNA ligation mix and transformants were selected for gentamycin resistance. The correct construct (pMMODEL-pUC-Gm (Figure 8.2D) was identified by restriction mapping with *Bam*HI (results not shown). It should have been possible to identify the orientation of the gentamycin cassette by restriction mapping with *EcoR*V/ *Xba*I, but an additional *EcoR*V site, not predicted in the sequence, made this impossible and so the orientation was not determined. If the *mmoC* start codon and ribosome binding site had not been removed it would have been necessary to confirm that the gentamycin gene was inserted in the reverse orientation since the gentamycin gene promoter could have initiated transcription of *mmoC*. However, the deletion of the 5'

– end of the *mmoC* gene should effectively prevent *mmoC* transcription even in the presence of a promoter.

8.2.3 The suicide vector pK18mobsacB-pMMODEL-Gm

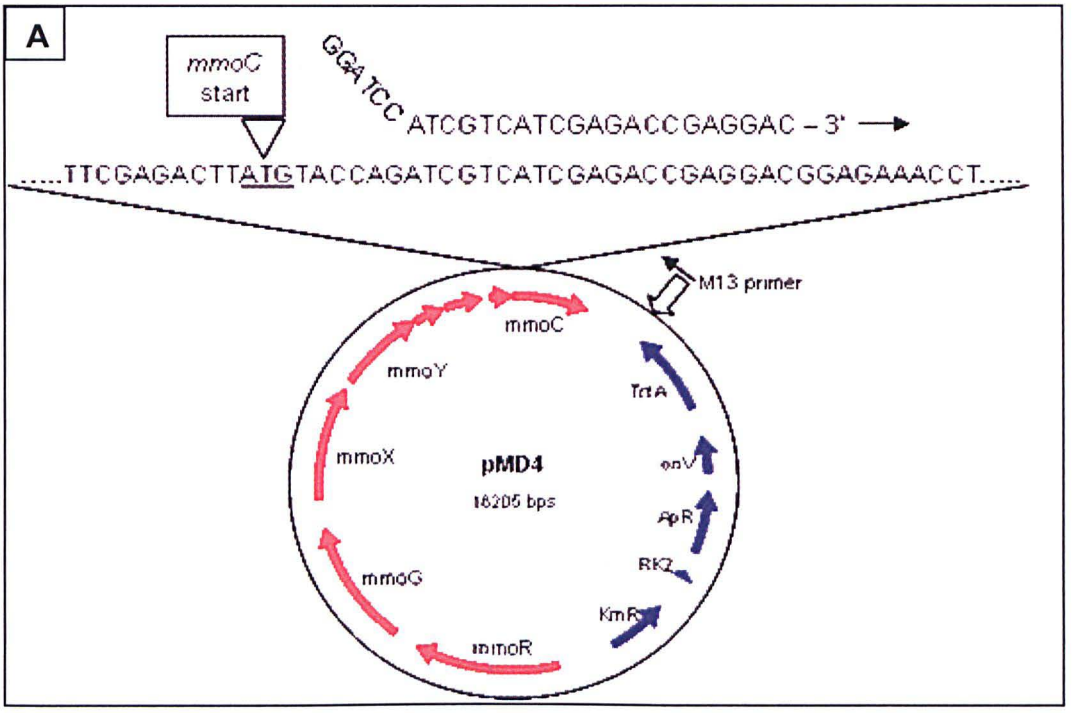
Once pMMODEL-pUC-Gm was made, it was necessary to change the vector backbone to one more suitable for marker-exchange mutagenesis. Although pUC19 would be a "suicide vector" in a methanotroph, since its ColE1 origin of replication is not recognized, it is not ideal because it requires selection with ampicillin which is too chemically unstable for the relatively long incubations required for growth of *Ms. trichosporium* (typically 7 – 10 days). The pK18mobsacB vector was used instead (Table 2.2); pK18mobsacB encodes resistance to kanamycin, which is adequately stable for use with methanotrophs, and also possesses the *sacB* gene which is useful, when necessary, to force rare double recombination events (Schäfer *et al.*, 1994). The pMMODEL-pUC-Gm plasmid was cut with *Hin*dIII to release the insert from the pUC19 backbone. The insert was ligated into pK18mobsacB (Table 2.2) linearized with *Hin*dIII. The construct, pK18mobsacB-mmodel-Gm (Figure 8.3A), was identified and its integrity confirmed by restriction mapping with *Hin*dIII, with *Bam*HI and with *Kpn*I (Figure 8.3B). The sizes of the DNA fragments are consistent with the predicted sizes, with the exception of the 4.3 kb fragment of the *Hin*dIII digest, which appears to be approximately 5 kb. It is possible that two copies of the gentamycin cassette were ligated when constructing pMMODEL-pUC-Gm. This is not a concern since it is unlikely that an additional copy of the gentamycin gene would affect the marker-exchange experiment.

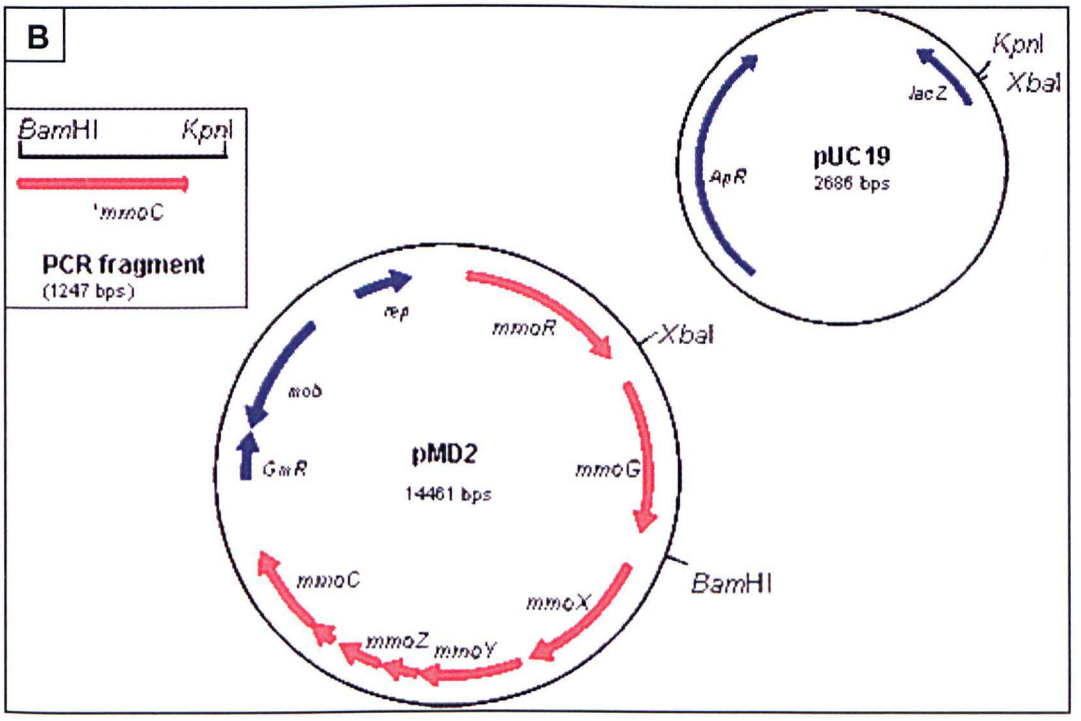
8.2.4 Marker-exchange mutagenesis

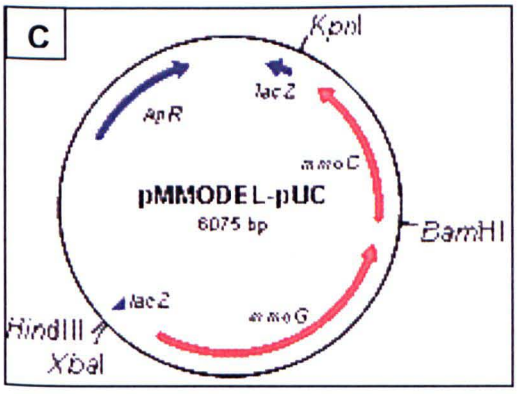
The pK18mobsacB-mmodel-Gm suicide vector was conjugated into *Ms. trichosporium* OB3b as described in Section 2.7.1. After mating, the cells were grown on NMS plates containing only gentamycin (and not kanamycin) for the selection of both single and double recombinants. Colonies (~ 5 mm diameter) appeared on the plates after three weeks incubation with methane/ air. 45 colonies were replica plated onto NMS-kanamycin and NMS-gentamycin agar plates. Six of the transconjugants did not grow with kanamycin. A naphthalene assay (Section 2.16) was performed with the transconjugants grown on NMS (0.8 μM copper) plates

containing gentamycin. The 39 gentamycin/ kanamycin resistant transconjugants gave positive reactions with the naphthalene assay and the six kanamycin sensitive transconjugants gave negative reactions (results not shown). Evidently, the six kanamycin-sensitive transconjugants had undergone a double recombination event, thereby losing the pK18mobsacB backbone and deleting the sMMO genes from the chromosome. One of the transconjugants was selected at random. The strain was named *Ms. trichosporium* SMDM (sMMO Deletion Mutant).

Figure 8.2 *Strategy 2*, used to replace the sMMO structural genes with an antibiotic resistance cassette (refer to the text for more information). Panel A indicates the location of primer binding and the direction of *Taq* DNA polymerase transcription (arrows). The exact annealing location of the mmodel1 primer, relative to the start codon of the *mmoC* gene, is indicated; the *Bam*HI recognition sequence (GGATCC) present on the 5' - end of the primer is shown. Panel B shows the molecules used to create pMMODEL-pUC, the map of which is in Panel C (the molecule sizes illustrated are not necessarily proportional). Panel D depicts the map of pMMODEL-pUC-Gm; the orientation of the gentamycin gene (*GmR*) relative to *mmoG* and *mmoC* is not known and therefore not indicated.


A
mmoC start
GGATCC ATCGTCATCGAGACCGAGGAC – 3'
.....TTCGAGACTTATGTACCAGATCGTCATCGAGACCGAGGACGGAGAAACCT.....
M13 primer
mmoC
mmoY
TetA
mmoX
pMD4
16205 bps
ori V
ApR
RK2
mmoG
KmR
mmoR


B
BamHI
KpnI
'mmoC
PCR fragment
(1247 bps)
KpnI
XbaI
lacZ
ApR
pUC19
2686 bps
rep
mmoR
XbaI
mob
GmR
pMD2
14461 bps
mmoG
mmoC
BamHI
mmoX
mmoZ
mmoY


C
KpnI
lacZ
ApR
mmoC
pMMODEL-pUC
6075 bp
BamHI
lacZ
mmoG
HindIII
XbaI

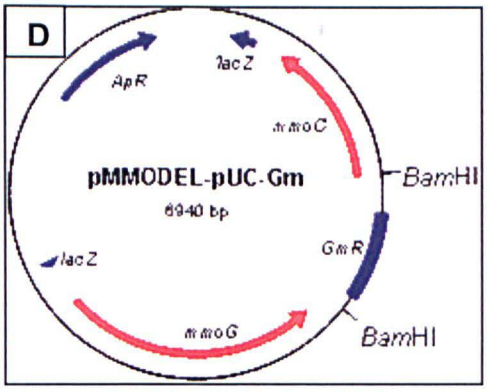
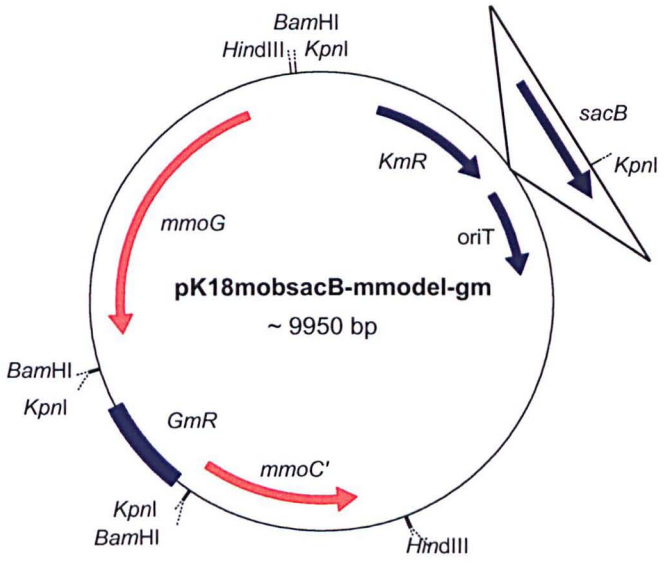
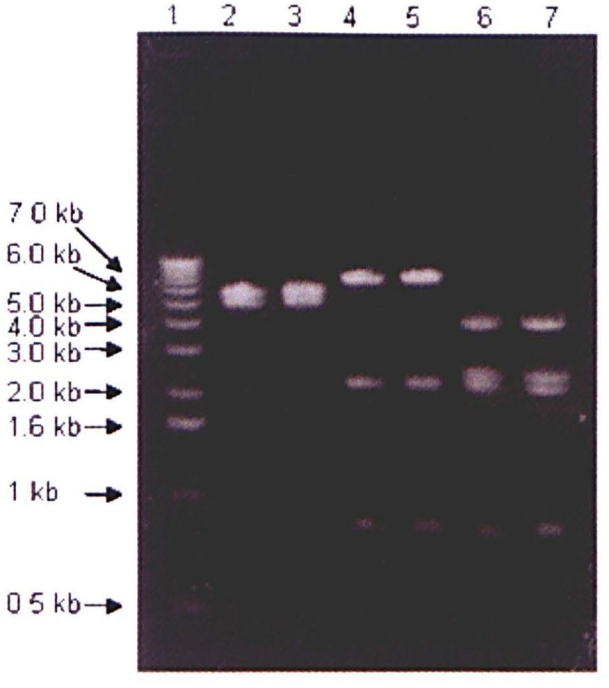

D
ApR
lacZ
mmoC
pMMODEL-pUC-Gm
6940 bp
BamHI
GmR
lacZ
mmoG
BamHI

Figure 8.3 (A) Map of the suicide plasmid pK18mobsacB-mmodel-Gm. The exact sequence of *sacB* is not available, and so the molecule size is approximate, as is the position of the *sacB* gene. (B) Agarose gel of a restriction digest of two pK18mobsacb-mmodel-Gm plasmids (a and b), tentatively selected by an initial screen with *Hin*dIII (results not shown). Lane 1, 1 kb DNA ladder (Invitrogen); lanes 2 and 3, clones a and b, respectively, cut with *Hin*dIII; lanes 4 and 5, clones a and b, cut with *Bam*HI; lanes 6 and 7, clones a and b, cut with *Kpn*I. The predicted sizes of the DNA fragments are: 5,660 bp and 4,290 bp (*Hin*dIII); 865 bp, 2,220 bp and 6,865 bp (*Bam*HI); and 847 bp, 2,220 bp, ~ 2.5 kb, and ~ 4.5 kb (*Kpn*I).

A


BamHI
HindIII
KpnI
sacB
KpnI
KmR
mmoG
oriT
pK18mobsacB-mmodel-gm
~ 9950 bp
BamHI
KpnI
GmR
mmoC'
KpnI
BamHI
HindIII

B


1 2 3 4 5 6 7
7.0 kb
6.0 kb
5.0 kb
4.0 kb
3.0 kb
2.0 kb
1.6 kb
1 kb
0.5 kb

8.2.5 Confirmation of the *Ms. trichosporium* SMDM *mmo*⁻ genotype

The deletion of the sMMO genes was confirmed by Southern blotting. Genomic DNA preparations from *Ms. trichosporium* OB3b, *Ms. trichosporium* mutant F and *Ms. trichosporium* SMDM were digested with *Cla*I and *Sph*I restriction endonucleases (Figure 8.4), and three blots prepared, as described in Section 2.10.1. Three probes were generated by labelling pMD2 restriction fragments: an *mmoR* gene fragment was used as template for one probe (prepared from a 2.2 kb *Kpn*I/ *Xba*I fragment from pMD2), an *mmoX* fragment and promoter region (prepared from a 1.6 kb *Bam*HI/ *Apa*LI fragment from pMD2) for a second and an *mmoY* fragment (prepared from a 1 kb *Sst*I fragment from pMD2) for a third (Section 2.10.3). The hybridizations were performed as described in Section 2.10.4, and the results are shown in Figure 8.5. The results confirmed that the *mmoX* and *mmoY* genes were absent in *Ms. trichosporium* SMDM, indicating that the marker-exchange mutagenesis was successful. As predicted, the *mmoR* gene was present in all three strains since it was not targeted for deletion. Hybridization intensity to the *mmoX* gene of *Ms. trichosporium* mutant F was weaker than to that of *Ms. trichosporium* OB3b, which is consistent with the fact that part of the gene was disrupted in *Ms. trichosporium* mutant F (Martin & Murrell, 1995). Also, the restriction patterns in the vicinity of the *mmoX* and *mmoY* genes were altered in *Ms. trichosporium* mutant F by the mutagenesis, which explains the difference in these fragment sizes between this organism and the wild-type.

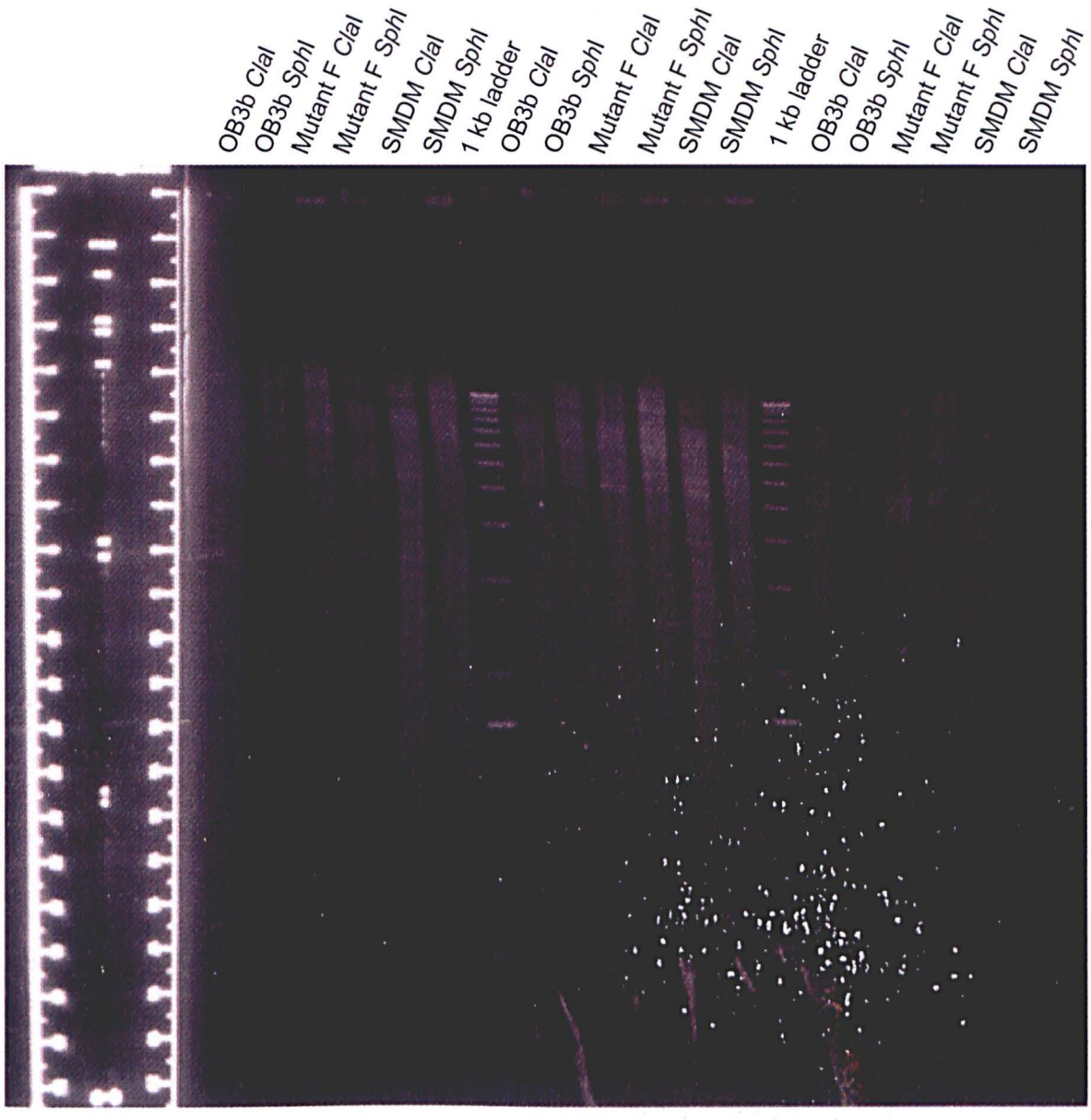


Figure 8.4 Agarose gel of genomic DNA digests of *Ms. trichosporium* OB3b, *Ms. trichosporium* Mutant F and *Ms. trichosporium* SMDM genomic DNA with *Cla*I and *Sph*I restriction enzymes. The lanes are labelled according to the *Ms. trichosporium* strain name and the restriction enzyme used. The digests were performed in triplicate for blotting and hybridization with three gene probes, as shown in Figure 8.5.

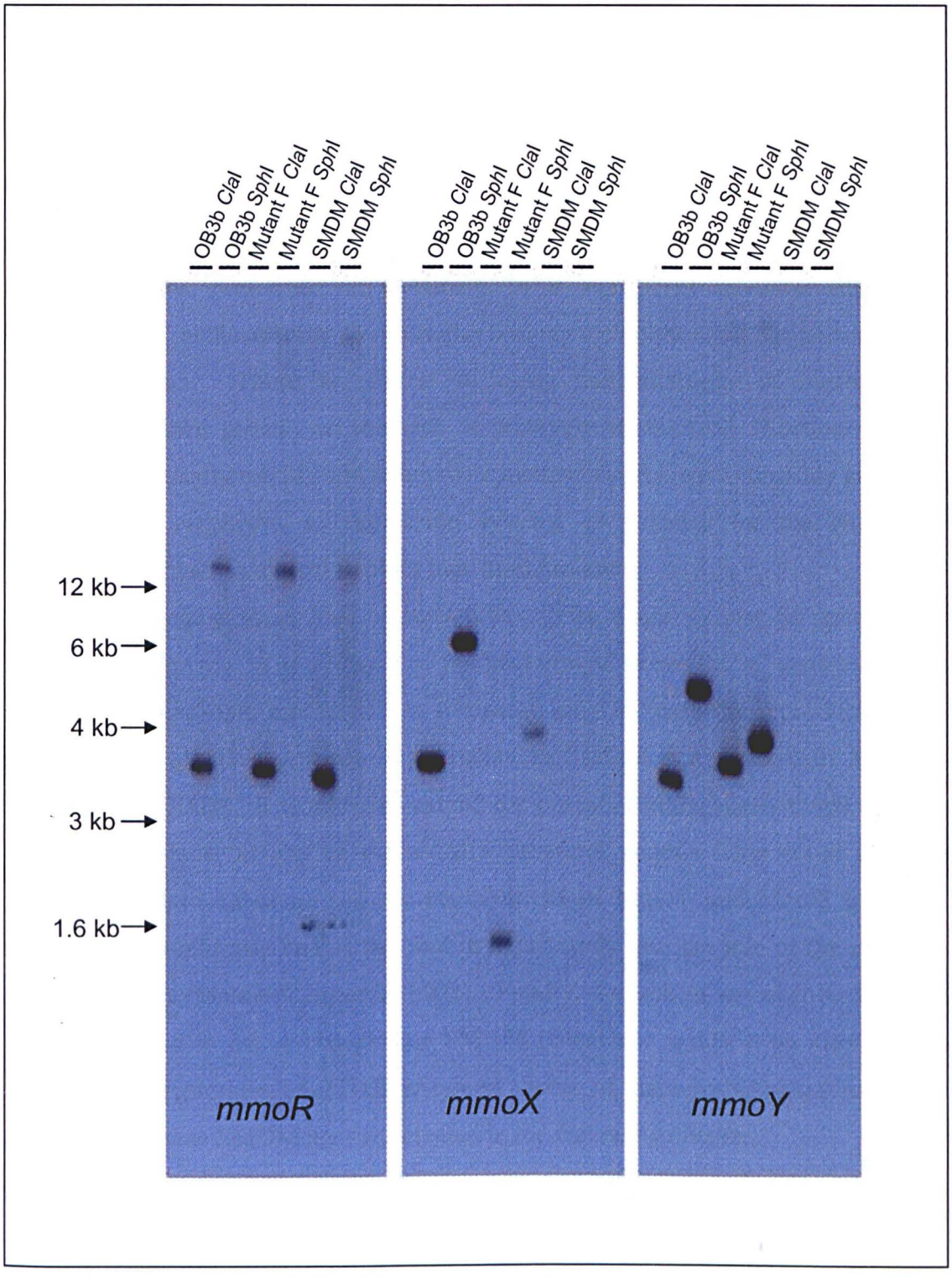


Figure 8.5 Southern blot confirming that the marker-exchange mutagenesis was successful and that the genes were deleted in *Ms. trichosporium* SMDM. DNA preparations were digested with *Cla*I and *Sph*I, as shown in Figure 8.4. *Ms. trichosporium* OB3b (OB3b), *Ms. trichosporium* Mutant F (Mutant F) and *Ms. trichosporium* SMDM (SMDM) were hybridized with *mmoR*, *mmoX* and *mmoY*, as indicated.

8.3 Discussion

The deletion of the sMMO structural genes from *Ms. trichosporium* OB3b performed in this study, in which 4.5 kb of the chromosome sequence was replaced with an 865 bp gentamycin resistance gene cassette. This is the largest allelic exchange performed on this organism to date. The relative ease with which this experiment was performed is testament to a significant advancement in the amenability of methanotrophs to molecular biology techniques that has been achieved in recent years. There is interest in using methanotrophs in biotechnology applications since growth on methane is relatively inexpensive (Section 1.7). An ability to genetically modify and manipulate methanotrophs may ultimately enable the use of these organisms as expression systems or vehicles for the large scale production of pharmaceutical or other industrial products.

Ms. trichosporium SMDM may prove to be a useful host for recombinant sMMO experiments. The absence of the background *mmoYBZDC* gene expression from the chromosome that occurs in Mutant F may increase the reproducibility of recombinant sMMO expression experiments in SMDM compared with Mutant F (Smith *et al.*, 2002). It would also extend the potential mutagenesis target from the *mmoX* gene alone, to any of the sMMO structural genes. This could allow an investigation into the some specific contributions of MmoY and MmoZ to sMMO catalysis. It would also enable the first in depth study into the role of the enigmatic MmoD subunit (Merkx & Lippard, 2002). Finally, the lack of any sMMO structural gene expression in *Ms. trichosporium* SMDM potentially makes it an ideal host for heterologous expression of sMMO enzymes. Some of these potential applications of *Ms. trichosporium* SMDM are investigated in the following chapter.

Chapter 9

sMMO Expression Experiments in *Ms. trichosporium* SMDM

9.1 Introduction

Ms. trichosporium SMDM is a sMMO-minus strain of *Ms. trichosporium* OB3b lacking the *mmoX* promoter region and the *mmoXYBZD* genes (Chapter 8). The strain was created by marker-exchange mutagenesis, replacing the sMMO genes with a gene encoding gentamycin resistance. As discussed in the previous chapter, there are several potential experiments that can be performed using *Ms. trichosporium* SMDM as a recombinant sMMO expression host. The recombinant expression of homologous and heterologous sMMO genes in *Ms. trichosporium* SMDM is described in this Chapter. Also, the recombinant expression of sMMO in the absence of the *mmoD* gene was performed.

9.2 Partial re-sequencing of the *Ms. trichosporium* OB3b sMMO gene cluster

9.2.1 Introduction

For the experiment described in Section 9.5, it was necessary to sequence the *mmoC* PCR product amplified to generate the pD-del construct (Figure 9.11) to ensure that the PCR had not introduced mutations. Several differences were seen between the PCR product and the Genbank sequence[6] (Cardy *et al.*, 1991a; Cardy *et al.*, 1991b), and it was therefore necessary to re-sequence the original clone to determine which of these were errors in the original sequence data and which were potentially errors introduced into the PCR product by the *Taq* DNA polymerase. Also, for this experiment it proved impossible to PCR amplify the *mmoD* gene, either from the cloned cluster or from *Ms. trichosporium* OB3b genomic DNA (results not shown) and the possibility existed that additional sequence data errors in this region had caused PCR primer mismatching. A 2.0 kb section of the *mmo* gene cluster, including *mmoZ*, *mmoD* and *mmoC*, was re-sequenced.

9.2.2 Sequencing the *Ms. trichosporium* OB3b *mmoZDC* genes

The plasmid pMD4 (Table 2.2) was used as template for the sequencing reactions. Although having been subcloned several times, the *mmoZDC* genes in pMD4 originate from the original clone fragment (Cardy *et al.*, 1991a). The sequencing primers used are shown in Table 9.1.

Table 9.1 Primers used for sequencing the *Ms. trichosporium* OB3b *mmoZDC* genes. The primer number corresponds to the position within the Genbank sequence (accession number X55394), and 'F' refers to forward and 'R' reverse orientation.

Primer	Sequence (5' to 3')
7985F	ACC AAA TTC ACC ATC ACC
8474F	CAA CAC CGA TTA TTA CG
9102R	ACC TTC ACA TCG ATC AGC
9542R	GAT CAT GCA GGA AGA ACG
9990R	ACC GAA GAG AAG AGT CG

[6] Genbank accession numbers X55394 and S81887

The re-sequencing of the *mmoZDC* genes has revealed numerous apparent errors in the original sequence. Six nucleotide residue differences within the *mmoZ* gene were revealed and result in four differences between the derived aminoacyl sequences of MmoZ (Figure 9.1). Eight errors within the intergenic region between *mmoZ* and *mmoD* were found (results not shown). Within the *mmoD* gene there were 15 gaps and four nucleotide residue differences resulting in 28 differences between the derived aminoacyl sequences (Figure 9.2). There are 11 gaps and eight nucleotide residue differences resulting in 19 aminoacyl differences between the derived MmoC sequences (Figure 9.3). The sequence entry in Genbank will be updated.

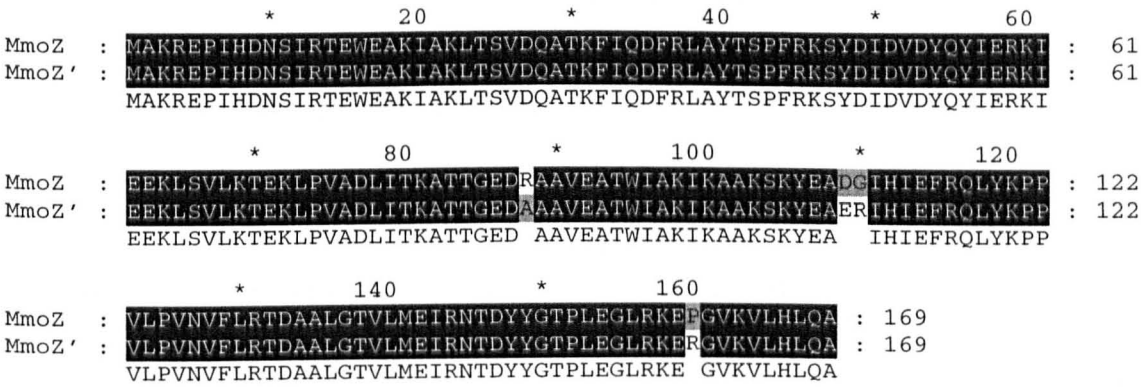


Figure 9.1 Alignment of the derived MmoZ sequences. MmoZ, the original sequence (Genbank accession CAA39071); MmoZ', data obtained in this study.

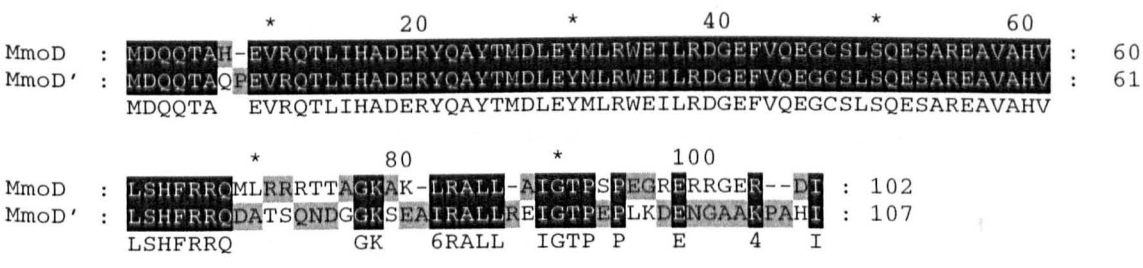


Figure 9.2 Alignment of the derived MmoD sequences. MmoD, the original sequence (Genbank accession CAD61957); MmoD', data obtained in this study. The alignment is shaded according to conservation of property (GeneDoc, http://www.psc.edu/biomed/genedoc/).

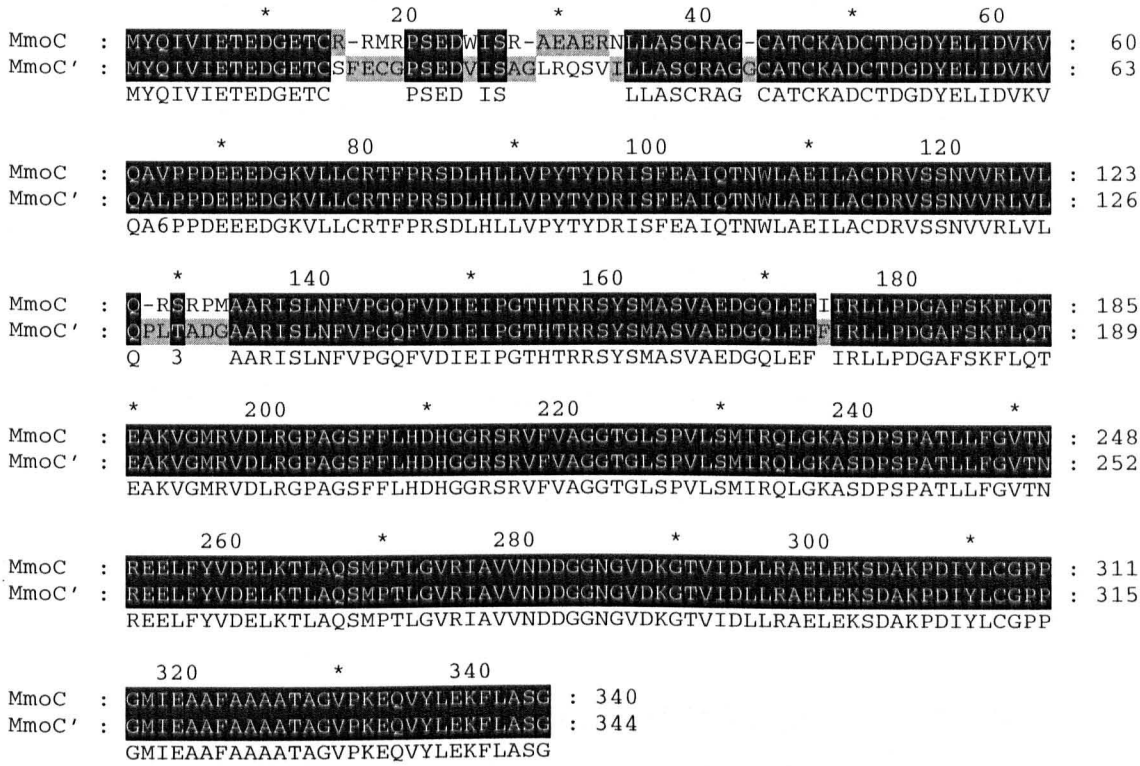


Figure 9.3 Alignment of the derived MmoC sequences. MmoC, the original sequence (Genbank accession CAB45257); MmoC', data obtained in this study. The alignment is shaded according to conservation of property (GeneDoc, http://www.psc.edu/biomed/genedoc/).

Sequencing technology and computer software for sequence data analysis has advanced greatly since these *mmo* genes were originally sequenced. Re-sequencing the *mmoZDC* genes of *Ms. trichosporium* OB3b has revealed numerous putative errors in the original sequence. The quality of the chromatogram trace data obtained in this study is good, and both DNA strands were sequenced where it demonstrated errors in the original data (the chromatogram trace data are shown in the appendix).

The derived peptide sequence of the corrected *mmoZ* sequence more closely matches the MmoZ sequence derived by X-ray diffraction of the sMMO hydroxylase (Elango *et al.*, 1997). The X-ray diffraction derived sequence (Genbank accession 1MHZG) differs only at position 110, which was suggested to be alanine instead of the glycine. The authors found that the glycine was sterically possible, but the electron density map indicated that the residue more closely resembled an alanine, as found in *Methylococcus capsulatus* (Bath). The nucleotide sequencing performed here indicates that the residue is in fact an arginine, as found also at this position in

the derived MmoZ sequences from *Methylocystis* spp. (Genbank accessions AAC45292 and AAF01271) (Grosse *et al.*, 1999; McDonald *et al.*, 1997).

The derived peptide sequences of the original and corrected *mmoD* genes share 72 % identity. The *mmoD* gene is the least conserved open reading frame within the *mmo* cluster and the closest relative of the corrected MmoD sequence is that of *Methylocystis* sp. strain M at 58 %; the MmoD sequence of *Methylococcus capsulatus* (Bath) shares 31 % sequence identity.

The corrections in the *mmoC* gene sequence improve the match between the derived MmoC sequences of *Ms. trichosporium* OB3b and *Methylocystis* sp. strain M (Figure 9.4). Many of the errors in the original *mmoC* sequence are gaps and resulted in frameshift errors when inferring the protein sequence.

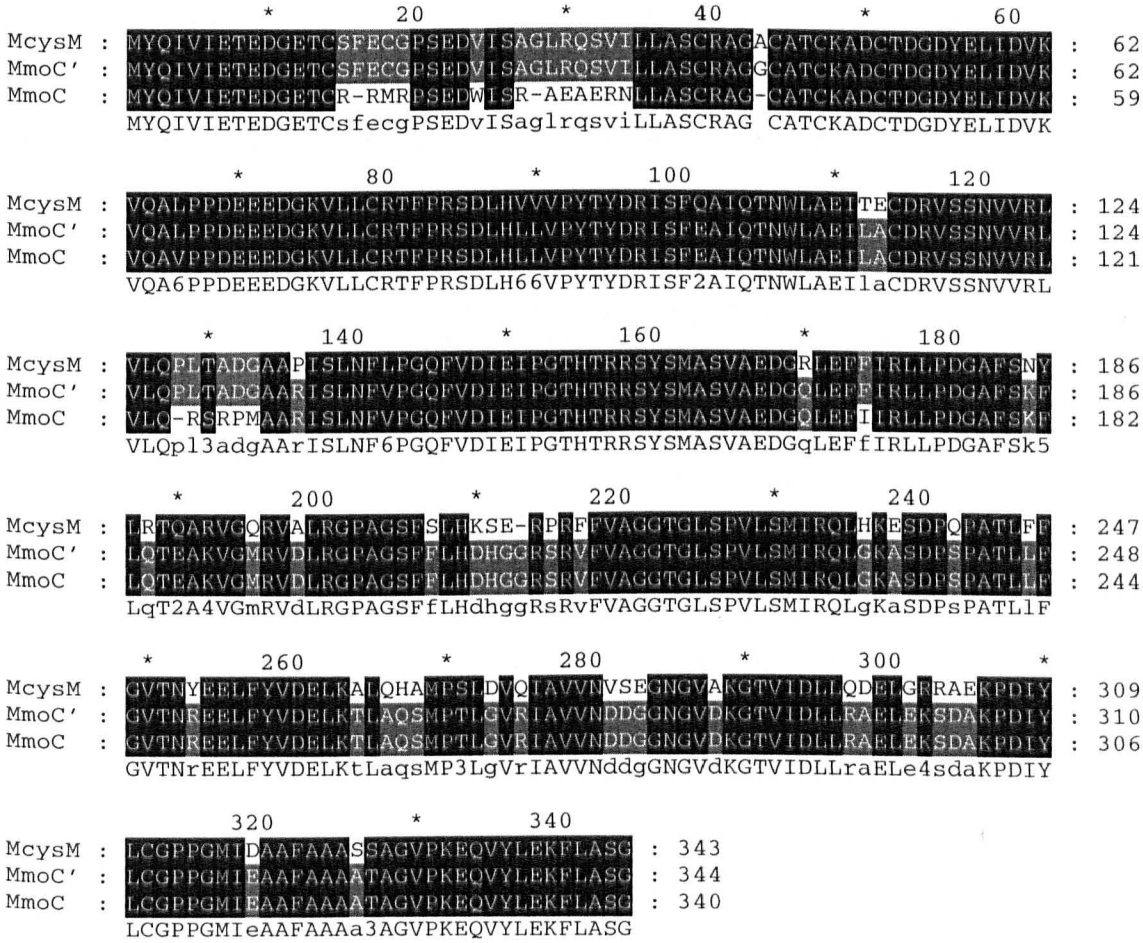


Figure 9.4 Alignment of the *Methylocystis* sp. strain M derived MmoC sequence (McysM; Genbank accession AAC45294) with the derived *Ms. trichosporium* OB3b MmoC sequences (labelled as in Figure 9.3). The alignment is shaded according to conservation of property (GeneDoc, http://www.psc.edu/biomed/genedoc/).

9.3 Homologous expression of recombinant sMMO

9.3.1 Introduction

The *Ms. trichosporium* Mutant F expression system has been used for the recombinant expression of wild-type sMMO (Lloyd *et al.*, 1999b) and also for mutagenized sMMO (Smith *et al.*, 2002). It should be possible to complement the sMMO$^-$ phenotype of *Ms. trichosporium* SMDM using the same sMMO expression plasmid used for recombinant sMMO expression in *Ms. trichosporium* Mutant F (Smith *et al.*, 2002). This is investigated here.

9.3.2 Complementation of *Ms. trichosporium* SMDM with pTJS175

The broad-host-range sMMO expression plasmid pTJS175 (Smith *et al.*, 2002) was introduced into *Ms. trichosporium* SMDM by conjugation (Section 2.7.1). Small colonies appeared after incubation with methane for one week on NMS agar containing 20 μg ml^{-1} streptomycin and 20 μg ml^{-1} spectinomycin. The colonies were subcultured onto NMS agar (0.8 μM copper) plates with 5 μg ml^{-1} gentamycin, streptomycin and spectinomycin selection. Substantial growth appeared on the plates after four weeks incubation with methane. A naphthalene plate assay was performed (Section 2.16) to test for sMMO activity and this was found to be strongly positive (Figure 9.5).

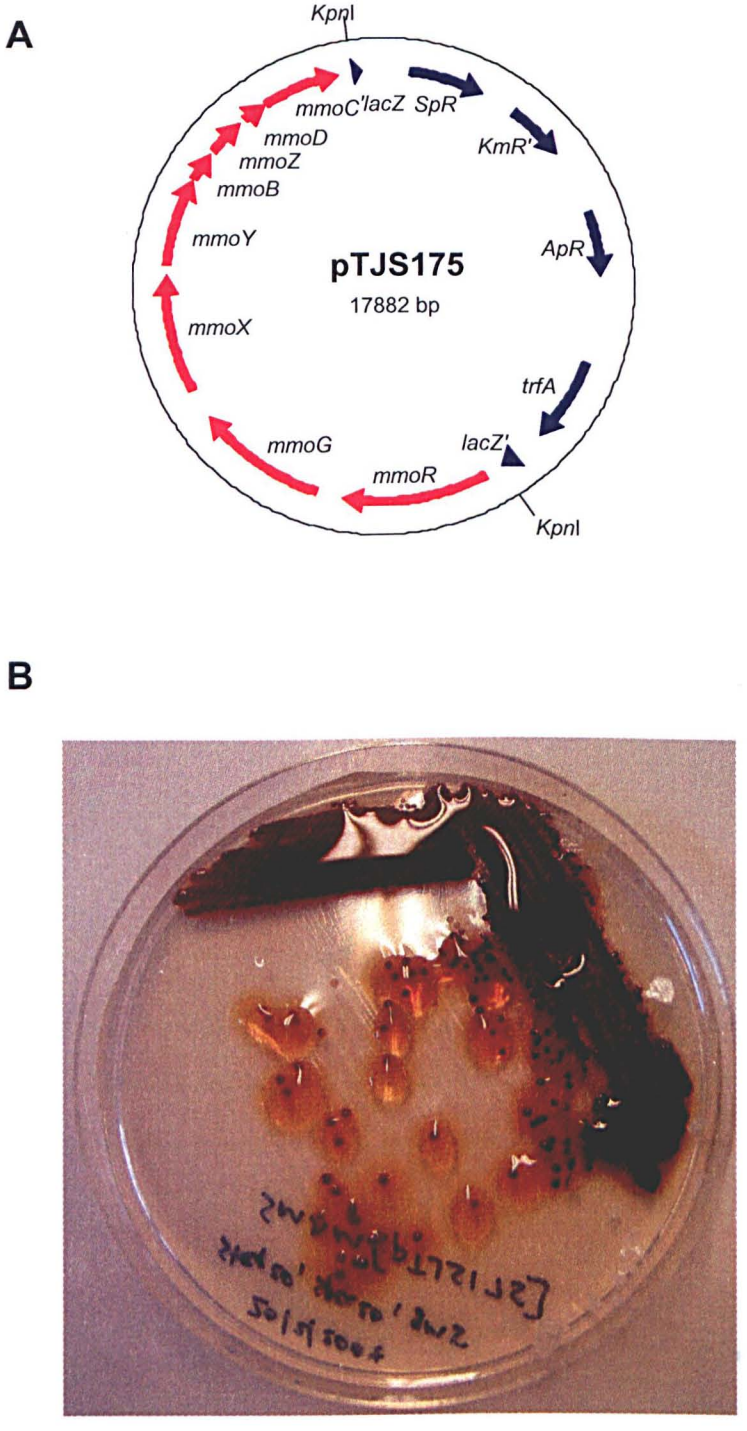


Figure 9.5 A) Map of the pTJS175 sMMO expression plasmid. B) Naphthalene assay of *Ms. trichosporium* SMDM carrying pTJS175; the medium used for the assay was NMS (0.8 µM copper) agar with 20 µg ml^{-1} spectinomycin, 20 µg ml^{-1} streptomycin and 5 µg ml^{-1} gentamycin antibiotics.

9.4 Heterologous expression of *Methylocella silvestris* BL2^T sMMO

9.4.1 Introduction

Mclla. silvestris BL2^T is a moderately acidophilic methanotroph isolated from a forest cambisol (Dunfield & Dedysh, 2003). *Methylocella* spp. are unique among methanotrophs in that they do not appear to possess a pMMO enzyme and *Mclla. silvestris* BL2^T can grow on acetate and ethanol (Dunfield & Dedysh, 2003). All *Methylocella* species possess a sMMO enzyme, but only partial sequence data of the *mmoX* gene have been published. The goal of this study was to clone the complete sMMO operon of *Mclla. silvestris* BL2^T, analyse the sequence data and attempt to express the cloned genes in *Ms. trichosporium* SMDM.

9.4.2 Cloning the *mmo* genes of *Methylocella silvestris* BL2^T

Mclla. silvestris BL2^T was kindly donated by Peter Dunfield (Max Planck Institute, Marburg) and the cloning of the *mmo* genes was performed by Stefan Radajewski (University of Warwick). The *mmoX* of *Mclla. silvestris* BL2^T was identified by PCR amplification using the 206F and 886R PCR primers (Chapter 4). The partial *mmoX* gene fragment was used as an homologous probe in a Southern blot (Section 2.10.1) of *Mclla. silvestris* BL2^T DNA digested with various restriction enzymes (Figures 9.6). The *Bam*HI (> 15 kb), *Eco*RI (~ 6 kb) and *Sac*I (~ 11 kb) fragments were targeted for cloning. Genomic DNA was digested with these restriction enzymes and resolved by agarose gel electrophoresis. Exposing the *Mclla. silvestris* BL2^T DNA to UV radiation was avoided by removing the lanes containing the DNA markers, which were stained by soaking in an ethidium bromide solution (0.5 μg ml^{-1} in water), photographed and used to generate a calibration curve. The positions of the target region for each restriction enzyme digest within the gel were estimated according to the calibration curve, and the agarose fragments were excised (± 1 kb) with a scalpel. The DNA was recovered using the Geneclean™II Kit (Section 2.8.9) and the DNA ligated into pUC18 vector that had been linearized with the same enzyme and dephosphorylated.

A *Sac*I clone (designated Clone 7 (pSMR1)) containing the *mmoX* gene was identified by PCR using the *mmoX* PCR primers (mmoX206F & mmoX886R). This clone was shotgun sequenced as described in Section 2.15.2, except that 1.5-2.5 kb fragments were subcloned. Contiguous sequences were joined and the ambiguous sequence data verified by directly sequencing the clone with several custom primers (Table 9.2).

Table 9.2 Primers used for sequencing Clone 7 (pSMR1).

Primer name	Sequence
Mclla-10700F	5'- CAGCCGGACGGCAATTGG -3'
Mclla-8150R	5'- CAACTTTGTCCACCTTCG -3'
Mclla-5560R	5'- TATTTCGTCCGGTTCTGC -3'
Mclla-4330R	5'- CATAGACGAAGATCTACG -3'
Mclla-2920R	5'- AATCGAGCGACGTCCAGC -3'

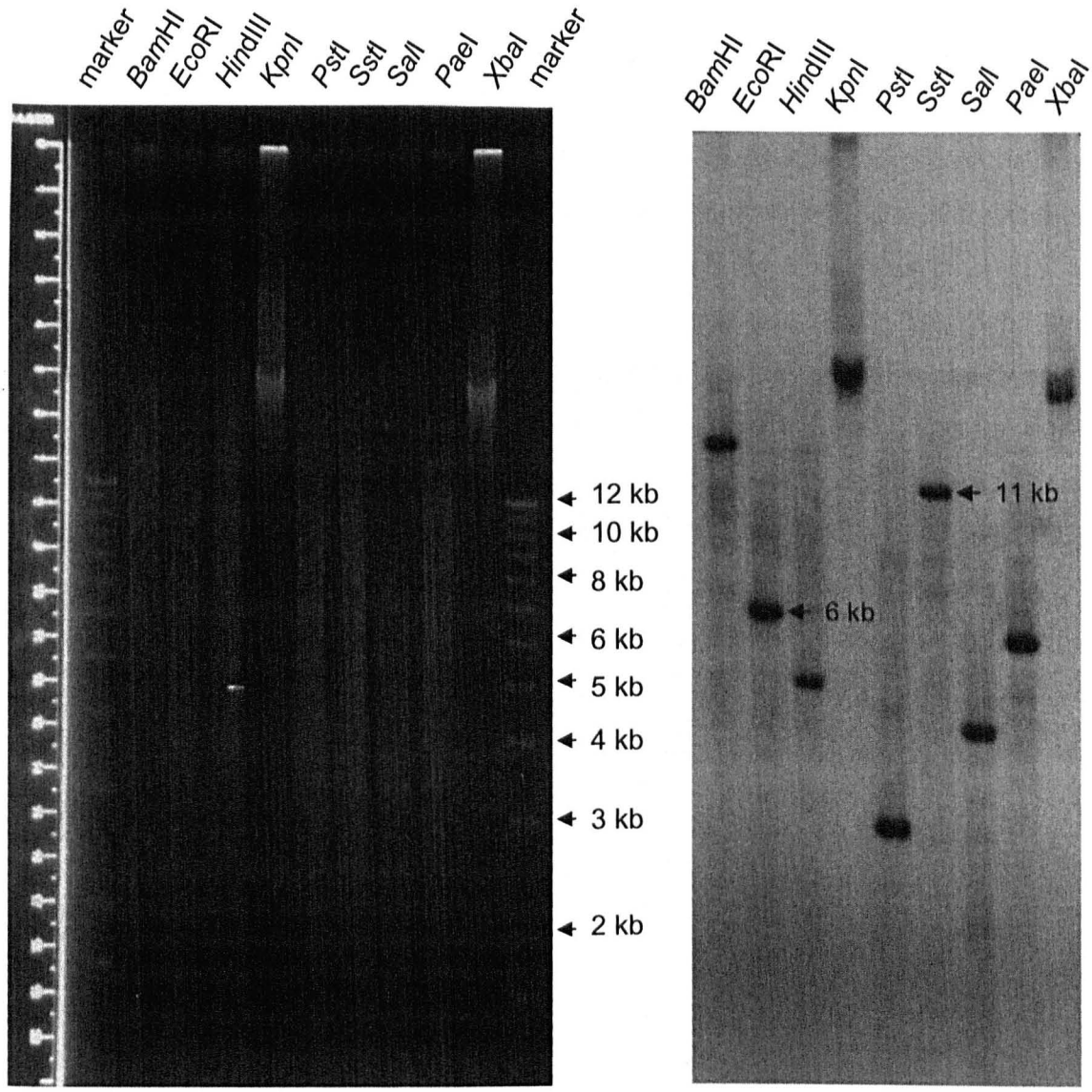


Figure 9.6 Agarose gel electrophoresis (left hand panel) and Southern blot hybridization (right hand panel) of *Mclla. silvestris* BL2^T genomic DNA digested with restriction enzymes (as indicated). The membrane was hybridized with a homologous *mmoX* probe, as described in the text.

A genetic map indicating the arrangement of putative genes on Clone 7 (pSMR1) is shown Figure 9.7. The *mmoXYBZDC* genes were all present and arranged in the same order as observed in all other sMMO operons (Section 1.5.2). The *mmoC* gene was immediately followed by the 5' end of a *mmoR* gene. The genes downstream (3') of the *mmo* operon were targeted by Southern blotting with a probe consisting of the 3' end of *mmoC* and the 5' *mmoR* gene fragment, that was generated by digesting pSMR1 with *Stu*I/*Sac*I (~ 750 bp). The bound *mmoX* probe was removed from the membrane filter and then re-hybridized with the radiolabelled *Stu*I/*Sac*I fragment (Figure 9.8). A *Hin*dIII fragment (~ 8 kb) was targeted and cloned in the same manner as used to clone the *mmoX* fragment. The clone that contained the *mmoC* and *mmoR* genes was identified by colony hybridization (Section 2.10.2) using the *Stu*I/*Sac*I probe. The clone was designated 'Clone 95' (or plasmid pSMR2) and was fully sequenced by Hanif Ali (University of Warwick). The arrangement of genes on Clone 95 is indicated in Figure 9.9. A Genbank search was done with the open reading frames on Clones 7 and 95. The results of the searches for genes with known functions in sMMO are shown in Table 9.3, and the results for genes with no known role in sMMO are summarized in Table 9.4. A putative σ^{54} binding site was found 126 bp upstream of the *mmoX* start codon, and a sequence comparison is shown in Figure 9.10.

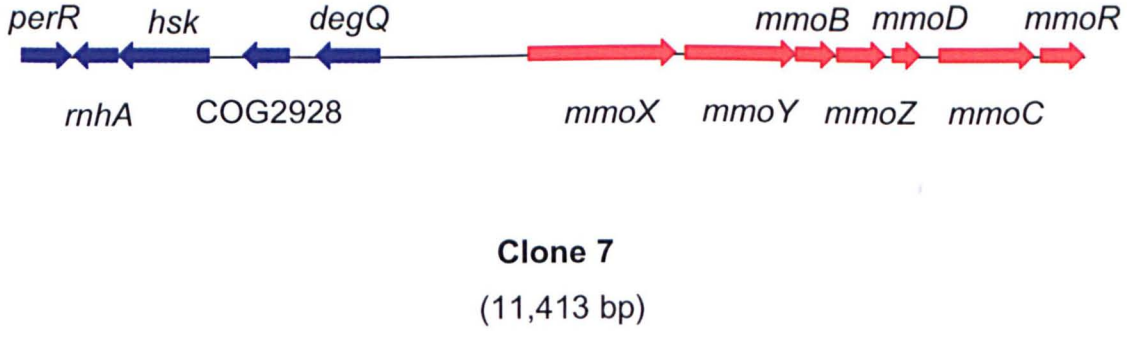


Figure 9.7 Genetic map of the *Mclla. silvestris* BL2^T 'Clone 7' on plasmid pSMR1. The genes with known structural or other roles in the sMMO are in red, and other putative genes are shown in blue.

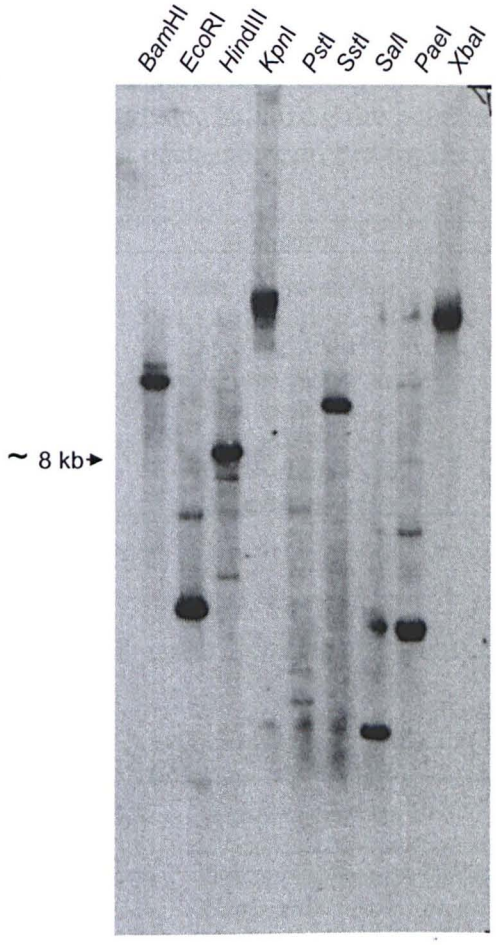


Figure 9.8 Southern blot hybridization of *Mclla. silvestris* BL2^T genomic DNA digested with restriction enzymes (as indicated). The membrane was hybridized with a homologous *mmoR* gene probe, as described in the text.

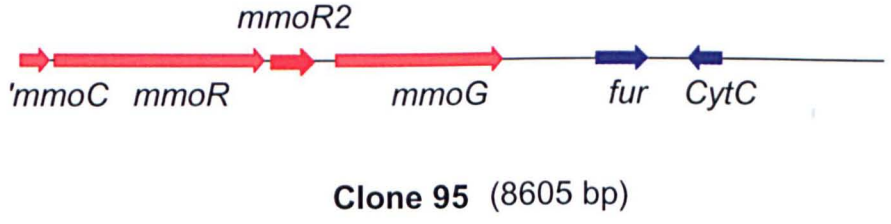


Figure 9.9 Genetic map of the *Mclla. silvestris* BL2^T Clone 95 (pSMR2). The genes with known structural or other roles in the sMMO are in red, and other putative genes are shown in blue. The *mmoR2* gene has no relatives in the database, but similar open reading frames are found in *Ms. trichosporium* OB3b and *Ms. sporium* (Section 1.5.3.1).

Table 9.3 A description of the *Mclla. silvestris* BL2^T structural and regulatory genes. A BLASTP search was performed using the derived aminoacyl sequences. The result of the top match are given, including the percent identity (BLAST ID), the organism name and the Genbank accession number for the retrieved sequence.

Gene	BLAST ID	Organism	Genbank Acc.
mmoX	85 %	*Ms. trichosporium*	CAA39068.2
mmoY	71 %	*Mcy.* sp. M	AAC45290.1
mmoB	67 %	*Mcy.* sp. WI14	AAF01270.1
mmoZ	63 %	*Mcy.* sp. WI14	AAF01271.1
mmoD	55 %	*Mcy.* sp. WI14	AAF01272.1
mmoC	54 %	*Mcy.* sp. M	AAC45294.1
mmoR	48 %	*Ms. trichosporium*	CAD61955.1
mmoG	49 %	*Ms. trichosporium*	CAD61956.1

Table 9.4 A description of genes flanking the *mmo* cluster of *Mclla. silvestris* BL2^T.

Gene (Enzyme)	BLAST ID	Organism	Genbank Acc.
ahpC (periredoxin)	57 %	*Mesorhizobium* sp. BNC1	ZP_00194129.2
rnhA (ribonuclease H)	63 %	*Rhodopseudomonas palustris*	NP_949605.1
hsk (homoserine kinase)	54 %	*Rhodopseudomonas palustris*	NP_949606.1
COG2928 (unknown function)	55 %	*Rhodopseudomonas palustris*	NP_948003.1
degQ (serine protease)	35 %	*Rhodopseudomonas palustris*	NP_947769.1
fur (transcriptional regulator)	50 %	*Rhodopseudomonas palustris*	NP_945777.1
cytC (cytochrome C precursor)	35 %	*Rhodopseudomonas palustris*	NP_946134.1

The arrangement of the sMMO structural genes in *Mclla. silvestris* BL2^T is the same as in all other sMMO gene clusters sequenced to date. The *mmo* genes are most closely related to those of other α-*Proteobacteria* methanotrophs (Table 9.3). As in *Ms. trichosporium* OB3b, the *mmoR* gene lies directly upstream (5') of *mmoG*; however these genes are positioned immediately upstream of *mmoXYBZDC* in *Ms. trichosporium* OB3b and immediately downstream in *Mclla. silvestris* BL2^T (Figure 1.8). There is only 43 bp separating the *mmoC* and *mmoR* genes in *Mclla. silvestris* BL2^T, with no evidence of a transcriptional stop sequence or a promoter consensus sequence within this region. Therefore, it is possible that *mmoR* (and *mmoG*) are co-transcribed with *mmoXYBZDC*. It would be necessary to perform Northern blots or RT-PCR to test the nature of *mmoRG* transcription in *Mclla. silvestris* BL2^T.

The presence of an *mmoR* gene, which encodes a σ^{54}-dependent transcriptional activator (Csáki *et al.*, 2003; Stafford *et al.*, 2003), in close proximity to the *mmoXYBZDC* cluster in *Mclla. silvestris* BL2^T suggests that *mmoX* gene transcription is initiated from a σ^{54}–dependent promoter in a manner similar to that of *Methylococcus capsulatus* (Bath) and *Ms. trichosporium* OB3b (Csáki *et al.*, 2003; Stafford *et al.*, 2003). A search upstream of *mmoX* located a putative σ^{54} binding site (Figure 9.10). The -24 box of the putative consensus differs at two residues compared with the consensus (Barrios *et al.*, 1999) and with other putative *mmoX* promoters. It will be necessary to locate the transcriptional start site by primer extension analysis (or a similar technique) to establish if this sequence is indeed the RNAP binding site.

	-24	-12
Bacterial σ^{54} consensus	**TGGCAC—N_5—TTGCW**	
Ms. trichosporium OB3b *mmoX*	**TGGCAC—N_5—TTGCC**	
Mcys. strain M *mmoX*	**TGGCAC—N_5—TTGCC**	
Mc. capsulatus (Bath) *mmoX*	**TGGCAC—N_5—CTGTA**	
Mm. strain KSWIII *mmoX*	**TGGCAC—N_5—TTGCA**	
Mclla. silvestris BL2^T *mmoX*	**CGGCGC—N_5—TTGCT**	

Figure 9.10 Comparison of a putative σ^{54} promoter binding site positioned 126 bp upstream of the *Mclla. silvestris* BL2^T *mmoX* gene with the bacterial consensus and with putative and known *mmoX* σ^{54} promoter consensus sequences. 'W' is A or T; '*Mcys.*' is *Methylocystis* and '*Mm.*' is *Methylomonas*. Residues are represented in bold-type when matching the bacterial consensus.

9.4.3 Expression of *Mclla. silvestris* BL2^T sMMO in *Ms. trichosporium* SMDM

Clone 7 (pSMR1) contains all the sMMO structural genes and the region 5' of the *mmo* operon (Figure 9.7). The genes were subcloned as an *Sst*I fragment into the pTJS140 broad-host-range vector linearized with the same enzyme. The resulting plasmid, pSMR1-TJS140, was suitable for conjugation into *Ms. trichosporium* SMDM.

The conjugation was performed as described in Section 2.7.1 and transconjugants were selected on NMS agar plates containing 20 μg ml^{-1} spectinomycin and 20 μg ml^{-1} streptomycin. Colonies appeared after one week incubation with methane. 25 colonies were subcultured onto NMS agar plates containing 5 μg ml^{-1} gentamycin, 20 μg ml^{-1} spectinomycin and 20 μg ml^{-1} streptomycin. All the colonies grew after 4 weeks incubation. The colonies were subcultured again onto NMS plates (0.8 μM copper) for the purpose of performing a naphthalene assay. The naphthalene assay has been used successfully to monitor sMMO expression in *Mclla. silvestris* BL2^T (Dunfield & Dedysh, 2003). *Ms. trichosporium* SMDM carrying pTJS175 (Section 9.3.2) was patched onto the same plate, to be used as a positive control for the assay. The naphthalene plate assay was performed on the colonies after 4 weeks incubation with methane. The control gave a positive reaction, but there was no colour change with *Ms. trichosporium* SMDM carrying pSMR1-TJS140 (results not shown). This indicated that there was no recombinant expression of the *Mclla. silvestris* BL2^T sMMO in *Ms. trichosporium* SMDM.

9.5 Recombinant sMMO expression in the absence of *mmoD*

9.5.1 Introduction

In Section 9.3 it was shown that *Ms. trichosporium* SMDM is a functional homologous sMMO expression host. The aim of this study was to attempt the recombinant expression of sMMO in *Ms. trichosporium* SMDM without the *mmoD* gene.

9.5.2 Plasmid construction

The construction of an sMMO expression vector lacking the *mmoD* gene was performed in relatively few steps. By serendipity, a unique *Not*I site was present in the intergenic region between *mmoZ* and *mmoD*. A second *Not*I site was introduced 3' of *mmoD* by PCR, using primers: forward, 5' ATA GCG GCC GCT AGC AGG AGG GAC TT*A TGT* ACC AGA TCG TCA TCG 3' (the artificial *Not*I site is underlined, and the *mmoC* start codon is italicized) and reverse, 5' ATA ACT AGT TCG ACC GAA GAG AAG AGT CG 3'. A *Nhe*I site, that partially overlaps the *Not*I site, was also included in the forward primer so that the *mmoD* gene could be subsequently inserted into the plasmid by directional cloning. The pMD4 plasmid was used as template and the PCR product (1,113 bp) was cloned into pCR2.1 (Section 2.8.7) and sequenced with the M13F and M13R primers. There were several unexpected sequence differences compared with the published data for the *Ms. trichosporium* OB3b *mmoC*, which prompted the re-sequencing work presented in Section 9.2. The sequence of the PCR product matched the re-sequenced data precisely (except for the differences introduced by the PCR primers). The pMD4 plasmid and the TA-cloned PCR product were digested with *Not*I and *Bst*API and the *mmoDC* region from pMD4 replaced with the PCR amplified *mmoC* gene; the resulting plasmid was called pD-del (Figure 9.11). The region of pD-del containing the *mmoZ* and *mmoC* genes was re-sequenced, and this confirmed that no errors were introduced during the cloning process.

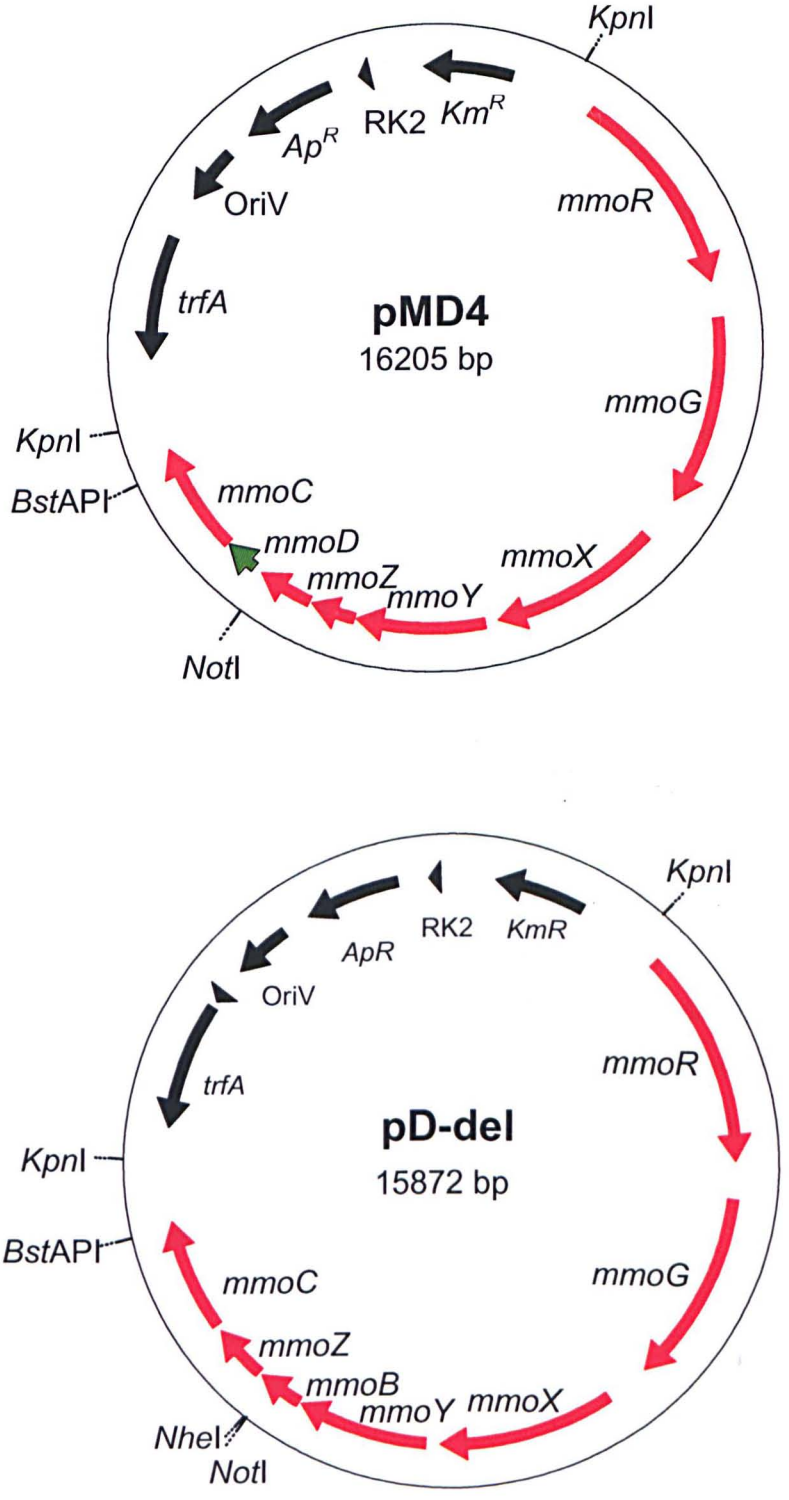


Figure 9.11 Plasmid maps of pMD4 and pD-del.

9.5.3 Conjugation and expression experiments

An attempt was made to conjugate the pD-del and pMD4 plasmids into *Ms. trichosporium* SMDM (Section 2.7.1). Selection of transconjugants was performed on NMS agar plates containing 5 $\mu g\ ml^{-1}$ gentamycin and 12.5 $\mu g\ ml^{-1}$ kanamycin. No transconjugants were obtained. The reason for this is unclear, but the conjugation of plasmids requiring selection with kanamycin into *Ms. trichosporium* are often more troublesome than with plasmids requiring spectinomycin/ streptomycin selection (Julie Scanlan, personal communication).

The vector backbone was changed by releasing the insert with *Kpn*I and cloning into the *Kpn*I site of pTJS140. The resulting plasmid, designated pD-del-140, was conjugated into *Ms. trichosporium* SMDM. Transconjugants were selected on NMS agar plates containing 20 $\mu g\ ml^{-1}$ spectinomycin and 20 $\mu g\ ml^{-1}$ streptomycin. Colonies appeared after one week of incubation with methane and 15 colonies were subcultured onto NMS agar plates containing 5 $\mu g\ ml^{-1}$ gentamycin, 20 $\mu g\ ml^{-1}$ spectinomycin and 20 $\mu g\ ml^{-1}$ streptomycin. The colonies were subcultured again onto NMS plates (0.8 μM copper) for the purpose of performing a naphthalene assay. *Ms. trichosporium* SMDM carrying pTJS175 (Section 9.3.2) was patched onto the same plate, to be used as a positive control. The naphthalene plate assay was performed on the plates after 4 weeks incubation with methane. The control gave a positive reaction, but there was no colour change with *Ms. trichosporium* SMDM carrying pD-del-140 (results not shown). SDS-PAGE analysis showed no subunits in *Ms. trichosporium* SMDM carrying pD-del-140 (results not shown). Although these are preliminary data suggesting that the *mmoD* gene is essential for sMMO expression, clearly more experiments are required.

9.6 Discussion and Future Perspectives

9.6.1 Recombinant expression of *Methylocella silvestris* BL2^T sMMO

For expression of the *Mclla. silvestris* BL2^T sMMO genes to occur in *Ms. trichosporium* SMDM, the *cis* elements required for transcription must be recognized in *Ms. trichosporium* SMDM. This includes the binding of the *Ms. trichosporium* RNA polymerase (RNAP) to the *Mclla. silvestris* BL2^T *mmoX* promoter. The -24, -12 boxes of a putative σ^{54} binding site for sMMO transcription in *Mclla. silvestris* BL2^T have been identified and three differences exist between the sequence and that of *Ms. trichosporium* (Figure 9.10). Hypothetically, these differences may be sufficient to prevent the *Ms. trichosporium* σ^{54}–RNAP binding to the *Mclla. silvestris* BL2^T *mmoX* promoter. Transcription from σ^{54}–dependent promoters also requires activation from an enhancer-binding transcriptional regulator (see Section 1.5.3.1). The putative σ^{54}–dependent transcriptional regulator for *mmoX* is called MmoR (Csáki *et al.*, 2003; Stafford *et al.*, 2003), and an *mmoR* homologue was identified downstream (3') of *mmoC* in *Mclla. silvestris* BL2^T. Because it was cloned separately from *mmoXYBZDC*, the *Mclla. silvestris* BL2^T *mmoR* gene was not included on the expression plasmid. Therefore, expression of the *mmoXYBZDC* genes in *Ms. trichosporium* required that the *Ms. trichosporium* MmoR initiate transcription from the *Mclla. silvestris* BL2^T promoter. The MmoR proteins of *Mclla. silvestris* BL2^T and *Ms. trichosporium* OB3b share 48 % sequence identity and the regions upstream of *mmoX* (the putative region for MmoR binding) in *Mclla. silvestris* BL2^T and *Ms. trichosporium* OB3b possess little similarity (results not shown). Because of the relatively high sequence difference between the MmoR proteins and their putative DNA binding sites, it seems likely that the *Ms. trichosporium* OB3b MmoR protein would not initiate transcription of the *Mclla. silvestris* BL2^T *mmoX* promoter. This is the most likely reason for not obtaining heterologous expression of *Mclla. silvestris* BL2^T sMMO in *Ms. trichosporium* SMDM.

There are additional experiments which need to be performed. The first will be to express the *Mclla. silvestris* BL2^T sMMO from the *Ms. trichosporium* OB3b *mmoX* promoter. If functional expression is obtained, this will confirm that the lack of expression obtained in this experiment did indeed result because the *Mclla. silvestris* BL2^T *cis*-acting elements for transcription were not compatible with the *Ms.*

trichosporium expression machinery. At this time, experiments designed to locate the MmoR binding site and identify the exact position of the *Ms. trichosporium* OB3b *mmoX* promoter is underway by Julie Scanlan in the JCM laboratory. Secondly, the heterologous expression experiment should be repeated using a *Mclla. silvestris* $BL2^T$ sMMO expression plasmid also carrying the *mmoRG* genes in addition to *mmoXYBZDC*.

9.6.2 Expression of sMMO in the absence of *mmoD*

A review of the literature relevant to the *mmoD* is present in Section 1.5.1.4. The experiment performed here is the first attempt at the expression of sMMO in the absence of *mmoD*, and the preliminary results suggest that *mmoD* is required for expression of sMMO.

Several of the necessary control experiments have yet to be performed. First, it may be necessary to probe the recombinant *Ms. trichosporium* SMDM with *mmo* gene probes to prove that the plasmid is present. However, experience has shown that *Ms. trichosporium* colonies that grow on NMS agar containing spectinomycin and streptomycin antibiotic selection following conjugation with a pTJS140-based plasmid are nearly always bona fide transconjugants.

The proper positive control for the experiment is to re-introduce the *mmoD* gene into pD-del-140 and show that the ability of the plasmid to drive sMMO expression in *Ms. trichosporium* SMDM is restored. There is little reason not to expect the plasmid to restore sMMO expression. The *Not*I and *Nhe*I sites present between *mmoZ* and *mmoC* make the construction of the plasmid a relatively simple cloning experiment; the *Not*I site is already present upstream (5') of *mmoD* and the *Nhe*I site can be introduced by incorporating the recognition site into the reverse PCR primer.

Once it has been proved that the *mmoD* is required for sMMO expression, the precise role of MmoD can be explored. There is a possibility that MmoD is involved in transcription of *mmo* genes, and this can be investigated using RT-PCR. The current hypothesis is that the MmoD is involved in inserting iron into the active site of the sMMO hydroxylase (Merkx & Lippard, 2002) and this role has been shown for the DmpK protein of the phenol hydroxylase (Powlowski *et al.*, 1997). Like the sMMO, the phenol hydroxylase is a non-haem iron monooxygenase and there are

some similarities between the MmoD and DmpK (Merkx & Lippard, 2002). The experiments of Powlowski and co-workers demonstrating that the DmpK is necessary for inserting the iron molecules into the active-site of the phenol hydroxylase were performed *in vitro*. Phenol hydroxylase and DmpK were over-expressed separately in *E. coli*, and it was shown that active enzyme was obtained only following incubation with both iron and DmpK. It is possible that a similar experiment can be performed with sMMO enzyme purified from *Ms. trichosporium* SMDM carrying pD-del-140 and incubated with MmoD over-expressed in *E. coli*.

Chapter 10

Synopsis

Chapter 3

i. A plasmid system for the mutagenesis of the *Ms. trichosporium* OB3b sMMO α-subunit was created. The sMMO expression plasmid (pMD2) could be conjugated into *Ms. trichosporium* Mutant F in a fashion similar to the expression plasmid (pTJS175) previously developed for this purpose (Smith *et al.*, 2002); the advantage of the new system is that the mutagenesis requires one less cloning step and site-directed modifications can be performed over a larger region of the *mmoX* gene.

ii. Chimaeric sMMO mutants were created by introducing gene sequence from the alkene monooxygenase enzyme into the *mmoX*. The mutant enzymes expressed in *Ms. trichosporium* Mutant F could not be detected by SDS-PAGE, which is a characteristic of unstable sMMO mutants previously analysed (Smith *et al.*, 2002).

Chapter 4

i. New PCR primers targeting the *mmoX* were developed. The primers amplify the active site region of the sMMO and match all the *mmoX* genes in Genbank.

ii. The primers were used in this study to generate libraries from a blanket bog peat and Movile Cave water. In a separate study the primers were used to generate an *mmoX* PCR library from a soda lake (Lin *et al.*, in press).

iii. Genes that group with the sequences from *Methylococcus* spp., *Methylomonas* spp. and *Methylosinus/ Methylocystis* spp. were obtained. A group of "novel *mmoX*" sequences that did not cluster with the *mmoX* of any known methanotroph was also obtained from Movile Cave water.

iv. A comparison of the derived MmoX aminoacyl sequences from this "novel *mmoX*" with the MmoX from *Ms. trichosporium* OB3b and *Methylococcus capsulatus* (Bath) indicated that the residue differences were conservative. A chimaeragenesis experiment (Chapter 3) was not performed since it was unlikely to alter the catalytic properties of the enzyme.

Chapter 5

i. A method was developed to clone ^{13}C-DNA from a DNA-SIP experiment into a BAC vector. A DNA purification method using agarose gel electrophoresis and Nile blue staining was developed.

ii. ^{13}C-DNA from a DNA-SIP experiment with forest soil that consumed 50 cc $^{13}CH_4$ was partially digested with *Bam*HI and cloned into the pCC1BAC^{tm} vector. A library of 2300 clones was generated.

iii. 48 clones were analysed by RFLP. > 95 % of plasmids contained inserts and ranged in size from approximately 10 – 30 kb.

iii. The library was screened for *mxaF*, *mmoX* and *pmoA* by colony hybridization. Two clones containing *pmoA* genes were identified and sequencing indicated that they were very similar (ie. from strains of the same species). One of the *pmoA* clones was sequenced and found to be 15 kb. Other genes encoding proteins with (potential) roles in methylotrophy were contained on the plasmid, including *pmoC, pmoB, folP, folK, mptG* and *moxF*.

Chapter 6

The objective of this study was to disrupt the Lon protease of *Ms. trichosporium* Mutant F and test if the expression of unstable sMMO mutants was improved.

i. The *lon* gene of *Ms. trichosporium* OB3b was identified by Southern blotting and probing with *Sinorhizobium meliloti* 1021 *lon* gene. The *Ms. trichosporium* OB3b *lon* was cloned and sequenced.

ii. Attempts were made to disrupt the *Ms. trichosporium* OB3b *lon* by marker-exchange mutagenesis using the suicide vectors pK18mob and pK18mobsacB. A mutant was not obtained, indicating that Lon may be essential for vegetative growth of *Ms. trichosporium* OB3b.

Chapter 7

i. Sequence analysis of the *mmoG* from methanotrophs was performed. The results suggest that translation of MmoG is initiated from a GTG codon in *Ms. trichosporium* OB3b and *Ms. sporium*.

ii. A sMMO expression vector (pMD30) was created by replacing the *mmoX* promoter with the promoter for neomycin phosphotransferase. A plasmid for the expression of the *Ms. trichosporium* OB3b *mmoG* was created (pMD38). The *mmoXYBZDC* and *mmoG* genes were co-expressed from the different plasmids in *E. coli* BLR(DE3); no naphthalene oxidation activity was detected.

iii. The pMD30 plasmid was conjugated into *Ms. trichosporium* Mutant F; no naphthalene oxidation activity was detected, nor were sMMO proteins detected by SDS-PAGE.

Chapter 8

i. Marker-exchange mutagenesis was used to knock-out the sMMO genes from the chromosome of *Ms. trichosporium* OB3b. The *mmoX, mmoY, mmoB, mmoZ, mmoD* genes and the first few codons of *mmoC* were removed by the mutagenesis. The *mmoX* promoter was also removed. The strain, designated *Ms. trichosporium* SMDM, displayed an sMMO-minus phenotype.

Chapter 9

i. The *mmoZ, mmoD* and *mmoC* genes from *Ms. trichosporium* OB3b were re-sequenced. The MmoZ sequence determined by X-ray crystallography (Elango *et al.*, 1997) corresponded well with the re-sequenced data.

ii. The wild-type sMMO genes carried on pTJS175 were conjugated into *Ms. trichosporium* SMDM. The transconjugants expressed sMMO, and the expression was detected by SDS-PAGE and by the naphthalene oxidation assay; this result indicated that *Ms. trichosporium* SMDM is a suitable expression system for sMMO.

Unlike *Ms. trichosporium* Mutant F whose suitability is limited to the analysis of sMMO α-subunit mutants, *Ms. trichosporium* SMDM is suitable for the analysis of mutations in any of the sMMO structural genes.

ii. The sMMO genes from *Methylocella silvestris* $BL2^T$ were sequenced and analysed. The *mmoXYBZDC* genes from *Methylocella silvestris* $BL2^T$ were conjugated into *Ms. trichosporium* SMDM. Transconjugants were analysed and no naphthalene oxidation activity was detected, nor were sMMO proteins detected by SDS-PAGE.

iii. A plasmid was created, designated pD-del, that was virtually identical to pTJS175 except lacked the *mmoD* gene. Unlike pTJS175, this plasmid was unable to drive the expression of sMMO in *Ms. trichosporium* SMDM. This suggests that *mmoD* is essential for sMMO expression.

References

Adkins, S. & Burmeister, M. (1996). Visualization of DNA in agarose gels as migrating colored bands: applications for preparative gels and educational demonstrations. *Anal Biochem* **240**, 17-23.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.

Alvarez Cohen, L., Mccarty, P. L., Boulygina, E., Hanson, R. S., Brusseau, G. A. & Tsien, H. C. (1992). Characterisation of a methane-utilizing bacterium from a bacterial consortium that rapidly degrades trichloroethylene and chloroform. *Applied and Environmental Microbiology* **58**, 1886-1893.

Amann, R. I., Ludwig, W. & Schleifer, K. H. (1995). Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiological Reviews* **59**, 143-169.

Amaral, J. A. & Knowles, R. (1995). Growth of methanotrophs in methane and oxygen counter gradients. *FEMS Microbiology Letters* **126**, 215-220.

Anthony, C. (1982). The bacterial oxidation of methane, methanol, formaldehyde and formate. In *The Biochemistry of the Methylotrophs*, pp. 152-194. London: Academic Press.

Anthony, C. (1986). Bacterial oxidation of methane and methanol. *Adv Microb Physiol* **27**, 113-210.

Anthony, C. & Dales, S. L. (1996). The biochemistry of methanol dehydrogenase. In *Microbial Growth on C1 Compounds*, pp. 213-219. Edited by M. E. Lidstrom & F. R. Tabita. The Netherlands: Kluwer Academic Publishers.

Aravind, L. & Ponting, C. P. (1997). The GAF domain: an evolutionary link between diverse phototransducing proteins. *Trends Biochem Sci* **22**, 458-459.

Arp, D. J., Sayavedra-Soto, L. A. & Hommes, N. G. (2002). Molecular biology and biochemistry of ammonia oxidation by *Nitrosomonas europaea*. *Archives of Microbiology* **178**, 250-255.

Auman, A. J., Stolyar, S., Costello, A. M. & Lidstrom, M. E. (2000). Molecular characterization of methanotrophic isolates from freshwater lake sediment. *Applied and Environmental Microbiology* **66**, 5259-5266.

Auman, A. J. & Lidstrom, M. E. (2002). Analysis of sMMO-containing type I methanotrophs in Lake Washington sediment. *Environ Microbiol* **4**, 517-524.

Balendra, S., Lesieur, C., Smith, T. J. & Dalton, H. (2002). Positively charged amino acids are essential for electron transfer and protein-protein interactions in the soluble methane monooxygenase complex from *Methylococcus capsulatus* (Bath). *Biochemistry* **41**, 2571-2579.

Barrios, H., Valderrama, B. & Morret, E. (1999). Compilation and analysis of σ^{54}-dependent promoter sequences. *Nucleic Acids Res* **27**, 4305-4313.

Basu, P., Katterle, B., Andersson, K. K. & Dalton, H. (2003). The membrane-associated form of methane mono-oxygenase from *Methylococcus capsulatus* (Bath) is a copper/iron protein. *Biochem J* **369**, 417-427.

Bédard, C. & Knowles, R. (1989). Physiology, biochemistry, and specific inhibitors of CH_4, NH_4^+, and CO oxidation by methanotrophs and nitrifiers. *Microbiological Reviews* **53**, 68-84.

Beja, O., Aravind, L., Koonin, E. V. & other authors (2000). Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**, 1902-1906.

Beja, O., Suzuki, M. T., Heidelberg, J. F., Nelson, W. C., Preston, C. M., Hamada, T., Eisen, J. A., Fraser, C. M. & DeLong, E. F. (2002). Unsuspected diversity among marine aerobic anoxygenic phototrophs. *Nature* **415**, 630-633.

Bender, M. & Conrad, R. (1993). Kinetics of methane oxidation in oxic soils. *Chemosphere* **26**, 687-696.

Bernhardt, R. (2004). Optimized chimeragenesis; creating diverse P450 functions. *Chem Biol* **11**, 287-288.

Blatny, J. M., Brautaset, T., Winther-Larsen, H. C., Haugan, K. & Valla, S. (1997). Construction and use of a versatile set of broad-host-range cloning and expression vectors based on the RK2 replicon. *Appl Environ Microbiol* **63**, 370-379.

Blazyk, J. L. & Lippard, S. J. (2002). Expression and characterization of ferredoxin and flavin adenine dinucleotide binding domains of the reductase component of soluble methane monooxygenase from *Methylococcus capsulatus* (Bath). *Biochemistry* **41**, 15780-15794.

Bodrossy, L., Holmes, E. M., Holmes, A. J., Kovacs, K. L. & Murrell, J. C. (1997). Analysis of 16S rRNA and methane monooxygenase gene sequences reveals a novel group of thermotolerant and thermophilic methanotrophs, *Methylocaldum* gen. nov. *Archives of Microbiology* **168**, 493-503.

Botos, I., Melnikov, E. E., Cherry, S. & other authors (2004). The catalytic domain of *Escherichia coli* Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J Biol Chem* **279**, 8140-8148.

Bourne, D. G., Holmes, A. J., Iversen, N. & Murrell, J. C. (2000). Fluorescent oligonucleotide rDNA probes for specific detection of methane oxidising bacteria. *FEMS Microbiology Ecology* **31**, 29-38.

Bourne, D. G., McDonald, I. R. & Murrell, J. C. (2001). Comparison of *pmoA* PCR primer sets as tools for investigating methanotroph diversity in three Danish soils. *Applied and Environmental Microbiology* **67**, 3802-3809.

Bowman, J. (2000). The Methanotrophs -The Families Methylococcaceae and Methylocystaceae. In *The Prokaryotes*. New York: Springer-Verlag.

Bowman, J. P., Skerratt, J. H., Nichols, P. D. & Sly, L. I. (1991). Phospholipid fatty acid and lipopolysaccharide fatty acid signature lipids in methane-utilising bacteria. *FEMS Microbiology Ecology* **85**, 15-22.

Bowman, J. P., Sly, L. I., Nichols, P. D. & Hayward, A. C. (1993). Revised taxonomy of the methanotrophs: Description of *Methylobacter* gen. nov., emendation of *Methylococcus*, validation of *Methylosinus* and *Methylocystis* species, and a proposal that the family *Methylococcaceae* includes only the group I methanotrophs. *International Journal of Systematic Bacteriology* **43**, 735-753.

Bowman, J. P., Sly, L. I. & Stackebrandt, E. (1995). The phylogenetic position of the family *Methylococcaceae. International Journal of Systematic Bacteriology* **45**, 182-185.

Bowman, J. P., McCammon, S. A. & Skerratt, J. H. (1997). *Methylosphaera hansonii* gen. nov., sp. nov., a psychrophilic, group I methanotroph from Antarctic marine-salinity, meromictic lakes. *Microbiology* **143**, 1451-1459.

Brandstetter, H., Whittington, D. A., Lippard, S. J. & Frederick, C. A. (1999). Mutational and structural analyses of the regulatory protein B of soluble methane monooxygenase from *Methylococcus capsulatus* (Bath). *Chem Biol* **6**, 441-449.

Bratina, B. J., Brusseau, G. A. & Hanson, R. S. (1992). Use of 16S rRNA analysis to investigate phylogeny of methylotrophic bacteria. *International Journal of Systematic Bacteriology* **42**, 645-648.

Brazeau, B. J., Austin, R. N., Tarr, C., Groves, J. T. & Lipscomb, J. D. (2001a). Intermediate Q from soluble methane monooxygenase hydroxylates the mechanistic substrate probe norcarane: evidence for a stepwise reaction. *J Am Chem Soc* **123**, 11831-11837.

Brazeau, B. J., Wallar, B. J. & Lipscomb, J. D. (2001b). Unmasking of deuterium kinetic isotope effects on the methane monooxygenase compound Q reaction by site-directed mutagenesis of component B. *J Am Chem Soc* **123**, 10421-10422.

Brazeau, B. J. & Lipscomb, J. D. (2003). Key amino acid residues in the regulation of soluble methane monooxygenase catalysis by component B. *Biochemistry* **42**, 5618-5631.

Bretz, J., Losada, L., Lisboa, K. & Hutcheson, S. W. (2002). Lon protease functions as a negative regulator of type III protein secretion in *Pseudomonas syringae*. *Mol Microbiol* **45**, 397-409.

Brusseau, G. A., Tsien, H. C., Hanson, R. S. & Wackett, L. P. (1990). Optimization of trichloroethylene oxidation by methanotrophs and the use of a colorimetric assay to detect soluble methane monooxygenase activity. *Biodegradation* **1**, 19-29.

Bukhari, A. I. & Zipser, D. (1973). Mutants of *Escherichia coli* with a defect in the degradation of nonsense fragments. *Nat New Biol* **243**, 238-241.

Bull, I. D., Parekh, N. R., Hall, G. H., Ineson, P. & Evershed, R. P. (2000). Detection and classification of atmospheric methane oxidizing bacteria in soil. *Nature* **405**, 175-178.

Bulygina, E. S., Galchenko, V. F., Govorukhina, N. I., Netrusov, A. I., Nikitin, D. I., Trotsenko, Y. A. & Chumakov, K. M. (1990). Taxonomic studies on methylotrophic bacteria by 5S ribosomal RNA sequencing. *Journal of General Microbiology* **136**, 441-446.

Bulygina, E. S., Govorukhina, N. I., Netrusov, A. I., Trotsenko, Y. A. & Chumakov, K. M. (1993). Comparative studies on 5S RNA sequences and DNA-DNA hybridization of obligately and restricted facultatively methylotrophic bacteria. *Systematic and Applied Microbiology* **16**, 85-91.

Burrows, K. J., Cornish, A., Scott, D. & Higgins, I. J. (1984). Substrate specificities of the soluble and particulate methane mono-oxygenases of *Methylosinus trichosporium* OB3b. *Journal of General Microbiology* **130**, 3327-3333.

Campbell, R. K., Bergert, E. R., Wang, Y., Morris, J. C. & Moyle, W. R. (1997). Chimeric proteins can exceed the sum of their parts: implications for evolution and protein design. *Nat Biotechnol* **15**, 439-443.

Canceill, D., Dervyn, E. & Huisman, O. (1990). Proteolysis and modulation of the activity of the cell division inhibitor SulA in *Escherichia coli lon* mutants. *J Bacteriol* **172**, 7297-7300.

Cardy, D. L. N., Laidler, V., Salmond, G. P. C. & Murrell, J. C. (1991a). Molecular analysis of the methane monooxygenase (MMO) gene cluster of *Methylosinus trichosporium* OB3b. *Molecular Microbiology* **5**, 335-342.

Cardy, D. L. N., Laidler, V., Salmond, G. P. C. & Murrell, J. C. (1991b). The methane monooxygenase gene cluster of *Methylosinus trichosporium*: cloning and sequencing of the *mmoC* gene. *Archives of Microbiology* **156**, 477-483.

Chambers, S. P., Prior, S. E., Barstow, D. A. & Minton, N. P. (1988). The pMTL *nic*-cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. *Gene* **68**, 139-149.

Chan, S. I., Chen, K. H., Yu, S. S., Chen, C. L. & Kuo, S. S. (2004). Toward delineating the structure and function of the particulate methane monooxygenase from methanotrophic bacteria. *Biochemistry* **43**, 4421-4430.

Chang, A. C. Y. & Cohen, S. N. (1978). Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *Journal of Bacteriology* **134**, 1141-1156.

Chang, S. L., Wallar, B. J., Lipscomb, J. D. & Mayo, K. H. (1999). Solution structure of component B from methane monooxygenase derived through heteronuclear NMR and molecular modeling. *Biochem J* **38**, 5799-5812.

Chistoserdova, L., Vorholt, J. A., Thauer, R. K. & Lidstrom, M. E. (1998). C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic archaea. *Science* **281**, 99-102.

Chistoserdova, L., Gomelsky, L., Vorholt, J. A., Gomelsky, M., Tsygankov, Y. D. & Lidstrom, M. E. (2000). Analysis of two formaldehyde oxidation pathways in *Methylobacillus flagellatus* KT, a ribulose monophosphate cycle methylotroph. *Microbiology* **146**, 233-238.

Chistoserdova, L., Chen, S. W., Lapidus, A. & Lidstrom, M. E. (2003). Methylotrophy in *Methylobacterium extorquens* AM1 from a genomic point of view. *J Bacteriol* **185**, 2980-2987.

Choi, D. W., Kunz, R. C., Boyd, E. S. & other authors (2003). The membrane-associated methane monooxygenase (pMMO) and pMMO-NADH:quinone oxidoreductase complex from *Methylococcus capsulatus* Bath. *J Bacteriol* **185**, 5755-5764.

Cicerone, R. J. & Oremland, R. S. (1988). Biogeochemical aspects of atmospheric methane. *Global Biogeochemical Cycles* **2**, 299-327.

Colby, J., Dalton, H. & Whittenbury, R. (1975). An improved assay for bacterial methane mono-oxygenase: some properties of the enzyme from *Methylomonas methanica*. *Biochem J* **151**, 459-462.

Colby, J., Stirling, D. I. & Dalton, H. (1977). The soluble methane mono-oxygenase of *Methylococcus capsulatus* (Bath). Its ability to oxygenate n-alkanes, n-alkenes, ethers, and alicyclic, aromatic and heterocyclic compounds. *Biochem J* **165**, 395-402.

Costello, A. M. & Lidstrom, M. E. (1999). Molecular characterization of functional and phylogenetic genes from natural populations of methanotrophs in lake sediments. *Applied and Environmental Microbiology* **65**, 5066-5074.

Costello, A. M., Auman, A. J., Macalady, J. L., Scow, K. M. & Lidstrom, M. E. (2002). Estimation of methanotroph abundance in a freshwater lake sediment. *Environ Microbiol* **4**, 443-450.

Cowing, D. W., Bardwell, J. C., Craig, E. A., Woolford, C., Hendrix, R. W. & Gross, C. A. (1985). Consensus sequence for *Escherichia coli* heat shock gene promoters. *Proc Natl Acad Sci U S A* **82**, 2679-2683.

Crameri, A., Raillard, S. A., Bermudez, E. & Stemmer, W. P. (1998). DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**, 288-291.

Csáki, R., Bodrossy, L., Klem, J., Murrell, J. C. & Kovács, K. L. (2003). Genes involved in the copper-dependent regulation of soluble methane monooxygenase of *Methylococcus capsulatus* (Bath): cloning, sequencing and mutational analysis. *Microbiology* **149**, 1785-1795.

Dalton, H., Prior, S. D., Leak, D. J. & Stanley, S. H. (1984). Regulation and control of methane monooxygenase. In *Microbial growth on C1 compounds*, pp. 75-82. Edited by R. L. Crawford & R. S. Hanson. Washington D.C.: Am.Soc.Microb.

Dedysh, S. N., Panikov, N. S. & Tiedje, J. M. (1998). Acidophilic methanotrophic communities from Sphagnum peat bogs. *Applied and Environmental Microbiology* **64**, 922-929.

Dedysh, S. N., Liesack, W., Khmelenina, V. N., Suzina, N. E., Trotsenko, Y. A., Semrau, J. D., Bares, A. M., Panikov, N. S. & Tiedje, J. M. (2000). *Methylocella palustris* gen. nov., sp. nov., a new methane-oxidizing acidophilic bacterium from peat bogs, representing a novel subtype of serine-pathway methanotrophs. *International Journal of Systematic Bacteriology* **50**, 955-969.

Dedysh, S. N., Derakshani, M. & Liesack, W. (2001). Detection and enumeration of methanotrophs in acidic Sphagnum peat by 16S rRNA fluorescence *in situ* hybridization, including the use of newly developed oligonucleotide probes for *Methylocella palustris*. *Applied and Environmental Microbiology* **67**, 4850 -4857.

Dedysh, S. N., Khmelenina, V. N., Suzina, N. E., Trotsenko, Y. A., Semrau, J. D., Liesack, W. & Tiedje, J. M. (2002). *Methylocapsa acidiphila* gen. nov., sp. nov., a novel methane-oxidizing and dinitrogen-fixing acidophilic bacterium from Sphagnum bog. *International Journal of Systematic Bacteriology* **52**, 251-261.

Dedysh, S. N., Berestovskaya, Y. Y., Vasylieva, L. V., Belova, S. E., Khmelenina, V. N., Suzina, N. E., Trotsenko, Y. A., Liesack, W. & Zavarzin, G. A. (2004). *Methylocella tundrae* sp. nov., a novel methanotrophic bacterium from acidic tundra peatlands. *Int J Syst Evol Microbiol* **54**, 151-156.

deMare, F., Kurtz, D. M., Jr. & Nordlund, P. (1996). The structure of *Desulfovibrio vulgaris* rubrerythrin reveals a unique combination of rubredoxin-like FeS4 and ferritin-like diiron domains. *Nat Struct Biol* **3**, 539-546.

Dennis, J. J. & Zylstra, G. J. (1998). Plasposons: modular self-cloning minitransposon derivatives for rapid genetic analysis of gram-negative bacterial genomes. *Appl Environ Microbiol* **64**, 2710-2715.

DiSpirito, A. A., Zahn, J. A., Graham, D. W., Kim, H. J., Larive, C. K., Derrick, T. S., Cox, C. D. & Taylor, A. (1998). Copper-binding compounds from Methylosinus trichosporium OB3b. *J Bacteriol* **180**, 3606-3613.

Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K. & Mattick, J. S. (1991). 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res* **19**, 4008.

Dunfield, P. F., Yimga, M. T., Dedysh, S. N., Berger, U., Liesack, W. & Heyer, J. (2002). Isolation of a *Methylocystis* strain containing a novel *pmoA*-like gene. *FEMS Microbiology Ecology* **41**, 17-26.

Dunfield, P. F. & Dedysh, S. N. (2003). A novel, facultatively methanotrophic *Methylocella* species isolated from an acidic forest soil. In *American Society for Microbiology 103rd General Meeting.*

Dunfield, P. F., Khmelenina, V. N., Suzina, N. E., Trotsenko, Y. A. & Dedysh, S. N. (2003). *Methylocella silvestris* sp. nov., a novel methanotroph isolated from an acidic forest cambisol. *Int J Syst Evol Microbiol* **53**, 1231-1239.

Eastgate, J. A., Taylor, N., Coleman, M. J., Healy, B., Thompson, L. & Roberts, I. S. (1995). Cloning, expression, and characterization of the *lon* gene of *Erwinia amylovora*: evidence for a heat shock response. *J Bacteriol* **177**, 932-937.

Elango, N., Radhakrishnan, R., Froland, W. A., Wallar, B. J., Earhart, C. A., Lipscomb, J. D. & Ohlendorf, D. H. (1997). Crystal structure of the hydroxylase component of methane monooxygenase from *Methylosinus trichosporium* OB3b. *Protein Sci* **6**, 556-568.

Enfors, S. O. (1992). Control of *in vivo* proteolysis in the production of recombinant proteins. *Trends Biotechnol* **10**, 310-315.

Ensign, S. A., Hyman, M. R. & Arp, D. J. (1993). *In vitro* activation of ammonia monooxygenase from *Nitrosomonas europaea* by copper. *J Bacteriol* **175**, 1971-1980.

Feinberg, A. P. & Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132**, 6-13.

Feinberg, A. P. & Vogelstein, B. (1984). "A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity". Addendum. *Anal Biochem* **137**, 266-267.

Felske, A., Wolterink, A., van Lis, R., de Vos, W. M. & Akkermans, A. D. (1999). Searching for predominant soil bacteria: 16S rDNA cloning versus strain cultivation. *FEMS Microbiol Ecol* **30**, 137-145.

Ferenci, T., Strom, T. & Quayle, J. R. (1975). Oxidation of carbon monoxide and methane by *Pseudomonas methanica. J Gen Microbiol* **91**, 79-91.

Fjellbirkeland, A., Torsvik, V. & Ovreas, L. (2001). Methanotrophic diversity in an agricultural soil as evaluated by denaturing gradient gel electrophoresis profiles of *pmoA*, *mxaF* and 16S rDNA sequences. *Antonie van Leeuwenhoek* **79**, 209-217.

Fox, B. G., Froland, W. A., Dege, J. E. & Lipscomb, J. D. (1989). Methane monooxygenase from *Methylosinus trichosporium* OB3b - purification and properties of a 3 component system with high specific activity from a type II methanotroph. *J Biol Chem* **264**, 10023-10033.

Fox, B. G., Borneman, J. G., Wackett, L. P. & Lipscomb, J. D. (1990). Haloalkene oxidation by the soluble methane monooxygenase from *Methylosinus trichosporium* OB3b - mechanistic and environmental implications. *Biochem J* **29**, 6419-6427.

Fox, B. G., Liu, Y., Dege, J. E. & Lipscomb, J. D. (1991). Complex formation between the protein components of methane monooxygenase from *Methylosinus trichosporium* OB3b. *J Biol Chem* **266**, 540-550.

Froland, W. A., Andersson, K. K., Lee, S. K., Liu, Y. & Lipscomb, J. D. (1992). Methane monooxygenase component B and reductase alter the regioselectivity of the hydroxylase component-catalyzed reactions. *J Biol Chem* **267**, 17588-17597.

Fuse, H., Ohta, M., Takimura, O., Murakami, K., Inoue, H., Yamaoka, Y., Oclarit, J. M. & Omori, T. (1998). Oxidation of trichloroethylene and dimethyl sulfide by a marine *Methylomicrobium* strain containing soluble methane monooxygenase. *Bioscience Biotechnology and Biochemistry* **62**, 1925-1931.

Galibert, F., Finan, T. M., Long, S. R. & other authors (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**, 668-672.

Gallagher, S. C., Cammack, R. & Dalton, H. (1997). Alkene monooxygenase from *Nocardia corallina* B-276 is a member of the class of dinuclear iron proteins capable of stereospecific epoxygenation reactions. *Eur J Biochem* **247**, 635-641.

Gallagher, S. C., George, A. & Dalton, H. (1998). Sequence-alignment modelling and molecular docking studies of the epoxygenase component of alkene monooxygenase from *Nocardia corallina* B-276. *Eur J Biochem* **254**, 480-489.

Garcia, J. L., Patel, B. K. C. & Ollivier, B. (2000). Taxonomic, phylogenetic and ecological diversity of methanogenic Archaea. *Anaerobe* **6**, 205-226.

Ghosh, M., Anthony, C., Harlos, K., Goodwin, M. G. & Blake, C. (1995). The refined structure of the quinoprotein methanol dehydrogenase from *Methylobacterium extorquens* at 1.94 A. *Structure* **3**, 177-187.

Gilbert, B., McDonald, I. R., Finch, R., Stafford, G. P., Nielsen, A. K. & Murrell, J. C. (2000). Molecular analysis of the *pmo* (particulate methane monooxygenase) operons from two type II methanotrophs. *Applied and Environmental Microbiology* **66**, 966-975.

Gill, R. E., Karlok, M. & Benton, D. (1993). *Myxococcus xanthus* encodes an ATP-dependent protease which is required for developmental gene transcription and intercellular signaling. *J Bacteriol* **175**, 4538-4544.

Goff, S. A. & Goldberg, A. L. (1987). An increased content of protease La, the *lon* gene product, increases protein degradation and blocks growth in *Escherichia coli. J Biol Chem* **262**, 4508-4515.

Goshima, N., Kohno, K., Imamoto, F. & Kano, Y. (1990). HU-1 mutants of *Escherichia coli* deficient in DNA binding. *Gene* **96**, 141-145.

Gottesman, S. & Zipser, D. (1978). Deg phenotype of *Escherichia coli lon* mutants. *J Bacteriol* **133**, 844-851.

Gottesman, S. (1990). Minimizing proteolysis in *Escherichia coli:* genetic solutions. *Methods Enzymol* **185**, 119-129.

Gottesman, S. (1996). Proteases and their targets in *Escherichia coli. Annu Rev Genet* **30**, 465-506.

Graham, D. W., Korich, D. G., LeBlanc, R. P., Sinclair, N. A. & Arnold, R. G. (1992). Applications of a colorimetric plate assay for soluble methane monooxygenase activity. *Appl Environ Microbiol* **58**, 2231-2236.

Graham, D. W., Chaudhary, J. A., Hanson, R. S. & Arnold, R. G. (1993). Factors affecting competition between type I and type II methanotrophs in two-organism, continuous-flow reactors. *Microbial Ecology* **25**, 1-17.

Gray, N. D. & Head, I. M. (2001). Linking genetic identity and function in communities of uncultured bacteria. *Environ Microbiol* **3**, 481-492.

Green, J. & Dalton, H. (1985). Protein B of soluble methane monooxygenase from Methylococcus capsulatus (Bath) - A novel regulatory protein of enzyme activity. *Journal Of Biological Chemistry* **260**, 5795-5801.

Green, P. N. (1992). Taxonomy of methylotrophic bacteria. In *Methane and Methanol Utilizers*, pp. 23-62. Edited by J. C. Murrell & H. Dalton. New York: Plenum Press.

Grosse, S., Laramee, L., Wendlandt, K. D., McDonald, I. R., Miguez, C. B. & Kleber, H. P. (1999). Purification and characterisation of the soluble methane monooxygenase of the type II methanotrophic bacterium *Methylocystis* sp. WI14. *Applied and Environmental Microbiology* **65**, 3929-3935.

Grossman, A. D., Burgess, R. R., Walter, W. & Gross, C. A. (1983). Mutations in the *lon* gene of *E. coli* K12 phenotypically suppress a mutation in the sigma subunit of RNA polymerase. *Cell* **32**, 151-159.

Guckert, J. B., Ringelberg, D. B., White, D. C., Hanson, R. S. & Bratina, B. J. (1991). Membrane fatty acids as phenotypic markers in the polyphasic taxonomy of methylotrophs within the proteobacteria. *Journal of General Microbiology* **137**, 2631-2641.

Gulledge, J., Ahmad, A., Steudler, P. A., Pomerantz, W. J. & Cavanaugh, C. M. (2001). Genus- and family-level 16S rRNA-targeted oligonucleotide probes for ecological studies of methanotrophic bacteria. *Applied and Environmental Microbiology* **67**, 4726-4733.

Han, J. I. & Semrau, J. D. (2004). Quantification of gene expression in methanotrophs by competitive reverse transcription-polymerase chain reaction. *Environ Microbiol* **6**, 388-399.

Hanson, R. S. & Hanson, T. E. (1996). Methanotrophic bacteria. *Microbiological Reviews* **60**, 439-471.

Henckel, T., Friedrich, M. & Conrad, R. (1999). Molecular analyses of the methane-oxidizing microbial community in rice field soil by targeting the genes of the 16S rRNA, particulate methane monooxygenase, and methanol dehydrogenase. *Applied and Environmental Microbiology* **65**, 1980-1990.

Henckel, T., Jackel, U., Schnell, S. & Conrad, R. (2000a). Molecular analyses of novel methanotrophic communities in forest soil that oxidize atmospheric methane. *Applied and Environmental Microbiology* **66**, 1801-1808.

Henckel, T., Roslev, P. & Conrad, R. (2000b). Effects of O_2 and CH_4 on presence and activity of the indigenous methanotrophic community in rice field soil. *Environ Microbiol* **2**, 666-679.

Heyer, J., Galchenko, V. F. & Dunfield, P. F. (2002). Molecular phylogeny of type II methane-oxidizing bacteria isolated from various environments. *Microbiology* **148**, 2831-2846.

Hiraga, K. & Arnold, F. H. (2003). General method for sequence-independent site-directed chimeragenesis. *J Mol Biol* **330**, 287-296.

Holmes, A. J., Costello, A. M., Lidstrom, M. E. & Murrell, J. C. (1995). Evidence that particulate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. *FEMS Microbiology Letters* **132**, 203-208.

Holmes, A. J., Owens, N. J. P. & Murrell, J. C. (1996). Molecular analysis of enrichment cultures of marine methane oxidising bacteria. *J Exp Mar Biol Ecol* **203**, 27-38.

Holmes, A. J., Roslev, P., McDonald, I. R., Iversen, N., Henriksen, K. & Murrell, J. C. (1999). Characterization of methanotrophic bacterial populations in soils showing atmospheric methane uptake. *Applied and Environmental Microbiology* **65**, 3312-3318.

Hooper, A. B., Vannelli, T., Bergmann, D. J. & Arciero, D. M. (1997). Enzymology of the oxidation of ammonia to nitrite by bacteria. *Antonie van Leeuwenhoek* **71**, 59-67.

Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. & Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61-68.

Horz, H. P., Yimga, M. T. & Liesack, W. (2001). Detection of methanotroph diversity on roots of submerged rice plants by molecular retrieval of *pmoA, mmoX, mxaF,* and 16S rRNA and ribosomal DNA, including *pmoA*-based terminal restriction fragment length polymorphism profiling. *Applied and Environmental Microbiology* **67**, 4177-4185.

Hugenholtz, P., Goebel, B. M. & Pace, N. R. (1998). Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *Journal of Bacteriology* **180**, 4765-4774.

Hutchens, E., Radajewski, S., Dumont, M. G., McDonald, I. R. & Murrell, J. C. (2004). Analysis of methanotrophic bacteria in Movile Cave by stable isotope probing. *Environ Microbiol* **6**, 111-120.

Hyman, M. R. & Arp, D. J. (1992). $^{14}C_2H_2$- and $^{14}CO_2$-labeling studies of the de novo synthesis of polypeptides by *Nitrosomonas europaea* during recovery from acetylene and light inactivation of ammonia monooxygenase. *J Biol Chem* **267**, 1534-1545.

IPCC (1995). *International Panel on Climate Change; Climate Change 1994: Radiative Forcing of Climate Change.* Cambridge: Cambridge University Press.

Jahng, D. & Wood, T. K. (1994). Trichloroethylene and chloroform degradation by a recombinant pseudomonad expressing soluble methane monooxygenase from *Methylosinus trichosporium* OB3b. *Appl Environ Microbiol* **60**, 2473-2482.

Jensen, S., Ovreas, L., Daae, F. L. & Torsvik, V. (1998). Diversity in methane enrichments from agricultural soil revealed by DGGE separation of PCR amplified 16S rDNA fragments. *FEMS Microbiology Ecology* **26**, 17-26.

Jeon, C. O., Park, W., Padmanabhan, P., DeRito, C., Snape, J. R. & Madsen, E. L. (2003). Discovery of a bacterium, with distinctive dioxygenase, that is responsible for *in situ* biodegradation in contaminated sediment. *Proc Natl Acad Sci U S A* **100**, 13591-13596.

Kaluzhnaya, M., Khmelenina, V. N., Eshinimaev, B. T. & other authors (2001). Taxonomic characterization of new alkaliphilic and alkalitolerant methanotrophs from soda lakes of the Southeastern Transbaikal Region and description of *Methylomicrobium buryatense* sp. nov. *Systematic and Applied Microbiology* **24**, 166-176.

Kalyuzhnaya, M. G., Khmelenina, V. N., Starostina, N. G., Baranova, S. V., Suzina, N. E. & Trotsenko, Y. A. (1998). A new moderately halophilic methanotroph of the genus *Methylobacter*. *Mikrobiologiya* **67**, 438-444.

Kalyuzhnaya, M. G., Khmelenina, V. N., Kotelnikova, S., Holmquist, L., Pedersen, K. & Trotsenko, Y. A. (1999a). *Methylomonas scandinavica* sp. nov., a new methanotrophic psychrotrophic bacterium isolated from deep igneous rock ground water of Sweden. *Systematic and Applied Microbiology* **22**, 565-572.

Kalyuzhnaya, M. G., Khmelenina, V. N., Suzina, N. E., Lysenko, A. M. & Trotsenko, Y. A. (1999b). New methanotrophic isolates from soda lakes of the Southern Transbaikal region. *Mikrobiologiya* **68**, 677-685.

Kawarasaki, Y., Griswold, K. E., Stevenson, J. D., Selzer, T., Benkovic, S. J., Iverson, B. L. & Georgiou, G. (2003). Enhanced crossover SCRATCHY: construction and high-throughput screening of a combinatorial library containing multiple non-homologous crossovers. *Nucleic Acids Res* **31**, e126.

Kazlauskaite, J., Hill, H. A., Wilkins, P. C. & Dalton, H. (1996). Direct electrochemistry of the hydroxylase of soluble methane monooxygenase from *Methylococcus capsulatus* (Bath). *Eur J Biochem* **241**, 552-556.

Khmelenina, V. N., Kalyuzhnaya, M. G., Starostina, N. G., Suzina, N. E. & Trotsenko, Y. A. (1997). Isolation and characterisation of halotolerant alkaliphilic methanotrophic bacteria from Tuva soda lakes. *Current Microbiology* **35**, 257-261.

Konieczny, I., Doran, K. S., Helinski, D. R. & Blasina, A. (1997). Role of TrfA and DnaA proteins in origin opening during Initiation of DNA replication of the broad host range plasmid RK2. *J Biol Chem* **272**, 20173-20178.

Kopp, D. A. & Lippard, S. J. (2002). Soluble methane monooxygenase: activation of dioxygen and methane. *Curr Opin Chem Biol* **6**, 568-576.

Kovach, M. E., Phillips, R. W., Elzer, P. H., Roop, R. M., 2nd & Peterson, K. M. (1994). pBBR1MCS: a broad-host-range cloning vector. *Biotechniques* **16**, 800-802.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lane, D. J. (1991). 16S/23S rRNA sequencing. In *Nucleic acid techniques in bacterial systematics*, pp. 115-175. Edited by E. Stackebrandt & M. Goodfellow. Chichester, UK: J. Wiley and Sons Ltd.

Latterich, M. & Patel, S. (1998). The AAA team: related ATPases with diverse functions. *Trends in Cell Biology* **8**, 65-71.

Leahy, J. G., Batchelor, P. J. & Morcomb, S. M. (2003). Evolution of the soluble diiron monooxygenases. *FEMS Microbiology Reviews* **27**, 449-479.

Lee, S. Y. & Lipscomb, J. D. (1999). Oxygen activation catalyzed by methane monooxygenase hydroxylase component: Proton delivery during the O-O bond cleavage steps. *Biochem J* **38**, 4423-4432.

Lelieveld, J., Crutzen, P. J. & Bruhl, C. (1993). Climate effects of atmospheric methane. *Chemosphere* **26**, 739-768.

Lidstrom, M. E. (2001). Aerobic methylotrophic prokaryotes. In *The Prokaryotes*. Edited by E. Stackebrandt. New York: Springer-Verlag.

Lieberman, R. L., Shrestha, D. B., Doan, P. E., Hoffman, B. M., Stemmler, T. L. & Rosenzweig, A. C. (2003). Purified particulate methane monooxygenase from *Methylococcus capsulatus* (Bath) is a dimer with both mononuclear copper and a copper-containing cluster. *Proc Natl Acad Sci U S A* **100**, 3820-3825.

Liles, M. R., Manske, B. F., Bintrim, S. B., Handelsman, J. & Goodman, R. M. (2003). A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Appl Environ Microbiol* **69**, 2684-2691.

Lin, J.-L., Radajewski, S., Eshinimaev, B. T., Trotsenko, Y. A., McDonald, I. R. & Murrell, J. C. (2004). Molecular diversity of methanotrophs in Transbaikal soda lake sediments and identification of potentially active populations by stable isotope probing. *Environ Microbiol* **6**, 1049-1060.

Lipscomb, J. D. (1994). Biochemistry of the soluble methane monooxygenase. *Annual Review of Microbiology* **48**, 371-399.

Liu, K. E. & Lippard, S. J. (1991). Redox properties of the hydroxylase component of methane monooxygenase from *Methylococcus capsulatus* (Bath). Effects of protein B, reductase, and substrate. *J Biol Chem* **266**, 12836-12839.

Liu, Y., Nesheim, J. C., Lee, S. K. & Lipscomb, J. D. (1995). Gating effects of component-b on oxygen activation by the methane monooxygenase hydroxylase component. *J Biol Chem* **270**, 24662-24665.

Lloyd, J. S. (1997). Heterologous Expression and Site-Directed Mutagenesis of Soluble Methane Monooxygenase. In *PhD Thesis*. University of Warwick, UK.

Lloyd, J. S., Bhambra, A., Murrell, J. C. & Dalton, H. (1997a). Inactivation of the regulatory protein B of soluble methane monooxygenase from *Methylococcus capsulatus* (Bath) by proteolysis can be overcome by a Gly to Gln modification. *Eur J Biochem* **248**, 72-79.

Lloyd, J. S., Bhambra, A., Murrell, J. C. & Dalton, H. (1997b). Inactivation of the regulatory protein B of soluble methane monooxygenase from *Methylococcus capsulatus* (Bath) by proteolysis can be overcome by a Gly to Gln modification. *Eur J Biochem* **248**, 72-79.

Lloyd, J. S., De Marco, P., Dalton, H. & Murrell, J. C. (1999a). Heterologous expression of soluble methane monooxygenase genes in methanotrophs containing only particulate methane monooxygenase. *Archives of Microbiology* **171**, 364-370.

Lloyd, J. S., Finch, R., Dalton, H. & Murrell, J. C. (1999b). Homologous expression of soluble methane monooxygenase genes in *Methylosinus trichosporium* OB3b. *Microbiology* **145**, 461-470.

Lueders, T., Manefield, M. & Friedrich, M. W. (2004a). Enhanced sensitivity of DNA- and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. *Environ Microbiol* **6**, 73-78.

Lueders, T., Wagner, B., Claus, P. & Friedrich, M. W. (2004b). Stable isotope probing of rRNA and DNA reveals a dynamic methylotroph community and trophic interactions with fungi and protozoa in oxic rice field soil. *Environ Microbiol* **6**, 60-72.

Ma, J., Campbell, A. & Karlin, S. (2002). Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures. *J Bacteriol* **184**, 5733-5745.

MacArthur, R., Sazinsky, M. H., Kuhne, H., Whittington, D. A., Lippard, S. J. & Brudvig, G. W. (2002). Component B binding to the soluble methane monooxygenase hydroxylase by saturation-recovery EPR spectroscopy of spin-labeled MMOB. *J Am Chem Soc* **124**, 13392-13393.

Makrides, S. C. (1996). Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev* **60**, 512-538.

Malashenko, Y. R., Romanovskaya, V. A., Bogachenko, V. N. & Shved, A. D. (1975). Thermophilic and thermotolerant bacteria assimilating methane. *Mikrobiologiya* **156**, 855-862.

Manefield, M., Whiteley, A. S., Griffiths, R. I. & Bailey, M. J. (2002a). RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. *Applied and Environmental Microbiology* **68**, 5367-5373.

Manefield, M., Whiteley, A. S., Ostle, N., Ineson, P. & Bailey, M. J. (2002b). Technical considerations for RNA-based stable isotope probing: an approach to associating microbial diversity with microbial community function. *Rapid Commun Mass Spectrom* **16**, 2179-2183.

Martin, H. & Murrell, J. C. (1995). Methane monooxygenase mutants of *Methylosinus trichosporium* constructed by marker-exchange mutagenesis. *FEMS Microbiology Letters* **127**, 243-248.

Marx, C. J. & Lidstrom, M. E. (2001). Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria. *Microbiology* **147**, 2065-2075.

Maurizi, M. R. (1992). Proteases and protein degradation in *Escherichia coli*. *Experientia* **48**, 178-201.

McDonald, I. R., Kenna, E. M. & Murrell, J. C. (1995). Detection of methanotrophic bacteria in environmental samples with the PCR. *Applied and Environmental Microbiology* **61**, 116-121.

McDonald, I. R. & Murrell, J. C. (1997a). The methanol dehydrogenase structural gene *mxaF* and its use as a functional gene probe for methanotrophs and methylotrophs. *Applied and Environmental Microbiology* **63**, 3218-3224.

McDonald, I. R. & Murrell, J. C. (1997b). The particulate methane monooxygenase gene *pmoA* and its use as a functional gene probe for methanotrophs. *FEMS Microbiology Letters* **156**, 205-210.

McDonald, I. R., Uchiyama, H., Kambe, S., Yagi, O. & Murrell, J. C. (1997). The soluble methane monooxygenase gene cluster of the trichloroethylene-degrading methanotroph *Methylocystis* sp. strain M. *Applied and Environmental Microbiology* **63**, 1898-1904.

McDonald, I. R., Upton, M., Hall, G., Pickup, R. W., Edwards, C., Saunders, J. R., Ritchie, D. A. & Murrell, J. C. (1999). Molecular ecological analysis of methanogens and methanotrophs in blanket bog peat. *Microbial Ecology* **38**, 225-233.

McTavish, H., Fuchs, J. A. & Hooper, A. B. (1993). Sequence of the gene coding for ammonia monooxygenase in *Nitrosomonas europaea*. *Journal of Bacteriology* **175**, 2436-2444.

Meerman, H. J. & Georgiou, G. (1994). Construction and characterization of a set of *E. coli* strains deficient in all known loci affecting the proteolytic stability of secreted recombinant proteins. *Biotechnology (N Y)* **12**, 1107-1110.

Merkx, M., Kopp, D. A., Sazinsky, M. H., Blazyk, J. L., Muller, J. & Lippard, S. J. (2001). Dioxygen activation and methane hydroxylation by soluble methane monooxygenase: a tale of two irons and three proteins. *Angew Chem Int Ed Engl* **40**, 2782-2807.

Merkx, M. & Lippard, S. J. (2002). Why OrfY? Characterization of MMOD, a long overlooked component of the soluble methane monooxygenase from *Methylococcus capsulatus* (Bath). *J Biol Chem* **277**, 5858-5865.

Merrick, M. J. (1993). In a class of its own--the RNA polymerase sigma factor σ^{54} (σ^{N}). *Mol Microbiol* **10**, 903-909.

Miguez, C. B., Bourque, D., Sealy, J. A., Greer, C. W. & Groleau, D. (1997). Detection and isolation of methanotrophic bacteria possessing soluble methane monooxygenase (sMMO) genes using the polymerase chain reaction (PCR). *Microbial Ecology* **33**, 21-31.

Miura, A. & Dalton, H. (1995). Purification and characterization of the alkene monooxygenase from *Nocardia corallina* B-276. *Bioscience Biotechnology and Biochemistry* **59**, 853-859.

Mizusawa, S. & Gottesman, S. (1983). Protein degradation in *Escherichia coli*: the *lon* gene controls the stability of SulA protein. *Proc Natl Acad Sci U S A* **80**, 358-362.

Morett, E. & Segovia, L. (1993). The sigma 54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J Bacteriol* **175**, 6067-6074.

Mori, E., Fulchieri, M., Indorato, C., Fani, R. & Bazzicalupo, M. (1996). Cloning, nucleotide sequencing, and expression of the *Azospirillum brasilense lon* gene: involvement in iron uptake. *J Bacteriol* **178**, 3440-3446.

Morris, S. A., Radajewski, S., Willison, T. W. & Murrell, J. C. (2002). Identification of the functionally active methanotroph population in a peat soil microcosm by stable-isotope probing. *Applied and Environmental Microbiology* **68**, 1446-1453.

Murby, M., Uhlen, M. & Stahl, S. (1996). Upstream strategies to minimize proteolytic degradation upon recombinant production in *Escherichia coli*. *Protein Expr Purif* **7**, 129-136.

Murrell, J. C. (1994). Molecular genetics of methane oxidation. *Biodegradation* **5**, 145-159.

Murrell, J. C., McDonald, I. R. & Bourne, D. G. (1998). Molecular methods for the study of methanotroph ecology. *FEMS Microbiology Ecology* **27**, 103-114.

Murrell, J. C., Gilbert, B. & McDonald, I. R. (2000). Molecular biology and regulation of methane monooxygenase. *Archives of Microbiology* **173**, 325-332.

Murrell, J. C. & McDonald, I. R. (2000). Methylotrophy. In *Encyclopedia of Microbiology, Vol 3*, pp. 245-255. Edited by J. Lederberg. London: Academic Press.

Murrell, J. C. & Radajewski, S. (2000). Cultivation-independent techniques for studying methanotroph ecology. *Res Microbiol* **151**, 807-814.

Murrell, J. C., Stafford, G. P. & McDonald, I. R. (2002). The role of copper ions in regulating methane monooxygenases in methanotrophs. In *Handbook of Copper Pharmacology and Toxicology*, pp. 559-569. Edited by E. J. Massaro. New Jersey: Humana Press.

Muyzer, G., De Waal, E. C. & Uitterlinden, A. G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology* **59**, 695-700.

Muyzer, G. & Smalla, K. (1998). Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie van Leeuwenhoek* **73**, 127-141.

Nesheim, J. C. & Lipscomb, J. D. (1996). Large kinetic isotope effects in methane oxidation catalyzed by methane monooxygenase: evidence for C-H bond cleavage in a reaction cycle intermediate. *Biochemistry* **35**, 10240-10247.

Newcomb, M., Shen, R., Lu, Y., Coon, M. J., Hollenberg, P. F., Kopp, D. A. & Lippard, S. J. (2002). Evaluation of norcarane as a probe for radicals in cytochome P450- and soluble methane monooxygenase-catalyzed hydroxylation reactions. *J Am Chem Soc* **124**, 6879-6886.

Newman, L. M. & Wackett, L. P. (1995). Purification and characterization of toluene 2-monooxygenase from *Burkholderia cepacia* G4. *Biochemistry* **34**, 14066-14076.

Nguyen, H. H. T., Elliot, S. J., Yip, J. H. K. & Chan, S. I. (1998). The particulate methane monooxygenase from *Methylococcus capsulatus* (Bath) is a novel copper-containing three-subunit enzyme. *J Biol Chem* **273**, 7957-7966.

Nielsen, A. K., Gerdes, K., Degn, H. & Murrell, J. C. (1996). Regulation of bacterial methane oxidation: transcription of the soluble methane monooxygenase operon of *Methylococcus capsulatus* (Bath) is repressed by copper ions. *Microbiology* **142**, 1289-1296.

Nielsen, A. K., Gerdes, K. & Murrell, J. C. (1997). Copper-dependent reciprocal transcriptional regulation of methane monooxygenase genes in *Methylococcus capsulatus* and *Methylosinus trichosporium*. *Molecular Microbiology* **25**, 399-409.

Nishii, W., Maruyama, T., Matsuoka, R., Muramatsu, T. & Takahashi, K. (2002). The unique sites in SulA protein preferentially cleaved by ATP-dependent Lon protease from *Escherichia coli*. *Eur J Biochem* **269**, 451-457.

Nordlund, I., Powlowski, J. & Shingler, V. (1990). Complete nucleotide sequence and polypeptide analysis of multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CF600. *Journal of Bacteriology* **172**, 6826-6833.

Nordlund, P., Dalton, H. & Eklund, H. (1992). The active site structure of methane monooxygenase is closely related to the binuclear iron center of ribonucleotide reductase. *FEBS Lett* **307**, 257-262.

Oakley, C. J. & Murrell, J. C. (1988). *nifH* genes in obligate methane oxidising bacteria. *FEMS Microbiology Letters* **49**, 53-57.

Omelchenko, M. V., Vasilyeva, L. V., Zavarzin, G. A., Savel'eva, N. D., Lysenko, A. M., Mityushina, L. L., Khmelenina, V. N. & Trotsenko, Y. A. (1996). A novel psychrophilic methanotroph of the genus *Methylobacter*. *Mikrobiologiya* **65**, 339-343.

Ormö, M., Regnström, K., Wang, Z., Que, L., Jr., Sahlin, M. & Sjöberg, B.-M. (1995). Residues important for radical stability in ribonucleotide reductase from *Escherichia coli*. *J Biol Chem* **270**, 6570-6576.

Otey, C. R., Silberg, J. J., Voigt, C. A., Endelman, J. B., Bandara, G. & Arnold, F. H. (2004). Functional evolution and structural conservation in chimeric cytochromes P450: calibrating a structure-guided approach. *Chemistry & Biology* **11**, 309-318.

Ouahrani-Bettache, S., Porte, F., Teyssier, J., Liautard, J. P. & Kohler, S. (1999). pBBR1-GFP: a broad-host-range vector for prokaryotic promoter studies. *Biotechniques* **26**, 620-622.

Patt, T. E., Cole, G. C., Bland, J. & Hanson, R. S. (1974). Isolation and characterization of bacteria that grow on methane and organic compounds as sole sources of carbon and energy. *Journal of Bacteriology* **120**, 955-964.

Patt, T. E., Cole, G. C. & Hanson, R. S. (1976). *Methylobacterium*, a new genus of facultatively methylotrophic bacteria. *International Journal of Systematic Bacteriology* **26**, 226-229.

Paulsen, K. E., Liu, Y., Fox, B. G., Lipscomb, J. D., Munck, E. & Stankovich, M. T. (1994). Oxidation-reduction potentials of the methane monooxygenase hydroxylase component from *Methylosinus trichosporium* OB3b. *Biochemistry* **33**, 713-722.

Pfeil, W., Nolting, B. O. & Jung, C. (1993). Apocytochrome P450cam is a native protein with some intermediate-like properties. *Biochemistry* **32**, 8856-8862.

Pfiffner, S. M., Palumbo, A. V., Phelps, T. J. & Hazen, T. C. (1997). Effects of nutrient dosing on subsurface methanotrophic populations and trichloroethylene degradation. *J Ind Microbiol Biotechnol* **18**, 204-212.

Phillips, T. A., VanBogelen, R. A. & Neidhardt, F. C. (1984). *lon* gene product of *Escherichia coli* is a heat-shock protein. *J Bacteriol* **159**, 283-287.

Pikus, J. D., Studts, J. M., Achim, C., Kauffmann, K. E., Munck, E., Steffan, R. J., McClay, K. & Fox, B. G. (1996). Recombinant toluene-4-monooxygenase: catalytic and Mossbauer studies of the purified diiron and rieske components of a four-protein complex. *Biochemistry* **35**, 9106-9119.

Ponting, C. P. & Aravind, L. (1997). PAS: a multifunctional domain family comes to light. *Curr Biol* **7**, 674-677.

Powlowski, J., Sealy, J., Shingler, V. & Cadieux, E. (1997). On the role of DmpK, an auxiliary protein associated with multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CF600. *J Biol Chem* **272**, 945-951.

Prior, S. D. & Dalton, H. (1985a). Acetylene as a suicide substrate and active site probe for methane monooxygenase from *Methylococcus capsulatus* (Bath). *FEMS Microbiology Letters* **29**, 105-109.

Prior, S. D. & Dalton, H. (1985b). The effect of copper ions on membrane content and methane monooxygenase activity in methanol grown cells of *Methylococcus capsulatus* (Bath). *Journal of General Microbiology* **131**, 155-163.

Pulver, S. C., Froland, W. A., Lipscomb, J. D. & Solomon, E. I. (1997). Ligand field circular dichroism and magnetic circular dichroism studies of component b and substrate binding to the hydroxylase component of methane monooxygenase. *Journal Of The American Chemical Society* **119**, 387-395.

Quaiser, A., Ochsenreiter, T., Klenk, H. P., Kletzin, A., Treusch, A. H., Meurer, G., Eck, J., Sensen, C. W. & Schleper, C. (2002). First insight into the genome of an uncultivated crenarchaeote from soil. *Environ Microbiol* **4**, 603-611.

Radajewski, S., Ineson, P., Parekh, N. R. & Murrell, J. C. (2000). Stable-isotope probing as a tool in microbial ecology. *Nature* **403**, 646-649.

Radajewski, S. & Murrell, J. C. (2002). Stable isotope probing for detection of methanotrophs after enrichment with $^{13}CH_4$. *Methods Mol Biol* **179**, 149-157.

Radajewski, S., Webster, G., Reay, D. S., Morris, S. A., Ineson, P., Nedwell, D. B., Prosser, J. I. & Murrell, J. C. (2002). Identification of active methylotroph populations in an acidic forest soil by stable-isotope probing. *Microbiology* **148**, 2331-2342.

Radajewski, S., McDonald, I. R. & Murrell, J. C. (2003). Stable-isotope of nucleic acids: a window to the function of uncultured microorganisms. *Curr Opinion Biotechnol* **14**, 296-302.

Reeburgh, W. S., Whalen, S. C. & Alperin, M. J. (1993). The role of methylotrophy in the global methane budget. In *Microbial Growth on C1 Compounds*, pp. 1-14. Edited by J. C. Murrell & D. P. Kelly. Andover, UK: Intercept Ltd.

Reitzer, L. & Schneider, B. L. (2001). Metabolic context and possible physiological themes of sigma(54)-dependent genes in *Escherichia coli*. *Microbiol Mol Biol Rev* **65**, 422-444.

Richards, A. O., Stanley, S. H., Suzuki, M. & Dalton, H. (1994). The biotransformation of propylene to propylene oxide by *Methylococcus capsulatus* (Bath). 3. Reactivation of inactivated whole cells to give a high productivity system. *Biocatalysis* **8**, 253-267.

Ricke, P., Erkel, C., Kube, M., Reinhardt, R. & Liesack, W. (2004). Comparative analysis of the conventional and novel *pmo* (particulate methane monooxygenase) operons from *Methylocystis* strain SC2. *Appl Environ Microbiol* **70**, 3055-3063.

Rondon, M. R., August, P. R., Bettermann, A. D. & other authors (2000). Cloning the soil metagenom: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Applied and Environmental Microbiology* **66**, 2541-2547.

Rosen, R., Biran, D., Gur, E., Becher, D., Hecker, M. & Ron, E. Z. (2002). Protein aggregation in *Escherichia coli*: role of proteases. *FEMS Microbiol Lett* **207**, 9-12.

Rosenzweig, A. C., Frederick, C. A., Lippard, S. J. & Nordlund, P. (1993). Crystal structure of a bacterial non-haem iron hydroxylase that catalyses the biological oxidation of methane. *Nature* **366**, 537-543.

Rosenzweig, A. C., Brandstetter, H., Whittington, D. A., Nordlund, P., Lippard, S. J. & Frederick, C. A. (1997). Crystal structures of the methane monooxygenase hydroxylase from *Methylococcus capsulatus* (Bath): Implications for substrate gating and component interactions. *Proteins* **29**, 141-152.

Roslev, P., Iversen, N. & Henriksen, K. (1998). Direct fingerprinting of metabolically active bacteria in environmental samples by substrate specific radiolabelling and lipid analysis. *Journal of Microbiological Methods* **31**, 99-111.

Roslev, P. & Iversen, N. (1999). Radioactive fingerprinting of microorganisms that oxidize atmopheric methane in different soils. *Applied and Environmental Microbiology* **65**, 4064-4070.

Saeki, H. & Furuhashi, K. (1994). Cloning and characterization of a *Nocardia corallina* B-276 gene cluster encoding alkene monooxygenase. *Journal of Fermentation and Bioengineering* **78**, 399-406.

Saeki, H., Akira, M., Furuhashi, K., Averhoff, B. & Gottschalk, G. (1999). Degradation of trichloroethene by a linear-plasmid-encoded alkene monooxygenase in *Rhodococcus corallinus* (*Nocardia corallina*) B-276. *Microbiology* **145**, 1721-1730.

SaiSree, L., Reddy, M. & Gowrishankar, J. (2001). IS186 insertion at a hot spot in the *lon* promoter as a basis for lon protease deficiency of *Escherichia coli* B: identification of a consensus target sequence for IS186 transposition. *J Bacteriol* **183**, 6943-6946.

Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*, 3 edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Santiago, B., Schubel, U., Egelseer, C. & Meyer, O. (1999). Sequence analysis, characterization and CO-specific transcription of the *cox* gene cluster on the megaplasmid pHCG3 of *Oligotropha carboxidovorans*. *Gene* **236**, 115-124.

Sarbu, S. M., Kinkle, B. K., Vlasceanu, L., Kane, T. C. & Popa, R. (1994). Microbiological characterization of a sulfide-rich groundwater ecosystem. *Geomicrobiol J* **12**, 175-182.

Sarbu, S. M. & Kane, T. C. (1995). A subterranean chemoautotrophically based ecosystem. *NSS Bull* **57**, 91-98.

Sarbu, S. M., Kane, T. C. & Kinkle, B. K. (1996). A chemoautotrophically based cave ecosystem. *Science* **272**, 1953-1955.

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G. & Pühler, A. (1994). Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145**, 69-73.

Schloss, P. D. & Handelsman, J. (2003). Biotechnological prospects from metagenomics. *Curr Opin Biotechnol* **14**, 303-310.

Schmidt, R., Decatur, A. L., Rather, P. N., Moran, C. P., Jr. & Losick, R. (1994). *Bacillus subtilis* Lon protease prevents inappropriate transcription of genes under the control of the sporulation transcription factor sigma G. *J Bacteriol* **176**, 6528-6537.

Scott, J. W. & Rasche, M. E. (2002). Purification, overproduction, and partial characterization of beta-RFAP synthase, a key enzyme in the methanopterin biosynthesis pathway. *J Bacteriol* **184**, 4442-4448.

Semrau, J. D., Chistoserdov, A., Lebron, J. & other authors (1995). Particulate methane monooxygenase genes in methanotrophs. *Journal of Bacteriology* **177**, 3071-3079.

Shiemke, A. K., Cook, S. A., Miley, T. & Singleton, P. (1995). Detergent solubilization of membrane-bound methane monooxygenase requires plastoquinol analogs as electron donors. *Archives of Biochemistry and Biophysics* **321**, 421-428.

Shigematsu, T., Hanada, S., Eguchi, M., Kamagata, Y., Kanagawa, T. & Kurane, R. (1999). Soluble methane monooxygenase gene clusters from trichloroethylene-degrading *Methylomonas* sp. strains and detection of methanotrophs during in situ bioremediation. *Applied and Environmental Microbiology* **65**, 5198-5206.

Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y. & Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci U S A* **89**, 8794-8797.

Sieburth, J. M., Johnson, P. W., Eberhardt, M. A., Sieracki, M. E., Lidstrom, M. & Laux, D. (1987). The first methane-oxidizing bacterium from the upper mixed layer of the deep ocean: *Methylomonas pelagica* sp. nov. *Current Microbiology* **14**, 285-293.

Simon, R., Priefer, U. & Pühler, A. (1983). A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram-negative bacteria. *Bio/Technology* **1**, 784-791.

Sluis, M. K., Sayavedra-Soto, L. A. & Arp, D. J. (2002). Molecular analysis of the soluble butane monooxygenase from *Pseudomonas butanovora*. *Microbiology* **148**, 3617-3629.

Smith, D. D. & Dalton, H. (1989). Solubilisation of methane monooxygenase from *Methylococcus capsulatus* (Bath). *Eur J Biochem* **182**, 667-671.

Smith, M. T. (2004). Characterisation of Novel Thermophilic Methanotrophs. In *PhD Thesis*. University of Warwick, UK.

Smith, T. J., Lloyd, J. S., Gallagher, S. C., Fosdike, W. L. J., Murrell, J. C. & Dalton, H. (1999). Heterologous expression of alkene monooxygenase from *Rhodococcus rhodochrous* B-276. *Eur J Biochem* **260**, 446-452.

Smith, T. J., Slade, S. E., Burton, N. P., Murrell, J. C. & Dalton, H. (2002). Improved system for protein engineering of the hydroxylase component of soluble methane monooxygenase. *Applied and Environmental Microbiology* **68**, 5265-5273.

Smith, T. J. & Dalton, H. (2004, in press). Commercial aspects of methane monooxygenase. In *Petroleum Microbiology*.

Sorokin, D. Y., Jones, B. E. & Kuenen, J. G. (2000). An obligate methylotrophic, methane-oxidizing *Methylomicrobium* species from a highly alkaline environment. *Extremophiles* **4**, 145-155.

Stafford, G. P., Scanlan, J., McDonald, I. R. & Murrell, J. C. (2003). *rpoN*, *mmoR* and *mmoG*, genes involved in regulating the expression of soluble methane monooxygenase in *Methylosinus trichosporium* OB3b. *Microbiology* **149**, 1771-1784.

Stainthorpe, A. C., Murrell, J. C., Salmond, G. P. C., Dalton, H. & Lees, V. (1989). Molecular analysis of methane monooxygenase from *Methylococcus capsulatus* (Bath). *Archives of Microbiology* **152**, 154-159.

Stainthorpe, A. C., Lees, V., Salmond, G. P. C., Dalton, H. & Murrell, J. C. (1990). The methane monooxygenase gene cluster of *Methylococcus capsulatus* (Bath). *Gene* **91**, 27-34.

Stanley, S. H., Prior, S. D., Leak, D. J. & Dalton, H. (1983). Copper stress underlies the fundamental change in intracellular location of methane monooxygenase in methane oxidizing organisms - studies in batch and continuous cultures. *Biotechnology Letters* **5**, 487-492.

Steele, L. P., Dlugokencky, E. J., Lang, P. M., Tans, P. P., Martin, R. C. & Masarie, K. A. (1992). Slowing down of the global accumulation of atmospheric methane during the 1980s. *Nature* **358**, 313-316.

Stemmer, W. P. (1994). Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* **370**, 389-391.

Stewart, B. J., Enos-Berlage, J. L. & McCarter, L. L. (1997). The *lonS* gene regulates swarmer cell differentiation of *Vibrio parahaemolyticus*. *J Bacteriol* **179**, 107-114.

Stolyar, S., Costello, A. M., Peeples, T. L. & Lidstrom, M. E. (1999). Role of multiple gene copies in particulate methane monooxygenase activity in the methane-oxidizing bacterium *Methylococcus capsulatus* Bath. *Microbiology* **145**, 1235-1244.

Stolyar, S., Franke, M. & Lidstrom, M. E. (2001). Expression of individual copies of *Methylococcus capsulatus* Bath particulate methane monooxygenase genes. *Journal of Bacteriology* **183**, 1810-1812.

Studholme, D. J. & Buck, M. (2000). The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences. *FEMS Microbiol Lett* **186**, 1-9.

Studholme, D. J., Buck, M. & Nixon, T. (2000). Identification of potential sigma(N)-dependent promoters in bacterial genomes. *Microbiology* **146**, 3021-3023.

Studholme, D. J. & Dixon, R. (2003). Domain architectures of σ^{54}-dependent transcriptional activators. *J Bacteriol* **185**, 1757-1767.

Summers, M. L., Botero, L. M., Busse, S. C. & McDermott, T. R. (2000). The *Sinorhizobium meliloti* Lon protease is involved in regulating exopolysaccharide synthesis and is required for nodulation of alfalfa. *J Bacteriol* **182**, 2551-2558.

Suzuki, C. K., Suda, K., Wang, N. & Schatz, G. (1994). Requirement for the yeast gene LON in intramitochondrial proteolysis and maintenance of respiration. *Science* **264**, 273-276.

Takaya, A., Tomoyasu, T., Tokumitsu, A., Morioka, M. & Yamamoto, T. (2002). The ATP-dependent lon protease of *Salmonella enterica* serovar Typhimurium regulates invasion and expression of genes carried on *Salmonella* pathogenicity island 1. *J Bacteriol* **184**, 224-232.

Takaya, A., Suzuki, M., Matsui, H., Tomoyasu, T., Sashinami, H., Nakane, A. & Yamamoto, T. (2003). Lon, a stress-induced ATP-dependent protease, is critically important for systemic *Salmonella enterica* serovar typhimurium infection of mice. *Infect Immun* **71**, 690-696.

Takeguchi, M., Miyakawa, K. & Okura, I. (1998). Properties of the membranes containing the particulate methane monooxygenase from *Methylosinus trichosporium* OB3b. *Biometals* **11**, 229-234.

Tchawa Yimga, M., Dunfield, P. F., Ricke, P., Heyer, J. & Liesack, W. (2003). Wide distribution of a novel *pmoA*-like gene copy among type II methanotrophs, and its expression in *Methylocystis* strain SC2. *Appl Environ Microbiol* **69**, 5593-5602.

Tellez, C. M., Gaus, K. P., Graham, D. W., Arnold, R. G. & Guzman, R. Z. (1998). Isolation of copper biochelates from *Methylosinus trichosporium* OB3b and soluble methane monooxygenase mutants. *Applied and Environmental Microbiology* **64**, 1115-1122.

Thöny, B., Auerbach, G. & Blau, N. (2000). Tetrahydrobiopterin biosynthesis, regeneration and functions. *Biochem J* **347**, 1-16.

Tojo, N., Inouye, S. & Komano, T. (1993a). The *lonD* gene is homologous to the *lon* gene encoding an ATP-dependent protease and is essential for the development of *Myxococcus xanthus*. *J Bacteriol* **175**, 4545-4549.

Tojo, N., Inouye, S. & Komano, T. (1993b). Cloning and nucleotide sequence of the *Myxococcus xanthus lon* gene: indispensability of *lon* for vegetative growth. *J Bacteriol* **175**, 2271-2277.

Tonge, G. M., Harrison, D. E., Knowles, C. J. & Higgins, I. J. (1975). Properties and partial purification of the methane-oxidising enzyme system from *Methylosinus trichosporium*. *FEBS Lett* **58**, 293-299.

Torres-Cabassa, A. S. & Gottesman, S. (1987). Capsule synthesis in *Escherichia coli* K-12 is regulated by proteolysis. *J Bacteriol* **169**, 981-989.

Torsvik, V., Goksoyr, J. & Daae, F. L. (1990). High diversity in DNA of soil bacteria. *Applied and Environmental Microbiology* **56**, 782-787.

Torsvik, V. & Ovreas, L. (2002). Microbial diversity and function in soil: from genes to ecosystems. *Curr Opin Microbiol* **5**, 240-245.

Tsien, H. C., Brusseau, G. A., Hanson, R. S. & Wackett, L. P. (1989). Degradation of trichloroethylene by *Methylosinus trichosporium* OB3b. *Applied and Environmental Microbiology* **55**, 3155-3161.

Tsien, H. C., Bratina, B. J., Tsuji, K. & Hanson, R. S. (1990). Use of oligodeoxynucleotide signature probes for identification of physiological groups of methylotrophic bacteria. *Applied and Environmental Microbiology* **56**, 2858-2865.

Venter, J. C., Remington, K., Heidelberg, J. F. & other authors (2004). Environmental genome shotgun sequencing of the Sargasso sea. *Science* **304**, 66-74.

Vlasceanu, L., Popa, R. & Kinkle, B. K. (1997). Characterization of *Thiobacillus thioparvus* LV43 and its distribution in a chemoautotrophically based groundwater ecosystem. *Applied and Environmental Microbiology* **63**, 3123-3127.

Voigt, C. A., Martinez, C., Wang, Z. G., Mayo, S. L. & Arnold, F. H. (2002). Protein building blocks preserved by recombination. *Nat Struct Biol* **9**, 553-558.

Vorholt, J. A., Chistoserdova, L., Stolyar, S. M., Thauer, R. K. & Lidstrom, M. E. (1999). Distribution of tetrahyromethanopterin-dependent enzymes in methylotrophic bacteria and phylogeny of methenyl tetrahyromethanopterin cyclohydrolases. *Journal of Bacteriology* **181**, 5750-5757.

Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *Embo J* **1**, 945-951.

Wallar, B. J. & Lipscomb, J. D. (1996). Dioxygen activation by enzymes containing binuclear nonheme iron clusters. *Chem Rev* **96**, 2625-2657.

Wallar, B. J. & Lipscomb, J. D. (2001). Methane monooxygenase component B mutants alter the kinetics of steps throughout the catalytic cycle. *Biochemistry* **40**, 2220-2233.

Walters, K. J., Gassner, G. T., Lippard, S. J. & Wagner, G. (1999). Structure of the soluble methane monooxygenase regulatory protein B. *Proceedings of the National Academy of Science USA* **96**, 7877-7882.

Wellington, E. M. H., Berry, A. & Krsek, M. (2003). Resolving functional diversity in relation to microbial community structure in soil: exploiting genomics and stable isotope probing. *Curr Opinion Microbiol* **6**, 295-301.

West, C. A., Salmond, G. P. C., Dalton, H. & Murrell, J. C. (1992). Functional expression in *Escherichia coli* of proteins B and C from soluble methane monooxygenase of *Methylococcus capsulatus* (Bath). *Journal of General Microbiology* **138**, 1301-1307.

Whistler, C. A., Stockwell, V. O. & Loper, J. E. (2000). Lon protease influences antibiotic production and UV tolerance of *Pseudomonas fluorescens* Pf-5. *Appl Environ Microbiol* **66**, 2718-2725.

Whitby, C. B., Hall, G., Pickup, R., Saunders, J. R., Ineson, P., Parekh, N. R. & McCarthy, A. J. (2001). ^{13}C incorporation into DNA as a means of identifying the active components of ammonia-oxidizer populations. *Letters in Applied Microbiology* **32**, 398-401.

Whittenbury, R., Phillips, K. C. & Wilkinson, J. F. (1970). Enrichment, isolation and some properties of methane-utilizing bacteria. *Journal of General Microbiology* **61**, 205-218.

Whittenbury, R. & Dalton, H. (1981). The methylotrophic bacteria. In *The Prokaryotes*, pp. 894-902. Edited by M. P. Starr, H. Stolp, H. G. Truper, A. Balows & H. G. Schlegel. Berlin: Springer-Verlag.

Whittenbury, R. & Krieg, N. R. (1984). Family IV. *Methylococcaceae* fam. nov. In *Bergey's Manual of Systematic Bacteriology*, pp. 256-261. Edited by N. R. Krieg & J. G. Holt. Baltimore: The Williams and Wilkins Co.

Wickner, S. & Maurizi, M. R. (1999). Here's the hook: similar substrate binding sites in the chaperone domains of Clp and Lon. *Proc Natl Acad Sci U S A* **96**, 8318-8320.

Wise, M. G., McArthur, J. V. & Shimkets, L. J. (2001). *Methylosarcina fibrata* gen. nov., sp. nov. and *Methylosarcina quisquiliarum* sp. nov., novel type I methanotrophs. *International Journal of Systematic Bacteriology* **51**, 611-621.

Wright, R., Stephens, C., Zweiger, G., Shapiro, L. & Alley, M. R. (1996). *Caulobacter* Lon protease has a critical role in cell-cycle control of DNA methylation. *Genes Dev* **10**, 1532-1542.

Xia, Z. X., Dai, W. W., Zhang, Y. F., White, S. A., Boyd, G. D. & Mathews, F. S. (1996). Determination of the gene sequence and the three-dimensional structure at 2.4 A resolution of methanol dehydrogenase from *Methylophilus* W3A1. *J Mol Biol* **259**, 480-501.

Yang, Y. I., Hong, H. Y., Lee, I. S., Bai, D. G., Yoo, G. S. & Choi, J. K. (2000). Detection of DNA using a visible dye, Nile blue, in electrophoresed gels. *Anal Biochem* **280**, 322-324.

Yu, S. S., Chen, K. H., Tseng, M. Y., Wang, Y. S., Tseng, C. F., Chen, Y. J., Huang, D. S. & Chan, S. I. (2003). Production of high-quality particulate methane monooxygenase in high yields from *Methylococcus capsulatus* (Bath) with a hollow-fiber membrane bioreactor. *J Bacteriol* **185**, 5915-5924.

Zahn, J. A., Duncan, C. & DiSpirito, A. A. (1994). Oxidation of hydroxylamine by cytochrome P-460 of the obligate methylotroph *Methylococcus capsulatus* Bath. *J Bacteriol* **176**, 5879-5887.

Zahn, J. A. & DiSpirito, A. A. (1996). Membrane-associated methane monooxygenase from *Methylococcus capsulatus* (Bath). *J Bacteriol* **178**, 1018-1029.

Zhao, S. & Hanson, R. S. (1984). Variants of the obligate methanotroph isolate 761M capable of growth on glucose in the absence of methane. *Appl Environ Microbiol* **48**, 807-812.

Zhou, J., Bruns, M. A. & Tiedje, J. M. (1996). DNA recovery from soils of diverse composition. *Applied and Environmental Microbiology* **62**, 316-322.

Appendices

APPENDICE NOT COPIED UNDER INSTRUCTION FROM THE UNIVERSITY

Appendix 2

GSC sequence data (Chapter 5)

>GSC357
TCGAGCTCGGTCCCGGGGATCCTGCTTCGGCTCGTATTTCCAGAGGATCTTGCCGTC
GTGATTGAGATCCAGCGCATAGACGATATTGGGGAACGGCGTATGAATATACATCAC
GTCGCCGAGCACCAGCGGCGCGCCTTCGTGGCCGCGCAGCACGCCGGTCGAGAAGCT
CCATGCCGGCGCCAGCTTGGAGACGTTTTCGGTGGTGATCTGCTTTAGGGTAGAATA
GCGCGTATTGGCGTAATCGCCGGTGGGAGACACCCATTGCTTGGGGTCTTTCTGCAG
GGTCAGCAATTCGTCGGCCTTCGCCGCGCCAACCGAAAGCAGCGCGAAAACACCGAC
CGAAGTCGACAACAACAGCTTATGCATGAGCTTTTCCTCCTGACGACATCGGCGCGC
GCCTATCTTGGCCTCGTCGCCGACATCGCGCCTCGATGGCGCCTCGACACTGGGAAG
GCGTTTTCCGTTCCTTGTCTCGCGTCGATCGGCTGGACCCTACAAAATGGCTCCATC
GACCTCCGAAGAGAAGCTTCCGGTTCCCGTCCAGAAGACATCTTTTAAGTCGGAAAC
GGCGGGATAATATGGCGCCGCCTTTTAGAGCGCAACCCGGAAAAAGCGCTCTACTGA
GTTTGCGGAGCCGGGACGCCGATTGGACGTGAGCAGCGAGGTCGAACATCATGGTGC
GCCGCATTATATTGTCGCTAGTCTTGGCGTTTTATTACATAAGCCCCTCCGCCGCGA
CGGAGTTCAAGCTCGAGGAAATCGCGCCCGGCGTATTCGCGCATCCGGGCCGGACCG
CGCTTATGGACCACGAAAATGATGGAGATATCGCGAATCTTGGAGCGATCGTCGGCG
ATGACGCCGTGGCGGTGATCGACACGGGCGGCAGCCTGATCGAAGCAAAGGCTTTTA
TTGCTGCGTTGCAGCAAGCGACATCGAAGCCGATTCGCTATGTGATTAACACCCACG
CCCATCCGGACCACGTCTTCGGCAATGGCGCGTTTGTCGGCAAAGGCGTCGTTTTCG
TCGGTCACAAGAATCTCCCGCGCGCCCTGGCGGCGCGCGGCGCCTTTTATCTGGAAT
CGTTTCGCGCGGCGATGGGCGACGCTCTCGACGGCGTGGTTATCGTCGGCCCAAGCG
TCACCGTCGACAAAACGACGACGCTCGACCTTGGCGGACGCAAGATCACCCTGCAGG
CCTGGAGCGCCGCGCATAGCGATTGCGATCTGACGGCGCTCGACGAGCAAACCGGCA
CGCTGTTCGCGGGCGACCTGTTGTTCCTGCGGCACGTCCCCATCGTCGACGGCAGCC
TGCTCGGATTTCTCGACGTGGCGGAAGAATTGGCGAAAATCCCGGCCGCGCGCGTCG
TGCCCGGTCATGGGCCGCTGGTCGCGGCCTGGCCCAAGGCGCTCGACGACGAACGGG
CCTATCTGACCCGATTGACCGCCGATTTGCGCGCCGCGATCAAGAAAGGCGACACTG
TCGTGGACGCCGCGAGCGCGGCGGCGCAAAGCGAGCGCCAGAAATGGCGTCTCTTCG
ACAATTACAACGGCCGCAACGCGACCGCCGGCTTCGCCGAATTGGAGTGGGAGACGC
CTTGAGGAATTTTCGCGCGAGGATTAGCCTCGCGCCCTTGTTCTGGAGACACCCATG
AAGATCGTCGTCCCGCGCCGCTGGCGCCCCGCCGGACTCCTATGCGTCGTACTTATG
CTCGGGGCGCTTGACGGAGCGCTCGCTCAAGCGCCCGCTCCGGTCGACCCCTGGCCC
GATCTAGCCCGCGATTTGTTTCAAGGCCGCGCGATCGTGAACGACGGCTCTGTCGTT
CTCGAAGCGCCGGTCCGCGCGGAGGACGCCGCCATCGTGCCGATGACGCTGCGGCTC
AAGGCGGCCGGATTGCGTCGGGCGACGCTCGTCGTCGATCAAAACCCGATGCCGATG
GCGGCGGCTTTCACGTTCGGCGAAACGGCGGGCGTGACCTTTATCTCGACTCGCGTG
CGGGTGAACTCATATACCAACGTGCACGCGGTCGGCGAAACGGGCGACGGCGCCCTT
CACGCGCAAGTCAAATTCGTCAAGGCCGCCGGGGGTTGTTCCGCGCCCGCCGTCAAA
AACATGGATGAAGCCGCGTCTCATCTTGGACAGATGAAGTATCGCGAATTCAAGAGC
GGGGAATCTGGACATCGCGAAGCGCAAATCATGATCCGACATCCCAATAATTCGGGA
ATGCAGATGGATCAACTCACGCATCTCTATATCCCCGCCCATTACATCGACAAGGTG
GTGGTCCGGCAAGGCGACACGCTGATTTTTTCAATGGAGGGCGGGATCTCCTTGTCG
GAAGACCCCAATTTCCGGTTTAGCTATGCGCCCAACGGAGCCAAAACGATTTCGGTC
GAGGCTCATGACACCAAAGGAAACGTATTTAAACAAGAATGGCCTATCGGCGACGCT
TCTTGACGCATGTTCGCCCGTGCTCTCGATCCATTCCACCAAATCAACTTCGCGCCG
AACCAAGGAATAAGCGCCCGCACGGGCCTCCACGAAAGGGTTCGCCCTACTTGGAGG
ACCGCGAGCGCGCCCGTCGACCCGGACTTTTGGGCGCGGAGGCCTTAACCGCCCCAG
ATCATCAGGGCGATGGAAATCAGCAGCGGCGTCGACGCCAGTAAGCAGATCGCCGCG
ATACAGAAGACTTTCTGTTCCGACGACGGAGCGCCAGCCGTCAGGCTCTCGTTATGC
TTAGGGAACATCGCGAAGGACATCGTTCAACCCTCTCGGATCACATACGCCGCAACT
CTGAGAATGACGCCGCGCACTGGTCGCGACGTTAGACCCGCAACCGTAAACGCAAGG
GCGTCGGGTTAAGTTCCAATTTAGCGGTCGCGAAGTAAAACTCTTGTTGACGCGCGA
AATTTCGGCGCCTCGCAACCCGTGAAAAATGATTCGCGATTAACGCGCCGCGATTGA
TCGCGCCGCTGGCGCGTTGTTTCGGCCGGGCATCGCCGGAGCGGCGTTCTCGCTGAA
CAATTCCGCCAGTTTTTCGGTCATGGCCCCGCCCAGTTCCTCGGCGTCCACAATGGT
CACGGCCCGGCGGTAATAACGCGTCACGTCGTGGCCGATGCCTATGGCGATCAGTTC
GACCGGCGACTTCGTTTCAATATCATGGATAATCTGGCGCAAATGGCGCTCGAGATA
ATTGCCGGGATTGACCGACAGCGTGGAATCGTCGACCGGCGCCCCGTCCGAGATCAT

CATCAGGATGCGACGCTGCTCGCTTCGAGCCATCAACCGGCGATGCGCCCAGTCCAG
CGCCTCGCCATCAATGTTTTCCTTCAGGAGACCTTCGCGCATCATCAGGCCCAGATT
GCGCCGAGCCCGCCGCCAGGGCGCGTCCGCCGCCTTATAGATAATGTGGCGAACATC
GTTGAGCCGGCCAGGGCTGGGTTTCTTGCCGTCGTTTATCCAGGCCTCGCGCGACTG
GCCGCCCTTCCAGGCTCGCGTCGTAAAGCCGAGAATCTCGACCTTGACCCCGCATCG
CTCCAATGTTCGCGCGAGAATATCGGCGCAGGTGGCTGCGACCGTGATCGGGCGGCC
GCGCATGGAGCCGGAATTGTCGAGCAGCAGCGTAACGACAGTGTCGCGAAAATTCGT
GTCCTTCTCGCGCTTGAACGAAAGGGGCTGCTGGGAATCGATGATGACCCGTGGCAG
GCGGGCCGGATCGAGCATTCCCTCTTCCAGGTCAAACTCCCAGGAGCGGCTCTGTTG
CGCCATTAGCCTGCGTTGCAGCCGATTGGCCAGCCGCCCGACCACCGAGGAAAGATG
CAAAAGCTGCTTGTCGAGGTAGGAACGCAGGCGGTTCAGTTCCTCGGCGTCGCACAG
TTCCTCGGCCCGGACGATCTCGTCGAATTTGGTCGTATAAACCTTGTAATCAGTGCC
GGAGCGATCGCTCGCCGGCAGGTTGGGGCGGCGCGCGCCCGAATCGTCGTCGGCTTC
GCCGCCGTCGGAGTCTTCGCGCGTTTCGCCGTAGGGGGCGTCGGCGGATTCGCTCTC
GCCTTCCCGCGACGAATCGTCCGAGACCATCGTGGTATCCATCTCGGACGCGCTCTG
GGCTTGATCCTTCTCGTCTTGATCGGCCTCGTTCCGGTCCGGATTGGGCGGCGCGTC
GTCACTCAAATCCTCGTCGGAGGCGTCATCGCTATTTTGTTCGGGCCCGGACATATC
GAGGCTGGCGAGCAAATGGCGCACGATCCGCGCGAAGTCCTGCTGATCCTCGATCGC
GCCGCTTAGCCGATCGAGATCGGCGCTGGCGCGCTCTTCGATAAATCCTTGCCAAAG
ATCGACAATTTTTTTTGCGTGCGGCGGCGGCGGCAGGCCGGTCAGCCGTTGGCGCGC
GATCAGCGCGAGCGCGTCTTCGAGCGGTGCGTCCTCTGGCCTCTCGATCGCGTCGAA
TTTGCCCCGCTGATAGCGATCGTCGAGCATGGCGGCGATATTGGCGGCGACGCCGCT
CATCCGCCGCGATCCGATCGATTCGACGCGGGCCTGCTCGATCGCGTCGAAGGCCGC
GCGAGCCTCCATGCTGGTCGGCGAAAGGCGGCGGTGCGCGGCTTCATTGTGACAGGC
GAGGCGTAGCGCGAGCGAATCCGCGCACCCGCGCGTGATCGCCGCCTCTTTTTGCGT
GATTTTGCGCGCCGGTTCGGGCAAGCGGGCCTTCGCTCCGTCCTCGCCAACGATCAG
GGATGGACGTTCGGGCGCGAAGCTGACTTCCAGCTTGGGCGTTTTCGCCATCGCCCG
CATCGTCCCGGCGACCGCGCGTTTAAACGGCTCCTGCGGAGATTCCTTTGGGGTGGC
GGCTTTGCGATTGGTGGGCGCTGACATGAGAAAGATATAGCTCCCGCGAGCGCCGCG
GCAAAGGCCGCCATTGGGGCGGCCCATGCCGGAGGCTCGCTTGTCGTTTCATGGCGG
AGACATGGGAGCGCGCGGTCCTTGCGTTTTCCTCAAGTCATGACGACGTTGACCGCG
CTCTCCGGCAACTCCTTGCCAAAGCAGCGTTGATAAAACTCCGCGACCAGAGGGCGC
TCCAACTCGTCGCATTTATTGAGGAAAGTGAGGCGAAAGGCGAATCCGATATCGCCG
AAAATTTCAGCGTTTTCCGCCCAGGTGATGACCGTGCGGGGGCTCATCACGGTCGAC
AGATCGCCCGCCATGAAAGCGTTGCGGGTCAAATCCGCGACCCGCACCATTTTATTG
GCGGCGTCGCGCCCTTCCTTGGTCGTTTGCAAATTCTTGACCTTGGCGAGAACGATA
TTGGTCTCGGCGTCGTGCGGCAAATAATTCAACGTCGCCACGATCGACCACCGGTCC
ATCTGGCCCTGATTTATCTGCTGAGTGCCATGGTAAAGTCCGGAAGTGTCGCCAAGG
CCAACGGTATTGGCGGTCGCGAACAGGCGAAAGGCGGGATGCGGACGAATAACCCGA
TTTTGATCCAACAATGTCAGCCGGCCCGACACCTCCAACACCCGCTGGATGACGAAC
ATGACATCGGGCCGACCCGCGTCATATTCATCGAAACATAGGGCGATGTTATTTTGC
AGCGCCCAGGGGAGGATGCCGTCGCGAAACTCGGTGACCTGCTTGCCGTCCTTCAGA
ACGATCGCGTCCTTGCCGACGAGATCGATGCGCGAGATGTGGCTGTCGAGGTTCACG
CGCACGCAAGGCCAATTGAGTCGCGCCGCGACCTGCTCGATATGGGTCGATTTGCCG
GTGCCGTGATAGCCCGTGACCATAACGCGCCGATTATGCGCGAAACCGGCCAGAATG
GCCAAGGTGGTCTCGCGATCGAAGAGATAGTCGGGATCTAGCTCGGGCACGTGCTCG
TCGGCGGCCGTAAAAGCTGGAACCATAAGTTCGGTGTCGATTTTGAAGACTTCGCGC
ACGGAAATTCGTTTGTCCGGCAGTCCCGACATTTGCGCGGCCTTTTCTACTGCGACC
AACACTCTTCCTCCGTTTCGCCAGCGGTTCGGCTCGCGCCGCCCAAATTTCGCGCGA
AACCATCAAACCAGGCGGGCGGCTCTCAACGTCTTATAGGCGTGGATGATCTCCCGC
AATTTTTCTTCGCAGGATCGGTCTCCCCCATTGGCGTCGGGATGGTTGCGTTTGACC
AGATCCTTATAGCGCGATTTAATTGTTTCCGCACTGGCGGCTTCGTCCAAGCCCAAA
GCGTCGAGAGCGCGCATCGTCACCGGCGAATAACGAGGGGACGCGGGTTCCGTCTCC
CTCGCGCGGTGGGGCCGTTTTCGATTTGCCCCGAGCGGATCGGAATAATTGCCGGCG
CTTTCTCCATCGGCGCGGAATCCATTGGCGCCGCGCTTAACGCCCATCGACCAGGTG
GGGCGATGGCCAACAGTCGCGTCCTTGAGATACCGAGCCACCGCCTCATCGCTCATC
CCATTGAAATAATTGTAGGAATTGTTGTATTGGCGAACATGTTCCAAACAAAAGCAG
AAATATTGGCCCTCGCGCAGGCGCCCCATCGGCGCTCGAAATTCTCCCGCCTCTTGG

CAGCCCGGAGAGTCGCATCGCAGCCGAGTTCCGTTCGGCTCCCGCTTAGGGGGGCGA
TTGACGCGGATGCGATCGAAGAGAGGTGAGTTAAGATCCATGGCGGCTCATTATGAA
GAACTAATATGATGCCGCAAGTGACGAAGCCGCATTTATTTTGCCTGCCCGCGGCGA
AATGTGGCGCCAACCCGGAACGTCCCAGGACTTGCTTGGGATTGAATATGAGGAACG
GAAAAATGACAAATCTCGGTAAAGGGGTCGTCAAGGCCGCGATTGAAAGCAAGATTT
CCGCAGCTCTATCGCCGACGCGGTTGATCGTTATCGACGAGTCGGCCCACCACGCCG
GGCACCGCGGTCATCACTCGGATGGAGAGAGTCACTTTCGCGTCAACGTAACGAGCG
CGGCTTTTCGAGGTAAAAGCCGAATTGAGCGGCATCGAATTATTCATGAGATTTTGG
CCGAGGAATTGGCGGGGCGCGTCCATGCTCTCGCGCTGGTGGCCTCGGCGCCGGAAG
AAGCGGAGTGAATGGCTCATCGCGGGCTGGCGGATGTGGCTTCGCGTGAAAGCGTGG
TTCTGCTTGGCTGGGCGTCGCGAAAATATCGGCCTATCGCCTTGTGTGTTTTCGGCA
AAAGTGGACGGCGGTTTTGCCCGGAGAGCACGACAAAATAAAGGAATCAGAGTCTTT
CCGGTTCAATCTAAACCGGAGCAACTCTAGCGTTGCTCGCTGACAAAAAGAAGCCGG
CCGGACTTTTGGTCCGGCCGGCTGTGTTTTAGGCTTTGATCAGAAGTCGATCAGGGC
ATGTCGCCCGCGACGAACTTCGGAATGACCGGTCCGCCGATTTCAGCCGCGAAGCGA
CGGCCGCTCGGCGTGAAGAAGAAGAACAGGCCGCCGACTTGGCTGTCGGTGTCGTAG
GCCAGGTCGGACAGACGCTCGATATCCCAGCGCGCGTCCTGGATCTTGACCTGGATT
TCCTTCGTCTCGCCCGGCGCGATCGGAGTCGCGTCGGTCGAAAGGCCGCGATCCGCG
AGCAGATAGTCGGGGAACTCGGGCTTCGTGGTGAACACGTCGGGGTTCAGGAAGCGC
AGACCCGCCGCGGTGTATTCGCCAAGACGCAAGGGCTCCGAGGTGTTGTTCTTCACC
TTGACGCTCAAGCTCATCTCGCGGCCCGGAACCTTGTAAACGCCACCCTGGAGGTCG
GCGGTCACCTGCTCCTTGCCGACGCCGACCGTGCCTTCGGTCGTGATCGGAGTCAGC
GGCTTCTGCAAACCAGCCTGAAGCGGGATGGTGCGCGGGAAGGTGCTGTTGGTCACG
GCGTAACCGACGATCGTCGCGAGGATCGTCAGCGCGAGAACGATCGCGCCAATGCGG
CGGTCCTCGTCCGTCACGGCCTCATCGGCGCGGCCATCGGCCACCGTCAGATAGGAC
GAGATGATGCCCTTGCGCACGAACCAGAAGAAGATCCACGCGGCGCCAACCGCCAGC
CAAGGCAAATGCCACGCATAAATGCGGTCAATGCCATAATGCTCGAGGTCGACGGTC
GTGCCGTCGAGGAGCGTCACCGGATCGGTGAAGTCCTTCATGTCGCCCTTGATCTCG
ATCCACTGGCCGGGTCCGATGATCGGGCCGCCGCCTTCGACGTTGACCTGAGCATGG
ACGTGCCAGCGGCCGGCGCGACGACCGCGCAGGTTGATCTGGAACGCATAGTCGTTG
CCCGGAACCAGCGACACCGAGCGAGGAGCGAAGTTCTCGCCGATGAACTGAGCCGTG
CGCACCAGAACCGGACCCGGCTCGCCGGCGTTCAGGAACGACGACTTCGGATTCGCA
ACAGCCTGCGGCCAGGCCGAGAACACATGGAACTTGCCCGAAAGCACCATGTCTTCG
TTGACGTTGACCGTCGTCTTGGACCAAGCCACATCATACCAGTTCAGCGTGCGCATC
CGCAAAAACGCCTGCTGCGACTTCTCGCCGTGCGCGGAGGCCGGAGCAATGGCTCCA
ACCGTCGCGATCACAGCCGCCGCGCCCATGGCGGCCAGCTTGACTAGGTTTCTCATG
AACAAACCTCCCAGTTTCTTCCTCTTCGCGGCTAACGCGACACATGCGCTTTCAACA
CAGGCCTTCGACGTCGCCGCAGCTTGTCTTCCCAAGCCCTTGTTTCTTCCCAAGAGC
TCTCGTTCTTTGTGAGTTCCCGATCTTCCGAGTTCTCGATTTTCTCCGGGGCGGCTT
CCCGTTCTCCTCGCGGAAAACCGCGCCGCCCCGTCCGGTCCTGCTCGATTTCGATCA
GATCTTGTCGATGACCTTGGTCGTCGAATACCAACGACCCATGAACCACCACAGGAA
ATACACCATCATCGAGACGAAGCCCGAGAAGAACGCGGCCACCGGCACGACGTCCTT
ACCGAAGGTGCGCAGCGTGCCGCGCTCGACCATGCGGATGTATTCCGGCATCGACGT
GCGGACGAAGTGCAGGCCGATCAGATCGGCGAGGGACATAAGCTGCCCATGCTGCTC
GGTCGCCTGATGGAACGCCGCGATCGCCGGCCAGTTGTTCGGATAGAACAGCAGACC
CCAGCCCAGCGAACCGACAACCGCCGTGATCACATAGCTGCCCGAAAGAAGCAGGAT
CACGTCGAGCCAAATCGCCGGAACGATCAAAGCGGACGGGAACACCAGCGAGATCGG
GAAATAGGTCCAACCCCAGAAGTTCACGTAGCGGTTGATCCACTCGCCGATCAGCAG
GCCGAGAGCCGCGAACACCGCGCCAAACGGAAGACGGAAGTTGACCCAGAAAAACGC
CTGCGCCGCGGCGCAGAACGTCACGCCCAGGATCGGAATAACCGTCGGCCACATACG
ACGATCCTTCCAGTCGACCCAGAAGTCCCAGTCGCCCGCCGTCAGCATGAAGTGGAC
GTGATAGCCGCCGAGAACGGCGAAAAACAACAGTGTCAGAAGCAACCAGTCGACGGT
CTGAACGCATCCCGCCGCCTCGGCGACGGAATGGAACGGACCGACCGCTCCCCCGCT
CTTCGATGATGAAGACATCTTTCTTCTCCTTGGTGTCCTTGACGATAGTCCCCGGCT
TTTCACTCGCCGGATTTTTTAGGCAAATGCGCGCGACCCATCCAAACCGCCACCTTC
AAGCAAGCCTCGCGAGAACATCCCTCCTCGCGGGAGCCGTCTTGTTTCGAGCATGGA
ACCGGCGCCTCGCGAAACGATCGCGCCTGATCCTAACTGTAAAGGGCTTGGTCTAAA
GCGGTCGGCGTCGTCAAAGCCGCGCGGACAATTCCGCCGCGCGGCTCGTTACGGAAT

ATTACGTCAGCAGAGCGACGCCTTCCTTGCCGACCAGAGCGTGAATGCGGGCGAGGA
TCTGGAGAACCACGCCGAACACGCCAAGGGCCATCCAGCCGAAGAACACGAAGCCCC
AATGCAGCGGAGCGACGAACAGTTCTTCCATGAACCAGAAGGTGTGGCCCCATTCAT
TCAGGCCGACGTTCGGGATGATCATGAACGGGCCAATCGCCACGATCAGGAACGCCA
GCGAGTAGCCATGAGCAAAAAACGGAATGCGGGTCTTCGCATAGAAGAACGCGCCAA
CCGCGATAACCGAATAGATCGGGTAGCTCATGTAGAACTCGATGATGTGCGACGGCG
TGAAGTCCGTGTCGCGAATCACCGTCATGTGCCAGGTGCCGTCCTGCTCCGTGAAGA
ACGAAGCGCCCCAGTAAATGGCGATGCCGTAAACCGTCAGCCATTGAACCAGGTTCA
CGAGGCGGCGCATCTCCTCGCGCGGGGCGACCGCGTCGACGTTGCGGTCGCGGGTCT
TCCACAGATAGCCGGCCAGGCCGATGCCCGAAACCAGCTCAAGCGGAATCTCCGTCC
ACAGAATGCTCAGCCAGTAGGTCTGGAACTCCGGAGCGAACGAGTCGAGACCCGCGC
GCCAGCCGAAAACCTGCTCGTAAATACGCACGATCAGATAGAAAATGTCCAGAACCA
CAAGGCCAATCCACATGCCTCGCAGATCGACAACCGTATCTGTGCCAGCAGCCGCGC
CCGCGGCTGTCTCGGTGGTGGAACTCATGCTTAGTCCTCCTATATTGCTCCCAGGGA
TTTGCGCGCGATCGGTGACGCCCCGCCTCCCAGCACGGAACTTGTTCGTCCTTTTTC
TTAACGTAACGTTCGCGGCGCTCGAAAGCGCCCAACCGCCCTGTCTCGTCGCGACTT
AACGCCCCCAAGGAGTCACGGTTCAACTGAATCAAGGATGGCTCAATCGAACCGGGG
CTCCGACGCCGCCGCCGTTGTTGCTTTTTGGACGAACACCGATGTCGCCGCCTAAAC
AGGCCCCGCCCCGTCCGTTGAAAGACCACGGCGGACCCTGTTTAGCGTGGCCTAACA
ATTCAGCTTCGCGCCCCTAAAAACAATACTCGATAGGATAGCAGATGACAATCTATT
GAAGGGGTCTATTAGCCCATATGGGTGGTAATGCCCCCACCCAAATTTCGAGACAAA
ATCCAGCCAAATCCACTCGCTTCGCTCTCGCAATAAAACGCAAATTGGCCAATTTCA
TAAGCAAACAGCTATCCCAACCAAAATCCAACACCGACCAGAAATCTCAACAAATAC
CCCCAACCCGAATGGCTGCGACATTTTGTGCTGCAGCGCGGTAAATGGGGTGGGGGC
GATTGGCGAAATAGGACGCGACAAACGACGAGCGCCATAATTCGAAACGCGGATTTT
GAAGTCTTATTGTTCCCAAACGGAAATAGCGGAGAGATCGAAACCCGCCGCCGCCTC
GGCGACGGACACCCCGTCGAGAACCAGATTGGGCGCGATTTCCGTCAACGGAGCCAG
AACAAATGCCCGTTGAAACAATTCCTTATGAGGAATCGTGAGTTCCGGATCGTCAAG
GCTGCGATCGCCCAGGAACAAAATATCGACGTCGATCAGCCTCGGACCCCAGCGCTT
GGTCGGCTCCCGCCCCATGTCCGCCTCGATTTTTTTCACCGCGGCGAGAAGTTCGTA
AGGAGCAAGGCTGGTCACGCCCACGGCGCAGGCGTTGGCGAAATCTCCTTGGTCGAG
AACGCCCCAGGGCGGCGTCCGGTATATTCGCGACACCAATTCCAGGCGGGCGACGCC
GCGCTCCTCGATAAGGCGCAACGCCTTGCGAATATTCCCCGGCTTGTCCCCAATGTT
GGAGCCCAGTCCAAAACCGACCTGCGCTATCGTCATTGCGTCACATCCGTTCAAAAT
CAACGCGGCTTCATCGCCGCCATGACAGCGAAGGCGGTCCTGTGCTCGAGCGCGTCA
TGCACTCTGAACACTCGCGCGCCGAGGCTCGCGGCGAAGACATTCGCCGCCAGAGTG
CCGACCAGCCGTCCCTCGATCAATCCGTCGGGCAGATCGCGAAAGATGGTTTTCCGC
GAAACGCCGATCAATAAAGGACAACCAAACAGCAGCAATTGCGGAATGGCGCGCAGC
GCCTGATAATTTTGCTCGCGGCTTTTACCAAAGCCGACGCCGGGGTCGAGCAGGATA
TGCCGGGCAGGAATGCCGGCGCGCCGAGCCAAATCCAGTGAGCGCTCAAAAAATCCC
AGCATGTCGGAGACGATGTCGACGCCAGCGTCCACCGTGTCGCGATTGTGCATGATC
ACAACGGCCCCGCCCGCTTCCGCGACCACGGACGCCATCGCGGGGTCTCGATGCAAA
CCCCAAACGTCGTTCACGACCGCCACGCCGGTTCGCAGCGCCCGACGCGCCGTCTCC
GCCTTGTAGGTGTCGATCGAGACCGGAGCCCCCGCCTGCGCGACCAGCGTCGCCAGC
AATGGGCGCAAGCGCTCCCATTCTTCCTCGGCGGAAATGGGCGTATGGCCCGGCCTC
GTCGACTCCGCGCCGACATCCACGATGCAGGCCCCGTCGGCCACTAGTTTTCGCGCC
TGGGCCAGCGCCGCGCGAGGCGAGGTGAAAAGACCGCCGTCCGAAAAAGAATCCGGG
GTCACGTTGACGATGCCCATGATGGCGGGTCCCGCCTCGACACTAGACAGAAACGCC
TCCCTCGCCGCGCTAATATCTTCAACTGGTACAGGCAAACCCTTTCCCCCCGTTCCA
CGTCGCTCTTGTCCAAGGGGGTTCTAATCGTCGTATTTCAGATAGGCGAATCGCAGG
TCGTTTTCAGGCGTCCCTTCGAGCGCGCCGCCCTGCTTGCCAGGTCTCCAGCCCACG
ACTTCCTCGATCTTGCGGGCGAGTTCGAAATCCTTGTCGGTGATCCCCTTGGCGGTA
TGCGTCATCAATTTGACCTCGACCCAGGCATAGGAGGCGGCGATATCGGGGTGGTGC
CAGGCAGCCTCGGCCAAATGGCCGACCGCATTTATGACCATGAGCGTTCCCTTCCAC
GAATGGGTTTTGAATTTGCGGCGAATCCAGCCGTTTTCATAACGCCAATGCGGCAAG
TCCCTGGCCAGCCGCTCGGCGATCTCGTTCTCGGAATAGGCGTGTTCCTTGGTGGTC
ATGACGCTCTCCTTGGGCTGGGGTTCTAGAAAGCCATTTGAACTGACGAGATAAGGC
GCGATGGCGCTAGAATAGCTAAATCCTCTTGCCATTATTGGTCAGGCGCGTTACTCG

CGCAACGGCGTCGCGACGAGCGGCGCCATCGCGTGATGTAGGGCGAGGCGGCGCCAA
GTCGCGCCGCGAACTCGCGCTTTGAAGTCTTGCTCGATCGAAACGGGGCGTCATGTC
GGCCCAATCTAGCGTCAACGAAGTTCGCGTCACGGCCGCGGCGCGCCTGCACCTGGG
CTTTCTCGACATGAACGGCGGCCTCGGCCGCAATTTTGGCGGACTCGGTCTGTCGGT
CGGCGGCCCAAGGACAAGGCTCACGCTGAAGCGGGCCGGCGAGACTCTCGTGGAAGG
CGCCGAATCCGAACGCGCCCGACGCCTTCTGGGAAAAGCCCAAGCGGCTTTGGCGCC
GCGTTCGGCCCATCATCTCACCATTCACCAAGCGATCCCCGCGCATTCCGGACTTGG
TTCTGGCACTCAATTGGCGCTGGCCATCGCGGCGGCGCTACGGCGGCTGGAAGATTT
GCCGCTCGACCCGAACGCCGACGCCGCCCTGTTGGAGCGCGGCACCCGCTCGGGACT
GGGCGCGGGCTTGTTTCAGGACGGCGGATTCGTGGTGGACGGCGGACGCGGGGCGAG
CGGCCGCACCCCGCCGGTCATCGCCCGCCTGCCGTTCCCGCCCGACTGGCGCATCCT
GCTTGTCATGGACAACAACGCCCAGGGCCTGAACGGCGAGCGCGAGCGCGAGGCTTT
CGCCGCGTTGCCGCCCTTCTCCGAGGCACGGGCGGGAGACATTTGCCGGCGCGTGCT
CATGCAAGCCTTGCCGGCCTTGGCCGAACGCGACATCGCCGCCTTTGGCGAGGCCAT
TACCCATATCCAGAACATCGTCGGCGATTATTTCGCGCCCGCGCAAAACGGGCGGCG
CTTCACCAGCGCCCGTGTCGAAGCCGTCGTCGCGCGGCTGTTGCGCGAAGGCGCCAC
GGGCGGCGGTCAGACTTCGTGGGGCCCCACGGGCTTCGCATTCGTCGCTGGCGAAGC
GGAAGCGCGGCGCCTGGCCGAGCGCGCGCGCTCCGACGCCGCCGATGGGCTGACCAT
CGAGATTGTCGAAGGGCTTGCGCATGGCGCGCGCATTGACGCGATTTATGCGCGGAT
AGCCTGAACGGCGCCTTGCCGCCGCGCTCGCCCGCGCCTCACCCGGGCCCGTCTTCC
TCTCCCTTGATGTCAGGGTTCGTCCGCGTGGCGCCTGCATCGTTGCGGGCCGAACCG
AGGCGGCGGCGCGCCATGATTCGTCCGCGCGTCGCGGCGCGCGCCACGTGACGCTGG
CGCGCGGCTTCCGATTTGGGTTTCAGCAATTCCCGATATTCGTCGCGTTTTTCGTGG
ATCGAGGCGATAACGAAGCCCATCGGCACGCCGATGTCGACGAGCACGGTTTCGGCC
AGTTGCAGGCTCGCCTCGATTGTTTCGGGGATCGCGTCGCTGGCGCCCGCTTCATAA
AGGCGCGTCGCATGACCCGCGTCGCGCGCTCTGGCGACGATGGTCATGTCGGCGCGC
GCCTCATGCGCGAGACGCACAATCTCCTCCGCCGCGCCCGGGTTTGGCACGGTGACG
ACCAATGCTCGCGCCTGCGCGACGCCGCAGGCCAGCAGAAGTTCGCGCCGGGTCGCG
CTGCCCCAATAGATGTCGACGCCTTCGTCGCGCCGCGACGCGACGAGTTCTACGGCG
TCGTCTATGGCGACGAAGGGGATGTCGTGACGCTTGAGCATCTCGCCGACCAGCGCG
CCGACGCGGCCAAAGCCGATGATGACGACACGGCCGGAGGCCACGTCGCCTTCGGGC
GCGAGATGCGCGAACTCGGCCTGGAGTTCGGCCGGCGTCTTCGCCCGCGCGAAGCTC
GCGCCGAGCCGCCCCAGCGCGGGAATCGCGAACATGGACAGGGTCACCACGACCATG
GCGTCGGCGCCTTGGCCGCCTGGCAACACGCCCGCGCCAATCGCCGAGGTGAGCAAC
ACAAAGGCGAATTCGCCGCCTGGCGCCAGCAGCAGGGCGACCTCGCCGCCGACCGCG
CCCGTCAGACGAAACGCTCGCGCCAGCGCGAAGATCACCAGCGCCTTGAGCGCCACG
AGCCCGCAGGCGTTCAGCGTGATCGGGATCGGAT

>GSC357 PmoC
MSSTTETAAGAAAGTDTVVDLRGMWIGLVVLDIFYLIVRIYEQVFGWRAGLDSFAPE
FQTYWLSILWTEIPLELVSGIGLAGYLWKTRDRNVDAVAPREEMRRLVNLVQWLTVY
GIAIYWGASFFTEQDGTWHMTVIRDTDFTPSHIIEFYMSYPIYSVIAVGAFFYAKTR
IPFFAHGYSLAFLIVAIGPFMIIPNVGLNEWGHTFWFMEELFVAPLHWGFVFFGWMA
LGVFGVVLQILARIHALVGKEGVALLT

>GSC357 PmoA
MSSSSKSGGAVGPFHSVAEAAGCVQTVDWLLLTLLFFAVLGGYHVHFMLTAGDWDFW
VDWKDRRMWPTVIPILGVTFCAAAQAFFWVNFRLPFGAVFAALGLLIGEWINRYVNF
WGWTYFPISLVFPSALIVPAIWLDVILLLSGSYVITAVVGSLGWGLLFYPNNWPAIA
AFHQATEQHGQLMSLADLIGLHFVRTSMPEYIRMVERGTLRTFGKDVVPVAAFFSGF
VSMMVYFLWWFMGRWYSTTKVIDKI

>GSC357 PmoB
MRNLVKLAAMGAAAVIATVGAIAPASAHGEKSQQAFLRMRTLNWYDVAWSKTTVNVN
EDMVLSGKFHVFSAWPQAVANPKSSFLNAGEPGPVLVRTAQFIGENFAPRSVSLVPG
NDYAFQINLRGRRAGRWHVHAQVNVEGGGPIIGPGQWIEIKGDMKDFTDPVTLLDGT
TVDLEHYGIDRIYAWHLPWLAVGAAWIFFWFVRKGIISSYLTVADGRADEAVTDEDR
RIGAIVLALTILATIVGYAVTNSTFPRTIPLQAGLQKPLTPITTEGTVGVGKEQVTA
DLQGGVYKVPGREMSLSVKVKNNTSEPLRLGEYTAAGLRFLNPDVFTTKPEFPDYLL
ADRGLSTDATPIAPGETKEIQVKIQDARWDIERLSDLAYDTDSQVGGLFFFFTPSGR
RFAAEIGGPVIPKFVAGDMP

>GSC357 MptG
MSAQSSVNEVRVTAAARLHLGFLDMNGGLGRNFGGLGLSVGGPRTRLTLKRAGETLV
EGAESERARRLLGKAQAALAPRSAHHLTIHQAIPAHSGLGSGTQLALAIAAALRRLE
DLPLDPNADAALLERGTRSGLGAGLFQDGGFVVDGGRGASGRTPPVIARLPFPPDWR
ILLVMDNNAQGLNGEREREAFAALPPFSEARAGDICRRVLMQALPALAERDIAAFGE
AITHIQNIVGDYFAPAQNGRRFTSARVEAVVARLLREGATGGGQTSWGPTGFAFVAG
EAEARRLAERARSDAADGLTIEIVEGLAHGARIDAIYARIA

>GSC357 FolC
MGIVNVTPDSFSDGGLFTSPRAALAQARKLVADGACIVDVGAESTRPGHTPISAEEE
WERLRPLLATLVAQAGAPVSIDTYKAETARRALRTGVAVVNDVWGLHRDPAMASVVA
EAGGAVVIMHNRDTVDAGVDIVSDMLGFFERSLDLARRAGIPARHILLDPGVGFGKS
REQNYQALRAIPQLLLFGCPLLIGVSRKTIFRDLPDGLIEGRLVGTLAANVFAASLG
ARVFRVHDALEHRTAFAVMAAMKPR

>GSC357 FolK
MTIAQVGFGLGSNIGDKPGNIRKALRLIEERGVARLELVSRIYRTPPWGVLDQGDFA
NACAVGVTSLAPYELLAAVKKIEADMGREPTKRWGPRLIDVDILFLGDRSLDDPELT
IPHKELFQRAFVLAPLTEIAPNLVLDGVSVAEAAAGFDLSAISVWEQ

>GSC357 MoxF (partial sequence)
MHKLLLSTSVGVFALLSVGAAKADELLTLQKDPKQWVSPTGDYANTRYSTLKQITTE
NVSKLAPAWSFSTGVLRGHEGAPLVLGDVMYIHTPFPNIVYALDLNHDGKILWKYEP
KQDPRDRAR

Appendix 3

Ms. trichosporium OB3b *lon* sequence data (Chapter 6)

>Ms. trichosporium OB3b lon gene; includes flanking sequence

GATCTCGCGCAAGTCCGACAATCCCTCGATCACCCGCGACGTGTCGGGCGAGGGCGT
GCAGCAGGCGCTGCTGAAGATCATGGAAGGCACCATCGCCTCCGTGCCGCCGCAGGG
CGGGCGCAAGCATCCGCAGCAGGAGTTTCTGCAGGTCGACACGACCAACATCCTGTT
CATCTGCGGCGGCGCCTTCGCCGGACTCGAGAAGATCATCTCCTCGCGCGGCCGCGG
CACCTCGATCGGCTTCGGCGCCACCGTGCAGGCGCCGGACGAGCGCCGCACCGGCGA
CATCTTCCGCCATGTGCAGCCGGAGGATCTGCTGAAGTTCGGCCTGATCCCGGAATT
CGTCGGCCGTCTGCCGGTGATCGCGACGCTCGAGGATCTGGACGAGGAGGCGTTGAA
GCGCATCCTCACCGAGCCGAAGAATGCGCTGGTGAAGCAATATCAGCGACTGTTCGA
GATGGAGAGCACCGAGCTGACCTTCCAGGACGAGGCGCTCTCCTCGGTCGCGCGCAA
GGCGATCGAGCGCCACACCGGCGCGCGCGGCCTGCGCTCGATCATGGAAGGCATTCT
GCTCGACACGATGTTCGAGCTGCCGGGCCTCGAAGGCGTCGAGCAGGTCGTCATCGG
TCCGGAAGTGGTCGAGGGCAAGGCGCGGCCGCTCTACATCTATGCCGAGCGCAACGA
GAAGAGCGGCGCCAGCGCGTGACGCGAATTTCGGCAGGAGCGGCCGCATTCTTCGCC
GCTCCCGCCCGCGACCAAGTCGCAGTCGGTTGTTTTTCCAGCGTTCTTGAAGTGCGG
CGCCGCCAGGCCCATTTGTTGCGTAGAGCCTTCAAAGCGCGAGCGGCGAAGGCGGCT
GTTTCAGTCGGCGACTAGCCCGGAACTTGCTTGCCTTCCATGCGTCATAGCCCCGCC
ACGACCCGCCCGTGCGGAGTGCGAAACGGAACAGGATACACGAAATGACGAACGAAA
AGCGGAATTCCATCGCGCCCGGCGCGGTCGAGAGTTACCCGGTTCTGCCGCTGCGCG
ACATCGTGGTCTTTCCTCATATGATCGTTCCGCTGTTCGTCGCGCGTGAAAAATCGA
TTCACGCGCTCGAGGAGGTGACGAAATCCGATCGGCTGATTCTGCTCGCGACGCAGA
AGAACGCCGGCGACGACGATCCGGCCGCCGATTCCATCTATCAGATCGGCACGCTCG
CCTCGGTGCTGCAGCTGCTGAAGCTGCCCGACGGCACTGTGAAGGTGCTGGTGGAAG
GCGTCGCACGCGCCAAGGTGCGCACCTACACGCGCACCGATGAATATTACGAGGCGG
ACGCCGAGACGCTGGGCGACGACACGGAAGCGCCCGTCGAGATCGAGGCGCTCGGCC
GCTCGGTCATCGCCGAGTTCGATTCCTATGTGAAGCTCAACAAGAAGGTCTCGCCCG
AGATCGCCAGCGCGGTCACGCAGATCGAGGATTTCTCCAAGCTCGCCGACACGGTCG
CCTCGCATCTCTCGGTGAAGATCGCCGAGAAGCAGGATGTGCTGGAGACCATCTCCG
TCGCCAAGCGTCTGGAGAAATGCCTCTCGCTGATGGAGAGCGAAATCTCCGTGCTGC
AGGTGGAGAAGCGCATCCGCACGCGCGTCAAGCGCCAGATGGAGAAGACGCAGCGCG
AGTATTATCTCAACGAGCAGATGAAGGCGATCCAGAAGGAGCTCGGCGACGAGGACG
GCAAGGACGATCTCGCCGAGCTCGAGGAGCGGATCAAGAACACCAAGCTCTCGAAAG
AGGCGCGCGACAAGGCGGTCGCCGAGTTCAAGAAGCTGCGTCAGATGTCGCCCATGT
CGGCGGAAGCGACCGTCGTGCGCAATTATCTCGACTGGCTGCTGGCGCTGCCCTGGG
GCAAGAAATCGAAGATCAAGCGCGATCTCGAGGCGGCGCAGGACGTGCTCGACACCG
ATCACTTCGGCCTCGACAAGGTCAAGGAGCGCATCCTCGAATATCTCGCCGTGCAGA
GCCGTGCCAATAAGCTGACCGGTCCGATCCTGTGCCTCGTCGGCCCGCCCGGCGTCG
GCAAGACATCGCTCGGCAAGTCGATCGCCAAGGCCACGGGCCGCGAGTTCGTGCGCA
TGTCGCTCGGCGGCGTGCGCGACGAGGCCGAGATCCGCGGCCACCGCCGCACCTATA
TCGGCTCCATGCCCGGCAAGATCATCCAGTCGATGCGCAAGGCCAAGACGTCCAACC
CGCTCTTCCTGCTCGACGAGATCGACAAGATGGGCATGGACTTCCGCGGCGATCCGT
CCTCCGCTCTGCTCGAGGTGCTCGATCCCGAGCAGAACGCCACTTTCGCCGATCACT
ATCTCGAGGTCGACTACGATCTGTCGAATGTCATGTTCGTGACGACGTCGAACACGT
TGAACATCCCTGCGCCCTTGATGGACCGCATGGAGATCATCCGCATCGCCGGCTACA
CCGAGGACGAGAAGGCCGAGATCGCGCGCAAGCATCTCATTCCCAATGCGGTCAAGA
AGCACGGCCTCTCCGCCGACGAATGGGCGATCGACGACGAAGCGCTGATGACGCTGA
TCCGCCGCTACACACGTGAAGCGGGCGTGCGCAGCCTCGAGCGCGAGATCTCCAATC
TCGCCCGCAAGGCGGTGAAGGAGATTCTCCTCTCCAAGGAGAAGGGCAAGAAGGTTC
GCGTCACCAACGACAACATCACCGATTACCTCGGCGTGCATAAGTTCCGCTATGGCG
AGGCGGAGCTCGAGGATCAGGTGGGCGTCGTCACCGGCCTCGCGGTCACCGGCGTCG
GCGGCGAGCTGCTGACGATCGAAGGCGTGATGATGCCGGGCAAGGGCAAGATGACCG
TCACCGGCAATCTGCAGGATGTGATGAAGGAGTCGATCTCCGCGGCCGCCTCTTATG
TGCGCTCCCGCGCCGTCGATTTCGGCATCGAGCCGCCGCTGTTCGACCGCCGCGACA
TCCACGTCCATGTGCCGGAAGGCGCGACGCCCAAGGACGGTCCCTCGGCCGGCACCG
CCATGGCGACGACGATCGTGTCGATCCTCACCGGCATTCCGGTGCGTCGCGACATCG
CCATGACCGGCGAGATCACGCTGCGCGGCCGCGTGCTGCCGATCGGCGGCCTGAAGG
AGAAGCTGCTCGCGGCGCTGCGCGGCGGCCTSAAGAAGGTGCTGATCCCGGAGGAGA

ACGCCAAGGATCTGGCCGATATTCCGGACTCTGTGAAGAACGGCCTCGAGGTGGTGC
CGGTCTCGCGCATGGACGAGGTGCTCGCGCATGCGCTCGTGCGTCAGCCGACGCCGA
TCGTCTGGGAGGAGCCCGCCCCGATCACGCGTGTGGTCGAGGAGGATGCGGCCGGCG
TTCGCGCGCACTGATCGCCGAGGCTCGAGGCTCGCGCCGTGAGACATGCGGCGCGGG
TCGATCAGCCGATCCGGGAATTCCAGAGCTTTGCGCGTAAAACAGCGCCCGCCGCCC
CTGTTTCACAGGGATTGCGGCGGGTCGCGGMGTTCTAACCCTTGCTTTTGCTGACAT
TCCGGCACATCACTCTCGAACGCCGGGTTTGACGATTCGTGGCCGCTCGCCCTTTCG
ACGCCCGGCTCCCCGGCGGAAAGAGAAGGGATCTGAGACCATGGTCAACAAATTGGA
GCTGGTCGAGCATGTCGCCGAGGCGACCGACACGTCGAAAGCCGCCGCCACCGCCGC
CATCGACGCCGTTCTGGAAGGCATCACCATGGCCCTGAAGAAGGGCGAGGAAGTGCG
GCTCGTCGGTTTCGGCACCTTTTCCGTCAAGAAGCGCGCCGCCGGCGTCGGCCGCAA
TCCGGCGACCGGCGAGGAGATCCAGATTCCCGAATCGACGAGCGCCCGCTTCAAGCC
CGGCGCAACGCTGAAGGCGGCCCTCAACAAGACCAAGTGAGGATTGGAGCCTCGCGC
CGCTCGTCGCCGAGCGGACCGATCTTCCACTCGCCGTCGGGGCGCGTCCCTCGCCGG
CGGATGCTCATGACGGCCGGCGCCATCTTGCGCCGAAGGGCTCCTGTCCGTAAACCT
GTCCCACGTTCTCCGGCTGCGGAGGACGCGACGGGCGGTTAGCTCAGTTGGTAGAGC
GTCTCGCTTACACCGAGAGGGTCGGCGGTTCGAGTCCGTCACCGCCCACCAATTCAA
CGCGGGATCCTCGGAGTCGACACAGCGGAGCGCCGAGCGGGCATCGAATGAACATCC
TCCCTGTCGTTCTCATCGCGCTGGCGGCGGCCGCCGCCGCAACAGCGCCGGTCCGCG
CCGACGAAATGTCGTTTCGTCTGGTGACGCTGGCCGATTCCCATTGCGGCGGGCGCT
GTCCGCAGGTGGTCGCGGCCGAGGGCGAGATCGTCGAGGACACGCCCGACGTCTTTT
TGCGATTCCTGCGCAATCGCGCCAGCCGGGAGCCGTTGCGGAGCATCTTGCTCATCG
ATTCCCCCGGCGGCAAAGTCGTCGCCTCCATGTCGCTCGGCCAGGAGCTGCGCAAGC
TCGGCATGGCGGTGGTGGTCGCGCGCGTCGGCGCCGAGGGCGGGCCGCTGGTGTCGG
GCCGCTGCTATTCGGCCTGCGTCTATGCGCTGATGGGCGGGCTGAAGCGGGTCATTC
CGCGCCAGAGCCGCGTCGGCGTCCACCGCATGTTCAATTATGTCGACGGCGTCGACC
CGATGAGCGGCGAAGCCTTGCGCGAGCGCCGCTACGCCGACGGCGGGATGCGCAAGG
CGCTGTCCCGCTACAGCGCCTCGATGGGGATCAGCCGCGACCTCATCGCGCTCGCCG
AGCGCACATCTCCCGATCGCGTGCATCTGCTCACGGCCGCCGAGATC

>Ms. trichosporium OB3b Lon protease
MTNEKRNSIAPGAVESYPVLPLRDIVVFPHMIVPLFVAREKSIHALEEVTKSDRLIL
LATQKNAGDDDPAADSIYQIGTLASVLQLLKLPDGTVKVLVEGVARAKVRTYTRTDE
YYEADAETLGDDTEAPVEIEALGRSVIAEFDSYVKLNKKVSPEIASAVTQIEDFSKL
ADTVASHLSVKIAEKQDVLETISVAKRLEKCLSLMESEISVLQVEKRIRTRVKRQME
KTQREYYLNEQMKAIQKELGDEDGKDDLAELEERIKNTKLSKEARDKAVAEFKKLRQ
MSPMSAEATVVRNYLDWLLALPWGKKSKIKRDLEAAQDVLDTDHFGLDKVKERILEY
LAVQSRANKLTGPILCLVGPPGVGKTSLGKSIAKATGREFVRMSLGGVRDEAEIRGH
RRTYIGSMPGKIIQSMRKAKTSNPLFLLDEIDKMGMDFRGDPSSALLEVLDPEQNAT
FADHYLEVDYDLSNVMFVTTSNTLNIPAPLMDRMEIIRIAGYTEDEKAEIARKHLIP
NAVKKHGLSADEWAIDDEALMTLIRRYTREAGVRSLEREISNLARKAVKEILLSKEK
GKKVRVTNDNITDYLGVHKFRYGEAELEDQVGVVTGLAVTGVGGELLTIEGVMMPGK
GKMTVTGNLQDVMKESISAAASYVRSRAVDFGIEPPLFDRRDIHVHVPEGATPKDGP
SAGTAMATTIVSILTGIPVRRDIAMTGEITLRGRVLPIGGLKEKLLAALRGG?KKVL
IPEENAKDLADIPDSVKNGLEVVPVSRMDEVLAHALVRQPTPIVWEEPAPITRVVEE
DAAG

Appendix 4

Ms. trichosporium OB3b *mmoZDC* sequence data (Chapter 9)

>re-sequenced Ms. trichosporium OB3b mmo; partial mmoZ, complete mmoD & mmoC

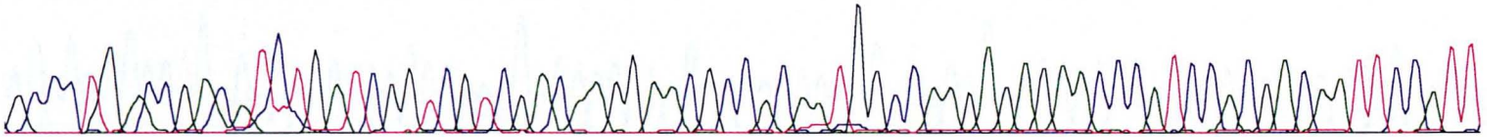
GCCCTGACCGACATCTGATCGGTCGTCGACAAGGAACGCCACAATGGCCAAGAGAGAACCCATCCA
CGACAATTCCATTCGCACCGAATGGGAAGCGAAGATCGCCAAGCTGACGAGCGTCGATCAGGCGAC
GAAGTTCATCCAGGACTTCCGCCTCGCCTACACGTCGCCCTTCCGCAAGAGCTACGATATCGACGT
CGATTATCAGTATATCGAGCGCAAGATCGAGGAGAAGCTCTCGGTGCTGAAGACCGAGAAGCTTCC
CGTCGCCGACCTCATCACCAAGGCGACGACCGGCGAGGACGCCGCCGCGGTGGAGGCGACCTGGAT
CGCCAAGATCAAGGCCGCCAAGAGCAAATATGAGGCGGAGCGCATCCACATCGAGTTCCGGCAGCT
CTACAAGCCGCCGGTTCTGCCGGTGAATGTGTTCCTGCGCACGGATGCGGCGCTCGGCACGGTGCT
GATGGAGATCCGCAACACCGATTATTACGGCACTCCGCTCGAAGGGCTGCGCAAGGAGCGCGGCGT
CAAGGTCCTGCATCTGCAGGCGTGACGCGCTACGCATTGCGAAAGACATGCGTTCGGGCGGCCGCG
TCGCCCGAACGACAAAGAAAAGCGGGAGGACGAAAATGGACCAACAGACGGCGCAGCCGGAAGTCC
GGCAAACCTTGATTCACGCCGACGAGCGCTATCAAGCCTATACGATGGACCTCGAATATATGTTGC
GATGGGAGATTCTGCGCGACGGCGAGTTCGTGCAGGAGGGCTGCTCTCTGTCGCAGGAGTCGGCGC
GCGAGGCGGTCGCTCATGTGCTGAGCCACTTTCGCCGACAGGATGCTACGTCGCAGAACGACGGCG
GGAAAAGCGAAGCGATCCGCGCGCTGCTGCGCGAGATCGGAACGCCCGAGCCCCTGAAGGACGAGA
ACGGCGCGGCGAAGCCGGCGCATATTTAGGGTTCGAGACTTATGTACCAGATCGTCATCGAGACCG
AGGACGGAGAAACCTGTTCGTTCGAATGCGGCCCCAGCGAGGATGTGATCTCCGCGGGGCTGAGGC
AGAGCGTAATTCTGCTCGCCTCCTGCCGAGCCGGCGGCTGCGCCACCTGCAAGGCCGATTGCACGG
ACGGCGATTATGAGCTGATCGATGTGAAGGTCCAGGCGCTGCCGCCGGACGAGGAGGAGGACGGCA
AGGTTCTGCTGTGCCGCACCTTTCCGCGCAGCGATCTGCATCTCCTCGTGCCTTACACCTATGATC
GCATCTCCTTCGAGGCGATTCAGACCAATTGGCTCGCCGAGATCCTCGCCTGTGATCGCGTGTCGT
CCAATGTCGTGCGTCTCGTGCTGCAGCCGCTCACGGCCGATGGCGCGGCGCGCATCTCGCTCAATT
TCGTTCCCGGCCAATTCGTCGACATCGAGATACCGGGCACGCATACACGGCGCTCCTACTCCATGG
CCTCGGTCGCGGAGGATGGGCAGCTCGAATTCTTCATCCGTCTGCTGCCGGACGGCGCCTTCTCGA
AATTCCTGCAAACGGAAGCGAAGGTCGGCATGCGCGTCGATCTGCGCGGACCGGCGGGCTCGTTCT
TCCTGCATGATCACGGCGGCAGATCGCGCGTGTTCGTCGCCGGCGGCACGGGATTGTCGCCGGTGC
TGTCGATGATCCGCCAGCTCGGCAAGGCGAGCGATCCGTCGCCGGCGACGCTTCTGTTCGGCGTCA
CCAATCGCGAGGAATTGTTCTATGTCGACGAGCTGAAGACTCTCGCGCAATCCATGCCGACCCTCG
GCGTGCGCATAGCGGTGGTCAATGACGACGGCGGCAATGGCGTCGACAAGGGAACGGTGATTGATC
TTCTGCGGGCCGAGCTCGAGAAGTCCGACGCCAAGCCGGACATCTATCTCTGTGGTCCTCCCGGCA
TGATCGAGGCCGCTTTCGCGGCGGCGGCGACGGCGGGCGTGCCCAAGGAGCAAGTCTATCTCGAGA
AATT

TEXT BOUND CLOSE TO THE SPINE IN THE ORIGINAL THESIS

10 20 30 40 50 60 70

GCCCTGACCGACATCTGATCGGTCGTCGACAAGGAACGCCACAATGGCCAAGAGAGAACCCATCCACGACAATTCCATT

22004-189_7959.ab1(1>755) →

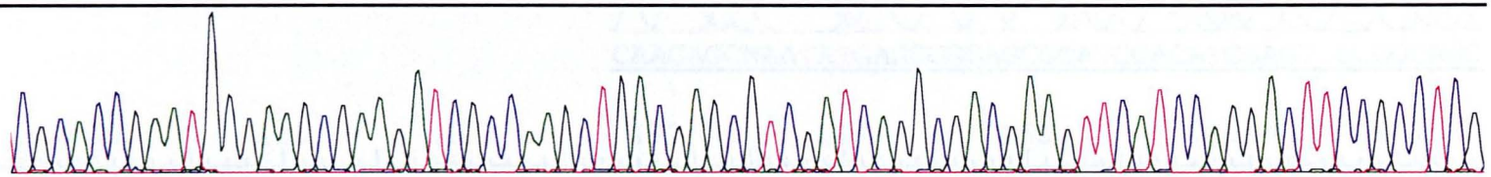


GCCCTGACCGACATCTGATCGGTCGTCGACAAGGAACGCCACAATGGCCAAGAGAGAACCCATCCACGACAATTCCATT

80 90 100 110 120 130 140 150

CGCACCGAATGGGAAGCGAAGATCGCCAAGCTGACGAGCGTCGATCAGGCGACGAAGTTCATCCAGGACTTCCGCCTCG

22004-189_7959.ab1(1>755) →

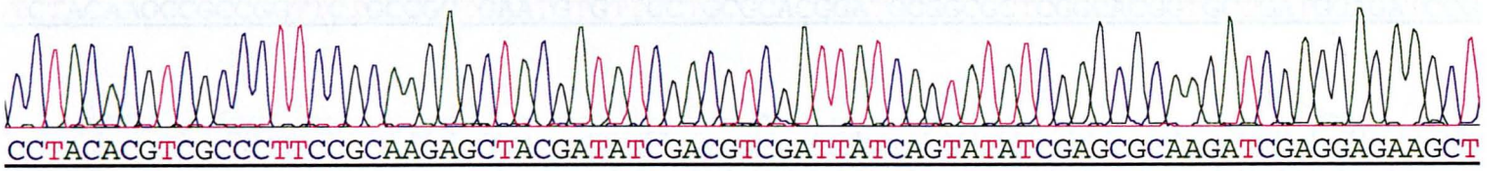


CGCACCGAATGGGAAGCGAAGATCGCCAAGCTGACGAGCGTCGATCAGGCGACGAAGTTCATCCAGGACTTCCGCCTCG

160 170 180 190 200 210 220 230

CCTACACGTCGCCCTTCCGCAAGAGCTACGATATCGACGTCGATTATCAGTATATCGAGCGCAAGATCGAGGAGAAGCT

22004-189_7959.ab1(1>755) →

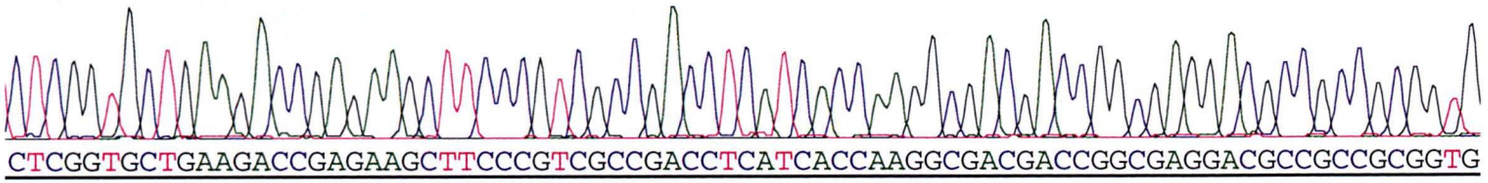


CCTACACGTCGCCCTTCCGCAAGAGCTACGATATCGACGTCGATTATCAGTATATCGAGCGCAAGATCGAGGAGAAGCT

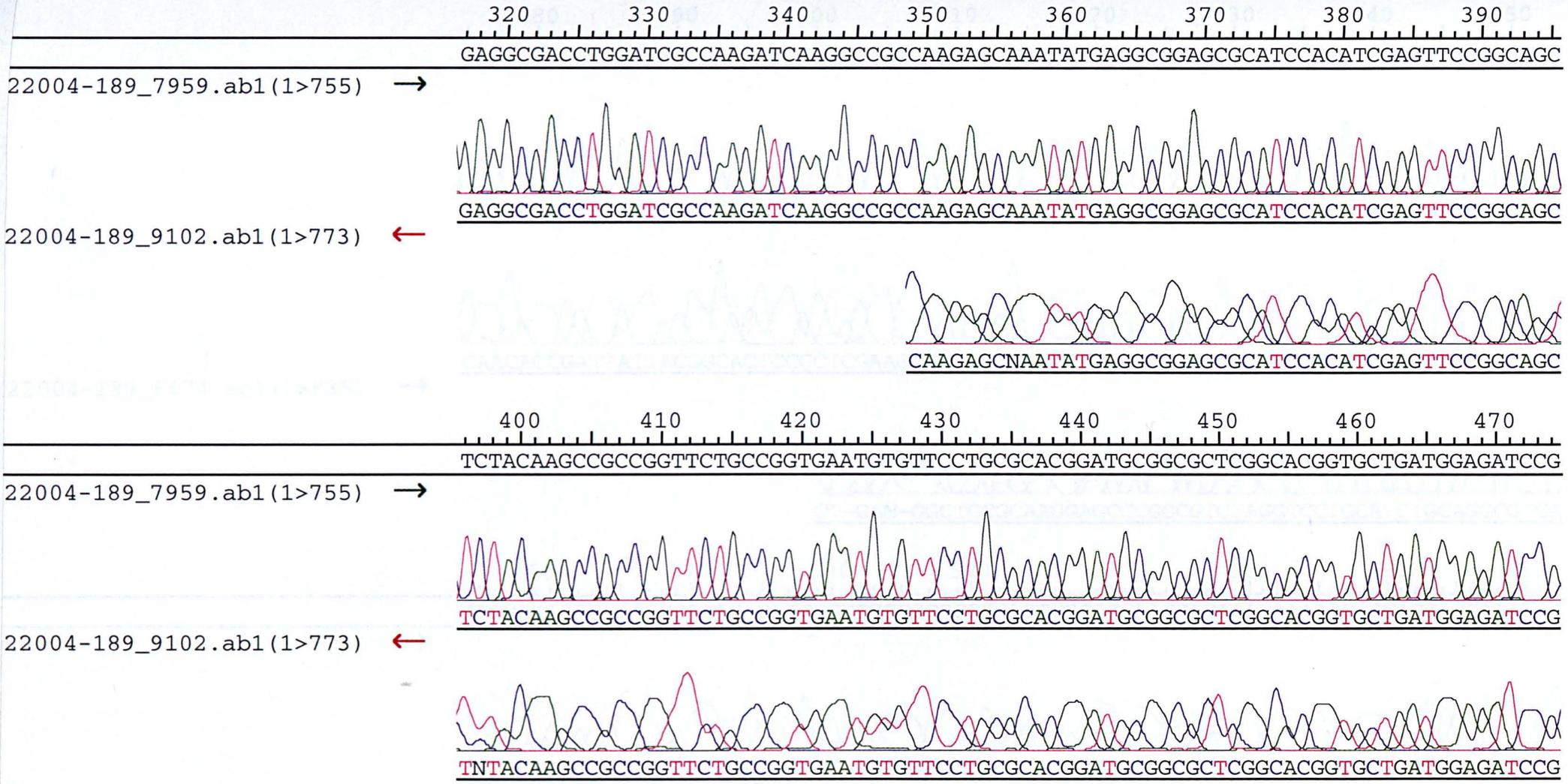
240 250 260 270 280 290 300 310

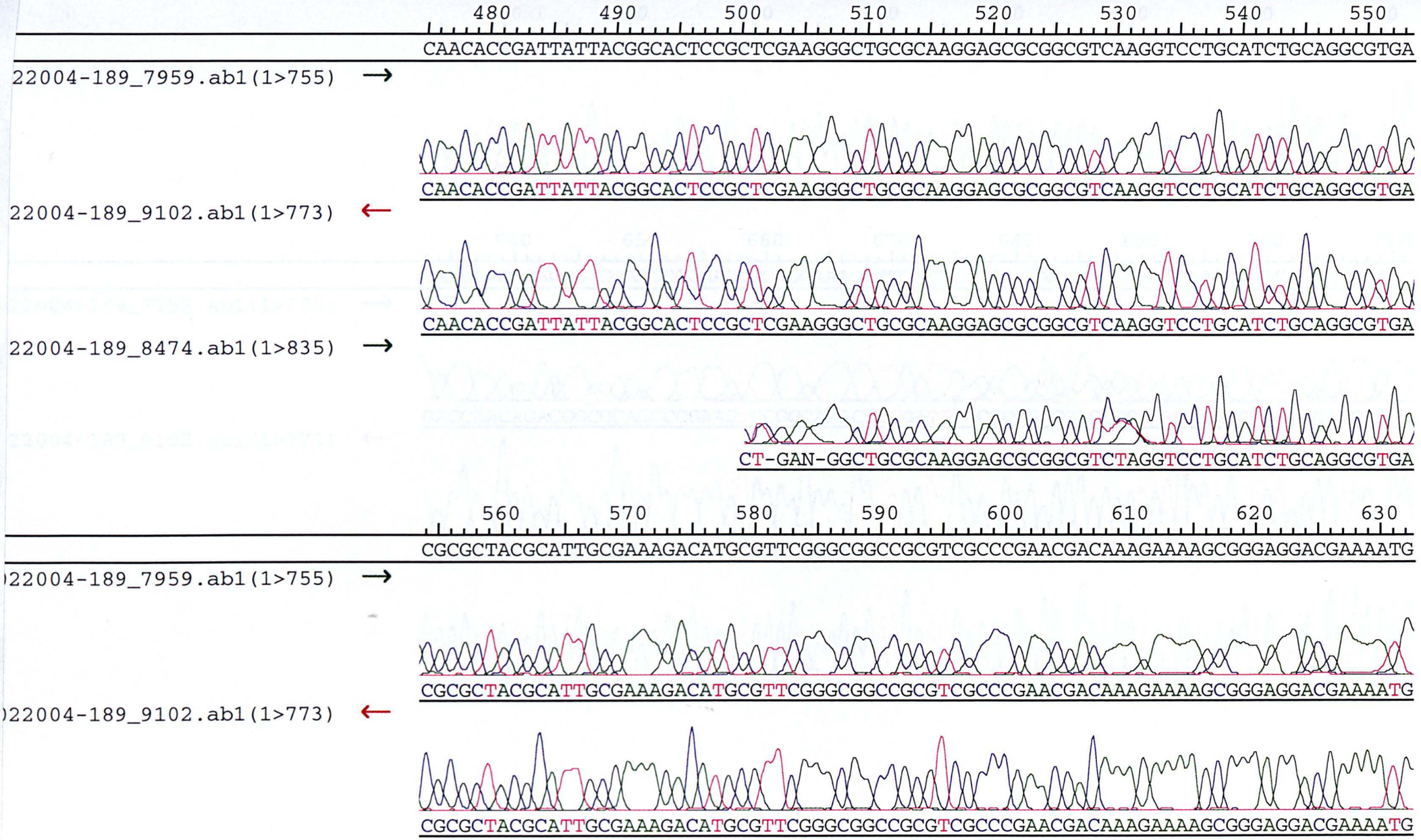
CTCGGTGCTGAAGACCGAGAAGCTTCCCGTCGCCGACCTCATCACCAAGGCGACGACCGGCGAGGACGCCGCCGCGGTG

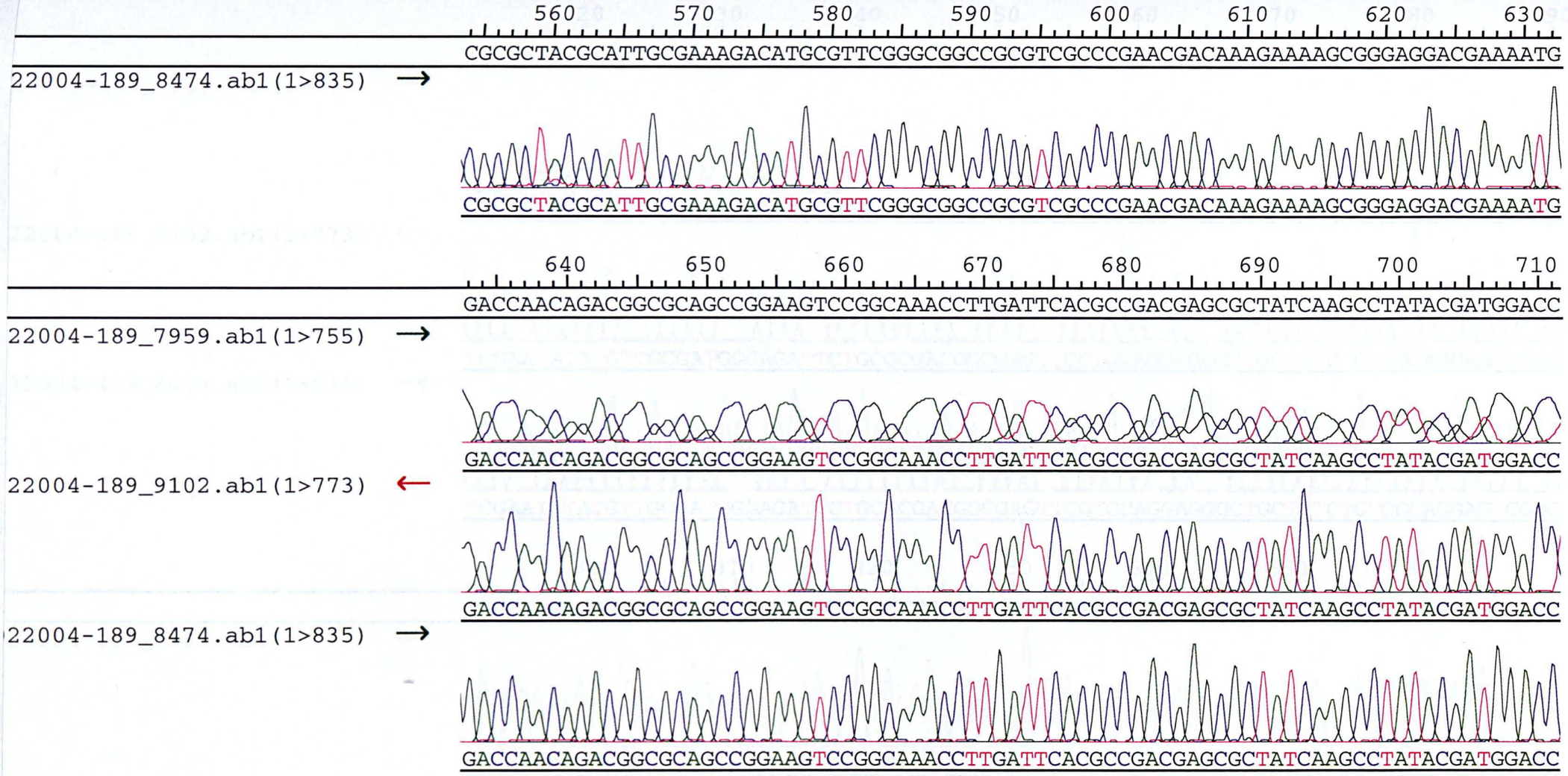
22004-189_7959.ab1(1>755) →

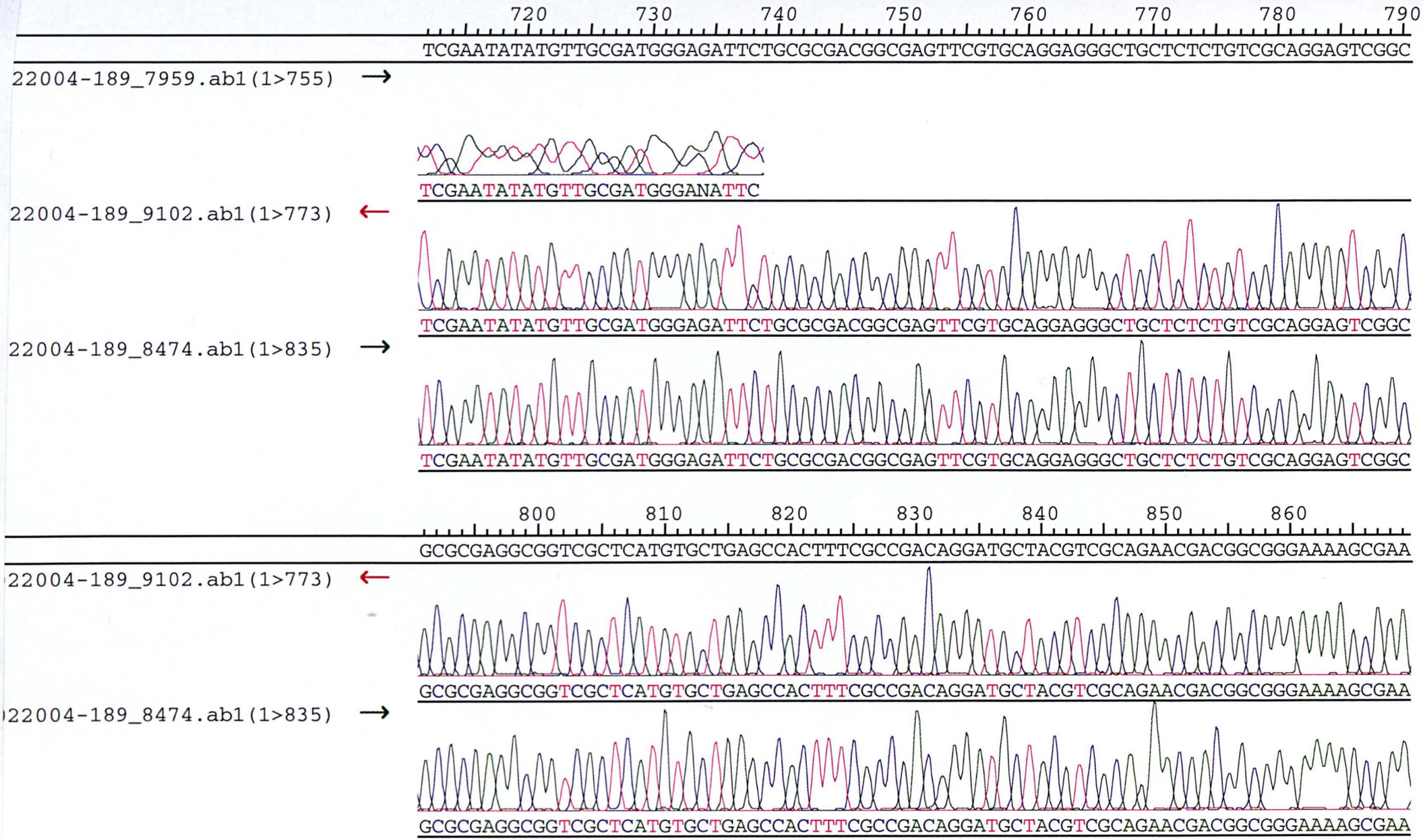


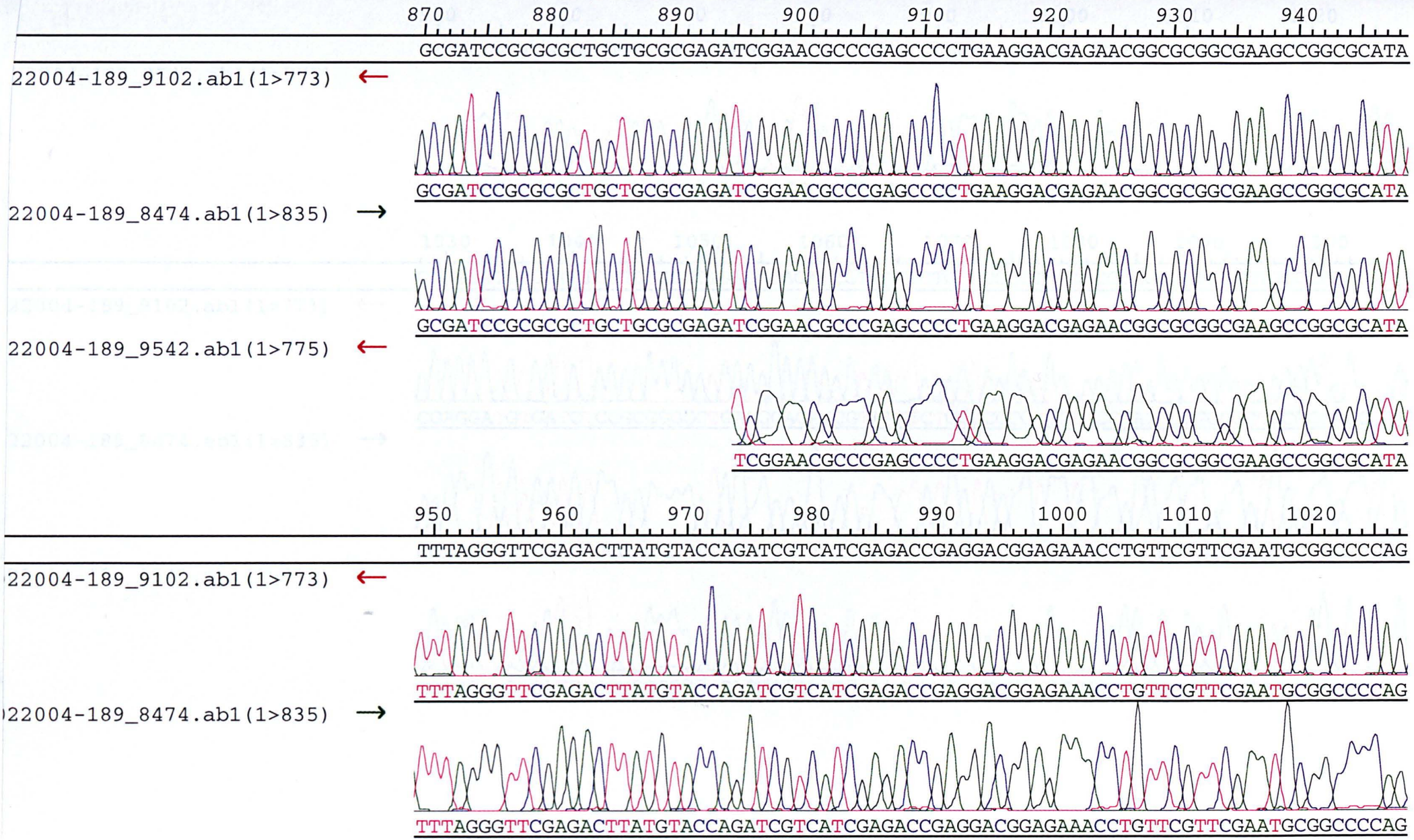
CTCGGTGCTGAAGACCGAGAAGCTTCCCGTCGCCGACCTCATCACCAAGGCGACGACCGGCGAGGACGCCGCCGCGGTG


320 330 340 350 360 370 380 390
GAGGCGACCTGGATCGCCAAGATCAAGGCCGCCAAGAGCAAATATGAGGCGGAGCGCATCCACATCGAGTTCCGGCAGC
22004-189_7959.ab1(1>755)
GAGGCGACCTGGATCGCCAAGATCAAGGCCGCCAAGAGCAAATATGAGGCGGAGCGCATCCACATCGAGTTCCGGCAGC
22004-189_9102.ab1(1>773)
CAAGAGCNAATATGAGGCGGAGCGCATCCACATCGAGTTCCGGCAGC
400 410 420 430 440 450 460 470
TCTACAAGCCGCCGGTTCTGCCGGTGAATGTGTTCCTGCGCACGGATGCGGCGCTCGGCACGGTGCTGATGGAGATCCG
22004-189_7959.ab1(1>755)
TCTACAAGCCGCCGGTTCTGCCGGTGAATGTGTTCCTGCGCACGGATGCGGCGCTCGGCACGGTGCTGATGGAGATCCG
22004-189_9102.ab1(1>773)
TNTACAAGCCGCCGGTTCTGCCGGTGAATGTGTTCCTGCGCACGGATGCGGCGCTCGGCACGGTGCTGATGGAGATCCG


480 490 500 510 520 530 540 550
CAACACCGATTATTACGGCACTCCGCTCGAAGGGCTGCGCAAGGAGCGCGGCGTCAAGGTCCTGCATCTGCAGGCGTGA
22004-189_7959.ab1(1>755)
CAACACCGATTATTACGGCACTCCGCTCGAAGGGCTGCGCAAGGAGCGCGGCGTCAAGGTCCTGCATCTGCAGGCGTGA
22004-189_9102.ab1(1>773)
CAACACCGATTATTACGGCACTCCGCTCGAAGGGCTGCGCAAGGAGCGCGGCGTCAAGGTCCTGCATCTGCAGGCGTGA
22004-189_8474.ab1(1>835)
CT-GAN-GGCTGCGCAAGGAGCGCGGCGTCTAGGTCCTGCATCTGCAGGCGTGA
560 570 580 590 600 610 620 630
CGCGCTACGCATTGCGAAAGACATGCGTTCGGGCGGCCGCGTCGCCCGAACGACAAAGAAAAGCGGGAGGACGAAAATG
22004-189_7959.ab1(1>755)
CGCGCTACGCATTGCGAAAGACATGCGTTCGGGCGGCCGCGTCGCCCGAACGACAAAGAAAAGCGGGAGGACGAAAATG
22004-189_9102.ab1(1>773)
CGCGCTACGCATTGCGAAAGACATGCGTTCGGGCGGCCGCGTCGCCCGAACGACAAAGAAAAGCGGGAGGACGAAAATG


560 570 580 590 600 610 620 630
CGCGCTACGCATTGCGAAAGACATGCGTTCGGGCGGCCGCGTCGCCCGAACGACAAAGAAAAGCGGGAGGACGAAAATG
22004-189_8474.ab1(1>835)
CGCGCTACGCATTGCGAAAGACATGCGTTCGGGCGGCCGCGTCGCCCGAACGACAAAGAAAAGCGGGAGGACGAAAATG
640 650 660 670 680 690 700 710
GACCAACAGACGGCGCAGCCGGAAGTCCGGCAAACCTTGATTCACGCCGACGAGCGCTATCAAGCCTATACGATGGACC
22004-189_7959.ab1(1>755)
GACCAACAGACGGCGCAGCCGGAAGTCCGGCAAACCTTGATTCACGCCGACGAGCGCTATCAAGCCTATACGATGGACC
22004-189_9102.ab1(1>773)
GACCAACAGACGGCGCAGCCGGAAGTCCGGCAAACCTTGATTCACGCCGACGAGCGCTATCAAGCCTATACGATGGACC
22004-189_8474.ab1(1>835)
GACCAACAGACGGCGCAGCCGGAAGTCCGGCAAACCTTGATTCACGCCGACGAGCGCTATCAAGCCTATACGATGGACC

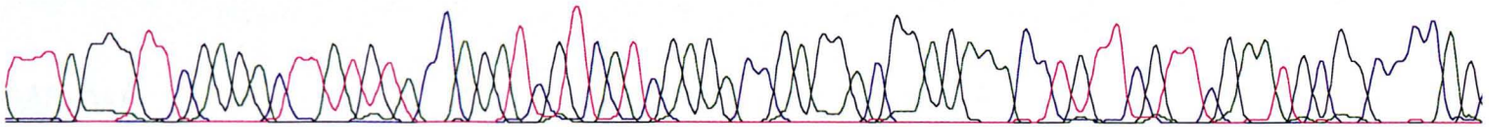

720 730 740 750 760 770 780 790
TCGAATATATGTTGCGATGGGAGATTCTGCGCGACGGCGAGTTCGTGCAGGAGGGCTGCTCTCTGTCGCAGGAGTCGGC
22004-189_7959.ab1(1>755)
TCGAATATATGTTGCGATGGGANATTC
22004-189_9102.ab1(1>773)
TCGAATATATGTTGCGATGGGAGATTCTGCGCGACGGCGAGTTCGTGCAGGAGGGCTGCTCTCTGTCGCAGGAGTCGGC
22004-189_8474.ab1(1>835)
TCGAATATATGTTGCGATGGGAGATTCTGCGCGACGGCGAGTTCGTGCAGGAGGGCTGCTCTCTGTCGCAGGAGTCGGC
800 810 820 830 840 850 860
GCGCGAGGCGGTCGCTCATGTGCTGAGCCACTTTCGCCGACAGGATGCTACGTCGCAGAACGACGGCGGGAAAAGCGAA
22004-189_9102.ab1(1>773)
GCGCGAGGCGGTCGCTCATGTGCTGAGCCACTTTCGCCGACAGGATGCTACGTCGCAGAACGACGGCGGGAAAAGCGAA
22004-189_8474.ab1(1>835)
GCGCGAGGCGGTCGCTCATGTGCTGAGCCACTTTCGCCGACAGGATGCTACGTCGCAGAACGACGGCGGGAAAAGCGAA


870
880
890
900
910
920
930
940
GCGATCCGCGCGCTGCTGCGCGAGATCGGAACGCCCGAGCCCCTGAAGGACGAGAACGGCGCGGCGAAGCCGGCGCATA
22004-189_9102.ab1(1>773)
GCGATCCGCGCGCTGCTGCGCGAGATCGGAACGCCCGAGCCCCTGAAGGACGAGAACGGCGCGGCGAAGCCGGCGCATA
22004-189_8474.ab1(1>835)
GCGATCCGCGCGCTGCTGCGCGAGATCGGAACGCCCGAGCCCCTGAAGGACGAGAACGGCGCGGCGAAGCCGGCGCATA
22004-189_9542.ab1(1>775)
TCGGAACGCCCGAGCCCCTGAAGGACGAGAACGGCGCGGCGAAGCCGGCGCATA
950
960
970
980
990
1000
1010
1020
TTTAGGGTTCGAGACTTATGTACCAGATCGTCATCGAGACCGAGGACGGAGAAACCTGTTCGTTCGAATGCGGCCCCAG
22004-189_9102.ab1(1>773)
TTTAGGGTTCGAGACTTATGTACCAGATCGTCATCGAGACCGAGGACGGAGAAACCTGTTCGTTCGAATGCGGCCCCAG
22004-189_8474.ab1(1>835)
TTTAGGGTTCGAGACTTATGTACCAGATCGTCATCGAGACCGAGGACGGAGAAACCTGTTCGTTCGAATGCGGCCCCAG

950 960 970 980 990 1000 1010 1020

TTTAGGGTTCGAGACTTATGTACCAGATCGTCATCGAGACCGAGGACGGAGAAACCTGTTCGTTCGAATGCGGCCCCAG

22004-189_9542.ab1(1>775) ←

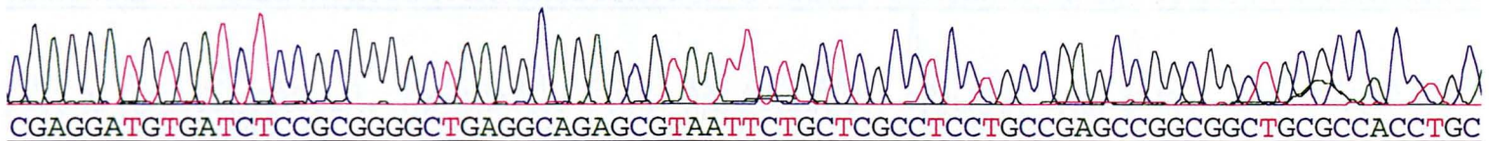


TTTAGGGTTCGAGACTTATGTACCAGATCGTCATCGAGACCGAGGACGGAGAAACCTGTTCGTTCGAATGCGGCCCCAG

1030 1040 1050 1060 1070 1080 1090 1100

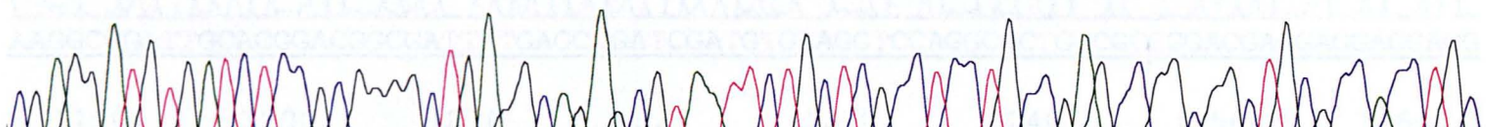
CGAGGATGTGATCTCCGCGGGGCTGAGGCAGAGCGTAATTCTGCTCGCCTCCTGCCGAGCCGGCGGCTGCGCCACCTGC

22004-189_9102.ab1(1>773) ←



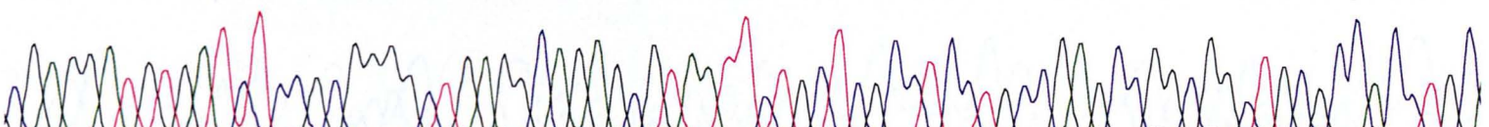
CGAGGATGTGATCTCCGCGGGGCTGAGGCAGAGCGTAATTCTGCTCGCCTCCTGCCGAGCCGGCGGCTGCGCCACCTGC

22004-189_8474.ab1(1>835) →

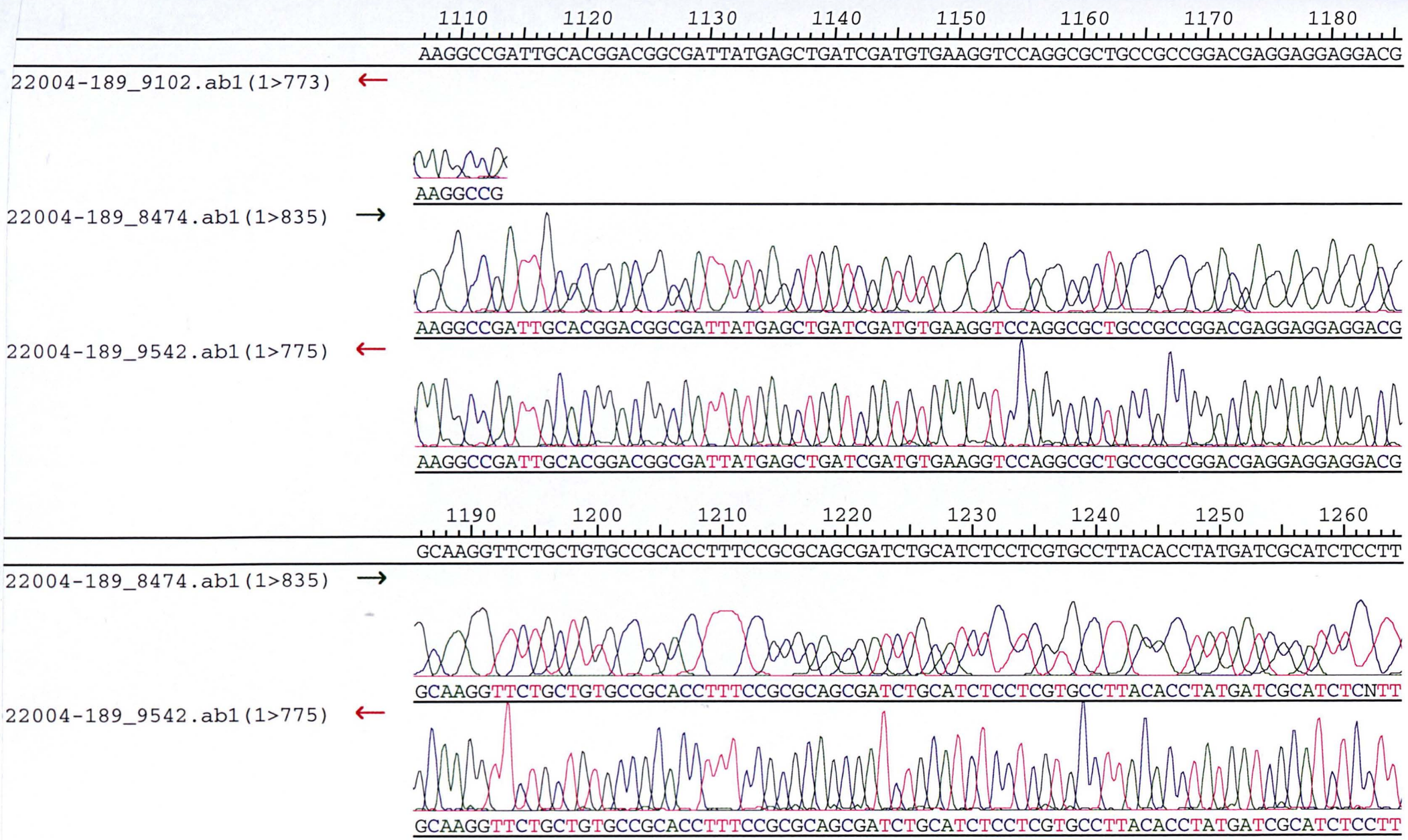


CGAGGATGTGATCTCCGCGGGGCTGAGGCAGAGCGTAATTCTGCTCGCCTCCTGCCGAGCCGGCGGCTGCGCCACCTGC

22004-189_9542.ab1(1>775) ←



CGAGGATGTGATCTCCGCGGGGCTGAGGCAGAGCGTAATTCTGCTCGCCTCCTGCCGAGCCGGCGGCTGCGCCACCTGC


1110 1120 1130 1140 1150 1160 1170 1180
AAGGCCGATTGCACGGACGGCGATTATGAGCTGATCGATGTGAAGGTCCAGGCGCTGCCGCCGGACGAGGAGGAGGACG
22004-189_9102.ab1(1>773)
AAGGCCG
22004-189_8474.ab1(1>835)
AAGGCCGATTGCACGGACGGCGATTATGAGCTGATCGATGTGAAGGTCCAGGCGCTGCCGCCGGACGAGGAGGAGGACG
22004-189_9542.ab1(1>775)
AAGGCCGATTGCACGGACGGCGATTATGAGCTGATCGATGTGAAGGTCCAGGCGCTGCCGCCGGACGAGGAGGAGGACG
1190 1200 1210 1220 1230 1240 1250 1260
GCAAGGTTCTGCTGTGCCGCACCTTTCCGCGCAGCGATCTGCATCTCCTCGTGCCTTACACCTATGATCGCATCTCCTT
22004-189_8474.ab1(1>835)
GCAAGGTTCTGCTGTGCCGCACCTTTCCGCGCAGCGATCTGCATCTCCTCGTGCCTTACACCTATGATCGCATCTCNTT
22004-189_9542.ab1(1>775)
GCAAGGTTCTGCTGTGCCGCACCTTTCCGCGCAGCGATCTGCATCTCCTCGTGCCTTACACCTATGATCGCATCTCCTT

1190 1200 1210 1220 1230 1240 1250 1260

GCAAGGTTCTGCTGTGCCGCACCTTTCCGCGCAGCGATCTGCATCTCCTCGTGCCTTACACCTATGATCGCATCTCCTT

22004-189_9990.ab1(1>800) ←

GCCTTACACCTATGATCGCATCTCCTT

1270 1280 1290 1300 1310 1320 1330 1340

CGAGGCGATTCAGACCAATTGGCTCGCCGAGATCCTCGCCTGTGATCGCGTGTCGTCCAATGTCGTGCGTCTCGTGCTG

22004-189_8474.ab1(1>835) →

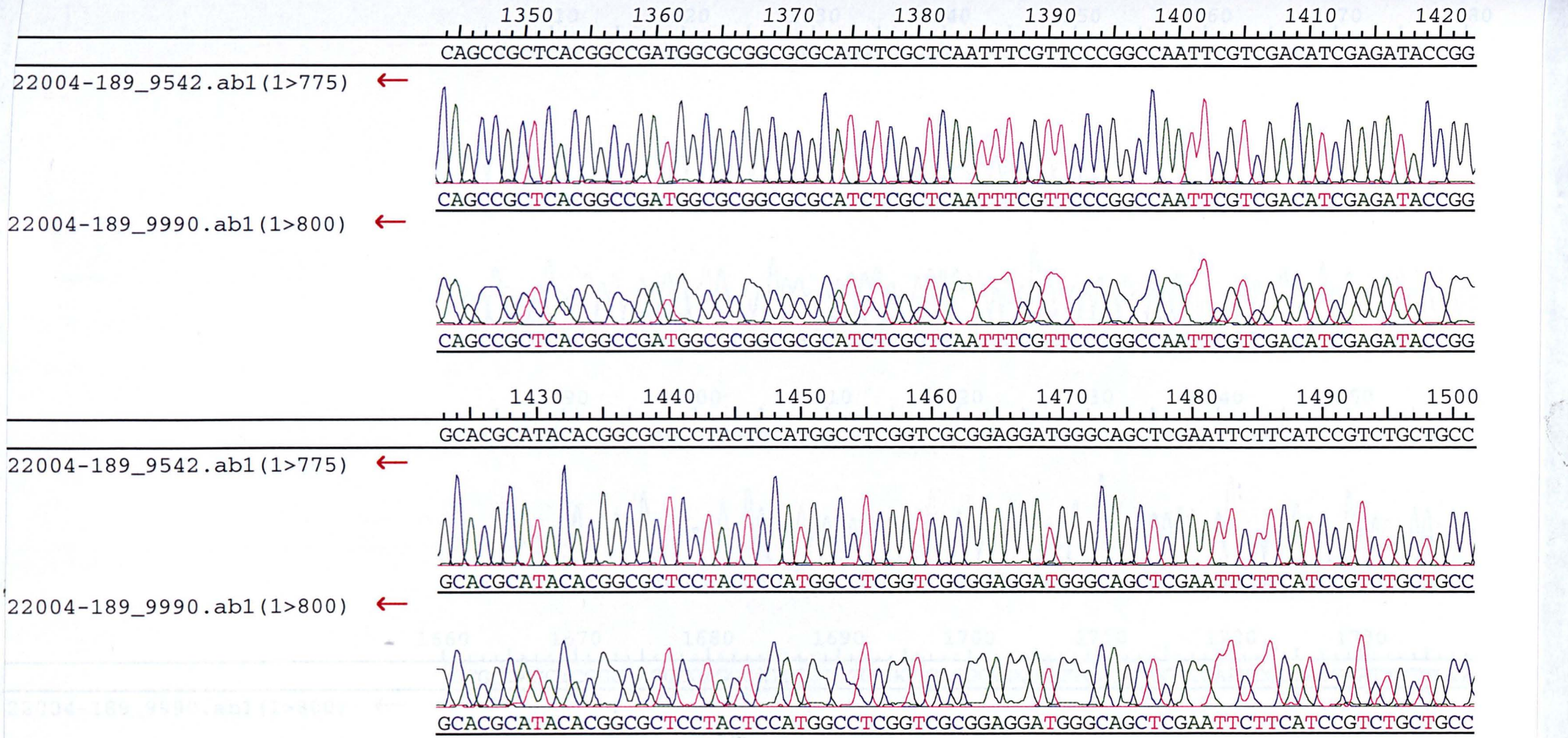
CGAGGCGATTC

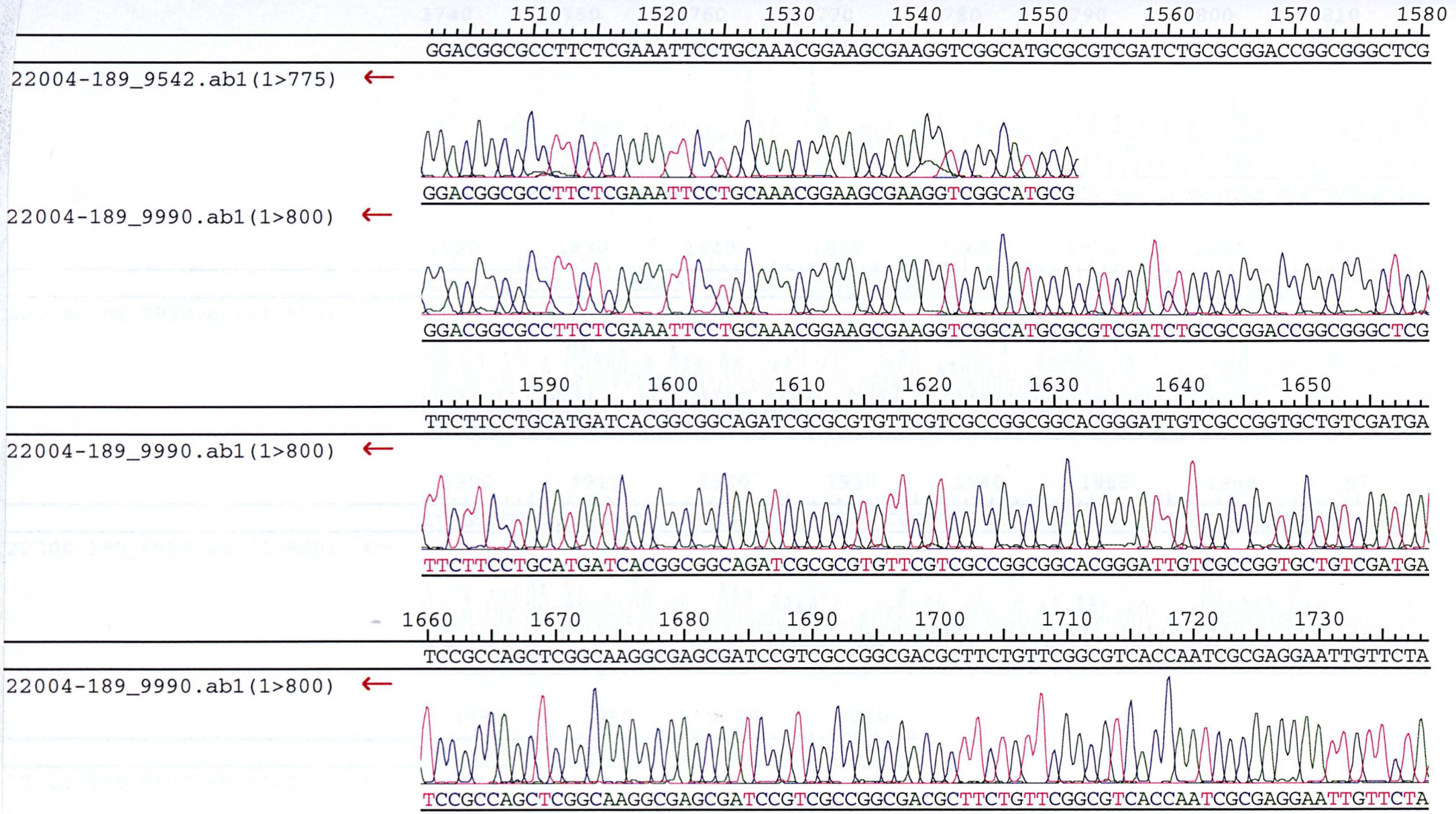
22004-189_9542.ab1(1>775) ←

CGAGGCGATTCAGACCAATTGGCTCGCCGAGATCCTCGCCTGTGATCGCGTGTCGTCCAATGTCGTGCGTCTCGTGCTG

22004-189_9990.ab1(1>800) ←

CGAGGCGATTCAGACCAATTGGCTCGCCGAGATCCTCGCCTGTGATCGCGTGTCGTCCAATGTCGTGCGTCTCGTGCTG

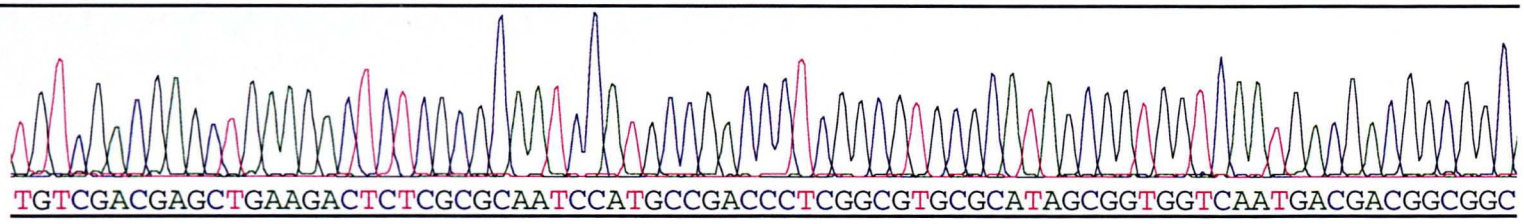

1350
1360
1370
1380
1390
1400
1410
1420
CAGCCGCTCACGGCCGATGGCGCGGCGCGCATCTCGCTCAATTTCGTTCCCGGCCAATTCGTCGACATCGAGATACCGG
22004-189_9542.ab1(1>775)
CAGCCGCTCACGGCCGATGGCGCGGCGCGCATCTCGCTCAATTTCGTTCCCGGCCAATTCGTCGACATCGAGATACCGG
22004-189_9990.ab1(1>800)
CAGCCGCTCACGGCCGATGGCGCGGCGCGCATCTCGCTCAATTTCGTTCCCGGCCAATTCGTCGACATCGAGATACCGG
1430
1440
1450
1460
1470
1480
1490
1500
GCACGCATACACGGCGCTCCTACTCCATGGCCTCGGTCGCGGAGGATGGGCAGCTCGAATTCTTCATCCGTCTGCTGCC
22004-189_9542.ab1(1>775)
GCACGCATACACGGCGCTCCTACTCCATGGCCTCGGTCGCGGAGGATGGGCAGCTCGAATTCTTCATCCGTCTGCTGCC
22004-189_9990.ab1(1>800)
GCACGCATACACGGCGCTCCTACTCCATGGCCTCGGTCGCGGAGGATGGGCAGCTCGAATTCTTCATCCGTCTGCTGCC


1510 1520 1530 1540 1550 1560 1570 1580
GGACGGCGCCTTCTCGAAATTCCTGCAAACGGAAGCGAAGGTCGGCATGCGCGTCGATCTGCGCGGACCGGCGGGCTCG
22004-189_9542.ab1(1>775)
GGACGGCGCCTTCTCGAAATTCCTGCAAACGGAAGCGAAGGTCGGCATGCG
22004-189_9990.ab1(1>800)
GGACGGCGCCTTCTCGAAATTCCTGCAAACGGAAGCGAAGGTCGGCATGCGCGTCGATCTGCGCGGACCGGCGGGCTCG
1590 1600 1610 1620 1630 1640 1650
TTCTTCCTGCATGATCACGGCGGCAGATCGCGCGTGTTCGTCGCCGGCGGCACGGGATTGTCGCCGGTGCTGTCGATGA
22004-189_9990.ab1(1>800)
TTCTTCCTGCATGATCACGGCGGCAGATCGCGCGTGTTCGTCGCCGGCGGCACGGGATTGTCGCCGGTGCTGTCGATGA
1660 1670 1680 1690 1700 1710 1720 1730
TCCGCCAGCTCGGCAAGGCGAGCGATCCGTCGCCGGCGACGCTTCTGTTCGGCGTCACCAATCGCGAGGAATTGTTCTA
22004-189_9990.ab1(1>800)
TCCGCCAGCTCGGCAAGGCGAGCGATCCGTCGCCGGCGACGCTTCTGTTCGGCGTCACCAATCGCGAGGAATTGTTCTA

1740 1750 1760 1770 1780 1790 1800 1810

TGTCGACGAGCTGAAGACTCTCGCGCAATCCATGCCGACCCTCGGCGTGCGCATAGCGGTGGTCAATGACGACGGCGGC

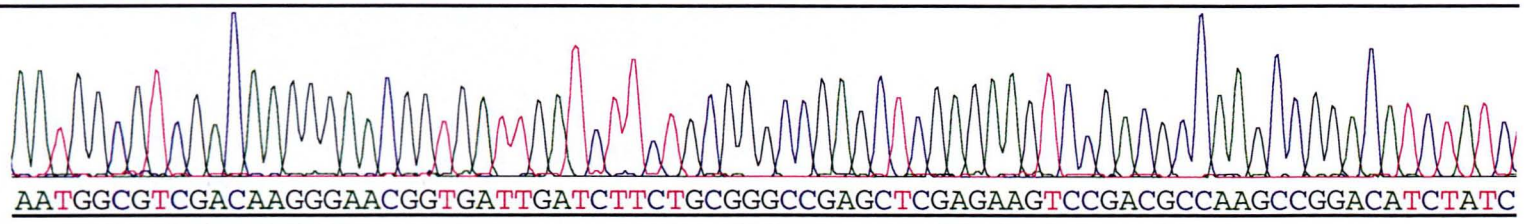
22004-189_9990.ab1(1>800) ←



1820 1830 1840 1850 1860 1870 1880 1890

AATGGCGTCGACAAGGGAACGGTGATTGATCTTCTGCGGGCCGAGCTCGAGAAGTCCGACGCCAAGCCGGACATCTATC

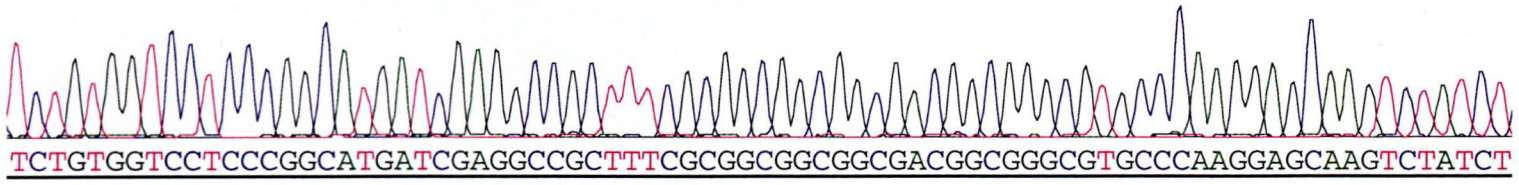
22004-189_9990.ab1(1>800) ←



1900 1910 1920 1930 1940 1950 1960 1970

TCTGTGGTCCTCCCGGCATGATCGAGGCCGCTTTCGCGGCGGCGGCGACGGCGGGCGTGCCCAAGGAGCAAGTCTATCT

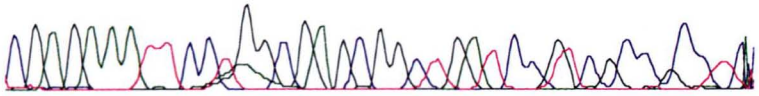
22004-189_9990.ab1(1>800) ←



1980 1990 2000 2010

CGAGAAATTCCTGGCGAGCGGCTGATCCGTCGCCGCCTC

22004-189_9990.ab1(1>800) ←



CGAGAAATTCCTGGCGAGCGGCTGATCCGTCGCCGCCTC

Appendix 5

Methylocella silvestris BL2^T sequence data (Chapter 9)

>Methylocella silvestris BL2 sMMO genes; includes flanking sequence

GCGCTGGACGCGCCGCCGCGGAGCGAACGGAGGGGGGAAAAAATGACAATTAAAGTCGGGGA
TCGTTTGCCAAACGCGAGCTTCGCCGTGATGACCCCGGAGGGTCCGGCCGTACGCTCGACGG
ACGATATTTTTAAAGGCCGGCGGGTCGTTCTGATCGGCGTGCCCGGCGCCTTTACGCCAACA
TGCAGCAATGTGCATCTGCCCGGTTTCGTCAACAGGCTCGACGATTTCAAACAGAAACGCAT
CGACGCCATCGCGGTCACCGCAGTCAACGACGTCTTTGTCATGAACGCCTGGGCGGCCTCCA
GCGACGCCGGCGCGCATATGTCCTTTCTCGCCGACGGCAATGGCGATTTCGCAAAGGCGCTC
GGCCTGACCCTTGACCTGACCGAGCGCGGCCTTGGCGTGAGGTCGCAGCGTTATTCGATGGT
GGTCGACGACGGCGTGGTGCAGCAGCTCAATGTCGAGGCGAGCGCCAGCAGGGCGGACGTCT
CCGGCGCCGAAGCGATTTTGCATCAGCTTTAACGCGGGGCGGAGCAGAACCTCTTGCGTCCG
CTGCAATCGTTACGGCGTGGCCGCTTTGCGGGCAGATAGAAACGGAGCCATGCCGAGCCGCG
CGAGTGCGTCGGCGCGCTCGTTCAATTCGTCGCCGGCATGGCCCTTGACCCAGCGCCATTCG
ATCGCATGGGCTTCCGCCGCGGCTTCGAGCCTTTGCCATAGATCGACATTCTTGACAGGTTT
CTTGTCCGCCGTGCGCCAGCCATTGCGCTTCCATCCGGCCAGCCATGTGGTGACGCCTCCTT
TCACATAGTTGGAATCTGTGACGAGCTGCACGGCGCAGGGCCGCGTCAGCGCTTCGAGCGCC
ATGATCGCCGCCATCAGCTCCATGCGATTGTTGGTCGTGGCCGCCTCGCCGCCGCAAAGCTC
CTTGAGGTGATCGCCGAACGTCATGACGGCGCCCCAGCCGCCGGGACCGGGATTTCCCGAAC
AGGCGCCGTCGGTGAAGATCACGACTGGCTTTGGCATCGACGCTTCCAATTCATATCAGGAC
GCCGAGTTCTCGCGCGTCGCCGATCGTCTGCAGCATTTGCAATTTGGCGAAATATTCGAGCG
GATTCTTCGGATTGACGAGCGCGCCCGGGGGAACGTTGAACCAGTCGACGAGGCGGGTCAGC
GCGAAGCGCAGGGCGGCGCCCCGGGCGAGCGTCGGGATGGCTTTCGCTTCCGCCGCGCTCAA
GGGGCGCAGGCTTGCATAAGCGCCGAACATCGCCGCGCCCTTGTCGATGTTAAACGCGCCAC
CCGGCTCGAAACACCAGGCGTTGAGGCAGATCGCAAGGTCATAGGAAAGATCGTCGACGCAG
GCGAAATAGAAGTCGATCACGCCGGAGATCTCGTCGCCGAGAAACAGGACATTGTCGGGAAA
GAGATCGGCGTGGATGATTCCCCGCGGCAGATCCCGCGGCCAGTTCGCTTGCAGGAATTTCA
GCTCGGCGGAGATTGTCGCCGCAAGGCCCGCTTTGACCTCGTCGGCGCGAAAGGCGATCGGC
GCGAAGAGTTTTGGCCAGGCTTCTAGGTTGAGAGCATTGGCGCGCGAGCCCTTGAAATCGGC
GGCCGCGAGGTGAAGCTGCGCCAGCGCGGCCCCGAGGCGGGCGCAATCGGCAATGCTCGGCC
GGTTGCGGCAAACGCCTTCGAGGAAGGTCACGATCGCCGCCGGACGCCCGGCGAGGCGCCCG
AGCGCCCCATCCGAACGCATGCGCACCGGCTGCGGACAGTTTAGGCCGCGCGCGGCCCGAGA
TGATCCAACAGCGAGAGGAAGAACGGCAGGCTCCCCTCCTCGACGCGCTTTTCATAGAGCGT
CAGGATGTAGACGCCCGTTTCGGTCGACAGGACGAAATTCGAATTCTCGACGCCTTCCGCGA
TGCCCTTCAGGGCGCGGACGCGGCCGAGATTGTAATGCTCAAGAAAGGCGGCGAGGTCGGCG
TCGGAGACTTCGGTATAGACTGCCATTGGGGTCCGCTGAAGACATTGAGTCCTGGCCACGGC
CAAGGACGCCGCAAAGGGCTGCTTCCCTTGGGCGCGGGAGGCGCTCCGGTGTTAGAGAGAAC
GGGCCGCAAGGTCAATTTGAGCCGGGCCAAGTGAGCCGGGCCGCCCCCTGCCGCGACATCAG
ACGGTCTCGCCCTGGGGATTGAGCGCAACGGCGGCGAGTACCGCCGGCGCCTCGGGCTGGAT
GACGCCGCAGGACATGATCAGCTTTGCGGCGGCCTCCGGCGTGATCGCCAGTTCGATCACCT
CGCGGGCCGGCACATAGAAATAGAAACCGGTGGTCGGATTCGGCGCGCAGGGCAGAAAAACT
GAAACATAGGGCTCGCCGTTGGCAAGATGAGAGCCAATGAGAGAGGACGGCGGCAGGGAAAT
AAAGACGATGGACCAGGAGCCCTTGCCCGGAAACTCGATCATCCCGACCTTGCGAAAACTTT
GCCCGGACTGCGAGAACAGCGTTTCGAAAATCTGCTTCACGCTTTTATAGATCGAGCGGACC
ACGGGCATTCGGGCGAGCAGCGCCTCGCCGATCTTGAAGAGGCCGAGCGCCGCGATGCTGTG
GGTGGCGCAGCCGAGCAGCGTCAGGCCGAGGAATGCCAGGACGACGCCGAGTCCCGGCACGC
GAAACGGCAAATAGAAATCCGGCACGACATTTTGCGGCAGCAGGCCGCGCACCCAGGCGTCG
ATGGTGTCGACGAACCACCAGACGATATAGGCGGTGACGGCGACCGGTCCGGCGACGACGAT
CCCGGTCAGGAACCAGTTGCGGATCTTTGCGCCGAACCCGGCCGACTTGGCGGGCCCCGGGC
GGCCGTCCGAGTCGGCGTCGATGAGGCTTTGCAGCGGCGGAAGTTGCGTCATGCCTCGCTCA
TGGCTCCACAGCTGGACGTCGCTCGATTCGCCGACGCAAAAAACGGGGCCGAGCCGCCTGCG
CGAGAGCGCCGATGCTTCGGCCGGCCGTCGTCTCTCTATATCACAGACCTCGCAACGGCAGC
GATTCCGGTCAGGGGACGCCGCCGAGCCGGTCAAATCTGGACATGAGGATAAGGGCGTTTCG
GCTTGGGAGCAGACAAAATGGCAGCGGCGCCAGCCGCTCGCCTGCTATCGGCATGAGACGGA
AACGCTCACGGCTCGAGCCTGAGCCGTCAGCGCCGGCGCCGACATCTCCTTCGACGCGGTCG
CGCCCGTGAAAGCGATATGGTTGCCCTCAAGAAAGCCCGCGAGGAGATTTCCTTTAAGCGCG
ACGTTCACCTTTTGCGGCGCATCGCCGCTCGTCGCTACAAGGGTCTTGAAGGCGTCGAGCCT
CGCCGAAACGACTCCGACAACATAGCCGCTCTGGTCGAAAAGCGGGCTTCCAGAATCGGCCG
GCTGGACCGGCGTCGAAATCTGCAGGAAGCGGTCGTCGCCCGCAACGCCAGTGAGGCCCGTG
ATGCGGCCGAAAGAAAACTCGCCCGCGTTCAGCGGCGCGGGAGGGAGCGGAAAGGCAAACGC
CTCGATCGTTTCGCCGAGCTGGACGTCGGAGCGAATATTGGCGACACCGTCCGGCTTCAACG
GCGTTTTGAGAAGGGCAAGATCCGCGGCGTCGTCGCGAGCGACAAGTTCGGCTTGGACCGCG
GCGCGTCCATTCTGCGCGACCGCGATGTTCGAGCAGGACGCGACGATTTGGGCGCTGGTTAC
AACAAATCCCTCGTCGTTCACAGAGAAGCCCATTCCGGGCATGTTCTCGCGCTGCTTTTGCG

CCGTCTCTGGCGCCAACGCAGCCGTTGGCCTGACGCCGGCGGGGTCCGCCGCCGGCGCCAAG
CCGACCGACGAGGGCAGCTGAGAGAGATCGGCGAATGGCGCGCCGCTCATCGAACTTGCGAG
CGACCCAGATATTAAAGCTGCCGCGCGCTCCAGATGCAGATCGGCGGCGGACTCCAGCCAAA
ACAGCGAGAGCCAAATCACGCCCGGGCGATCGCTTTGGCAGCGAATGTAACTGTCGAGCCCA
TTGGCGCGCGACGACACGATGAAGAAGTCATCCGTCAAAATCTTGAAAGAGATCTGCCCGCC
GTCGGCAATGACTTTCTCGACCACCGATTTATAGGCGTTGGCGAGATTGACGTCGCGGAGGT
AGTCATAGGTCAGCCAAACCCGCTTGGCCGGATCTCGCCAGGTGAGGCCATCCTTGTCCCGC
TCGGCCACCAGGCCAAGCCCCATCGGCGCCCAGATCTTGTAACCGCGCGATGGATGCGGGAT
CGCCCGAAATCCCCACATCTCGAACAGCGGCGCGCCAGCGCGAAAAAGCCGCGCGCGTAGAT
CTTCGTCTATGGCGCCGGTCGCCGAAAGCCCGTTGTCGCTTTGATACTGCGCCGTCGCTTCA
AAAAGCTGGCGGCTGAAATTGGCGTCCGCGGTCACCGGCCAATATCCTGCCGCCGTCAGCAG
GACTTTGAATTTGAGCCTTTCGCTGATCGGCAGCTGACTGAAGGCGCGCGCCGCCTCGACGT
AGCCCGGCGTTGGCGGCCCCGCAGCGACCGGCAGGCTCCAGATCATCAGAAGCGCAACCGCC
GCGAGTCCCTTCATGCTGTCCTCAAGCAAACCCATCAAGCCTGAACCCACGCGGAATCCGGC
GCCTGTAAAAAATCAATACAGTTCAAAACTGATACTGTCTGATATCGATACAGTTGAAGATG
AAAAACAGAAGGCTGCGCGAGCGGCTGATCCCCTGATCTGGGGGTGCGCGCTACAAAGCGAA
TGGAGGATCAGAGCGGAGGTTAACGGAGAGCGCCATTGTCGATCTGTATCGCAATTATACAT
ATGTATTATTCTTATACAGATTGCGATGGAGAAGAAATTTTCTAGAGGCGCAACTCAAGCGG
AGTGCATTTTATCCTGCCCCAGTTCCATTCGTAAGACTTAACGCGCCCCGACACGGAAACAA
CCTATCAATTCCCACGACTTTGGCAAAGAAAATTCGTCCTATAATCATAAACACATCATTGC
AGCGCACAACAGAGCTGCCGCACTGCAACGGAAACGGCGGAGACGACGCGCTTCTCCCCCCT
GCAGAAAATAGATCTCTACACAGGAAAGACAACAATCGACGGCGCGCTGCGTTGCAAAAAGC
CCGCTCTATTGGCCCGCGTCGAGTTCAGTCTGACGCTTGTCAAATTTGCGGACAGATGGCTG
GATTAAAATAATCCGAAGGAGCCGGAGAGCGCCAGCCAAGCATGATTAGAGAATTTCTACAT
CATATTGAAATAACATGTCTTTTTAAATTGGCTGGCATTCATCTCAAGCGCGGCGCCACACT
TGCTGATAGGGTAGCGCCACATCGAGCCCATCTCAATCGCTCGATCGATAAGGCGCTGCAGA
TCAAGGTTGCTGGCTCTGCCGGCGTCCAATGTGAGCACCAAGGAGGAGACATATGGCATTAA
GCACCGCGACGAAAGCTGCGTCGGACGCACTCGGAGCCAACCGCGCGCCGACAAGTGTAAGC
CCTCAGGAAGTGCATAGGTGGTTGCAAAGCTTCAACTGGGACTTCGCGCAGAACCGGACGAA
ATACCCCACCAAGTACCACATGGCGAACGACACCAAGGAGCAGTTCAAGCTCATCGCCAAGG
AATACGCGCGGATGGAATCGGTCAAGGATGAGCGCCAGTTCGGCACCCTGCTTGACGGATTG
ACCCGCCTTGAAGCTGGAAACCGCGTACATCCGCGCTGGGGCGAGACGATGAAGGTCGCCTC
CAACTTCCTCGAAGTCGGTGAATACAACGCCATCGCGGCGTCCGCGATGTTGTGGGATTCAG
CGAGCGCCGCCGAGCAGAAGAACGGCTATCTCGCGCAGGTCCTCGACGAAATTCGCCACACC
CATCAATGCGGCTTCGTGAATTACTATTTCTCGAAGCACTACCATGATCCGGCCGGCCACAA
TGACGCGCGCCGCACGCGCGCCATCGGACCGCTTTGGAAGGGCATGAAGCGCGTCTTCGCCG
ACGGCTTCATTTCCGGCGACGCGGTCGAATGCTCGGTCAATCTCCAGCTCGTCGGCGAAGCT
TGCTTCACCAACCCGCTGATCGTGGCCGTGACCGAATGGGCTTCGGCGAACGGCGACGAAAT
CACGCCGACGGTGTTTCTCTCGATCGAGACCGACGAGCTGCGCCATATGGCGAACGGCTACC
AGACGGTCGTGTCGATCGCCAATGATCCGGCGGCGCAGAAATATCTCAATACAGATCTCAAC
AATGCGTTCTGGACGCAGCAGAAATATTTCACCCCTGTGCTCGGAATGCTGTTCGAATATGG
CTCGAAATTCAAGGTCGAGCCGTGGGTCAAGACCTGGAATCGTTGGGTCTATGAAGATTGGG
GCGGCATCTGGATTGGTCGTCTCGCAAAATATGGCGTCAATTCCCCGCCGAGCCTGCGCGAC
GCCAAAAAGGACGCCTATTGGGCGCACCACGACCTCTTCCTGTTGGCCTACGCATTGTGGCC
GACCGGCTTCTTCCGGCTGTCGCTGCCGGATGAAGAAGACATGGAATGGTTCGAGGCGAACT
ATCCGGGATGGGACGCGCATTACGGCAAGATCCTGCGCGAATGGAAGGCTCTGGGCTGCGAG
GATCCGAAGAGCGGATTCCTGCCGATCCAGTGGCTCGTCCAGAACGGCCATCAGGTCTATGT
CGACCGCGTGTCTCAGGTGCCGTTCTGCCCGACGCTCGCAAAGTGCTCGGGTTCGCTGAGGG
TGCATGAGTTCAACGGCCAGAAGCACTCCTTCAGCGATGATTGGGGCGAGCGCATGTGGTTG
TCTGAGCCCGAGCGCTACGAGTGCCAGAGCGTCTTCGAGCAATACAGCGGCCGCGAGTTGTC
GGATGTGATCGTCGAAGGACATGGCGTTCGCGCCGATGGCAAGACCCTGATCGGTCAGCCCC
ATGTCGCAGGCAGCAATTTGTGGACCGTCGAAGATCTGAAGCGGGCAAATTGTGTGTTCGCC
GATCCGCTCGCAGGCTTCTAAAAAACAGCTTTCGTCGGATCGGCCCAGGTCGATCCGACGGC
TTCCGATCGACGCCGAACAACATGGCCTCTGTCGCGCCATGAAATAATATCAACGGAGGAAA
CTATCTATGGCTATCGCCTCGACGACCAAGCGCGGGCTAACCGATCCTGACAAAGCAGCGCA
AATTCTCGCCGCCGTTCCCGATCATGAACTCGACACGCAACGCCGGATGAATTATTTCGTGA
CGCCGCGCTGGAAGCGGCTCAGCGAATATGAAATCCTGACGCTCTACACGCAACCCAATCCC
GACTGGATCGCCGGCGGCCTCGATTGGGGCGACTGGACCCAGAAGTTTCATGGCGGCCGGCC
GTCCTGGGGCAACGAAAGCACCGAGCTTCGCACGACCGATTGGTATCGCCACCGCGATCCCG
CGCGCCGCTGGCATGCGCCCTATGTCAAGGACAAAGCGGAAGAGTGGCGCTACACGACGCGT
TTCCTGGAAGGCTATTCGGCCGAAGGAGCTGTGCGCTCGATCGATCCGACATGGCGCGACGA
GATTCTCGATAAATATTGGGGCGCCTTGCTGTTCAGCGAATACGGGCTGTTCAACGCGCATT
CGTCGGTCTCGCGCGACGCGCTGTCCGACACGATCCGTTCGACCGCGACCTTTGCGGCGCTC

GACAAGGTCGACAATGCCCAGATGATTCAACTTGAACGCAATTTTCTGTCGAAGATCGTTCC
GGGCTTTCCGGAGTCGACCGACGGCCCGAAAAAAGTGTGGCTGACCGACCCGATCTACCGCG
GCGCGCGCGAGACCGTGGAAGAAGCCTGGCAGGGAATTCAGGACTTCAATGAAATCCTCTGG
GCCGTGCATGGCGTCTATGATCCGCTGTTCGGTCAATTTGCCCGGCGCGAATTTTTCGGCCG
CCTCGCGGCGCACTATGGCGACGGCCTGACGCCGTTCTTCCTCAATCAAACGCAGACCTACT
TCCAGACCACAAAGGCTGCGATGAGCGATCTGTTCTTCTATTCGCTCGGAGACGATCCGGAA
TTCGGAGACCACAACCGGACCTGGCTTCGCGCCTGGACCAACAAATGGCTGCAGAAGACCGC
CGAGGCCCTCCATGACTTCCTTGGCATATACGCGAAGGTGGACAAAGTTGCGGGCGTTAGCG
ACCCCGCCGCCATCAAGGCGGCCGTCGGCCGGGTCGTGAACGATTGGGTCGAGGATTACGCG
AAGAAAATCGACTTCAAGGTCGACGCCGGTCAGCTCATCGCCAGCATTACTCGTGACGTCAA
GTGAAGGTGTAGGAACATGACAGCAAAGAACGCATATAACGCTGGCATCATGAAGAAATCCG
GCGAGGCTTTCGCGGCGGAGTTTTTCGCCGAGGAAAACCAGGTCGTTCATGAGTCGAATACG
GTGGTCCTCGTGCTCATGAAAAGCGACGAAATCGACGCGATTGTCGAAGACATCATCATGCG
CGAGGAAACGAAGCGCAATCCGACCCTCGTTGTCGAGGATCGCGGCGGCTTCTGGTGGATCA
AGGCCGATGGCAAGATCCAGATCGATACGGAAAAGGCGAGCGACTTGCTCGGGAAGACCTAC
AGCATTTATGATTTTCTTGTGAATGTGTCGTCGACCATCGGTCGCGCCTACACGCTCGGAAA
CACCTTTACGATTACATCCGAGCTGATGGGCCTCGATCGCAAGCTCACCGACGTTTGAGCGA
TCGAAAGGGAGGAACAAGAAATGCCCAACTACAAAATTCATGAAAATCCCGTGCGGTCGGAC
TGGCTCGAAAAAATCGCCGAGCTGAAGTCGGTGAAGGACGCCACAGCCTTCATTCAGGATTT
CCGCAAAAAGAATACCTCGCCGTTCCGCACGAGCTATGCGCTTGACGTCGATTATCTTTTTA
TCGAAGCGAAGATCGAGGAGCGGCTCGCCGTCCTGAAATCGTCGACCTTCTCCGCCGCCGAT
CTCTTCACGAAGGCGACCACGGGGGAGACCGCTCAGGCTGTCTCCGAAGACTGGATCGCCAA
GATTGACGCGGAAAAGGACAAGTTCGCCGCCGAGAAAATCCTGATCACCTTCCGGCAGCTCT
ACAAGCCGCCGGTGCTCCCGGTGAATTTGTTCTTCAAAGTGGATACGTATCTTGGCAGCCGC
CTGATGGAGCTGCGCAATACCGACTACTATGCGGATTCGCTTGACGACTTGCGCAAGAAGCG
CGGCGTCAAGGTGTTGAGGCTCGGCAACGTGGTTTGACATTGCAGTGCGTTTTGGGCGCCGG
CCGGTTTTCCCGCCGGCGCTCGGACTAAATCGGGGAGCGCTTTTGCAAGGGCTGCGATAATG
AACGTTGATCAAAAAGAAGAGAATGTTTTGACGCCCTCCCCGATCGATCGGGGCTCATTCAC
ACAACAAGCGACACGCGTGGAGATCTTCAGCGAAGGCCGATACCGCGCGTTCGTCCAGGATC
TCGAATGCATGTGGCGGTGGGAGATCCATCGGGACGGCGACTTCGTTCAGGAGGGATGTTCG
CTCAGCGAATCGTCTTCCCGGGAAGCTGTTGGCCACGTTGTGTCGTTTTATCAGCACCGCGA
TCGCGGCGACGTCAGCAGAGCAGACGCGCAGTATTAAGCCCTGCGGATCGGAAACGGGCGTT
GCGGATCGCCCCGGGCGGCAGCCTTCGTCCGGCGTTCGAAGGCGGATCGAGCATTCGGGTGG
ACCTGGGGCGTGCACGTTTGCGCAAGAGGCAGAACGCCTGGCCGCCGGCGCGGGGGACGGCG
CTGAACTTTTCTCATGAAACTGATCGCAAACGCGATCCTCCATTCGGGGATTTGGGGACGAA
AACATGTTCAAGGTGCGCGCGATAACTGAAGATCAGCACGATCTGACCTTCGAATGTTCGCC
GAGCGAGGACGTCATCTCTGCGGGTCTGAAGAGGGACGTAATTTTGCTCGCCTCCTGCCGCG
AAGGCGGCTGCGCGACGTGTAAGGCGGAGATTGTCGACGGCGATTACGAACTTGGCGGCTGC
AGCGTGCAGGCCCTCCCGCCGGATGAAGAAGAGGCGGGCGTGGTTCTGCTCTGCCGAACCTT
TCCAAGAAGCGATCTCGTTCTTCAGCTGCCCTACACTTTCGATCGCATTTCCTTTCATAAGG
TCAATACGGACTGGCAGGGCGAGATTGTCGCCGTGGAGCGCATCTCCTCCAATGTGGCGCGG
CTGCAAATCGAGCCGAAAGATCCCGAGACGGGAGCCGCGATCAGCATTCCTTTCGTCCCCGG
CCAATATGTCGATATCGAGATTCCGGGCTCTTCCGTCAGCAGATCCTATTCCATGGCGACGA
CGAGCACGCAGTCCCGGCTTGATTTCCTGATCCGTCTTTTGCCCGACGGGCAGTTTTCAAAT
TTCCTGACGATGGCGGCGAAGCCAGGACTGACCGTCAAACTGCGCGGTCCATTCGGGGCGTT
CAACCTGCGCGAGAATGGCTTTCGGCCTCGCTATTTCGTCGCTGGCGGAACGGGCCTGTCTC
CGGTCCTGTCCATGATCCGCTACATGCAGCAGGAGCAGCACCCGCAGGAAGCGAAGCTTTTC
TTCGGCGTCACCCACCAGCATGAGTTGTTTTATCTGGAAGAGTTGAAGAAGCTGGAGGAATC
GATGCCGAACTTCAGCGCCCATGTCGCGGTGATGCAGCCGGACGGCAATTGGCAGGGCAGCC
GGGGTACTGTCGTCGATGATCTTCTCAAACATCTTGAAGGGACGAAAGCGGCGCCGGATATT
TACATGTGCGGGCCGCCGGGAATGATCGACGCCACCTTCGCGGCTGCAGCAAATTATGGCGT
GCCGAAAGATCACGTCTATGTAGAAAAGTTCCTTGCAACAGGGCAGAATCAGGCCGCAGAGT
AGCAGACTTCCGAGGCCTTTCGCCTCTCAGGATGGCGCTTAGCTATGTACACCGTGGTCCAA
ATGCCGAAACACCTGAATCCGGAGACTATTGCGCCTGAACAGGACGCGGGCCTGTTCGGCGT
CGATCGGCCGGGAGTCGCGTGGGATTGGATGATGCAGACCGGTCAGCCGCCCGAAGGCGACG
ACTGGGTGCGGCCGCAAGTTTCGGAAGCCTGGTACAGATGCATCGAAGAATACAGGCTCTCG
CCGCGCACGAACCTTCTTCGTCCGCATGCGATCATGGATTTCGAGAACGGCGCCCGCGCGAC
CTCGGCGACCAATCTCGACGTCAGAACGGCACTGGCGACCATGGCGTTCAATCTGCAGCCGG
CGCTGCGGGACACAAGCGTCAGCCTGCTTCTGGCCGACAGCGCCGGCACGCTCATTCACGCC
CTCGACGCGGGCTCCAATCTCGGCCCGATGGGGCGTCGCCTGGTGCGGCTCGGCGAGAGCTG
GAACGAGCTCATCATCGGCAACAATGGGTTGGGGACGGCCGCCGTGTTGCGCGAGCCTGTCG
CTTTCGACGGCAAGGAGCATTTCTCGTCCATCCTGCATCCCTTCGCAACGGTTGGCCATCCG
CTTTTCGCGCATGACGGAAGCCTCGTCGCCCTGCTTGGCCTCATTACCGATCAGCGTTCGTC

GGCGCAGACGCTGCTCGGCTTCGTTCGGATCGCCGGCTATCTGGTCGAAACCAATTTGTTCG
AGTGTCAGGCTCCCGGCGCCTTCATGCTGCGAATGCGCCTGAACAATATAAGCGCCGGTCTG
ACCGGCCAGGATTGTCTGCTCGACGGCCTCATCTCACTCGATGAACAGGGCCGGATCGTCGG
CGCGACGCGCACGGGGCTCAATCTGCTCCGCGCCGAGCGTCATTCGGATATTCTTTCCAAGA
AAGTGGAGGCGGTGCTTGGGGTCACGCTTGGCGATCTGCGCGCCTGCGCGCGCCAGGGCGAA
CCGATTGAACTCGAGATCCCACACGGCTGGCGGCTCAAGGCGGAATTCATCGTCAGGCGGCG
CCCCGAGAAGGACATTGTTCAATCCAGCCGCGTCTCTACGGTCAAGGGGGGAAGCCGCGCCG
CCCCGGGCGCGGCGCCGGGCCGCGGCGCCAAGGAAGGAAATGGGGAGGAATGGCGCGACGCC
GTTCTCGAAGCCGCGCAGCAAAAAGCCATCAATCTCCAGAAACAGAAGATACCGATCCTGAT
CACTGGCGAGTCAGGCGTCGGCAAGGATCATTTGGTGCGGATGATTCACGCGCAGGGCATCC
GCAAGGGGCGGCCGCTCGTGCTGGTGAATTGCGCGGCGATCCCGCGCGATCTGATTTCGAGC
GAGCTGTTTGGCTACGTGCCCGGCAGTTTCACGGGCGCGCGCACCGGCGGCAAGGCGGGCAA
GTTCGTCGAAGCGCATACAGGCGTTCTGTTCCTCGACGAAATCGGCGATATGGCGCTGGATC
TCCAGACCGCGCTGCTCTGCGCGCTCGACAGTTCGGAAATCGTGCCGGTCGGCGGCTCGAAG
CCGGTCCGGGTCGACGTTCAAGTCATCGCGGCGACCAACAGCGATCTGCCCGACTGCGTCAG
AAAAGGCGCGTTTCGACGCGATCTCTACTATCGCCTGAACGGCGCGCAATTCTGGTTGCCGC
CGCTGCGGGAGCGTCCCGACAAGCTCGGTTTGATCGAGCATTTGTGGGAGCTCGAACTTAAG
GCGCAGGGAGTCGACGAGCGCAAGACGTTGAGCGCGGAAGTTTGGGATATTTTCGAGCGACA
TCCATGGCCGGGCAATATTCGCGAATTGCTCAATGTGCTGCGCTCCTGTTTGGCCATGACCA
CAGGCCGCGAAACCCAGGTTTCCGATCTCCCGATCGACTTTCTTAAGGAGATGAGCGCTTCG
GCCCAGGGCGACGAAGAGGGAGAGGCCGCGCTGTCGACGCTTGCCAAATGGCGGCCCTCGCT
CGAGGCCAAGGCGCTCGCCGACTGGGAGGCTGACGCGATCCGGTCGGCGCTCGCCAAGACGG
CGGGCAATATTTCGGAATCGGCCCGACAGCTCGGGATCACGCGCGCGACCCTGTACCATAAG
ATGGCGCGGCTCGGTTTGCGTAAATGAGGGACAAGACGGCGCCATAGCTGGGAGAGAAAAAT
GGACGAAGACGTCGAACAATTTCTCGCCGCGATGGCAAAAAGCGCCGACGCCACAACCCATG
AATTGGAGATCGAGGAAGCGGCGCTGCGCGCCCTGATCGGCGGCGCGCGCGCCGAGGAGCTG
GAGCTGCTTTGGACGCAGCAGCTTGATCCCGCCGATGAAGAAGACATCAAGCGCTACATGGA
TTGGAACGACAAAAAGTTGATCTGGATCTGGCGCCGCCTCGAGCGTTCGCGCGAGCGGCGGG
TCGCCGCGGGACGCGCCTATATGATCAACAGCCAGGGCACCGCGGCGCCGGCGATTTCCGCT
CCAGCGAAAGCCGGCGCGCGGCCTCGCAAACCTCTTTAAGCCGGCGTCGCTGCGCCGCGCTT
TTTTGACGCCGCCCTGAGCCGGGCGGGGCTTTTCGCCGGCCAACCAGGCTGCCGGATGGAGC
ATGACGCCGAATAGCATCACGGCTTTCGGACCGGCATCATGCTCTAAACACGAAGGTCGATC
ACGGTCATGGTCTTGGATTTAGGCGAAATCCAAACGCATCGTGGCCCAACAAGGCTTTGCGC
GTCGCCGCGCGAAGAGGCTGTGACGCGTTGCAAGCCGAATACAACACGCGCAACGAACAAGA
ACACAAGCGACGAAAGAGAACGCGGTCGGCGTTCCAGGTGAGGAAACAATGGCAAAAGATAT
TCGCTATGGCGACGCCGCGCGCCGCAGGATGCTCGCGGGCGCGAACCTACTGGCCGATGCGG
TTCAGGTGACTCTCGGACCGCGCGGCCGCAATGTGGTGATCCAGCATCGCACCTCGGGGATT
CTGCCGGTCGTAACCAAGGATGGGGTCACCGTCGCCCGATCGATCGCCGTCGACGATCGATT
CGAAAGCGCCGGAATCAACATGTTCAAGGAAATGGCGGGCCGCGTTTCCAAGGAATGCGGCG
ACGGCACGACGACGACCATCGTGCTCGCCCGCTTCATCGCCCGCAAGGTTCTGCGGGCCATG
AGCTCCGGACTTGATCCGAACGGGCTGCGGACGGGGCTGGAGCTTGCGACGCAGACCGCGGT
CGAGGACCTGAAGCGACGGGCGAAGAGCTGCGCCGATGAGCGCTCGATCGTTCACATCGCCG
CGACCGCGTCGAACGGCGACTTATCTGTCGGCGAGCTGCTGGCCTCCGCCTTCAAAAAGGTC
GGCCCGAACGGCATCGTCAACGTCAGCCTCGGCAATGGGACGAGCGACGAAATCGCGTTTCA
GGAAGGCGCGCATTGGGAACAGGGCTGGCTCTCGCCCTATTTCATGACGGACAAGACGCGCC
GCATTGCAGAGCTGATAAACCCCTACGTGCTCCTCTACGATCGCCCCATCAAACAGTTCGAC
GAGTTGATCCCGATCCTCGATCAGGTGCGGCAGGCTGAAGGATCTCTGTTGATCATCGCCGA
CGACATCGAGGAAGCGGCGCTCGGCGGGATTTTGCTCAACCATATCCGTTGCGTCCTGAAGG
CCGTTGCGGTGAAGCCGCCGGCCTATGGCGACCGCCGCAAGGAGACCCTTGCGGATCTCGCC
TGCCTTCTCGGCGGTCGGGCGATCCTGGAGGACAATGGCGACGAACTGAGCCATGTAAAGCT
CGCCGATCTCGGCCGCGCCAACCGCGCTGAAGTCACCGAGAGTGAGACGACCCTGTTCGGCG
GGGCGGGCGACGCCGACAAGATTGCCGAAAGGGTGCATGCGCTCCGGTTCGAGGCGGAACGT
CTGCAAAAAAACGATCGGGGCTCTCCCACGGGAAAGCTGCACGACCTCGAGGAATTCGACGA
ACGAATCGGCAATCTGTCCAGCGTCACCGCAACAATCCATGTTGGCGGCGGGACGGAAACGG
AGATGAAAGAGCGTCTCCAGCGCGTCGAGAACGCAAGAAACGCCGTCGCCGCGGCGCTAGCC
GAAGGAGGGCTTCCTGGCGGCGGCGCGGGGCTTCTGCGCTGCCGAAAAGCCCTGAGCGGCCT
CTCGTCCACGGATATCGCCGTCAGGCACGGTATTCAGATCATCGCTGACGCGGTGGGCGAAC
CCTTGCGCCGGATCGCCGACAATTCCGGCAAGGACGCGCTGGCCGTGGCGTATGAGACGCTG
GCCTCGAGCGATGAAGGCTTCGGCTGCGATGCGAGAACGGGACGGTTCGGCGATTTATTCGA
GGCCGGCGTGATCGATCCGCTGCGCGTGACGCGCCTTGCCTTGCAGCACGCCACCATCACAG
CGAGCACGCTCATGACGACGGAATGCGTCGTCGCCAATTTGCCGCCGGACGATCCGACATTC
GGCTATACGGGAGAATGGGCGGCGGCGACGCGCGAAGATCCGCGTTTGTGAGATCCATGTGA
CGGCGCGATAAAGCCTGCGCCGCGCCGATATGTTCGCGGAGGCGCAGGCAAACTCGGCGCCG

CGATCATGCAGCAGGGCGCGCGCGATGCGACGGCGCCCTGTTGTGATTCTGCCGGAATGCCA
TTCGCGCTTAATCCAAAGGTCGCCGTCACTGTGCATATCGCCTGCGAACCGCTCGACCGGAC
GCATCGGGTCCCTGTTCAGATGACATACGCTGTATAGAAAAGCGCAGGTCAATTCCCGCTAA
ATGCCTGTAGCAATGTTTTCGTCAACCTTTTTGCCATTGCAACGTATACATCTAACATACTG
TTTAAGCGTGTGTTTTTTAAGATTCCGAGTCGATGTTCCGATTGCTGTTTGATCCTGCCGTC
GACTGGCACGCCTCCTGCGATTGCTTTGTCAGCAAGGCGAACGCGAAATATGCGGGCGCCGA
AATTCCGACAAAAGCGGGAGGCGCTTCAATGAGCAATACCAGGAGTTTTTTGTTTGTCGGGG
CCGCCATCATGGCCGGGCTCCTTTGCCGCGAAGCGAGCGCGGAAGATCCTTCGCGCGCATTC
ATCGACCGCTTGTCCGGGGAGCCGTTCAGAACCAATAGCGCCGGTCCGGTTTCGACCAAGAT
CGAAACTAATCCTGCGCAAGCCTATATTGACAGGCTGGCGGGCGCCACGCAGGTCGCCGAAC
ACCTTGGGCCCGGCACGACGGCGATCACAACCGACACGACACAAGACTTCATCGATCGCGTG
GCTGCGCCACATGCTTTCACCACCCGCGGCGCAAGCACAGGCCGGCAACGCTCCGCAGCTGC
CCATTGAGCCCTTTCTTGAACTGCCCGCATGAAGATCAATCGACCACATGAAGGCCGATCGT
TCGCTACTCCAAAAAATACTTTCAATGCCAAGAAAAAAGAAAAATCAAAGGGACATGCGTCA
TGAACGATTCAAGGAGGCAGTATCTTGCCGGAACTCCGGCCCTATCCAAGGCTCAGATCGGC
GTTTTTTTGAAGAGCTACGGTCTGCGGGCGACTCGTCAGCGTCGAGGCCTCGTGAGGCTGCT
CTTCGGCAAGAACGATCGGCATGTAACAGCCGAATCTTTGGCTAATGAGGCCAGAGCCGCAA
ATACCCACACATCAATGGCCACAGTCTACAACGTTCTCAATCTGTTTGCGCAGGTGGGACTG
ATCCGCAGTCTGGCAATCGAGGGAGACAGGGCGATTTTTGACACGAACACAAGCGACCATTC
TCACTATTTCTTTGAGGATACAAAGGAGGTTCGCGATATCGAATCGAGCGAAATGAACCTCG
CCGCTCATATTGAGGCGCCAGACGGCTACGAGTCAGGTGGATGTTGTCGTGACGTTGAGGCG
AAAGGACGCGACGCCGCCTGCGCGGGCCGGTCGGTGACTCGTTCCAGTAGCGTATCGAATTC
AAATTGAAGCGTAAGTATTCGCCTTCGGTAGATCGACCGGCGCGCTAATTGGCGGCGCTATT
GGAGCCCGAACGGCTCAAACGATGCAGGCCGGTATCATTTCTGACGCTTTCCTTCGGACGTC
CTGCGGAAGCCCCGTCCCGCAGGACGTCTTTTTTTGCGCCTAGCATGATGAGCTGTAAGTTT
GCCCTTTAACGCGATGAGCCGCCTCCATGCGCGTCGCCGGGCCGCTACAATCGGCGAGGTCT
TGAAGTCCGACATATCTCCATTTTAAACACTTTGCTTAAGCAGACGTCGGGCGCCTCCGCGC
AGCCGTTCTCGGCGATGCGCCCCGCGGCGTCGAGGCGGCCGGCACGTCAGGCTCACACCCGA
TTCCCGCATGATGAACCCGGTGCGCCAAAAGCGGACCCACATGCGGTTGCGGATTAATGCTT
TGCGAGACTTAGGATGTAAGCCGCGATGGAATCGGTTTCTTGTTGATCCAATATGATATTCG
GCATGGTCGGATGGGACGACCGAAGGAAGACTTTGATCGCGAGTTCGCTTGTCGATGGCATG
CGGCTGATGTCGACGAAGCTGGGGGCCTTTGAGTTCGCAGGTTGCGGCGTTGGAACGGGGCC
AACCTGATGACAGGCGCTGCAGACGGTCTCGGCGAGCTGCCGCCCGATGGCTGGATCTGCCG
TTTGGGCGTTGGCGTTCGAAAAGGAGACGACTGCGGCGAGAAGCGCCGTGGGTGAGTTGAAT
TTTATCATCGACCAGCTTCTTTTGCTCACTTCGGCTTCCCCTCGTTAAACACACGCGCCGAT
TTTCTTAGGTACGTCCGCTGATTTTGGGCGGATGCCTTTTCGAGCGCTCTCACGCGCGCCGA
GACGCAGCGGCGCGCGAATCCATTTTACATCGCACGGGCAATAATTTCGTTACCAGATTGCG
ATTTAGTCTGGGGGCTGCGGGCTCGGAGATCGTAGCGCCTTTACAAATTCACGGACAGGGGA
ATACTCCCAAGGCCATTCCTGTTCCGGCGGCCGCTGCGGGCTTGGAATTCAGAGACGAAACT
GCCAATCAGGACCGACGGCCTTCCGCCTGCGAAGCCGTTAGCCCGCTGCGGCCTTTCCGCCG
CCGGACGACGCATGAGCCGCAGGGCCGTCGACGCAACAGCGCCCTGACGCTGAATGGCCTCG
CAGACAGCAGGTCTGTATGGCTCTCTGCTTCAGTCTGGACGAGGAGACGAATGGCTACTTTG
CGTTTGTAGGGGCGTGTTCTCCCGCGCCTTCGGGCGGCGGCGTGACTCAGGTTGGCGCGCGA
CGCTTGCACGTGCGGCAAACGAACCAAATTACGCTGATGCGCCGGAGCCGGCCGTTGGCGCC
CGGGACCCCAGACGCCTCACTCGTCATTGCTCGCGCGCCGGCCGCTGGCTGGCGCGCGGACC
ACGCAAGGTCGATTCGCTTAAGCGATCTTGAGCCGATCGGCGTTGGCCGCGACGTGTTCGCG
ATATCGGCCAATGACTTCGTCGAAAGAGCCGCCCGCGATCAAGGTGGACATCGCCTCAAATC
CAGCGTCGATCTTTTCGTTGATCATGAGGCTGGTTTCGTCGCTTGCCGAAGAGCCGCCCATG
GCCAGTTTCAGTAACCGGAGGCCGATCACGCTGTTTGATTCGATCAAAAGTGAGGTGAAGGC
GTGTGTAGACCTAAACATTGAAGCTGCTCCGGGCAGCCTGCATGCTGCAATGAGTGAAGTCT
GAATGTGTCACGAAAGTTCCACTAAAGCACACCCCACGCGTTTCGGCGGCGCGTGGCGTGCT
TGGCAGCGTCGAGCCGCCGGCGCCGGAGCATCGAGAGCGAGGCGCATGGAAAAATGGAGCCG
GCGCGCAAACCGGCTCCATGTTCAAGGCAGGAATTTGCCGCAGCCTGCGCAGCGAGGGGAGG
CAAGTCGAAGGCCGACAGCAAGGTCCAGATATCCAGGATGACAGGTAATCCACCGCCCGAGC
TATAGCCAATATCGACTAGTTCGGCGCGGGCGCCGAGGTCGGGACACGGAGCGCATTCGGGT
CCCAGCTTCAGCCGCCCTTCAGAAAACGATGTGGGTCAGCGAGAATCACATCCCGCACGGCC
TCATGATATTTGACGTAGCCTGTACAGCGGCAAAGGTGACCGTCGAGCGCATCGGCAATGGT
TTTCTCGAGATCGGCGCGGGCCACCGGCGTTCTCGCCAATTGCTCCAGCAGAACCTGACCCT
CGTTGAGAAAACCCGGAGTGCAATAGCCACACTGGAACGCAAAATGCTCGATGAAGGCTTTC
TGCAACACGGTCAGCTGGCCGTTCTCGGCGTGTCCCTCGATGGTGCGGATCGTCTTTCCGTC
AAAGCTT

>Methylocella silvestris BL2 MmoX
MALSTATKAASDALGANRAPTSVSPQEVHRWLQSFNWDFAQNRTKYPTKYHMANDTKEQFKL
IAKEYARMESVKDERQFGTLLDGLTRLEAGNRVHPRWGETMKVASNFLEVGEYNAIAASAML
WDSASAAEQKNGYLAQVLDEIRHTHQCGFVNYYFSKHYHDPAGHNDARRTRAIGPLWKGMKR
VFADGFISGDAVECSVNLQLVGEACFTNPLIVAVTEWASANGDEITPTVFLSIETDELRHMA
NGYQTVVSIANDPAAQKYLNTDLNNAFWTQQKYFTPVLGMLFEYGSKFKVEPWVKTWNRWVY
EDWGGIWIGRLAKYGVNSPPSLRDAKKDAYWAHHDLFLLAYALWPTGFFRLSLPDEEDMEWF
EANYPGWDAHYGKILREWKALGCEDPKSGFLPIQWLVQNGHQVYVDRVSQVPFCPTLAKCSG
SLRVHEFNGQKHSFSDDWGERMWLSEPERYECQSVFEQYSGRELSDVIVEGHGVRADGKTLI
GQPHVAGSNLWTVEDLKRANCVFADPLAGF

>Methylocella silvestris BL2 MmoY
MAIASTTKRGLTDPDKAAQILAAVPDHELDTQRRMNYFVTPRWKRLSEYEILTLYTQPNPDW
IAGGLDWGDWTQKFHGGRPSWGNESTELRTTDWYRHRDPARRWHAPYVKDKAEEWRYTTRFL
EGYSAEGAVRSIDPTWRDEILDKYWGALLFSEYGLFNAHSSVSRDALSDTIRSTATFAALDK
VDNAQMIQLERNFLSKIVPGFPESTDGPKKVWLTDPIYRGARETVEEAWQGIQDFNEILWAV
HGVYDPLFGQFARREFFGRLAAHYGDGLTPFFLNQTQTYFQTTKAAMSDLFFYSLGDDPEFG
DHNRTWLRAWTNKWLQKTAEALHDFLGIYAKVDKVAGVSDPAAIKAAVGRVVNDWVEDYAKK
IDFKVDAGQLIASITRDVK

>Methylocella silvestris BL2 MmoB
MTAKNAYNAGIMKKSGEAFAAEFFAEENQVVHESNTVVLVLMKSDEIDAIVEDIIMREETKR
NPTLVVEDRGGFWWIKADGKIQIDTEKASDLLGKTYSIYDFLVNVSSTIGRAYTLGNTFTIT
SELMGLDRKLTDV

>Methylocella silvestris BL2 MmoZ
MPNYKIHENPVRSDWLEKIAELKSVKDATAFIQDFRKKNTSPFRTSYALDVDYLFIEAKIEE
RLAVLKSSTFSAADLFTKATTGETAQAVSEDWIAKIDAEKDKFAAEKILITFRQLYKPPVLP
VNLFFKVDTYLGSRLMELRNTDYYADSLDDLRKKRGVKVLRLGNVV

>Methylocella silvestris BL2 MmoD
MNVDQKEENVLTPSPIDRGSFTQQATRVEIFSEGRYRAFVQDLECMWRWEIHRDGDFVQEGC
SLSESSSREAVGHVVSFYQHRDRGDVSRADAQY

>Methylocella silvestris BL2 MmoC
MFKVRAITEDQHDLTFECSPSEDVISAGLKRDVILLASCREGGCATCKAEIVDGDYELGGCS
VQALPPDEEEAGVVLLCRTFPRSDLVLQLPYTFDRISFHKVNTDWQGEIVAVERISSNVARL
QIEPKDPETGAAISIPFVPGQYVDIEIPGSSVSRSYSMATTSTQSRLDFLIRLLPDGQFSNF
LTMAAKPGLTVKLRGPFGAFNLRENGFRPRYFVAGGTGLSPVLSMIRYMQQEQHPQEAKLFF
GVTHQHELFYLEELKKLEESMPNFSAHVAVMQPDGNWQGSRGTVVDDLLKHLEGTKAAPDIY
MCGPPGMIDATFAAAANYGVPKDHVYVEKFLATGQNQAAE

>Methylocella silvestris BL2 MmoR
MYTVVQMPKHLNPETIAPEQDAGLFGVDRPGVAWDWMMQTGQPPEGDDWVRPQVSEAWYRCI
EEYRLSPRTNLLRPHAIMDFENGARATSATNLDVRTALATMAFNLQPALRDTSVSLLLADSA
GTLIHALDAGSNLGPMGRRLVRLGESWNELIIGNNGLGTAAVLREPVAFDGKEHFSSILHPF
ATVGHPLFAHDGSLVALLGLITDQRSSAQTLLGFVRIAGYLVETNLFECQAPGAFMLRMRLN
NISAGLTGQDCLLDGLISLDEQGRIVGATRTGLNLLRAERHSDILSKKVEAVLGVTLGDLRA
CARQGEPIELEIPHGWRLKAEFIVRRRPEKDIVQSSRVSTVKGGSRAAPGAAPGRGAKEGNG
EEWRDAVLEAAQQKAINLQKQKIPILITGESGVGKDHLVRMIHAQGIRKGRPLVLVNCAAIP
RDLISSELFGYVPGSFTGARTGGKAGKFVEAHTGVLFLDEIGDMALDLQTALLCALDSSEIV
PVGGSKPVRVDVQVIAATNSDLPDCVRKGAFRRDLYYRLNGAQFWLPPLRERPDKLGLIEHL
WELELKAQGVDERKTLSAEVWDIFERHPWPGNIRELLNVLRSCLAMTTGRETQVSDLPIDFL
KEMSASAQGDEEGEAALSTLAKWRPSLEAKALADWEADAIRSALAKTAGNISESARQLGITR
ATLYHKMARLGLRK

>Methylocella silvestris BL2 MmoR2
MDEDVEQFLAAMAKSADATTHELEIEEAALRALIGGARAEELELLWTQQLDPADEEDIKRYM
DWNDKKLIWIWRRLERSRERRVAAGRAYMINSQGTAAPAISAPAKAGARPRKPL

>Methylocella silvestris BL2 MmoG
MAKDIRYGDAARRRMLAGANLLADAVQVTLGPRGRNVVIQHRTSGILPVVTKDGVTVARSIA
VDDRFESAGINMFKEMAGRVSKECGDGTTTTIVLARFIARKVLRAMSSGLDPNGLRTGLELA

TQTAVEDLKRRAKSCADERSIVHIAATASNGDLSVGELLASAFKKVGPNGIVNVSLGNGTSD
EIAFQEGAHWEQGWLSPYFMTDKTRRIAELINPYVLLYDRPIKQFDELIPILDQVRQAEGSL
LIIADDIEEAALGGILLNHIRCVLKAVAVKPPAYGDRRKETLADLACLLGGRAILEDNGDEL
SHVKLADLGRANRAEVTESETTLFGGAGDADKIAERVHALRFEAERLQKNDRGSPTGKLHDL
EEFDERIGNLSSVTATIHVGGGTETEMKERLQRVENARNAVAAALAEGGLPGGGAGLLRCRK
ALSGLSSTDIAVRHGIQIIADAVGEPLRRIADNSGKDALAVAYETLASSDEGFGCDARTGRF
GDLFEAGVIDPLRVTRLALQHATITASTLMTTECVVANLPPDDPTFGYTGEWAAATREDPRL