Genes involved in the methyl *tert*-butyl ether (MTBE) metabolic pathway of *Mycobacterium austroafricanum* IFP 2012

Nicolas Lopes Ferreira,^{1,2} Diane Labbé,² Frédéric Monot,¹ Françoise Fayolle-Guichard¹ and Charles W. Greer²

Correspondence

Françoise Fayolle-Guichard
francoise.fayolle@ifp.fr

¹Institut Français du Pétrole, 1–4, avenue de Bois-Préau, F-92852 Rueil-Malmaison, France

²Biotechnology Research Institute, NRCC, 6100 Royalmount Avenue, Montreal, Quebec, Canada H4P2R2

Methyl *tert*-butyl ether (MTBE) is a persistent pollutant of surface and groundwater, and the reasons for its low biodegradability are poorly documented. Using one of the rare bacterial strains able to grow in the presence of MTBE, *Mycobacterium austroafricanum* IFP 2012, the protein profiles of crude extracts after growth in the presence of MTBE and glucose were compared by SDS-PAGE. Ten proteins with molecular masses of 67, 64, 63, 55, 50, 27, 24, 17, 14 and 11 kDa were induced after growth in the presence of MTBE. Partial amino acid sequences of N-terminal and internal peptide fragments of the 64 kDa protein were used to design degenerate oligonucleotide primers to amplify total DNA by PCR, yielding a DNA fragment that was used as a probe for cloning. A two-step cloning procedure was performed to obtain a 10 327 bp genomic DNA fragment containing seven ORFs, including a putative regulator, *mpdR*, and four genes, *mpdC*, *orf1*, *mpdB* and *orf2*, in the same cluster. The MpdB protein (64 kDa) was related to a flavoprotein of the glucose–methanol–choline oxidoreductase family, and the MpdC protein (55 kDa) showed a high similarity with NAD(P) aldehyde dehydrogenases. Heterologous expression of these gene products was performed in *Mycobacterium smegmatis* mc2 155. The recombinant strain was able to degrade an intermediate of MTBE biodegradation, 2-methyl 1,2-propanediol, to hydroxyisobutyric acid. This is believed to be the first report of the cloning and characterization of a cluster of genes specifically involved in the MTBE biodegradation pathway of *M. austroafricanum* IFP 2012.

Received 12 October 2005

Revised 11 January 2006

Accepted 16 January 2006

INTRODUCTION

Methyl *tert*-butyl ether (MTBE) has been extensively incorporated into gasoline to meet the requirements of the octane index and to allow more efficient fuel combustion in engines. The removal of leaded fuels and, later, the new standards on the maximum aromatic content of gasoline (42 % from January 2000, and 35 % from January 2005, in the European Union), has led to a decreased octane index of gasoline that could be counteracted by the addition of low levels of MTBE (2–5 %, v/v). The US Clean Air Act recommends the use of oxygenated gasolines containing up to 2.7 % oxygen (i.e. 15 % MTBE, v/v) in metropolitan areas. In the European Union, MTBE is the most frequently used fuel oxygenate, with a mean concentration in gasoline of 2.3 % (w/w).

Abbreviations: CTAB, cetyltrimethylammonium bromide; DEPC, diethyl pyrocarbonate; gDNA, genomic DNA; GMC, glucose–methanol–choline; HIBA, 2-hydroxyisobutyric acid; 2-M1,2-PD, 2-methyl 1,2-propanediol; MTBE, methyl *tert*-butyl ether; NDMA, *N,N'*-dimethylnitrosoaniline; TBA, *tert*-butyl alcohol; TBF, *tert*-butyl formate.

The GenBank/EMBL/DDBJ accession number for the genomic DNA sequence reported in this paper is DQ147773.

The widespread use of MTBE has resulted in accidental releases from leaking storage tanks and spills into ground-water and surface water (Johnson *et al.*, 2000; US Environmental Protection Agency, 1999). The persistence of MTBE in aquifers is attributed to its high aqueous solubility and its recalcitrance to biodegradation under aerobic and anaerobic conditions. Because of the possible toxic effect of MTBE, the US Environmental Protection Agency (1999) has classified this compound as a potential carcinogen and set the maximal threshold of MTBE allowed in drinking water at 20–40 µg l^{−1} (US Environmental Protection Agency, 1997). The environmental and health concerns related to MTBE have led to a progressive phase-out in the US. However, there is a need for remediation technologies when aquifers are impacted by MTBE, but the usual physico-chemical techniques are generally neither efficient nor cost-effective.

The biodegradation potential of MTBE has been investigated and a few strains, such as *Rubrivivax gelatinosus* PM1 (Deeb *et al.*, 2001; Hanson *et al.*, 1999), *Hydrogenophaga flava* ENV735 (Hatzinger *et al.*, 2001) and *Mycobacterium austroafricanum* IFP 2012 (François *et al.*, 2002), have been

isolated. More recently, other strains able to grow in the presence of MTBE have been isolated from different environments: strain UC1, closely related to *R. gelatinosus* PM1, and UC3, closely related to *M. austroafricanum* IFP 2012, were isolated from an MTBE-degrading biofilter (Pruden & Suidan, 2004); strain L108 was isolated from an MTBE-contaminated aquifer in Germany (Rohwerder *et al.*, 2004); and *M. austroafricanum* IFP 2015 was isolated from the drain water of an MTBE-supplemented gasoline storage tank (Lopes Ferreira *et al.*, 2005). The MTBE biodegradation pathway of *M. austroafricanum* IFP 2012 has been partially characterized by identifying several degradation intermediates, such as *tert*-butyl alcohol (TBA), which frequently accumulates in aquifers contaminated by MTBE (François *et al.*, 2002). We recently detected 2-methyl 1,2-propanediol (2-M1,2-PD) as the product of TBA oxidation in *M. austroafricanum* IFP 2012 and IFP 2015 (Lopes Ferreira *et al.*, 2005), and demonstrated its transformation to 2-hydroxyisobutyric acid (HIBA) in resting-cell experiments. 2-Methyl 1,2-propanediol (2-M1,2-PD) has been reported by Steffan *et al.* (1997), using propane-oxidizing bacteria able to degrade MTBE by cometabolism, and by Smith *et al.* (2003), using *Mycobacterium vaccae* JOB5 as the product of TBA degradation.

Using individual bacterial isolates, growth in the presence of MTBE is generally slow and the cellular yield is poor (Fayolle & Monot, 2005; Fortin *et al.*, 2001). Several possible limitations to growth in the presence of MTBE have been suggested by Salanitro (1995). In the case of *M. austroafricanum* IFP 2012, the crucial role of cobalt divalent cations in the biodegradation of HIBA, one of the MTBE degradation intermediates, has been shown (François *et al.*, 2002). François *et al.* (2003) also demonstrated that *tert*-butyl formate (TBF), the intermediate between MTBE and TBA, exerted inhibitory effects on MTBE and TBA oxidation.

Inefficient MTBE biodegradation could also be the result of a loss of energy (Fortin *et al.*, 2001), probably related to a lack of reduced co-enzymes and protons, thus de-coupling the proton pump and leading to a shortage of ATP. Also, the MBTE pathway has an unusually high requirement in reducing equivalents (Salanitro, 1995). The characterization of the enzymes involved in the MTBE degradation pathway is crucial to understand the limitations to MTBE degradation, but it has not yet been documented. To the best of our knowledge, we are the first to describe the isolation and characterization of a gene cluster, designated the *mpd* cluster, containing genes encoding an alcohol dehydrogenase and an aldehyde dehydrogenase, which are both involved in the MTBE metabolic pathway of *M. austroafricanum* IFP 2012. Functional heterologous expression was carried out in *Mycobacterium smegmatis* mc2 155 and the involvement of these enzymes in the specific degradation of 2-M1,2-PD to HIBA is demonstrated.

METHODS

Micro-organisms, growth media and culture conditions. All the micro-organisms used in this study are listed in Table 1. *M. austroafricanum* IFP 2012 was grown at 30 °C in Luria–Bertani (LB) medium, or in the mineral medium (MM) previously described (Piveteau *et al.*, 2001), supplemented with MTBE, TBA, 2-M1,2-PD, HIBA or glucose as the sole carbon source. The headspace volume was sufficient to prevent any oxygen limitation during growth. Growth was followed by measuring the OD₆₀₀. The substrate or metabolite concentrations were measured using filtered samples by GC or HPLC. *M. smegmatis* mc2 155, *Escherichia coli* DH10B and respective recombinant clones were grown in LB supplemented with kanamycin (20 µg ml⁻¹) or ampicillin (100 µg ml⁻¹), as required.

Preparation of crude cellular extracts and SDS-PAGE analysis. *M. austroafricanum* IFP 2012 was grown to an OD₆₀₀ of approximately 1·0 in MM medium supplemented with either MTBE or glucose. Cells were harvested by centrifugation at 8000 g for

Table 1. Bacterial strains and plasmids

Strain or plasmid	Description	Reference or source
Strains		
<i>M. austroafricanum</i> IFP 2012	Wild-type; MTBE ⁺ , ETBE ⁺ , TAME ⁺ , methanol ⁺	François <i>et al.</i> (2002)
<i>M. smegmatis</i> mc2 155	ept-1, highly transformable mutant	Snapper <i>et al.</i> (1990)
<i>E. coli</i> DH10B	Host for recombinant plasmid	Grant <i>et al.</i> (1990)
Plasmids		
pBluescript II KS+	Ap ^R	Stratagene
pCL4D	<i>Mycobacterium</i> – <i>E. coli</i> shuttle expression vector, Km ^R	Picardeau <i>et al.</i> (2000)
pKS1	5676 bp <i>Sma</i> I <i>M. austroafricanum</i> IFP 2012 genomic DNA fragment cloned in pBluescript II KS+ <i>Sma</i> I	This study
pKS2	4411 bp <i>Not</i> I fragment from pKS1 cloned in pBluescript II KS+ <i>Not</i> I	This study
pKS3	5574 bp <i>Pst</i> I fragment of <i>M. austroafricanum</i> IFP 2012 genomic DNA cloned in pBluescript II KS+ <i>Pst</i> I	This study
p4D1	4411 bp <i>Not</i> I filled-in fragment from pKS2 cloned in pCL4D <i>Hinc</i> II	This study
p4D2	5574 bp <i>Pst</i> I fragment from pKS3 cloned in p4D1 <i>Pst</i> I to generate a 9060 bp insert containing the functional <i>mpd</i> gene cluster	This study

15 min at 4 °C, washed twice in phosphate buffer (20 mM, pH 7.0) and resuspended in 5 ml Tris DTT buffer (50 mM Tris/HCl, 0.1 M DTT, pH 8.0). Cells were broken by three passages through a French press (138 000 kPa) and kept on ice at all times. Cell debris was removed by two centrifugation steps at 10 000 *g* for 2 min and the supernatant was used for electrophoresis. The total protein concentration was determined using the Bio-Rad protein assay. Electrophoresis (12.5 µg total protein was loaded in each well) in 10 and 15 % polyacrylamide gels containing SDS was carried out using the standard method (Laemmli, 1970), except for the migration conditions (150 V or 7.5 V cm⁻¹) during the first hour, and a constant current of 33 mA for the last 5 h.

Peptide sequencing of proteins induced during growth in the presence of MTBE. For internal sequencing, MTBE-induced protein bands from the MTBE profile on SDS-PAGE were cut out. An 'in-gel' reduction (DTT) and an alkylation by iodoacetamide of the disulfide groups (Jenö *et al.*, 1995) were performed prior to modified sequencing grade trypsin digestion (Hellman *et al.*, 1995) (Promega). Edman degradation was performed according to Hewick *et al.* (1981). The N-terminal sequences of the isolated proteins, previously transferred onto a Problott membrane (Applied Biosystems), were obtained by automated Edman degradation, using the sequencing protocol described above. For internal peptide sequencing, about 1 pmol of the peptide was loaded onto a model 494 cLC Procise sequencer, employing the general protocol (Hewick *et al.*, 1981), and a standard program using liquid-phase trifluoroacetic acid was employed for further sequencing.

Design of degenerate oligonucleotide primers. Degenerate primers were designed from N-terminal or internal peptide sequences of specific MTBE-induced proteins. The deduced DNA sequence was obtained according to universal codon usage and primers were created in the least degenerate part of each DNA sequence. Correlations between genome composition (in terms of G + C content) and usage of particular codons and amino acids have been widely reported and Giegel *et al.* (1990) demonstrated that the presence of an adenine at the third position in codons is rare in coding sequences of mycobacteria. When adenine was the base at this position, it was randomly

replaced with one of the other three nucleotides to reduce primer degeneracy. Annealing temperatures used for PCR experiments were calculated at 2 °C lower than the lowest melting temperature found in the primer pool.

Total DNA extraction. Total DNA was isolated from *M. austroafricanum* IFP 2012 according to Pospiech & Neumann (1995), with the following modifications: a 50 ml culture grown in LB medium to an OD₆₀₀ of 0.7 was centrifuged (8000 *g*, 4 °C, 15 min) and resuspended in 5 ml SET buffer (75 mM NaCl, 25 mM EDTA, 20 mM Tris/HCl, pH 8.0). Lysozyme (final concentration 1 mg ml⁻¹) prepared in 10 mM Tris/HCl, pH 8.0, and 500 µl lysostaphin at 50 µg ml⁻¹ (Sigma), prepared in 10 mM Tris/HCl, pH 7.5, were added to the cell suspension before incubation at 37 °C for 1 h. After SDS/proteinase K treatment, an extraction step with NaCl/cetyltrimethylammonium bromide (CTAB) replaced the addition of 0.33 vols 5 M NaCl; 0.25 vols 5 M NaCl (1 M final concn) and 0.13 vols 10 % CTAB/0.7 M NaCl (pre-heated to 65 °C) were added sequentially and the mixture was incubated for 10 min at 65 °C. Then, 1 vol. CHCl₃ was added and the subsequent steps were carried out according to the original protocol.

Cloning and sequence analysis of the *M. austroafricanum* IFP 2012 *mpd* gene cluster. The plasmids used for the cloning are described in Table 1 and the primers pairs used at the different steps are listed in Table 2. Southern blot analysis was carried out on total DNA of *M. austroafricanum* IFP 2012 restricted with different enzymes and probed with a 204 bp PCR fragment (*mpd* probe) labelled with digoxigenin-dUTP (DIG DNA Labelling and Detection Kit; Roche Diagnostics). The size of the *Sma*I-positive band was estimated to be between 5 and 6 kbp; this region of the gel was excised, the DNA was purified and then used to construct a DNA library in pBluescript II KS+ (Stratagene) cut with *Sma*I. For transformation, *E. coli* DH10B cells were treated chemically according to Hanahan *et al.* (1991) and transformants were screened by colony hybridization using the *mpdB* probe. Plasmid (pKS1) from one probe-positive transformant was extracted by a mini-prep procedure (QIAprep Spin Miniprep kit; Qiagen). The presence of the *Sma*I insert was checked with a new digestion with *Sma*I and this revealed a unique

Table 2. Degenerate primers and primers used for PCR and RT-PCR

Primer	Sequence (5'–3')	Amplified region
<i>mpd</i> probes		
MadF1	GGNTGGGCNTAYGAYCC	<i>mpdB</i> internal
MadR1	GCRTCRTTYTCRTCSTAST	<i>mpdB</i> internal
MadF2	GCGGBGGBWCSDCSGG	<i>Orf1</i> internal
MF1	TTCACCTTGTTGGAACCGCTGGG	<i>mpdC</i> internal
MR1	TCATTACCGAGCCGACCTGC	<i>mpdC</i> internal
p4D plasmid constructions		
F-M2	TGAGAAGCCTCGTGATTAC	<i>orf2-mpdB</i>
R-M2	GAGATAAGGCGTGGTGAA	<i>orf2-mpdB</i>
F-M3	AGTCGCTGCCAGGAATCCCAGAT	<i>mpdR-mpdC-orf1</i>
R-M3	CGCTTTGCTTGTTGGCCTCCGT	<i>mpdR-mpdC-orf1</i>
RT-PCR analysis		
RT-FI	AGTGACGGCACCCATAAGTG	<i>mpdR</i> internal
RT-RI	TCGAGGTGTTGAGGTCCGAAT	<i>mpdR</i> internal
RT-FII	GCAGGTCGGCTCGGTAATGA	<i>mpdC-orf1</i>
RT-RII	GTAATACACGAGGCTTCTCA	<i>mpdC-orf1</i>
RT-FIII	ACGGTCTCGTCGGCAAATAC	<i>mpdB-orf2</i>
RT-RIII	GCACATCCCAGGTCTGAT	<i>mpdB-orf2</i>

band corresponding to the size of the linearized plasmid. This could be explained by the presence of a damaged DNA insert containing only one *SmaI* site.

To sequence the 5.676 kbp insert, which was unstable in the pKS1 plasmid, a 10 µg digestion of the plasmid by *SmaI*, followed by phenol extraction of the resulting enriched linear plasmid, was performed for each sequencing reaction (500 bp of DNA sequence per reaction); the band was cut from a 0.7% low-melting-point gel (Seaplaque; FMC) and warmed (65 °C) to completely melt the agarose. One volume of Tris-equilibrated phenol, pH 8.0, was added, the mixture was vortexed, then centrifuged (13 000 g, 20 °C, 5 min) and the supernatant was recovered. Two additional phenol extractions were carried out as described above and these included re-extraction of the previous phenol extraction with 0.5 vols TE buffer (10:0.1, 10 mM Tris/HCl, 0.1 mM EDTA, pH 7.5). Supernatants collected after the extraction and re-extraction steps were pooled and precipitated at -20 °C by the addition of 6.4 µl per 100 µl 5 M NaCl, 1 µl per 450 µl glycogen (20 mg ml⁻¹) and 2.5 vols pure ethanol, centrifuged (13 000 g, 4 °C, 15 min), washed twice with 70% ethanol and resuspended in the appropriate volume of TE buffer (10:0.1) for sequencing.

To obtain the upstream *mpdC* gene sequence, a 904 bp PCR fragment, spanning part of the *mpdC* gene obtained by PCR using primers MF1 and MR1, was used as a probe (*mpdC* probe) for Southern hybridization to identify the 5574 bp *PstI* fragment. The 5–6 kbp *PstI* fragments were cloned in pBluescript II KS+, as described above, to obtain plasmid pKS3. The entire inserts of the resulting positive recombinant plasmids were sequenced on both strands by primer walking with a T7 DNA-sequencing kit (Applied Biosystems), using the ABI Prism 377 automated fluorescence sequencer (Applied Biosystems). To minimize artifacts that are often associated with high G + C content sequences, the following modifications were incorporated. Prior to amplification and the addition of the Big Dye Terminator, the DNA was denatured at 98 °C for 5 min. A mixture of Big Dye Terminator v3.1 and dGTP was also used in a ratio of 3:1. The nucleotide sequences, the ORFs determined by the ORF finder software (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>), and the corresponding amino acid sequences were compared with the EMBL, SWISS-PROT and GenBank databases using BLASTN and BLASTX from the National Center for Biotechnology Information (Altschul *et al.*, 1997).

Construction of an *mpd* gene expression vector and transformation in *M. smegmatis* mc2 155. The *mpd* gene cluster was introduced into pCL4D (Picardeau *et al.*, 2000) using a two-step procedure. In the first step, the 4411 bp *NotI* fragment of pKS2 (Table 1) was blunt-ended with T4 DNA polymerase (NEB) and cloned into the *HincII* site of pCL4D. Transformation was carried out in *E. coli* DH10B and the transformants were selected on LB agar containing kanamycin (20 µg ml⁻¹). The screening was performed by PCR amplification with primers F-M2 and R-M2 (Table 2, Fig. 2) to identify the transformants harbouring the positive recombinant plasmid p4D1. In the second step, the 5574 bp *PstI* fragment of pKS3, carrying *orf3*, *mpdR* and part of the *mpdC* gene, was introduced into p4D1 cut with *PstI*, from which a 950 bp fragment was removed to finally generate p4D2 (Fig. 2). The orientation of the *PstI* fragment in the recombinant plasmid was also screened by PCR using primers F-M3 and R-M3 (Table 2, Fig. 2) to isolate the final p4D2 recombinant plasmid carrying the entire *mpd* gene cluster. The *mpd* gene cluster inserted into the p4D2 plasmid was introduced by electroporation into *M. smegmatis* mc2 155 as follows. Cells of *M. smegmatis* mc2 155 were prepared for electroporation according to Parish & Stoker (1998), with some modifications: *M. smegmatis* mc2 155 was grown at 37 °C to exponential phase (OD₆₀₀ ~0.7) in LB medium supplemented with 0.1% Tween 80 to reduce formation of aggregates. Pellets were washed five times, sequentially, with 1, 0.5, 0.05 and 0.01 vols washing buffer (10%, v/v, glycerol/

0.1% Tween 80), resuspended in 4 ml 10% glycerol solution and 80 µl aliquots were stored at -80 °C. The competent cells (80 µl) were mixed with 1 µg pCL4D or p4D2 in a 0.2 cm gap cuvette. The electroporation conditions were 2.5 kV, 25 µF and 200 V in a Gene Pulser (Bio-Rad Laboratories). Electroporated cells were recovered in 0.4 ml SOC liquid medium (Ausubel *et al.*, 1990) and incubated for 2 h at 37 °C before plating. A transformant, *M. smegmatis* mc2 155(p4D2), was isolated on LB medium containing kanamycin (20 µg ml⁻¹). *M. smegmatis* mc2 155 was also electroporated with pCL4D without any insert and the resulting transformant *M. smegmatis* mc2 155(pCL4D) was used in control experiments.

Functional expression of the *mpd* genes in *M. smegmatis* mc2 155 resting cells. The strains *M. smegmatis* mc2 155(p4D2) and *M. smegmatis* mc2 155(pCL4D) were maintained in LB medium containing 20 µg kanamycin ml⁻¹. The two strains were grown for 72 h at 30 °C in 200 ml LB medium containing 20 µg kanamycin ml⁻¹. Cells were harvested by centrifugation (13 000 g for 15 min), washed twice and resuspended in 40 ml phosphate buffer (20 mM, pH 7.0) containing the test substrate (MTBE, TBA, 2-M1,2-PD or HIBA) in 160 ml sealed flasks. After inoculation, the flasks were incubated at 37 °C on an orbital shaker. When required, chloramphenicol was added from an aqueous solution sterilized by filtration (0.22 µm) to a final concentration of 400 mg l⁻¹. Filtered samples were analysed by GC or HPLC. Substrate degradation was followed over a 24 h period. Specific activities (µmol of substrate or product per min per g dry cell weight) were calculated from the maximal degradation or production rates.

RNA extraction. All glassware and solutions were made RNase-free according to standard procedures (Sambrook & Russell, 2001). Total RNA of *M. austroafricanum* IFP 2012 was extracted from a 20 ml culture grown to an OD₆₀₀ ~0.7 in MM medium supplemented with the required substrate. After incubation on ice for 30 min, cells were harvested (8000 g, 10 min, 4 °C) and kept on ice. Lysozyme (Roche Diagnostics) (600 µl of 3 mg ml⁻¹), prepared in 10 mM Tris/HCl, pH 8.0, and lysostaphin (Sigma) (600 µl of 50 µg ml⁻¹), prepared in 10 mM Tris/HCl, pH 7.5, were added to the cell pellet and incubated at 37 °C for 10 min. RNAwiz (10 ml) (Ambion) was added and the mixture was vigorously mixed for 15 s in a vortex mixer equipped with a vortex adapter (Ambion). The resulting mixture was divided into 2 ml tubes (1400 µl per tube) containing 250 mg zirconium-silica beads (0.1 mm diameter) and mixed for 10 min on the vortex mixer. After centrifugation (13 000 g, 4 °C, 5 min), the bacterial lysates were recovered from each tube, transferred to a clean 2 ml tube and 0.2 vols CHCl₃ was added. The tubes were then vortexed for 30 s, incubated at room temperature (10 min), centrifuged (13 000 g, 4 °C, 5 min) and the aqueous supernatants were transferred to another 2 ml tube. RNA extracts were precipitated by sequentially adding 0.5 vols H₂O treated with diethyl pyrocarbonate (DEPC), 1 vol. 2-propanol and 0.02 vols RNase-free glycogen at 5 mg ml⁻¹ (Ambion), mixed well and incubated at room temperature for 10 min. The RNA was pelleted (13 000 g, 4 °C, 15 min), washed twice with 70% ethanol, dried for 30 s (DNA Speed Vac; Savant) and resuspended in sterile tubes in 25 µl DEPC-treated H₂O. The quality of the pooled RNA was assessed by agarose gel electrophoresis (1% agarose gel in 0.5 × TBE). RNA was stored at -80 °C, precipitated with 0.1 vols 3 M sodium acetate, pH 7.0, and 2.5 vols 95% ethanol.

DNase treatment and RT-PCR. A 2 µg quantity of RNA extracted from *M. austroafricanum* IFP 2012 was treated with DNA-free (Ambion), according to the manufacturer's instructions. DNA removal was assessed by PCR using primers specific for the *M. austroafricanum* IFP 2012 *mpd* gene cluster (Table 2). The final concentration of RNA was quantified at 260 nm using NanoDrop (NanoDrop Technologies). RT-PCR was performed using the Qiagen OneStep RT-PCR kit, with 120 ng RNA. A positive PCR

control reaction, i.e. with 100 ng *M. austroafricanum* IFP 2012 total DNA, and a negative RT-PCR control for each set of primers used on every RNA preparation, were also prepared. The three RT-PCR primer pairs tested were: RT-PCR FI/RI, RT-PCR FII/RII and RT-PCR FIII/RIII (Table 2).

RT-PCR conditions were as follows. Reverse transcription step: 50 °C, 30 min. Positive, negative and (–) RT-PCR controls were then added, at 50 °C, before PCR amplification cycles. PCR amplification step: 95 °C, 15 min; 35 cycles of 94 °C, 30 s; 55 °C, 1 min; 72 °C, 1 min, followed by a final extension at 72 °C for 10 min. Samples were then stored at 4 °C until they were analysed.

Analytical procedures. MTBE, TBA and 2-M1,2-PD were quantified by flame-ionization detection on a Varian 3300 gas chromatograph (Varian) equipped with a 0.32 mm × 25 m Porabond-Q capillary column (J&W Scientific), using a gradient ranging from 105 to 210 °C at 10 °C min^{−1}, followed by a steady temperature of 210 °C for 20 min. Helium (1.6 ml min^{−1}) was used as the carrier gas. Samples, filtered through 0.22 µm filters (Prolabo), were injected without further treatment. HIBA was quantified in filtered samples by HPLC (Metrohm), using an anion column (Dual 2, 75 mm × 4.6 mm) and a conductivity detector (Metrohm 732). The eluent, a mixture of 1.8 mM Na₂CO₃ and 1.7 mM NaHCO₃, was previously flushed by a Spectra-Physics SCM 400 vacuum flusher and the flow rate was 1 ml min^{−1}.

Chemicals. MTBE, TBA and HIBA were of the highest purity available (Sigma-Aldrich) and 2-M1,2-PD was a kind gift from Dr Sergio Revah (Universidad Iztapalpa, Mexico City, Mexico).

RESULTS

Protein profiles of *M. austroafricanum* IFP 2012 and identification of proteins induced in the presence of MTBE

After growth of *M. austroafricanum* in medium containing MTBE or glucose, crude cytoplasmic extracts were prepared and analysed by SDS-PAGE using 10 and 15 % polyacrylamide concentrations and the protein profiles are shown in Fig. 1(a) and (b). Five overexpressed proteins, with molecular masses of 67, 64, 62, 55 and 50 kDa, were detected on the 10 % polyacrylamide gel by comparing the protein patterns after growth in the presence of MTBE and glucose. Five other induced proteins, with molecular masses of 27, 23, 17, 14 and 11 kDa, were detected on the 15 % polyacrylamide gel by comparing the protein patterns after growth in media containing MTBE and glucose.

Fig. 1. SDS-PAGE profiles of cellular crude extracts from *M. austroafricanum* IFP 2012 grown in the presence of glucose and MTBE; (a) 15 % acrylamide gel, (b) 10 % acrylamide gel. Identification of the induced proteins, performed by comparing N-terminal and internal peptide sequences to the deduced amino acid sequences of the isolated gDNA fragment of *M. austroafricanum* IFP 2012 as described in the text, is indicated on the right of each gel. The migration of molecular size markers is indicated on the left of each gel.

Similar protein profiles were observed after growth in medium containing TBF or TBA (results not shown), with one difference: a 23 kDa protein, which was induced only after growth in medium containing MTBE or TBF.

N-terminal and internal protein sequencing was carried out on induced proteins recovered from MTBE profiles, as described in Methods. Peptide sequences obtained were analysed using 'BLAST search for short, nearly exact matches' (Altschul *et al.*, 1997). The N-terminal sequences of the 62 and 11 kDa bands KIIAYDEEAR and ASVNIKPLEDKILV were 95 % similar to the molecular chaperones of mycobacteria GroEL and GroES, respectively. A band, KNP-LFLDEQLTRAEFQ, corresponding to an internal sequence of the 67 kDa protein, was homologous to the DnaK molecular chaperone of *Mycobacterium tuberculosis*. Two internal sequences of 16 and 30 aa, KQRGWAYDPNVRLPE and STEHGLEGTTIDWPISYEELAPYYDENDAIY (François *et al.*, 2002), belonging to the 64 kDa protein, were obtained. The GenBank homology search revealed that the 30 aa sequence had a significant similarity (93 %) with several enzymes of the (glucose-methanol-choline) GMC oxidoreductase family (Cavener, 1992), which includes oxidases and dehydrogenases.

The N-terminal sequences of the 23 kDa protein (SRT-DILAATQAAGAPIEQVV) and the 17 kDa protein (NX-VAVQQDTREAVGE) had similarities with a ribulose-1, 5-bisphosphate carboxylase (or Rubisco) of *Chloromonas* spp. (63 % similarity) and the pyrene dioxygenase small subunit of *Mycobacterium* sp. PYR-1 (80 % similarity), respectively. An internal amino acid sequence from the 50 kDa induced protein (KSFHPFPRAMMGPGAHEMVG-PEALK; François *et al.*, 2002) matched well (61 % similarity) with the partial sequence of a protein from *Mycobacterium gastri*, known as methanol: *N,N'*-dimethylnitrosoaniline (NDMA) oxidoreductase (Bystrykh *et al.*, 1993).

Peptide sequences obtained from the other MTBE-induced proteins did not show any significant similarity with known enzymes.

Amplification of a DNA sequence of the 64 kDa protein (*mpdB* gene)

The 16 and 30 aa sequences obtained from the 64 kDa protein were used to design the degenerate oligonucleotide primers MadF1 and MadR1 (Table 2). A 204 bp PCR product, designated the *mpd* probe, was obtained and sequenced. The peptide sequence correlated with the deduced amino acid sequence of the PCR product, which revealed a partial ORF corresponding to 83 aa. BLAST analysis of this new peptide sequence confirmed the similarity of the 64 kDa protein with flavoproteins of the GMC oxidoreductase family, which includes a variety of enzymes, such as choline dehydrogenase (CDH) and methanol oxidase (MOX). In the case of the 64 kDa protein, the best alignment was obtained (61 % similarity) with a putative choline dehydrogenase from *Bradyrhizobium japonicum* USDA110. The amino acid

sequence of this choline dehydrogenase was investigated to detect conserved domains that could be used to obtain a larger DNA sequence. We were able to detect the Walker-type-A motif GXGXXG, which was first detected by Cavener (1992), described by Wierenga *et al.* (1985) and now known to be an ATP/GTP binding motif. This motif was used to design a new forward degenerate primer MadF2 (Table 2). PCR amplification yielded an unexpected 604 bp DNA fragment, resulting from the new primer annealing upstream of the expected annealing position. The PCR fragment, designated the *mpdB* probe, was subsequently used for colony hybridizations.

Cloning of the *M. austroafricanum* IFP 2012 genomic DNA (gDNA) fragment containing the *mpd* gene cluster

Southern blot analysis of *Sma*I-digested total DNA of *M. austroafricanum* IFP 2012, using the 204 bp *mpdB* probe, revealed a 6 kbp probe-positive fragment (data not shown). To clone the *mpdB*-positive fragment, a DNA library was constructed in pBluescript II KS+ and screened by colony hybridization using the *mpdB* probe. The plasmid pKS1 was isolated from a positive transformant. All plasmid preparations of purified clones showed several bands on agarose gels. For each gel, we observed a major band of 3 kbp, corresponding to the empty linearized plasmid, and minor bands below the expected size, suggesting that the insert introduced into pBluescript II KS+ was highly unstable. The sequencing was carried out after gel extraction of the *Sma*I-linearized pKS1. Preliminary sequence analysis of the insert in pKS1 revealed that this 5676 bp gDNA fragment did not contain the 5' upstream part of the *mpdC* gene. Further Southern blot and colony hybridization analyses were then performed to obtain a 5574 bp *Pst*I fragment, using a 904 bp *mpdC* probe. This probe was obtained using the MF1/MR1 primer pair designed in the *mpdC* gene fragment included in the *Sma*I insert (see Fig. 2 for localization). The pKS3 plasmid containing the 5574 bp gDNA fragment was isolated from a transformant showing positive hybridization.

Analysis of the 10 327 bp *Pst*I-*Sma*I gDNA fragment

Compilation of the *Sma*I and *Pst*I inserts, both containing the *mpdC* probe, resulted in a 10 327 bp gDNA sequence of *M. austroafricanum* IFP 2012. The DNA and deduced amino acid sequences of this overlapping *Pst*I-*Sma*I fragment revealed seven ORFs. We identified six complete genes: *orf3*, *mpdC* (encoding a hydroxyisobutyraldehyde dehydrogenase), *orf1* (encoding an MTBE-induced polypeptide), *mpdB* (encoding a 2-M1,2-PD dehydrogenase), *orf2* (encoding a membrane protein) and *mpdR* (encoding a putative regulatory protein). The genes were all in the same orientation, except *mpdR*, which was located between *orf3* (one transmembrane domain protein) and *mpdC* (Fig. 2). BLAST comparisons of each putative coding sequence with known proteins are presented in Table 3.

Fig. 2. Physical map of gene organization and predicted *mpd* gene products from the 10327 bp cloned gDNA fragment from *M. austroafricanum* IFP 2012. The open bar and the hatched bar below the restriction map represent the *mpdC* and *mpdB* probes, respectively. The horizontal lines below the ORFs indicate the expected intergenic or internal segments of the *mpd* gene cluster obtained by primer pairs I, II and III used for RT-PCR experiments (the primers are described in Table 2). Restriction site abbreviations: P, *Pst*I; S, *Sma*I; N, *Not*I. The different gDNA fragments obtained from cloning experiments, and the different constructions used for heterologous expression analyses, are represented by filled bars, with the respective sizes given below the ORFs. The names of the recombinant plasmids are indicated on the right of each segment.

The *mpdR* gene of the cluster encoded a 47 kDa protein which, based on amino acid sequence alignment, exhibited similarity (43 % identity/58 % similarity) with a putative positive σ^{54} -dependent transcriptional regulator, but only in the N-terminal moiety of the protein.

The *mpdC* gene, encoding a 55 kDa protein, had a similar size and the highest amino acid sequence identity (45 % identity/61 % similarity) with an NAD-dependent aldehyde dehydrogenase of *Burkholderia fungorum*. The CD-search analysis revealed a strong alignment with the 'pfam00171.11 aldedh' domain (E value = $1e^{-123}$), which has been described in the Pfam database of protein families (Bateman *et al.*, 2002), and corresponds to an NADP-dependent aldehyde dehydrogenase family. To predict the secondary structure of the MpdC protein sequence, we used the two prediction servers PSIPRED (McGuffin *et al.*, 2000) and SSPro (Pollastri *et al.*, 2002). Both of these servers are known for their comparatively good performance of over 75 % accuracy on average (Rost & Eyrich, 2001), although they are based on different computational methods. Comparison of MpdC with the well-described structure of the betaine aldehyde dehydrogenase (BetB) of cod liver (Johansson *et al.*, 1998) revealed the presence of 27 of the 33 secondary structures previously determined in BetB of cod liver (Fig. 3a).

The next ORF sequence, *orf1*, encoding a 27 kDa protein, did not show any significant alignment in either the protein or DNA databases. However, the N-terminal (PDEDA-LSAPA) and internal (KAKHFPYLEEDKLKSA) sequences, obtained from the protein pattern induced after growth in the presence of MTBE, corresponded to the deduced amino acid sequence of the *orf1* gene product. This confirmed that the *orf1* gene product was specifically produced during growth of *M. austroafricanum* IFP 2012 in the presence of MTBE.

The *mpdB* gene, located immediately downstream of *orf1*, encoded a 64 kDa protein. Primary DNA sequence analysis correlated with the N-terminal peptide sequence obtained (MTTSADQTDVLVIG; this study) and confirmed the presence of a GTG start codon. This start codon is not frequently found, but has been described in coding sequences of mycobacteria (Hu & Coates, 1999; Kamalakannan *et al.*, 2002; Misra *et al.*, 1996; Revel-Viravau *et al.*, 1996). The amino acid sequence, which includes those of the peptide fragments directly sequenced from the 64 kDa protein induced in the presence of MTBE, had the best alignment (42 % identity/56 % similarity) with a putative GMC oxidoreductase of *B. japonicum*. The CD-search analysis revealed a strong alignment with a cluster of orthologous domains (COG; Tatusov *et al.*, 1997): the 'COG23031 BetA' (E value = $4e^{-49}$). The amino acid consensus sequence of COG2303 characterizes choline dehydrogenases and related flavoproteins, which are involved in the metabolism of choline. The alignment result of the most conserved region of MpdB with BetA of *E. coli* (Lamark *et al.*, 1991), *Sinorhizobium meliloti* (Pocard *et al.*, 1997) and COG2303 is shown in Fig. 3(b). We were able to detect the four highly conserved representative regions of the FAD-binding domain of choline dehydrogenases (Kiess *et al.*, 1998). This FAD-binding domain contains the Walker-type-A motif previously detected and a characteristic β -sheet domain which was revealed in MpdB by the secondary structure prediction.

The last ORF of the *mpd* cluster, *orf2*, encoded a 25 kDa protein, most closely related to the 'COG3104' domain (37 % identity/48 % similarity), including a putative di/tri-peptide permease of *R. gelatinosus* PM1. Hydropathy analysis of the protein revealed the presence of five transmembrane peptide segments, suggesting that it is located in the membrane. This protein was not detected by SDS-PAGE of the

Table 3. Putative *mpd* gene products

ORF	Position in sequence (nt)	Theoretical pI/ molecular mass (Da)	Predicted function	Similar protein	Organism	Amino acid identity (%)/positives (%)	Protein accession no.
<i>orf3</i>	1552–2304	5·01/28 214	No prediction	pnf 2500*	<i>Nocardia farcinica</i> IFM 10152	37/51†	YP_122100
<i>mpdR</i>	2764–3996	7·86/43 559	Putative regulator	SAV2124* SC06105* nfa29470	<i>Streptomyces avermitilis</i> MA-4680 <i>Brevibacterium linens</i> BL2 <i>Nocardia farcinica</i> IFM 10152	46/60 37/51 43/57	BAC69835 ZP_00378859 YP_119158
<i>mpdC</i>	4334–5848	4·91/53 848	Aldehyde dehydrogenase	BetB NsaF ThmS	<i>E. coli</i> K12 <i>Sphingomonas</i> sp. <i>Pseudonocardia</i> sp. K1	43/60 42/58 41/59	AAC73415 AAD45415 CAC10505
<i>orf1</i>	5917–6564	4·44/23 418	No prediction	TK0944	<i>Thermococcus kodakaraensis</i> KOD1	24/42†	BAD85133
<i>mpdB</i>	6590–8248	5·98/60 181	Alcohol dehydrogenase	G3DH BetA	<i>Halomonas</i> sp. alpha-15 <i>E. coli</i> K12	29/43 27/40	JC7628 AAC73414
<i>orf2</i>	8338–9009	4·41/23 411	Putative permease	Rgel02002828	<i>Rubrivivax gelatinosus</i> PM1	34/47	ZP_00242947
<i>tnpA</i> (inc.)	9202–10327	ND	Putative transposase	Rgel02003306 IS4	<i>Rubrivivax gelatinosus</i> PM1 <i>Frankia</i> sp. EAN1pec	38/48† 43/55	ZP_0022457 ZP_00567542

ND, Not determined
*Hypothetical protein.
†Partial alignment.

cytoplasmic extract of *M. austroafricanum* IFP 2012, which did not include any membrane polypeptides.

Nucleotide sequence analysis of the *mpd* gene cluster showed three short intercistronic regions of 68 (*mpdC*–*orf1*), 25 (*orf1*–*mpdB*) and 89 (*mpdB*–*orf2*) nt, including, in each case, a potential Shine–Dalgarno sequence.

An incomplete ORF, the *tnpA* gene, was found at the 3' end of the cloned gDNA fragment (Fig. 2). The ORF encodes an amino acid sequence which strongly matched (47 % identity/59 % similarity) with two putative transposases of *Rhodococcus erythropolis*. Analysis of the *orf2*–*tnpA* intergenic region showed that it was divided into two parts: the upstream sequence (110 nt) had an abnormally high A + T content (50 %), while the 3' region showed the high G + C content (66 %) usually found in mycobacteria.

Finally, *orf3*, which was upstream of *mpdR* (Fig. 2), did not show significant homology with any known protein in either the protein or the DNA database, but hydropathy analysis revealed a single C-terminal transmembrane domain.

Heterologous expression of the *mpd* gene cluster in *M. smegmatis* mc2 155

The isolated 9060 bp fragment of *M. austroafricanum* IFP 2012 gDNA, deleted for the *tnpA* incomplete gene, was cloned in the pCL4D shuttle vector to yield the p4D2 expression plasmid, which was used to transform *M. smegmatis* mc2 155. The empty plasmid pCL4D was also used to transform *M. smegmatis* mc2 155. Degradation capacities of the resulting recombinant strains *M. smegmatis* mc2 155(p4D2) and *M. smegmatis* mc2 155(pCL4D) were determined using resting cells after growth on LB medium containing kanamycin. No degradation of MTBE, TBA and HIBA was observed with either strain. *M. smegmatis* mc2 155(p4D2) degraded 2-M1,2-PD and a stoichiometric production of HIBA was observed (Fig. 4). The maximum rate of 2-M1,2-PD degradation by strain *M. smegmatis* mc2 155(p4D2) was 2·34 ± 0·41 µmol min^{–1} (g dry weight)^{–1} and the maximum rate of HIBA production was similar at 1·89 ± 0·26 µmol min^{–1} (g dry weight)^{–1}. This result clearly showed that the function of the cloned cluster was the oxidation of 2-M1,2-PD to HIBA by two enzymic steps, which were catalysed by 2-M1,2-PD dehydrogenase (MpdB) and hydroxyisobutyraldehyde dehydrogenase (MpdC).

Since the cells were grown in LB, induction of the genes was required to produce the corresponding enzymes and this could explain why 2-M1,2-PD degradation was observed only after 4 h incubation in the presence of this substrate. No degradation was observed in the presence of chloramphenicol, an inhibitor of protein translation, and this confirmed the necessity of *de novo* protein synthesis in *M. smegmatis* mc2 155(p4D2) for degradation of 2-M1,2-PD (Fig. 4). No degradation was observed either in resting cells of *M. smegmatis* mc2 155(pCL4D), containing the vector without insert, or in the abiotic control, demonstrating the

Fig. 3. Structural alignments of MpdC with betaine aldehyde dehydrogenases, and MpdB with choline dehydrogenases. (a) The deduced amino acid sequence of MpdC was structurally aligned with the sequences of betaine aldehyde dehydrogenase from *E. coli* (Eco; Lamark *et al.*, 1991), *Sinorhizobium meliloti* (Sme; Pocard *et al.*, 1997) and cod liver (cod; Johansson *et al.*, 1998). The secondary structure for betaine aldehyde dehydrogenase from cod liver is indicated below the sequences (TOPO). The secondary structures in parentheses were either partially found or not found in the MpdC secondary structure prediction using either the PSIPRED (McGuffin *et al.*, 2000) or the SSPro (Pollastri *et al.*, 2002) program. Active site residues are shown as **X**, whereas **x**, **x** and **X** are used for identical amino acids, conserved substitutions and semi-conserved substitutions, respectively, according to the CLUSTALW alignment tool (Thompson *et al.*, 1994). (b) The deduced amino acid sequence of MpdB was structurally aligned with the sequences of choline dehydrogenase from *E. coli* (Eco; Lamark *et al.*, 1991), *Sinorhizobium meliloti* (Sme; Pocard *et al.*, 1997) and the COG2303 consensus sequence, which includes choline dehydrogenases (*betA* genes) and related flavoproteins. Motifs were detected by similarity with consensus sequences of the GMC oxidoreductase family, described by Kiess *et al.* (1998), and secondary structure (TOPO) predicted as described above.

specific role of the *mpd* gene cluster in the conversion of 2-M1,2-PD to HIBA via hydroxyisobutyraldehyde (Fig. 5). The detection of hydroxyisobutyraldehyde by GC/MS was not successful, probably because the rate of 2-M1,2-PD degradation was equal to the rate of HIBA production, preventing accumulation of this intermediate.

Expression of the *mpd* gene transcript

Total RNA of *M. austroafricanum* IFP 2012 cells was analysed by RT-PCR, after growth on TBA or glucose (Fig. 6a, b, respectively). We used three different primer pairs designed to amplify (i) an internal sequence of *mpdR* (RT-FI/RT-RI primer pair), (ii) the overlapping sequence

mpdC-orf1 (RT-FII/RT-RII primer pair), and (iii) the overlapping sequence *mpdB-orf2* (RT-FIII/RT-RIII primer pair). These primer pairs are described in Table 2. Strong RT-PCR amplification with both *mpdC-orf1* and *mpdB-orf2* primer pairs was observed using RNA extracted from *M. austroafricanum* IFP 2012 cells grown on TBA. In addition, we also obtained amplification using the *mpdR* primers (Fig. 6a). These results clearly showed that transcription of the *mpd* gene cluster was induced after growth on TBA.

Using RNA extracted from *M. austroafricanum* IFP 2012 cells grown on glucose, no RT-PCR product was obtained with the *mpdR* primer set and only very weak bands were detected using the *mpdC-orf1* and *mpdB-orf2* primer pairs

Fig. 4. Degradation of 2-M1,2-PD by the recombinant strains *M. smegmatis* mc2 155(p4D2) and *M. smegmatis* mc2 155(pCL4D). Degradation of 2-M1,2-PD (filled symbols) and production of HIBA (open symbols) with strains *M. smegmatis* mc2 155(p4D2) (squares) and *M. smegmatis* mc2 155(pCL4D) (triangles), were followed using resting cells. Controls without cells (diamonds). Degradation activity was not observed when chloramphenicol (400 mg l^{-1}) was added to resting cells of *M. smegmatis* mc2 155(p4D2) (crosses). Cell masses of *M. smegmatis* mc2 155(p4D2) and *M. smegmatis* mc2 155(pCL4D) were 233.3 ± 14.1 and $214.6 \pm 22.0 \text{ mg l}^{-1}$, respectively. The data are means $\pm$ SD of three replicates.

(Fig. 6b). This could suggest a basal transcription level of the *mpd* gene cluster under non-selective conditions.

DISCUSSION

In this study, and to the best of our knowledge, we report the first cloning, sequencing and functional expression of a gene cluster from *M. austroafricanum* IFP 2012 that is involved in growth in the presence of MTBE. This cluster included an alcohol dehydrogenase and an aldehyde dehydrogenase. This is, to the best of our knowledge, the first report of the detailed genetic characterization of genes involved in an MTBE metabolic pathway.

We characterized proteins specifically induced after growth in the presence of MTBE (Fig. 1) and determined the

internal or N-terminal amino acid sequences of these proteins. The DNA sequences deduced from these peptide sequences were used to design degenerate oligonucleotide primers to obtain PCR amplification products. The products were labelled and used as probes in Southern analysis on digested total DNA of *M. austroafricanum* IFP 2012 to isolate and clone a 10 327 bp gDNA fragment containing seven ORFs (six complete and one incomplete, Fig. 2). A 9060 bp fragment was then cloned in pCL4D and transferred into *M. smegmatis* mc2 155. The functional heterologous expression of these genes isolated from *M. austroafricanum* IFP 2012 in *M. smegmatis* mc2 155 was also shown for the first time, to the best of our knowledge, by determining that the products of the *mpd* genes, including an alcohol dehydrogenase (MpdB) and an aldehyde dehydrogenase (MpdC), catalysed the conversion of 2-M1,2-PD to HIBA (Fig. 4). This activity was inducible by 2-M1,2-PD, as it occurred after a 4 h lag phase and did not occur when chloramphenicol, which inhibits translation, was present. The 2-M1,2-PD was previously identified as the product of TBA oxidation in *M. austroafricanum* IFP 2012 (Lopes Ferreira *et al.*, 2005) and our latest results show that the *mpd* genes are involved in MTBE catabolism in *M. austroafricanum* IFP 2012. In addition, we also detected the correlated overexpression of 64, 55 and 27 kDa proteins in *M. austroafricanum* IFP 2012 grown in the presence of MTBE, as shown in Fig. 1(a) and (b), and these proteins corresponded to the products of the *mpdB*, *mpdC* and *orf1* genes, respectively.

Sequence alignment of MpdB and MpdC with protein databases showed high similarities with BetA and BetB, respectively, which are two enzymes responsible for the synthesis of glycine betaine from choline in several bacteria (Csonka & Epstein, 1996; Lamark *et al.*, 1991; Landfald & Strom, 1986; Pocard *et al.*, 1997). The MpdB protein showed the highest similarity with a choline dehydrogenase of *B. japonicum*, which is a strain unable to degrade choline (Boncompagni *et al.*, 1999), and this was also the case in *M. austroafricanum* IFP 2012 (data not shown).

Structural alignment of MpdB with choline dehydrogenases is described in Fig. 3(b). This alignment showed the

Fig. 5. Schematic representation of the MTBE catabolic pathway in *M. austroafricanum* IFP 2012. The enzymes responsible for the steps described in the text are indicated. Compounds shown in bold were identified by GC or HPLC.

Fig. 6. *mpd* gene expression in *M. austroafricanum* IFP 2012 determined by RT-PCR. The total RNA of *M. austroafricanum* IFP 2012 cells grown in MM supplemented with TBA (a) and glucose (b) was reverse transcribed. The reverse transcripts were subjected to PCR amplification using the RT-FI/RT-RI (I) primer pair to amplify the internal sequence of *mpdR*, the RT-FII/RT-RII (II) primer pair to amplify the *mpdC-orf1* intergenic region and the RT-FIII/RT-RIII (III) primer pair to amplify the *mpdB-orf2* intergenic region (see Table 2, Fig. 2). RNA samples were analysed by PCR, preceded by the reverse transcription step (RT/+). To check for the absence of contaminant DNA in the RNA preparations, the samples were also analysed by PCR without the reverse transcription step (RT/-) and a PCR-positive control (100 ng total DNA) was included to test all primer sets (C+).

conservation of several consensus patterns required in the FAD-binding domain, such as a β - α - β -fold domain, a flavin-attachment loop, β -sheet domains and amino acids involved in the active sites. Structural alignment of MpdC with betaine aldehyde dehydrogenases (Fig. 3a) showed a strong conservation of both the amino acid sequences and the secondary structures of all the related proteins that were examined. Among the putative proteins encoded by the *mpd* genes, only MpdB and MpdC were closely related to functional alcohol dehydrogenase and aldehyde dehydrogenase enzymes, respectively.

Since the conversion of 2-M1,2-PD to HIBA requires the sequential action of an alcohol dehydrogenase and an aldehyde dehydrogenase, we can infer from these results that these two enzymes are specifically involved in this conversion in the MTBE metabolic pathway (Fig. 5).

Two distinct phases in the MTBE pathway can be identified during the degradation from MTBE to HIBA. The first one, from MTBE to 2-M1,2-PD, includes two oxidation steps catalysed by monooxygenases: (i) MTBE to *tert*-butoxy-methanol, and (ii) TBA to 2-M1,2-PD. This part of MTBE catabolism can be written as the sum of five enzymic reactions, including the mineralization of formate generated as a co-product by the hydrolysis of TBF to TBA: $C_5H_{12}O + 2O_2 \rightarrow C_4H_{10}O_2 + H_2O + CO_2$.

This upper part of the MTBE degradation pathway consumes reduced coenzyme equivalents. Even though NADH is regenerated by formate dehydrogenase during the conversion of formate to carbon dioxide, no ATP can be produced, resulting in a loss of energy. That is one of the reasons why MTBE-degrading micro-organisms have to be methylotrophic, which is the case for *M. austroafricanum* IFP 2012 (François *et al.*, 2002). In the lower part of the pathway, the

oxidation of 2-M1,2-PD to HIBA by the action of two successive dehydrogenases [$C_4H_{10}O_2 + H_2O + 2NAD(P)^+ \rightarrow C_4H_8O_3 + 2NAD(P)H + 2H^+$] produces reduced coenzyme equivalents that can be used in the above oxidation reactions. The requirement for reducing equivalents could explain (i) the strong expression of the enzymes encoded by the *mpd* cluster observed in the MTBE protein profile (Fig. 1), and (ii) the decrease in the MTBE degradation rate observed during growth of *M. austroafricanum* IFP 2012 (François *et al.*, 2002). The efficiency of the degradation of 2-M1,2-PD to HIBA explains why hydroxyisobutyraldehyde was not be detected as a degradation intermediate, since HIBA was produced at the same rate as 2-M1,2-PD was degraded [$1.89 \pm 0.26 \mu\text{mol min}^{-1} (\text{g dry weight})^{-1}$ and $2.34 \pm 0.41 \mu\text{mol min}^{-1} (\text{g dry weight})^{-1}$, respectively]. Moreover, the production of reduced co-enzyme equivalents is linked to the production of ATP and it has been proposed that the energy loss during growth in the presence of MTBE could be a key limiting factor in its catabolism (Fayolle *et al.*, 2003; Fortin *et al.*, 2001; Salanitro, 1995). It is known that the imbalance in reduced co-factors generated during the oxidation steps of xenobiotic biodegradation can limit growth of micro-organisms on these compounds (VanBriesen, 2001).

A high level of mRNA transcription was observed by RT-PCR following growth on TBA, clearly showing the specific induction of the *mpd* genes. This correlated with the results obtained from SDS-PAGE analysis.

The *mpd* genes appeared to be regulated by *mpdR*, which encodes a putative σ^{54} -dependent transcriptional regulator (Burgess *et al.*, 1969). However, alignments were found only in the N-terminal part of the corresponding protein. Moreover, two motifs that are usually conserved, i.e. the N-terminal ATP-binding motif (Austin *et al.*, 1991; Morett

& Segovia, 1993), and the C-terminal helix–turn–helix DNA-binding domain (Austin & Dixon, 1992), were not detected. Nevertheless, it is noteworthy that genes encoding σ^{54} were not found in the genomes of several recently sequenced Gram-positive strains that had a high-G + C content (Buck *et al.*, 2000). DNA analysis of the *mpd* cluster also revealed short intergenic regions, including Shine–Dalgarno sequences (Draper, 1996). A GUG start codon was present in the coding sequence of the *mpdB* gene and this is unusual (Ma *et al.*, 2002). Due to the short intergenic sequences separating the regulatory *mpdR* gene from the other *mpd* genes, promoters are probably overlapping, which should contribute to the complexity of the regulation. Røkenes *et al.* (1996) have shown that the *betIBA* operon of *E. coli* contains similar overlapping promoters.

Of interest was the specific expression of the molecular chaperones DnaK and GroEL/ES which were induced during growth in the presence of MTBE. Molecular chaperones are involved in a myriad of functions, including nascent protein folding and assembly, refolding of protein quaternary structure (Lopes Ferreira & Alix, 2002), synthesis of mycolic acids, translocation across membranes, or modulation of heat-shock responses, which can be considered as house-keeping functions, and a defence strategy against stress (for review see Alix, 2004). The combination of the high folding demand of newly synthesized proteins and the above-mentioned imbalance in reduced co-factors during growth in the presence of MTBE could explain the stimulation of these molecular chaperones in *M. austroafricanum* IFP 2012.

From a global perspective, our results showed that growth in the presence of MTBE induces ten cytoplasmic proteins. Seven of these proteins were characterized: (i) proteins of 64, 55 and 27 kDa encoded by the *mpdB*, *mpdC* and *orf1* genes, respectively, described in this study; (ii) proteins of 67, 63 and 11 kDa corresponding to DnaK and GroEL/ES molecular chaperones, respectively; and (iii) a protein of 50 kDa, which is the methanol:NDMA oxidoreductase detected in *Mycobacterium gastri* (Bystrykh *et al.*, 1993) and several other mycobacteria (Park *et al.*, 2003). This latter protein was detected in crude extracts of *M. austroafricanum* IFP 2012 after growth on HIBA and 2-propanol (François, 2002), suggesting its involvement in assimilation of 2-propanol, a putative intermediate of the MTBE metabolic pathway. The remaining induced proteins corresponded to: (i) a 24 kDa protein, which matched with the Rubisco involved in carbon dioxide fixation; (ii) a 17 kDa protein, which was closely related to the naphthalene dioxygenase small subunit from *Mycobacterium* sp. PYR-1 (Khan *et al.*, 2001); and (iii) an uncharacterized 14 kDa protein. The role of the latter three proteins in the MTBE catabolic pathway has not yet been determined.

At present, the nature of the monooxygenases catalysing the oxidation of MTBE and TBA has not been elucidated, as none of the strongly induced soluble proteins have shown similarities with known monooxygenases. The hypothesis that the monooxygenase responsible for MTBE and TBA

oxidation might be located in the membrane deserves further investigation, since we did not analyse this cell compartment. In this regard, the membrane-bound alkane hydroxylase could be a candidate, since several bacterial strains are able to degrade MTBE by cometabolism after growth on alkanes (Garnier *et al.*, 1999; Smith & Hyman, 2004; Smith *et al.*, 2003).

The products of the *mpd* genes involved in the biodegradation of 2-M1,2-PD play a key role in the growth of *M. austroafricanum* IFP 2012 in the presence of MTBE. The *mpdB* and *mpdC* sequences could provide new molecular tools to undertake studies on the distribution of MTBE-degrading bacteria. These tools should help in estimating the capacity for natural attenuation of MTBE-contaminated ecosystems.

ACKNOWLEDGEMENTS

Nicolas Lopes Ferreira was supported by a Convention Industrielle de Formation par la Recherche (C.I.F.R.E.) fellowship provided by both the Institut Français du Pétrole (IFP) and the Association Nationale de la Recherche Technique (ANRT). We thank Dr Pierre Béguin from the Institut Pasteur (Paris) for providing us with the pCL4D plasmid, and Dr Sergio Revah for providing us with 2-M1,2-PD. We also thank: (from the Biotechnology Research Institute) Sylvie Berardi for oligonucleotide synthesis, and Manon Lalumière and France Dumas for their technical help in DNA and protein sequencing, respectively; (from the Institut Français du Pétrole) Hugues Mathis for technical help with GC and HPLC analysis.

REFERENCES

- Alix, J. H. (2004). The work of chaperone. In *Protein Synthesis and Ribosome Structure*, vol. 1, pp. 529–553. Edited by K. H. Nierhaus & D. N. Wilson. Weinheim: Wiley-VCH.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, Z. J., Zhang Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
- Austin, S. & Dixon, R. (1992). The prokaryotic enhancer binding protein NTRC has an ATPase activity which is phosphorylation and DNA dependent. *EMBO J* **11**, 2219–2228.
- Austin, S., Kundrot, C. & Dixon, R. (1991). Influence of a mutation in the putative nucleotide binding site of the nitrogen regulatory protein NTRC on its positive control function. *Nucleic Acids Res* **19**, 2281–2287.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1990). *Current Protocols in Molecular Biology*. New York: Greene Publishing and Wiley Interscience.
- Bateman, A., Birney, E. & 8 other authors (2002). The Pfam protein families database. *Nucleic Acids Res* **30**, 276–280.
- Boncompagni, E., Østerås, M., Poggi, M. C. & Le Rudulier, D. (1999). Occurrence of choline and glycine betaine uptake and metabolism in the family *Rhizobiaceae* and their roles in osmoprotection. *Appl Environ Microbiol* **65**, 2072–2077.
- Buck, M., Gallegos, M. T., Studholme, D. J., Guo, Y. & Gralla, J. D. (2000). The bacterial enhancer-dependent σ^{54} (σ^N) transcription factor. *J Bacteriol* **182**, 4129–4136.

- Burgess, R. R., Travers, A. A., Dunn, J. J. & Bautz, E. K. (1969). Factor stimulating transcription by RNA polymerase. *Nature* **221**, 43–46.
- Bystrykh, L. V., Vonck, J., van Bruggen, E. F., van Beeumen, J., Samyn, B., Govorukhina, N. I., Arfman, N., Duine, J. A. & Dijkhuizen, L. (1993). Electron microscopic analysis and structural characterization of novel NADP(H)-containing methanol: *N,N'*-dimethyl-4-nitrosoaniline oxidoreductases from the Gram-positive methylotrophic bacteria *Amycolatopsis methanolica* and *Mycobacterium gastri* MB19. *J Bacteriol* **175**, 1814–1822.
- Cavener, D. R. (1992). GMC oxidoreductases: a newly defined family of homologous proteins with diverse catalytic activities. *J Mol Biol* **223**, 811–814.
- Csonka, L. N. & Epstein, W. (1996). Osmoregulation. In *Escherichia and Salmonella: Cellular and Molecular Biology*, 2nd edn, pp. 1210–1223. Edited by F. C. Neidhardt, R. Curtiss, J. L. Ingraham, E. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter & H. E. Umbarger. Washington, DC: American Society for Microbiology.
- Deeb, R. A., Hu, H.-Y., Hanson, J. R., Scow, K. M. & Alvarez-Cohen, L. (2001). Substrate interactions in BTEX and MTBE mixtures by an MTBE-degrading isolate. *Environ Sci Technol* **35**, 312–317.
- Draper, D. E. (1996). Translational initiation. In *Escherichia and Salmonella: Cellular and Molecular Biology*, 2nd edn, pp. 902–908. Edited by F. C. Neidhardt, R. Curtiss, J. L. Ingraham, E. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter & H. E. Umbarger. Washington, DC: American Society for Microbiology.
- Fayolle, F. & Monot, F. (2005). Biodegradation of fuel ethers. In *Petroleum Microbiology*, pp. 301–316. Edited by M. Magot & B. Ollivier. Washington, DC: American Society for Microbiology.
- Fayolle, F., Francois, A., Garnier, L., Godefroy, D., Mathis, H., Piveteau, P. & Monot, F. (2003). Limitations in MTBE biodegradation. *Oil Gas Sci Technol* **58**, 497–504.
- Fortin, N. Y., Morales, M., Nakagawa, Y., Focht, D. D. & Deshusses, M. A. (2001). Methyl *tert*-butyl ether (MTBE) degradation by a microbial consortium. *Environ Microbiol* **3**, 407–416.
- François, A. (2002). *Biodegradation of a recalcitrant xenobiotic compound: metabolism of methyl tert-butyl ether (MTBE) by Mycobacterium austroafricanum* IFP 2012. PhD thesis, Institut National Agronomique de Paris-Grignon, Paris.
- François, A., Mathis, H., Godefroy, D., Piveteau, P., Fayolle, F. & Monot, F. (2002). Biodegradation of methyl *tert*-butyl ether and other fuel oxygenates by a new strain, *Mycobacterium austroafricanum* IFP 2012. *Appl Environ Microbiol* **68**, 2754–2762.
- François, A., Garnier, L., Mathis, H., Fayolle, F. & Monot, F. (2003). Roles of *tert*-butyl formate, *tert*-butyl alcohol and acetone in the regulation of methyl *tert*-butyl ether degradation by *Mycobacterium austroafricanum* IFP 2012. *Appl Microbiol Biotechnol* **62**, 256–262.
- Garnier, P. M., Auria, R., Augur, C. & Revah, S. (1999). Cometabolic biodegradation of methyl *t*-butyl ether by *Pseudomonas aeruginosa* grown on pentane. *Appl Microbiol Biotechnol* **51**, 498–503.
- Giegel, D. A., Williams, C. H., Jr & Massey, V. (1990). L-Lactate 2-monooxygenase from *Mycobacterium smegmatis*. Cloning, nucleotide sequence, and primary structure homology within an enzyme family. *J Biol Chem* **265**, 6626–6632.
- Grant, S. G., Jessee, J., Bloom, F. R. & Hanahan, D. (1990). Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc Natl Acad Sci U S A* **87**, 4645–4649.
- Hanahan, D., Jesse, J. & Bloom, F. R. (1991). Plasmid transformation of *Escherichia coli* and other bacteria. *Methods Enzymol* **204**, 63–113.
- Hanson, J. R., Ackerman, C. E. & Scow, K. M. (1999). Biodegradation of methyl *tert*-butyl ether by a bacterial pure culture. *Appl Environ Microbiol* **65**, 4788–4792.
- Hatzinger, P. B., Mac Clay, K., Vainberg, S., Tugusheva, M., Condee, C. W. & Steffan, R. J. (2001). Biodegradation of methyl *tert*-butyl ether by a pure bacterial culture. *Appl Environ Microbiol* **67**, 5601–5607.
- Hellman, U., Wernstedt, C., Góñez, J. & Heldin, C.-H. (1995). Improvement of an 'In Gel' digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. *Anal Biochem* **224**, 451–455.
- Hewick, R. M., Hunkaliller, M. W., Hood, L. E. & Dreyer, W. J. (1981). A gas-liquid solid phase peptide and protein sequenator. *J Biol Chem* **256**, 7990–7997.
- Hu, Y. & Coates, A. R. (1999). Transcription of two sigma 70 homologue genes, *sigA* and *sigB*, in stationary-phase *Mycobacterium tuberculosis*. *J Bacteriol* **181**, 469–476.
- Jenö, P., Mini, T., Moes, S., Hintermann, E. & Horst, M. (1995). Internal sequences from proteins digested in polyacrylamide gels. *Anal Biochem* **224**, 75–82.
- Johansson, K., El-Ahmad, M., Ramaswamy, S., Hjelmqvist, L., Jorvall, H. & Eklund, H. (1998). Structure of betaine aldehyde dehydrogenase at 2.1 Å resolution. *Protein Sci* **10**, 2106–2117.
- Johnson, R., Pankow, J., Bender, D., Price, C. & Zogorsky, J. (2000). MTBE. To what extent will past releases contaminate community water supply wells? *Environ Sci Technol* **34**, 210A–217A.
- Kamalakkannan, V., Ramachandran, G., Narayanan, S., Vasan, S. K. & Narayanan, P. R. (2002). Identification of a novel mycobacterial transcriptional regulator and its involvement in growth rate dependence and stringent control. *FEMS Microbiol Lett* **209**, 261–266.
- Khan, A. A., Wang, R. F., Cao, W.-W., Doerge, D. R., Wennerstrom, D. & Cerniglia, C. E. (2001). Molecular cloning, nucleotide sequence and expression of genes encoding a polycyclic aromatic ring dioxygenase from *Mycobacterium* sp. strain PYR-1. *Appl Environ Microbiol* **67**, 3577–3585.
- Kiess, M., Hecht, H. J. & Kalisz, H. M. (1998). Glucose oxidase from *Penicillium amagasakiense*. Primary structure and comparison with other glucose-methanol-choline (GMC) oxidoreductases. *Eur J Biochem* **252**, 90–99.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Lamark, T., Kaasen, I., Eshoo, M. W., Falkenberg, P., McDougall, J. & Strom, A. R. (1991). DNA sequence and analysis of the *bet* genes encoding the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. *Mol Microbiol* **5**, 1049–1064.
- Landfald, B. & Strom, A. R. (1986). Choline-glycine betaine pathway confers a high level of osmotic tolerance in *Escherichia coli*. *J Bacteriol* **65**, 849–855.
- Lopes Ferreira, N. & Alix, J.-H. (2002). The DnaK chaperone is necessary for α -complementation of β -galactosidase in *Escherichia coli*. *J Bacteriol* **184**, 7047–7054.
- Lopes Ferreira, N., Maciel, H., Mathis, H., Monot, F., Fayolle-Guichard, F. & Greer, C. W. (2005). Isolation and characterization of a new *Mycobacterium austroafricanum* strain, IFP 2015, growing on MTBE. *Appl Microbiol Biotechnol* (in press). doi:10.1007/s00253-005-0074-y
- 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.
- McGuffin, L. J., Bryson, K. & Jones, D. T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404–405.

- Misra, N., Habib, S., Ranjan, A., Hasnain, S. E. & Nath, I. (1996). Expression and functional characterisation of the *clpC* gene of *Mycobacterium leprae*: ClpC protein elicits human antibody response. *Gene* **172**, 99–104.
- 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.
- Parish, T. & Stoker, N. G. (1998). Electroporation of *Mycobacteria*. In *Mycobacteria Protocols: Methods in Molecular Biology*, vol. 101, pp. 129–144. Edited by T. Parish & N. G. Stoker. Totowa, NJ: Humana Press.
- Park, S. W., Hwang, E. H. & 8 other authors (2003). Growth of mycobacteria on carbon monoxide and methanol. *J Bacteriol* **185**, 142–147.
- Picardeau, M., Le Dantec, C. & Vincent, V. (2000). Analysis of the internal replication region of a mycobacterial linear plasmid. *Microbiology* **146**, 305–313.
- Piveteau, P., Fayolle, F., Vandecasteele, J.-P. & Monot, F. (2001). Biodegradation of *tert*-butyl alcohol and related xenobiotics by a methylotrophic bacterial isolate. *Appl Microbiol Biotechnol* **55**, 369–373.
- Pocard, J. A., Vincent, N., Boncompagni, E., Smith, L. T., Poggi, M. C. & Le Rudulier, D. (1997). Molecular characterization of the *bet* genes encoding glycine betaine synthesis in *Sinorhizobium meliloti* 102F34. *Microbiology* **143**, 1369–1379.
- Pollastri, G., Przybylski, D., Rost, B. & Baldi, P. (2002). Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. *Proteins* **47**, 228–235.
- Pospiech, A. & Neumann, B. (1995). A versatile quick-prep of genomic DNA from Gram-positive bacteria. *Trends Genet* **6**, 217–218.
- Pruden, A. & Suidan, M. (2004). Effect of benzene, toluene, ethylbenzene and *p*-xylene (BTEX) mixture on biodegradation of methyl *tert*-butyl ether (MTBE) and *tert*-butyl alcohol by pure culture UC1. *Biodegradation* **15**, 213–227.
- Revel-Viravau, V., Truong, Q. C., Moreau, N., Jarlier, V. & Sougakoff, W. (1996). Sequence analysis, purification, and study of inhibition by 4-quinolones of the DNA gyrase from *Mycobacterium smegmatis*. *Antimicrob Agents Chemother* **40**, 2054–2061.
- Rohwerder, T., Cenini, V., Held, C., Martienessen, M., Lechner, U. & Müller, R. H. (2004). Novel MTBE-degrading bacterial isolate from Leuna groundwater (Germany): characterization of the degradation pathway with focus on HIBA oxidase. In *Proceedings of the Second European Conference on MTBE*. Edited by D. Barcelo & M. Petrovic. Barcelona: Consejo Superior de Investigaciones Científicas.
- Røkenes, T. P., Lamark, T. & Strom, A. R. (1996). DNA-binding properties of the BetI repressor protein of *Escherichia coli*: the inducer choline stimulates BetI-DNA complex formation. *J Bacteriol* **178**, 1663–1670.
- Rost, B. & Eyrich, V. A. (2001). EVA: large-scale analysis of secondary structure prediction. *Proteins* **45**, 192–199.
- Salanitro, J. P. (1995). Understanding the limitations of microbial metabolism of ethers used as fuel octane enhancers. *Curr Opin Biotechnol* **6**, 337–340.
- Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: a Laboratory Manual*, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Smith, C. A. & Hyman, M. R. (2004). Oxidation of methyl *tert*-butyl ether by propane-grown *Pseudomonas putida* GPo1. *Appl Environ Microbiol* **70**, 4544–4550.
- Smith, C. A., O'Reilly, K. T. & Hyman, M. R. (2003). Characterization of the initial reactions during the cometabolic oxidation of methyl *tert*-butyl ether by propane-grown *Mycobacterium vaccae* JOB5. *Appl Environ Microbiol* **69**, 796–804.
- Snapper, S. B., Melton, R. E., Mustafa, S., Keiser, T. & Jacobs, W. R., Jr (1990). Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* **4**, 1911–1919.
- Steffan, R. J., McClay, K., Vainberg, S., Condee, C. W. & Zhang, D. (1997). Biodegradation of the gasoline oxygenates methyl *tert*-butyl ether, ethyl *tert*-butyl ether, and *tert*-amyl methyl ether by propane-oxidizing bacteria. *Appl Environ Microbiol* **63**, 4216–4222.
- Tatusov, R. L., Koonin, E. V. & Lipman, D. J. (1997). A genomic perspective on protein families. *Science* **278**, 631–637.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTALW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.
- United States Environmental Protection Agency (1997). *Drinking Water Advisory: Consumer Acceptability Advice and Health Effects Analysis on Methyl Tertiary Butyl Ether (MtBE)*. EPA/822/F-97/008, December 1997. Washington, DC: Office of Water.
- United States Environmental Protection Agency (1999). *Achieving Clean Air and Clean Water: the Report of the Blue Ribbon Panel on Oxygenates in Gasoline*. EPA 420-R-99-021, 15 September, 1999. Washington, DC: Office of Water.
- VanBriesen, J. M. (2001). Thermodynamic yield predictions for biodegradation through oxygenase activation reactions. *Biodegradation* **12**, 265–281.
- Wierenga, R. K., De Maeyer, M. C. H. & Hol, W. J. G. (1985). Interaction of pyrophosphate moieties with alpha-helices in dinucleotide binding proteins. *Biochemistry* **24**, 1346–1357.