The molecular basis for the biotic degradation of metaldehyde

John Christopher Thomas

PhD

University of York

Biology

September 2016

Abstract

Metaldehyde is a molluscicide that is used to prevent crop damage by slugs and snails. It is frequently found in drinking water sources at concentrations above the legal maximum, and causes >90% of water quality failures in England. There is currently no economical, widely used treatment to remove metaldehyde pollution from water. Prior to the work presented here, it was known that metaldehyde is degraded biotically in soil, but the nature and identity of organisms responsible was unknown. Two novel metaldehyde degrading bacterial strains were isolated from domestic soil that can utilise metaldehyde as a sole carbon and energy source; *Acinetobacter calcoaceticus* E1 and *Variovorax* E3. Evidence that acetaldehyde is the primary metabolite of metaldehyde is presented: the metaldehyde dependent expression of an aldehyde dehydrogenase in *A. calcoaceticus* E1 was found, and the maximum metabolic flux of acetaldehyde was found to be greater than that of metaldehyde in cells grown using metaldehyde. The genomes of the isolated organisms were acquired. Comparative genomics of the two isolates showed that their novel catalytic ability was not dependent on genes shared between them, despite being isolated from the same soil sample. Comparison of the *A. calcoaceticus* E1 genome against other *Acinetobacter* that cannot utilise metaldehyde yielded a strong candidate for the primary metaldehyde degrading enzyme that has characteristics consistent with an enzyme that catalysed ether hydrolysis of isochorismate in its ancestral form and has evolved to hydrolyse ether bonds in metaldehyde.

Table of contents

Table of figures

Table of tables

Acknowledgements

I would like to thank my supervisors; Dr James Moir, Dr Thorunn Helgason, Chris Sinclair and Richard Thwaites for the guidance, knowledge, time and encouragement they gave me during this undertaking.

I would also like to thank my Thesis Advisory Panel members; Prof Neil Bruce and Dr Kelly Redeker, for their guidance.

Thanks to the fine people of L1, for making the lab environment so pleasant.

Thanks to Dr Peter Ashton for advice regarding matters bioinformatic. And to Keith Meech and Steve Jones for training me to use, and for help with troubleshooting the LCMS.

Thank you to my wonderful friends for some of the best times in my life so far.

And finally, thanks to all my family for their love and encouragement, and to all the academics and technical staff who cared to give me their time and thoughts.

Author's declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1: General Introduction

1.1 The molecular diversity of soil microorganisms

Soils are formed from weathered rock and decomposed or decomposing organic material. The properties of a soil depend on: (i) its composition (the organic and inorganic constituents); (ii) the nature of the inorganic constituent, which affects its physical structure and chemical properties; (iii) the physical environment, for example temperature and rainfall; and (iv) the activities of organisms that live in and above it, which are intimately involved in the input, removal and transformation of compounds in the soil. Soils are vertically stratified with greater organic content and nutrient availability at the surface (Elsas et al., 2006). The horizontal variability of the soil environment, in terms of, for example, species present, nutrient availability, and chemistry, can vary on scales of hundreds of metres; if, for example, a carbon gradient results from cultivation practices (Fromm et al., 1993) or a pH gradient is produced by the underlying geology. Soils provide a habitat for microorganisms, and these types of variations impact the types and identities of microorganisms that inhabit the substrate (Rousk et al., 2010). On a smaller scale, across tens of centimetres, physical parameters such as nitrogen availability, moisture content and the respiration rate of microorganisms present in soil can vary by one or two-fold (Stoyan et al., 2000). On the micro scale at which microorganisms operate (Raynaud & Nunan, 2014) there is also great heterogeneity in soils, and bacterial populations are found to be partitioned by ecological variation (Cordero & Polz, 2014). This can lead to differences in metabolic capacities of microbial populations across millimetre and centimetre distances (Gonod et al., 2003).

Microbes are numerous; the available surface area for bacterial habitats on soil particles is immense, and the population density has been given as between 4×10^7 and 2×10^9 cells per gram for different soils (Whitman et al., 1998). Microbial populations are inherently capable of generating greater genomic and phenotypic diversity than Metazoa, even discounting their number. Their short generation times, large number of reproductive units, and haploid genomes lead to the frequent generation of the novel phenotypes upon which natural selection acts. These properties, combined with a highly heterogeneous environment, lead to significant genetic diversity. The genetics of prokaryotes, the manner of their reproduction and propensity for horizontal gene transfer, make the definition of species in this domain difficult and controversial. In practice, in recent decades, operational taxonomic units (OTUs) are defined, using set values for genomically derived statistics to assign isolates to OTUs. Values such as <97% ribosomal 16S subunit ribonucleic acid (16S rRNA) sequence identity (Stackebrandt & Goebel, 1994) or <95% average nucleotide identity (Rodriguez-R

& Konstantinidis, 2014) are used to distinguish isolates at the species level. Estimates for the number of bacterial species are constantly evolving as more data become available; a recent meta-analysis using many sources of environmental 16S rRNA data estimates that there are between 10^{11} and 10^{12} bacterial species on the Earth (Locey & Lennon, 2016). Within individual species there is significant genomic diversity. Ozer et al. (2014) show that 6.9-18.0% of the genomes of 12 *Pseudomonas aeruginosa* strains are not conserved among all the strains. Tettelin et al. (2002) found that 18% of the genome in a reference strain of *Streptococcus agalactiae* was not present in at least one of 19 other strains. Estimating the frequency with which new genes will be discovered in a bacterial species as different strains of that species are sequenced can be done using an application of Heap's law (Tettelin et al., 2008). This will sometimes lead to the prediction of an "open pan-genome"; new genes would be expected with some frequency in every new strain sequenced. In a study of 3 *Aeromonas* species; within 6, 11 and 16 genomes; 6884, 7214 and 9181 genes were identified in total with 3344, 3380 and 3508 genes being core, respectively (Ghatak et al., 2016). All *Aeromonas* species studied were predicted to have open pan-genomes.

This genomic diversity, supplied by the very nature of bacterial genetics and sustained by the pressures of a heterogeneous environment, is a source of evolvability (Cordero & Polz, 2014). Bacteria can evolve new metabolic activities, such as the ability to catabolise synthetic compounds that are not found in nature (xenobiotics), through the molecular evolution of enzymes (O'Brien & Herschlag, 1999).

1.2 The anthropogenic influence on soil ecology

Human beings have shaped the earth to suit our needs. Human selection of the plant and animal species present in an area – directly, by planting in, or paving over the soil; and indirectly by the environments we create – is a major determinant of the ecology in that area (Ellis, 2015). The above ground environment will impact the soil environment, and agricultural soil is often intensively managed. In the United Kingdom (UK), in 2010, 15.3 million hectares (Mha) of a total land area of 24.3 Mha (62.9%) were used for agriculture. Agricultural practice in the UK typically involves intensive management of the land, with frequent tilling and the application of fertilisers and pesticides. Crops receive around 140 kg/ha of nitrogen, 29 kg/ha of phosphate and 39 kg/ha of potash per year as fertiliser (Department for Environment, Food & Rural Affairs, 2016). Crops also receive multiple applications of xenobiotic fungicides, herbicides, insecticides, molluscicides and growth regulators – often multiple formulations of each (Garthwaite et al., 2015). Many of these compounds will interact with the soil microbiome and be a source of selective pressure on the evolution of organisms; as potential food sources or by causing the inhibition of growth.

1.3 Metaldehyde is a widely used molluscicide

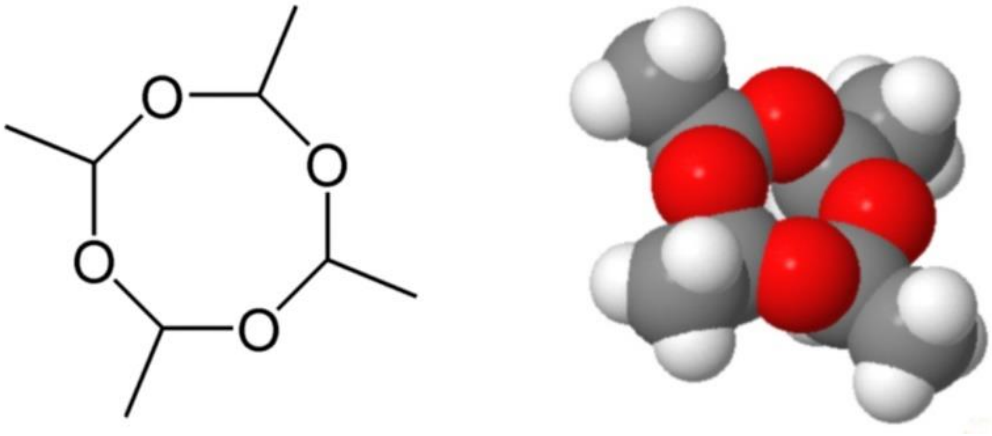


Figure 1-1. Skeletal structure (left) and 3-dimensional rendering (right) of metaldehyde. 3D rendering based on structural redetermination by Barnett et al. (2005), following Pauling & Carpenter (1936).

Metaldehyde is a cyclic tetramer of acetaldehyde with the formula $(CH_3CHO)_4$. Its structure was determined first by Pauling & Carpenter (1936), using X-ray crystallography (Figure 1-1). It has a molecular mass of 176.2 and is stable in water at around neutral pH. Sublimation of metaldehyde, by depolymerisation to acetaldehyde, begins at 80°C and is catalysed by the presence of acids such as dilute H_2SO_4 or H_3PO_4 (Eckert et al., 2000). Its melting point in a closed capillary is 246.2°C. It is soluble in water to 222 mg/L and soluble in methanol to 1730 mg/L. It has a log Kow of 0.12 (National Center for Biotechnology Information, n.d.).

Metaldehyde is not found in nature and is synthesised by the polymerisation of acetaldehyde at temperatures below 0°C in the presence of an acidic catalyst (Kekulé & Zincke, 1872). Metaldehyde crystallises and precipitates from the reaction giving yields of 7-15% (w/w). The majority product of the process is paraldehyde, a cyclic trimer of acetaldehyde. Aliphatic ethers are used to inhibit the precipitation of paraldehyde (Wilder, 1947) so that the solid metaldehyde can be removed and the remaining liquid paraldehyde recycled by being depolymerised to acetaldehyde. The preferential formation of paraldehyde suggests it is more thermodynamically favourable than metaldehyde.

Metaldehyde initially found use as a firelighter sold as “Meta-Fuel” (Miller, 1928); its sublimation to flammable acetaldehyde gas gave it this function. The application of metaldehyde in controlling slugs has been known about since as early as 1934 (Gimingham, 1940) and it is now widely used in both agricultural fields and domestic gardens. It is typically applied as a pelleted bran bait that inhibits slug feeding and causing death after exposure (Wedgwood & Bailey, 1988), on a cellular level the effects

include the distention and disintegration of the Golgi apparatus and endoplasmic reticulum in the mucus cells of slugs (Triebskorn et al., 1998).

At the time of writing, the latest report on pesticide usage in the United Kingdom available from Department for Environment Food & Rural Affairs covers 2014. In that year metaldehyde accounted for 87% of all recorded molluscicide applications on agricultural fields, by application area. 112 tonnes were applied over 920 thousand hectares (21% of all surveyed arable land used to grow crops), primarily on wheat, oilseed rape and potato crops (Garthwaite et al., 2015). The number of hectares to which the pesticide was applied to particular crops is given in Table 1-1.

Wheat	Winter barley	Spring barley	Oats	Rye	Oilseed rape	Linseed	Potato	Beans	Sugar beet	All crops
366,618	16,228	4,725	772	1,402	460,375	1283	64,810	976	3,128	920,317

Table 1-1. Hectares of crops that received metaldehyde treatments in the UK in 2014 – adapted from (Garthwaite et al., 2015)

The quantity used varies and is heavily influenced by the weather, with more frequent applications required to replace that which is washed away by rain (Wilson et al., 2014). For example, in 2008, 466 tonnes were applied over 1.5 million hectares (Garthwaite et al., 2010). Metaldehyde pellets applied to fields are broken up by weathering, and dissolved metaldehyde is washed into watercourses that contain wildlife and may be used for drinking water abstraction. The transfer from field to watercourse can peak at 1-4 days after rainfall (Lazartigues et al., 2012).

1.4 The toxicity of metaldehyde

The lethal effect of metaldehyde on slugs and snails is due to a specific mode of action that damages the mucus cells of the organisms (Triebskorn et al., 1998). The substance is also acutely toxic to vertebrates, but to a lesser degree. The dose required for 50% mortality (LD_{50}) in some terrestrial mammals is 227-690 mg/kg (metaldehyde/body mass) for rats, 290-1250 mg/kg for rabbits, 100-1000 mg/kg for dogs and 181 mg/kg for the Japanese quail. For rainbow trout, the 50% lethal concentration in water was 69 mg/L (Jones & Charlton, 1999; United States Environmental Protection Agency, 2006). In dogs the signs of metaldehyde poisoning indicate that it affects the central nervous system (Bates et al., 2012). Metaldehyde may be expected to depolymerise to acetaldehyde in the acidic environment of a vertebrate gut, and so the toxicity of metaldehyde could be a result of acetaldehyde poisoning. However, the LD_{50} values of acetaldehyde in rats are greater than those of metaldehyde, being reported as 634-687 mg/kg (Sprince et al., 1974) or 1620-2240 mg/kg (Smyth et al., 1951), compared to the value of 227-690 mg/kg given for metaldehyde above. Additionally, metaldehyde was found to be significantly more toxic to dogs than the equivalent dose of acetaldehyde (Booze & Oehme, 1986). Metaldehyde is therefore toxic to mammals in its own right.

Surprisingly little research has been published into the effect of metaldehyde exposure on aquatic invertebrates. Metaldehyde was not found to be toxic to water fleas (United States Environmental Protection Agency, 2006). Some evidence for harm has been suggested in the case of the Pacific oyster. Oysters from two families (i.e. oysters originating from two pairs of parents) that were exposed to 0.1 µg/L metaldehyde for 24 hours showed significant differences in 4 reported attributes. One family showed increased rate of haemocyte mortality measured at 48 and 72 hours and the other showed significant differences in cell size & complexity, non-specific esterase activities, and phagocytic activity (Moreau et al., 2014). However, the significance of these differences appears to have been determined by the repeated application of Wilcoxon-Mann Whitney test, and the differences were not consistent between the two families discussed. This latter observation may be, as the authors suggest, due to genetic differences between the two families, or it may be that the observed differences have stochastic origins. It is not clear if the lack of published studies, beyond those discussed above, on the effect of metaldehyde on aquatic organisms results from insufficient attention being paid to this pesticide thus far, or as a result of the bias against the publishing of "negative" results that do not show an effect (Jennions & Møller, 2002; Young et al., 2008).

Metaldehyde is a common cause of poisoning in domestic pets, with 772 cases reported to the Veterinary Poisons Information Service between 1985 and 2010 (Bates et al., 2012). In Northern Italy, 17.2% of suspected or confirmed cases of animal poisoning (mostly dogs, cats, birds and foxes) were thought to be due to metaldehyde (Bille et al., 2016).

1.5 Pollution of drinking water caused by metaldehyde

Each year approximately 11.5 billion m^3 of water is abstracted for domestic and commercial uses in the UK (Water UK, 2015; Kowalski et al., 2013). Water for domestic use must be ensured to be potable before use. Treatment for drinking water typically involves flocculation of the water by the addition of coagulant and allowing the particles to settle out, followed by passing the water through a single sand filter, or a series of sand filters.

The Drinking Water Directive (European Union Council Directive 98/83/EC) sets the maximum permitted concentration in water for human consumption at 0.1 µg/L for any one pesticide, or 0.5 µg/L total pesticides, and sets a legal requirement for sufficient monitoring to detect failures. Metaldehyde was first identified by water companies in the UK as being present in drinking water above the legal limit in 2008, and so monitoring of metaldehyde levels in reservoirs, water courses and tap water began in 2009, with hundreds of thousands of tests being conducted each year (Chief Inspector of Drinking Water, 2009). Since then, the majority of failures due to pesticides have been

the result of metaldehyde pollution (Figure 1-2). In 2013, 110 of 2269 raw water abstraction points were confirmed to have actual risk of metaldehyde contamination (Colbourne, 2013).

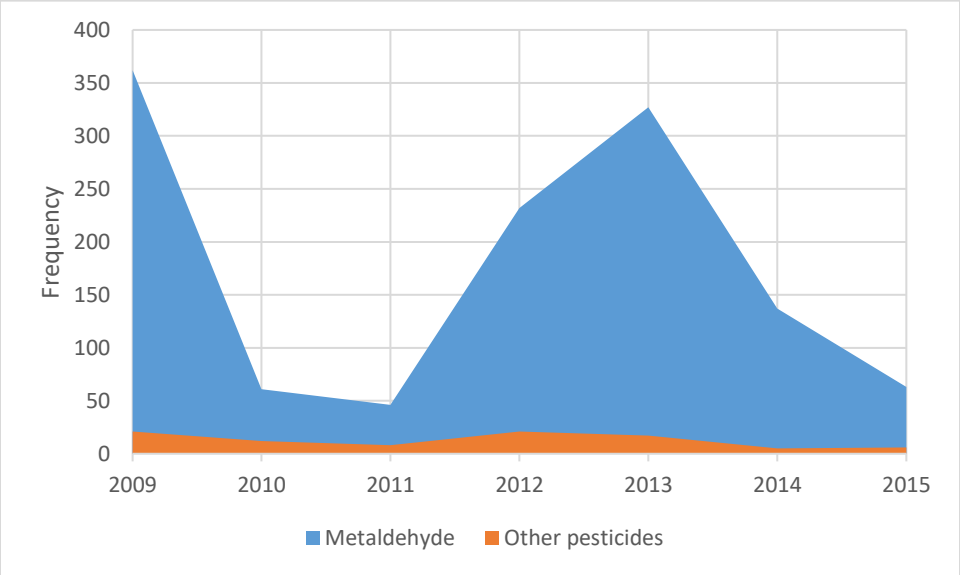


Figure 1-2. Frequency of water quality failures per year due to metaldehyde or all other pesticides. Compiled from the Drinking Water Inspectorate annual regional reports, available from http://www.dwi.gov.uk/about/annual-report.

A survey of 9 water treatment works and various water channels in the catchment area of the river Ouse showed that metaldehyde was frequently present in concentrations greater than the legal limit and that water treatment works did not measurably decrease its concentration (Figure 1-3; Kay & Grayson, 2014). Figure 1-3 also shows the seasonal variation in metaldehyde pollution.

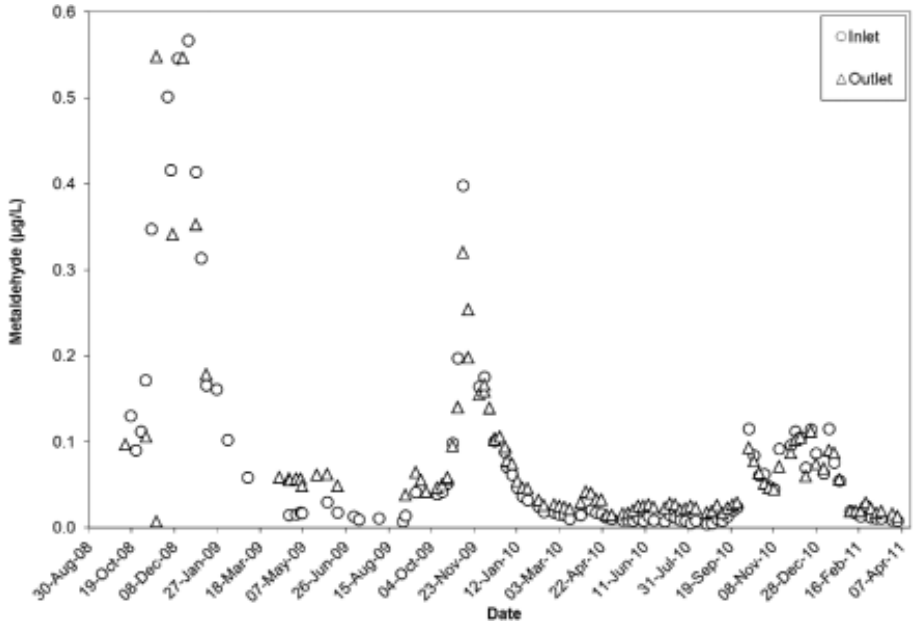


Figure 1-3. Metaldehyde concentration at the inlet and outlet of water treatment works. Taken from (Kay & Grayson, 2014).

The pesticide-manufacturer backed Metaldehyde Stewardship Group was formed in 2009 to publicise the problem and offer best practice advice to farmers to reduce the amount of the substance being transported to water courses. In 2010 a maximum legal application rate of 700 g/ha/year was brought into force by the Chemicals Regulation Directorate. In response to legally binding agreements with the Drinking Water Inspectorate (DWI), water companies have been pursuing a variety of strategies to tackle the problem. These include surveys of the water catchment areas to establish which fields are likely to leach the compound into water courses, communication with farmers to make them aware of the problem, funding the calibration of metaldehyde application equipment, awarding money to farmers if the levels of metaldehyde in water courses of a catchment are reduced below set levels, and paying for farmers to use more expensive alternative molluscicides such as methiocarb or ferric phosphate (Chief Inspector of Drinking Water, 2016). The DWI has instructed that a report on the effectiveness of these measures should be made in 2018, and the DWI has stated that if the problem still exists in 2020 it will assume that “more robust operational measures are required." (Purcell, 2014).

1.5.1 Metaldehyde use in other EU countries

A search of the literature did not retrieve any publications discussing metaldehyde pollution of drinking water sources in other EU countries. A reference to a Swedish water treatment works trialling metaldehyde degrading technologies was made by Tang et al. (2016), indicating that the problem exists there. Studies described above include one on the subject of poisoning by metaldehyde in Northern Italy (Bille et al., 2016) and metaldehyde in French fish ponds (Lazartigues et al., 2012), which show that the compound must be used in those countries.

The quantity of molluscicides, including metaldehyde, sold in 10 EU countries is given in Figure 1-4. The quantity of metaldehyde sold or used in other EU countries may not deviate too far, proportionally, from the amount seen in the UK of 87% of all molluscicides applied. If this is the case, then it is likely that the problem of metaldehyde pollution of drinking water sources exists in other countries covered by the Drinking Water Directive.

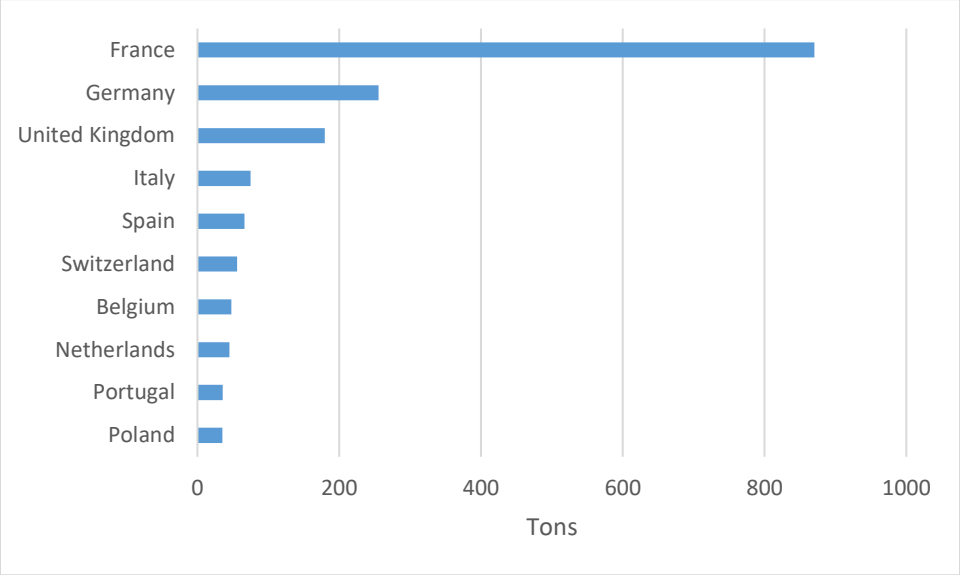


Figure 1-4. Tons of molluscicide sold in 2014 in the 10 European Union countries with highest sales (Eurostat, 2016).

1.6 Chemical and physical methods of removing metaldehyde pollution

The strategies of ameliorating metaldehyde pollution of drinking water sources discussed above all concern the prevention of metaldehyde from entering water courses in the first place. Metaldehyde has been found to be hydrolytically and photolytically stable. Hydrolysis was assessed at 25°C in buffers at pH 5, 7 and 9 over 32 days (Carpenter, 1989a). Photolysis was assessed using dissolved metaldehyde irradiated by a xenon arc lamp at pH 7 and 25°C for 30 days, with and without the addition of acetone (a photolysis sensitiser) (Carpenter, 1989b). No degradation of metaldehyde was observed in either study. Its hydrolytic half-life has been given as 16.8 years (Kegley et al., 2016). Currently no economical method exists to degrade or remove metaldehyde from water. Water companies try to avoid abstracting contaminated water, and when this fails must mix contaminated water with uncontaminated to avoid exceeding the legal limit (Allison, 2015). However, the only methods of quantifying metaldehyde with sufficient sensitivity is solid phase extraction followed by gas chromatography (Environment Agency, 2009), or liquid-chromatography mass-spectrometry (Schumacher et al., 2016) which is expensive and cannot be done *in situ,* leading to delays in the cessation of abstraction.

Researchers are pursuing methods to remove or degrade metaldehyde with the aim that these methods be used in water treatment. Advanced oxidation processes, that use ultraviolet light (UV)/H_2O_2 to produce hydroxyl radicals, could be used to degrade metaldehyde, but dissolved organic matter strongly inhibits the process by scavenging hydroxyl radicals and the method may be

uneconomical (Autin et al., 2013a). Autin et al. (2013b) predicted that the improvement of UV-light emitting diodes may make the process economical by 2020. The legal requirement for new water treatment plants to be effective meant that, despite the above discussed problems, Anglian Water installed a UV/H_2O_2 system with amalgam lamps at the Hall water treatment works in Newton-on-Trent in 2014, with the explicit aim of removing metaldehyde (Waste Water Treatment Online, 2014).

Alternative adsorption materials to activated carbon, which is commonly used in the treatment of drinking water, have been investigated. A graphite intercalating material called Nyrex can adsorb metaldehyde with sufficient affinity to reduce its concentration < 0.1 µg/L and can be regenerated by passing an electrical current through it, converting the metaldehyde to carbon dioxide (Nabeerasool et al., 2015). Phenolic carbon was found to have significantly higher affinity and adsorption capacity for metaldehyde than activated carbon and it was demonstrated to be capable of reducing metaldehyde concentration to < 0.1 µg/L (Busquets et al., 2014), though the possibility of phenolic compounds being released from the material may lessen its desirability as a treatment solution. Tao & Fletcher (2013) demonstrated a commercially available ion-exchange resin, S947 (Purolite, UK), is capable of adsorbing significant quantities of metaldehyde.

Doria et al. (2013) have shown that nano-sized zinc oxide/laponite composites could catalyse the photolysis of metaldehyde with UV treatment down to 350 µg/L. The use of tetraamidomacrocyclic ligands as catalysts for the reaction of peroxide with metaldehyde without UV irradiation results in up to 31% reduction of metaldehyde, with initial concentrations of 56 mg/L. Acetic acid and acetaldehyde are produced in a 3:1 ratio, in a process that is efficient in the use of materials but very slow (Tang et al., 2016). A resin that contains sulfonic acid functional groups (SA-SBA-15) has been found to adsorb and cause the depolymerisation of metaldehyde to acetaldehyde (Tao & Fletcher, 2014). Dissolved metaldehyde treated with SA-SBA-15 results in only acetaldehyde and residual metaldehyde. Tao & Fletcher (2014) therefore propose a mechanism whereby the protonation of an oxygen results in the sequential reduction of the ring's bonds, causing the release of acetaldehyde (Figure 1-5).

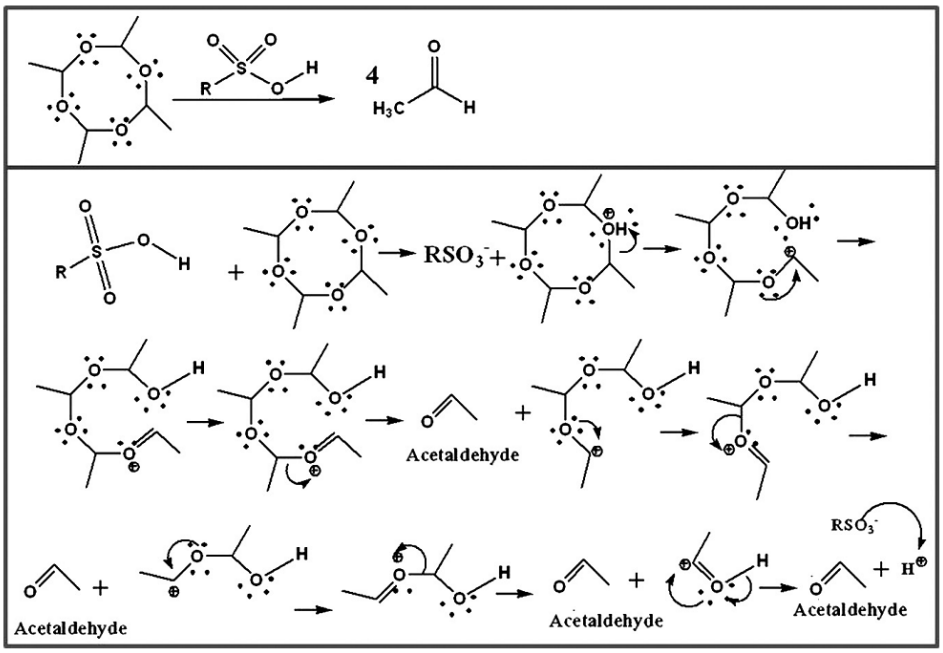


Figure 1-5. Proposed mechanism of sequential depolymerisation of metaldehyde following lysis of a single bond by a sulphone residue. Protonation of an ethyl oxygen results in the scission of the C-O bond and recruitment of an electron *from a lone pair from the next oxygen. This process repeats, releasing acetaldehyde. Taken from Tao & Fletcher (2014).*

The problem of metaldehyde pollution is attracting a lot of interest from researchers and water companies looking for a solution. It is not clear that an optimal solution has yet been achieved as the effective solutions involve high energy inputs or specialised materials that may be expensive to fabricate and have limited life spans.

1.7 The biotic degradation of metaldehyde

Metaldehyde is degraded by soil microorganisms. Most of the available studies that deal with metaldehyde degradation are part of the European Union Draft Assessment Report on metaldehyde (European Food Safety Authority, 2006, Vol. 3, B8). The studies discussed in the report were conducted by the corporation Lonza (Basel, Switzerland) to satisfy regulatory demands.

Cranor (1990a) treated agricultural soil in a microcosm with 10 mg of ^{14}C-labeled metaldehyde per kg of soil and incubated it for a year at 25°C. Acetaldehyde, paraldehyde (0.4 % applied radioactivity at 59 days only) and CO_2 were the identified products of metaldehyde degradation in this study (Figure 1-6). The time taken for 50% degradation of the compound (DT_{50}) was calculated to be 67.2 days, assuming first order kinetics.

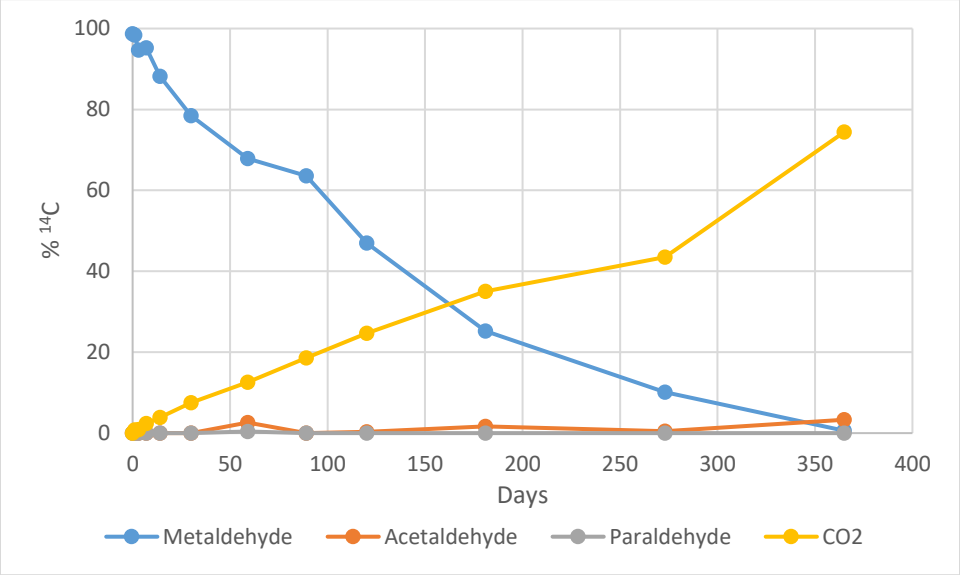


Figure 1-6. Amount of ^{14}C recovered as different compounds, as a percentage of applied radiation, when a soil microcosm was treated with radio labelled metaldehyde. Produced using data from Cranor (1990a).

A second experiment was conducted as described above, but the chamber was flooded with degassed water after 30 days to produce anaerobic conditions (Cranor, 1990b). The degradation of metaldehyde was reported to cease and acetaldehyde accumulates after anaerobic conditions were induced at 30 days (Figure 1-7), compared to the degradation seen in Figure 1-6. The cessation of metaldehyde degradation is not clear in this shorter experiment. The accumulation of labelled acetaldehyde suggests that some metaldehyde was depolymerised after anaerobic conditions were induced.

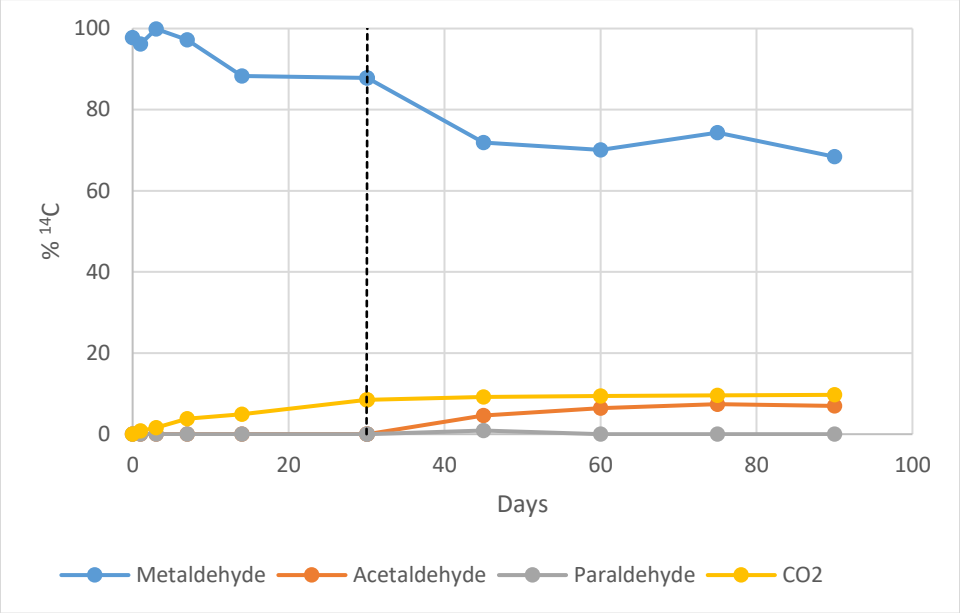


Figure 1-7. Amount of ^{14}C recovered as different compounds, as a percentage of applied radiation, when a soil microcosm was treated with radio labelled metaldehyde and flooded with degassed water at 30 days (flooding indicated by a dashed line). Produced using data from Cranor (1990a).

Möllerfeld et al. (1993) took samples from creeks in the German agricultural areas of Bickenbach and Unter-Widdensheim, which are approximately 100 km apart. Sediment and water samples were treated with 117 µg/cm^3 ^{14}C-labelled metaldehyde and incubated in the dark at 20°C (Figure 1-8). Extraction of metabolites from sediment used dichloromethane and methanol. Acetaldehyde and CO_2 were the detected metabolites. The DT_{50} was calculated as 12.4 and 11.9 days for the Bickenbach and Unter-Widdensheim systems respectively.

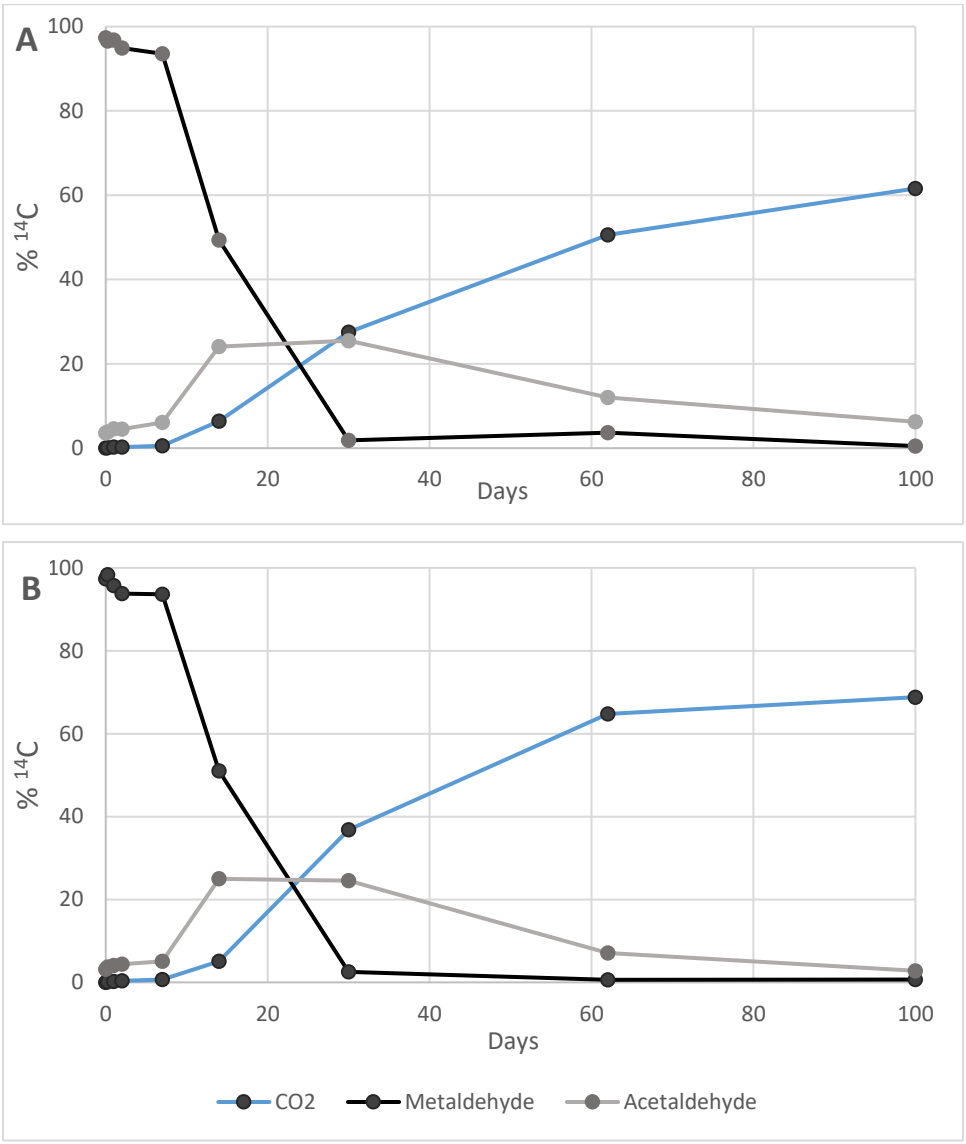


*Figure 1-8. Percentage of applied radioactivity that could be extracted from samples, does not show unextractable applied radiation that remained in the sediment. Chart **A** shows microcosms originating from Bickenbach and **B** shows those from Unter-Widdensheim. Produced using data from (Möllerfeld et al., 1993)*

Two field studies are available. These were not part of the European Union regulatory process and were published in peer reviewed journals. Granules containing 6% metaldehyde were applied to a tobacco field that had a history of metaldehyde treatment. The study was conducted twice in subsequent years. The initial concentrations of metaldehyde were approximately 2 and 1 mg/kg of soil for each year, and the DT_{50} values obtained were 3.5 and 3.9 days (Zhang & Dai, 2006). Metaldehyde powder was sprayed onto 3 different cabbage plots in Beijing, Zhejiang and Hubei resulting in initial concentrations in soil samples of 1.5-6 mg/kg. The determined DT_{50} values were 0.75 – 1.02 days (Zhang et al., 2011a). However, in both of these studies metaldehyde may have been lost due to dissipation into deeper soil or elsewhere, though there was little rain in the 2011

study. One site experienced no rain for the first 10 days of the study, during which time >90% of metaldehyde dissipated.

The biotic degradation of metaldehyde in treated substrates (soil or river sediment) proceeds at very different rates in different experiments; with DT_{50} values of between 67 and <1 days (or a minimum of 12 days if only microcosm studies are considered). Similar results were seen when similar, but geographically separated substrates were tested using the same methods. The substrates will vary in their physical properties, but the most significant determining factor is likely to be the ability of the microbial biomass in the samples to degrade metaldehyde. The abundance of microbes capable of a metabolic activity and the efficiency of the pathway by which they perform that activity will depend on the microbial population's history of exposure to the substrate (Krutz et al., 2010). The rapidity of biotic degradation compared to abiotic recalcitrance suggests the existence of a specifically evolved degradation pathway for this xenobiotic that is present in soil organisms.

1.8 Bioremediation and biosensors

Uncovering the molecular basis for the biotic degradation of metaldehyde could enable the development of technologies and techniques that would aid in the treatment and quantification of metaldehyde in water sources.

The expression of enzymes with the ability to degrade a problematic compound can be used to remediate that compound. For example, purified horseradish peroxidases may be used to treat water or industrial effluent that is contaminated with phenolic compounds (Klibanov et al., 1983), or anthraquinonic dyes (Šekuljica et al., 2015). The cloning and expression of cyanuric acid hydrolases in *Escherichia coli*, followed by the encapsulation of non-viable cells in porous silica beads, produces a material that can be used to remove cyanuric pollution from water, without purifying the enzymes from the other cellular material (Yeom et al., 2015). Enzymes can be improved, in terms of stability, turnover (k_{cat}), and affinity (K_M) for the substrate, by rational mutation of the amino acid sequence of the enzyme. This was performed by Scott et al. (2009) on the atrazine hydroxylase AtzA, yielding an enzyme with a 3-fold reduction in K_M and a 10-fold increase in k_{cat} compared to the wild-type enzyme.

Alternatively, live cells can be used for bioremediation. Organisms capable of the desired metabolism, either naturally or by genetic manipulation, can be introduced to remediate a contaminated substrate (Cordova-Rosa et al., 2009). Alternatively, the population of microorganisms in a substrate can be shifted towards degraders by incubating a sample of it with the compound of interest, possibly combined with the introduction of degrading organisms to the sample (Gentry et al., 2004). Such methods could be used in water treatment works, reservoirs, or at field edges which

have high risk levels of cause metaldehyde pollution in nearby watercourses. The potential for such treatments would depend on the nature of any enzyme discovered; a requirement for co-factors such as adenosine triphosphate to be available for a reaction to occur would limit the enzyme's usefulness in remediation, as would properties such as a short half-life or a narrow set of physical conditions in which the enzyme will function.

To monitor metaldehyde concentrations, most analyses are done following the Environment Agency (2009) methods that require concentration of a sample by solid phase extraction before mass spectrometry. As previously stated, this creates delays between the acquisition of a sample and determination of its metaldehyde content, hindering efforts to prevent metaldehyde contaminated water from entering reservoirs, as well as being a financial burden. Enzymes can be used to construct biosensors that can be used to perform a quantification in real-time and on-site.

Biosensors can utilise changes in an electrical current to produce a measurement of an analyte. Where the catalysed transformation of a target compound releases electrons, enzymes embedded in a conductive material can produce a current when the target compound is oxidised by the embedded enzymes (Wallace et al., 1999; Rather et al., 2012). The oxidation or reduction of an analyte that is coupled with the complementary change in oxidation of a nicotinamide adenine dinucleotide cofactor (β-NAD^+) can be applied in a biosensor. The β-NAD^+ can be attached to the conductive material and the change in current generated, as it is regenerated to its initial oxidation state, can be recorded to quantify the analyte (Rahman et al., 2009). Calorimetric biosensors can be used where the chemical transformation of an analyte results in sufficient change in enthalpy to be measured. Such devices have been miniaturised sufficiently to fit on a chip that requires only 200 nl reaction volumes (Davaji et al., 2014). As with the potential for the development of remediation technologies based on an understanding of the molecular basis for metaldehyde degradation, the potential for the development of biosensors for the compound, and the nature of any biosensor, would depend on the nature of the reactions that degrade it.

1.9 Aims of the project

A biotechnological approach utilising organisms or enzymes that have evolved the specific ability to degrade metaldehyde could provide the basis for a reliable and low cost solution to the detection of metaldehyde pollution and its mitigation. In parallel to the applied outcome, identifying the novel molecules responsible for the transformation of a xenobiotic would increase our understanding of how these capacities evolve in microorganisms.

These outcomes require the isolation of metaldehyde degrading organisms and the identification of the molecular actors that perform the degradation, and therefore the primary aims of the work

described in this thesis are the identification and characterisation of these. For this work to be performed a reliable method of quantifying metaldehyde was required, and so the improvement of an existing quantification method was undertaken. Hypotheses (in italics) and approaches are given:

- *The sensitivity and precision of a liquid chromatography-mass spectrometry method can be improved by optimising the ionisation conditions of the analyte.* Different ionisation voltage and temperatures were used in the quantification of metaldehyde; the strength and consistency of the recorded response were analysed to determine the best conditions (Chapter 2).
- *Organisms that are able to utilise metaldehyde as the sole carbon and energy source are present in soil and other substrates that have been in contact with metaldehyde.* Minimal media containing only metaldehyde as the carbon source were used to culture any such organisms under laboratory conditions (Chapter 3).
- *The molecular basis of metaldehyde degradation in the isolated organisms exists in their genetic material.* The genomic sequencing of isolated organisms was undertaken (Chapter 3).
- *Enzymes involved in the catabolism of metaldehyde are more highly expressed when cells are grown with metaldehyde.* Protein gels were used to look for more highly expressed proteins in metaldehyde grown cells (Chapter 4).
- *The kinetic nature of enzymes in the metaldehyde degrading pathway can be determined by quantifying the rate of substrate removal* in vivo. The kinetics of the pathway were investigated using direct quantification of metaldehyde, and oxygen utilisation. Step-wise models were employed in an attempt to determine the kinetics of individual enzymes in the pathway (Chapter 4).
- *The presence of metaldehyde catabolism in isolated organisms results from the acquisition of foreign DNA; this DNA can be identified by comparative genomics.* Candidates for the primary metaldehyde degrading enzyme were identified by comparing the sequences of predicted proteins in two isolates and closely related strains that do not degrade metaldehyde (Chapter 5).

Chapter 2: Optimisation of a liquid-chromatography mass-spectrometry method for quantifying metaldehyde

2.1 Introduction

Chromatography is used to separate complex mixtures of compounds in a sample. Compounds are separated by their retention time along a pathway. This retention time is dependent on the compounds' differing affinity to a mobile phase, that carries the sample along the path, and to a stationary phase, that retards this progress. Reverse-phase liquid chromatography (LC) is the predominant type discussed in this chapter. This uses a mostly aqueous mobile phase that flows through a column packed with a solid, hydrophobic, stationary phase. The retention time of compounds will therefore be dependent on their overall hydrophobicity and hydrophilicity. The separated mixture that flows from an LC column can then be analysed by a variety of methods to identify and quantify compounds in the mixture.

Metaldehyde may be quantified by mass-spectrometry (MS) following LC. MS, of the type discussed here, records the charge produced by ions striking a detector. The method of producing a stream of ions that is used in this chapter is electrospray ionisation (ESI). The sample solution passes through a positively or negatively electrically charged capillary tube, exiting the tube as a fine spray of charged droplets. The mobile phase liquid is caused to evaporate by a flow of heated inert gas, concentrating the charge in the remaining droplet and causing the ejection of ions from the droplet due to their mutual repulsion. Uncharged analytes can be given a charge by forming an adduct with ions in the mobile phase.

Mass selection can be used to filter the ion stream for a particular mass-charge ratio (*m/z*) allowing quantification of a specific compound with known *m/z* ratio against a standard of known concentration. Mass selection is achieved by altering the velocity of the ions using an electric field, the degree of deflection of an ion is dependent on its *m/z*, meaning that the strength of the electric field can determine which ions reach the detector. Tandem MS can be used to positively identify the compound being quantified. After the initial mass selection, the ion stream is directed through a collision chamber where the compounds are fragmented by collision with inert gas in the chamber. One or more fragments generated thusly are then mass selected and quantified. Different compounds that have similar masses will most likely have different fragmentation profiles, allowing them to be distinguished by this second step.

Several methods have been employed to quantify metaldehyde. Brown et al. (1996) describe heating a sample to depolymerise the metaldehyde to acetaldehyde, which is then reacted with 1,3-cyclohexanedione and forms a fluorescent product that can be quantified by UV excitation following

LC. The direct quantification of metaldehyde, without first derivatising it, requires chromatography followed by MS. Gas chromatography has been used to quantify metaldehyde, for example by Saito et al. (2008). Gas chromatography also depolymerises metaldehyde. Liquid chromatography-tandem mass spectrometry (LCMS) following solid phase extraction (SPE) to concentrate a water sample allows for detection of metaldehyde at 3 ng/L and a limit of quantification (LOQ) down to 10 ng/L (Li et al., 2010).

For the purposes of this project it was important to be able to distinguish between metaldehyde and the acetaldehyde that may be produced as a metabolic intermediate. It was also desirable that the method involves as little pre-processing as possible. An LCMS method that does not require SPE or derivatisation had been developed at the Food and Environment Research Agency (FERA). It was reported by FERA staff that the method being used there suffered from high levels of multiplicative noise leading to high relative standard deviation of quantifications compared to the analysis of other compounds, and a decreasing response over time leading to problems where the known standards give different responses at different parts of the run and insufficient sensitivity later in the run. In an attempt to ameliorate these problems, method development was undertaken in collaboration with FERA, with the aim that the method be used in further experiments that make up this thesis.

2.2 Materials and methods

2.2.1 LC and MS protocols

2.2.1.1 LC columns

In most experiments LC was performed using the Kinetex XB-C18 50 × 2.1mm column (Phenomenex, Macclesfield, UK) with 2.6 µm particles. Where specified, the Atlantis T3 2.1 mm × 100 mm, 3 µm C18 particles (Waters UK, Elstree, UK) was used.

Unless otherwise specified, LC was performed with a flow rate of 0.5 ml/min using a mixture of 1 mM ammonium acetate prepared with ultrapure water, and methanol (J. T. Baker brand; VWR, Leicester, UK) using a gradient flow (Figure 2-1 and Table 2-1). Metaldehyde has a much greater affinity for methanol than water and so increasing methanol concentration reduces the amount of time it takes to be eluted from the column. Prepared samples were stored at 10°C prior to injection and injected onto the LC column by an autosampler.

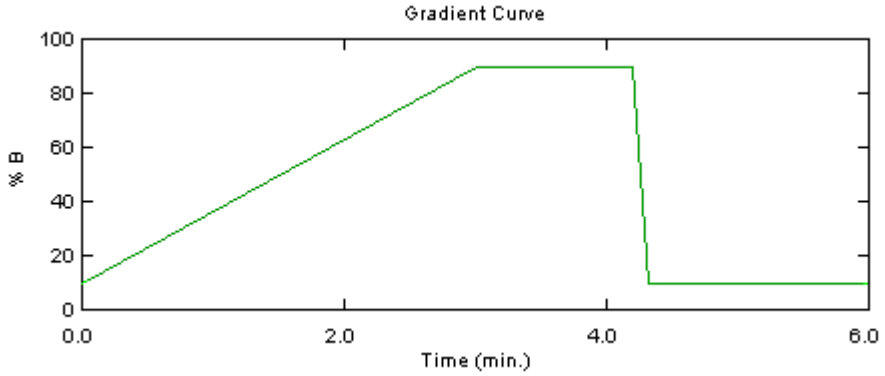


Figure 2-1. The proportion of methanol (% B) present in the LC mobile phase when quantifying metaldehyde. Times shown here are only accurate when flow rate is 0.5 ml/min.

	Gradient step time (minutes) when flow rate is:			
% methanol	**0.6 ml/min**	**0.5 ml/min**	**0.4 ml/min**	**0.3 ml/min or double length**
10	0	0	0	0
90	2.5	3	3.75	5
90	3.5	4.2	5.25	7
10	3.6	4.32	5.4	7.2
10	5	6	7.5	10

Table 2-1. Proportion of methanol in the mobile phase time gradient at all flow rates used when quantifying metaldehyde.

2.2.1.2 Default MS conditions

The default method developed by FERA staff used a flow rate of 0.6 ml/min, ESI capillary voltage of 5500 V and desolvation temperature of 500°C. During chromatography, metaldehyde is given a charge via an NH_4^+ adduct encountered in the mobile phase and so the precursor ion is selected at the mass of metaldehyde + ammonium: *m/z* = 194.1. Fragmentation is achieved using a collision energy of 12 eV. The product ions are observed at *m/z* = 106 and *m/z* = 62 using multiple reaction monitoring with a dwell time of 160 ms. These product ions were determined empirically, prior to the work described here, by scanning the post-collision mass spectrum for the strongest responses. Fragments with formulae $C_4H_8O_2$ and C_2H_4O with ammonium ion adducts have molecular masses of 106 and 62 respectively. A possible scheme for these ion transitions are shown in Figure 2-2.

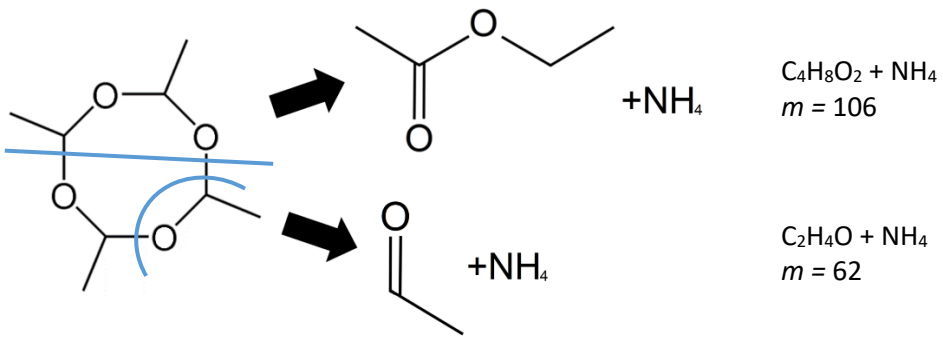


Figure 2-2. Possible ion transitions for metaldehyde following fragmentation during tandem-MS that would give m/z *values of 106 and 62. Blue lines show the portions of metaldehyde that could form the fragments indicated by arrows.*

2.2.1.3 Preparation of standards

To quantify metaldehyde from culture samples it was decided, after consultation with FERA staff, that no more than the equivalent of 0.1 µl of filtered culture should be injected for each measurement to avoid putting too much salt into the MS, which can damage it. Matrix matched standards were prepared using filtered cultures diluted with DI water with the proportion of culture filtrate depending on the injection volume used in the particular experiment to match the 0.1 µl maximum. For example, a 50-fold dilution of culture would be required for a 5 µl injection. *A. calcoaceticus* E1 cultures were grown in standard conditions with 5.9 mM methanol as the carbon sources. The cultures were filtered through polyethersulfone membranes with 0.22 µm pore size (Millipore, Billerica, USA) and the filtrate mixed with different proportions of 100 mg/L metaldehyde (Acros Organics, New Jersey, USA) and DI water to make the standards.

Prepared standards and samples were stored in amber vials with 8 mm screw caps with silicone/ polytetrafluoroethylene septa (Thermo Fisher Scientific, Waltham, USA)

2.2.1.4 Conditioning injections

When specified, to stabilise response prior to measurements being made, 100 µl injections of high metaldehyde standards (typically 100 µg/L) were made onto the LC column. Typically, 8-10 injections would be made.

2.2.1.5 Data analysis

Data were recorded using Analyst V1.5.2 (SCIEX, Warrington, UK). Data were smoothed across 5 points. Integration of the peaks was done by the software using the IntelliQuan Automatic IQA II method. Peaks were reviewed manually and where the analyte peak was not properly integrated, it was integrated manually. Calibration curves using known standards used linear regression with 1/x weighting.

2.2.1.6 Kolmogorov-Smirnov tests

Kolmogorov-Smirnov (KS) tests are used to determine if differences seen in the relative standard deviation (RSD; standard deviation/mean expressed as a percentage) of measurements made with different assay conditions are significant. The KS test determines if two distributions are significantly different. Individual values from an ordered set of observations are charted against the cumulative fraction of observations charted. Two sets of observations are charted and the maximum deviation on the y-axis between the two lines is determined, to give the statistic *D.* This statistic, along with the degrees of freedom in the dataset, is then used to obtain the probability (p) that the two sets of observations come from the same distribution. An example KS chart is shown in Figure 2-3.

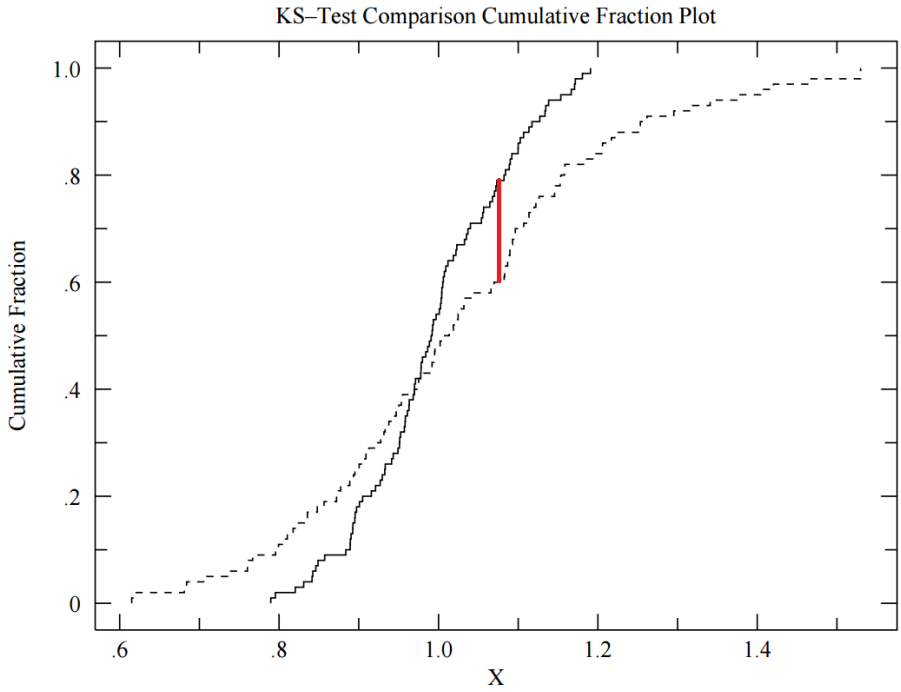


Figure 2-3. A KS chart showing two randomly generated sets of observations (n = 100) with different standard deviations. Each generated value (X_n) is plotted in order along the x-axis with the y-axis giving the fraction of X values that are less than or equal to X_n. The red line indicates the largest vertical deviation between the two sets, and therefore the D *statistic. The D statistic and the degrees of freedom in the data are used to calculate the likelihood that the two sets of data are from the same distribution. Figure generated using http://www.physics.csbsju.edu/stats/KS-test.html (Kirkman, 1996).*

So that the variance of the sets of measurements with different means could be tested, the sets were normalised by dividing each observation by the mean for those measurements. This is so that only the variance from the mean is tested. Both product ions are included in a single distribution, after normalisation, in order to increase the power of the test. The assumption is that the two product ions represent separate measurements of the same primary ion. They are derived from the same stream of ions produced by previous steps in the analysis, and while they will have their own separate sources of noise, any difference in noise resulting from changes to the conditions will affect both equally.

2.3 Results

2.3.1 Decreasing response during, at least, the first 40 injections when analysing metaldehyde

A preliminary run was conducted using default conditions (Section 2.2.1.2). 50 µl injections were made using 10 µg/L metaldehyde, 0.2% culture matrix standard aliquoted into several vials. The peak areas of the two product ions are shown in Figure 2-4.

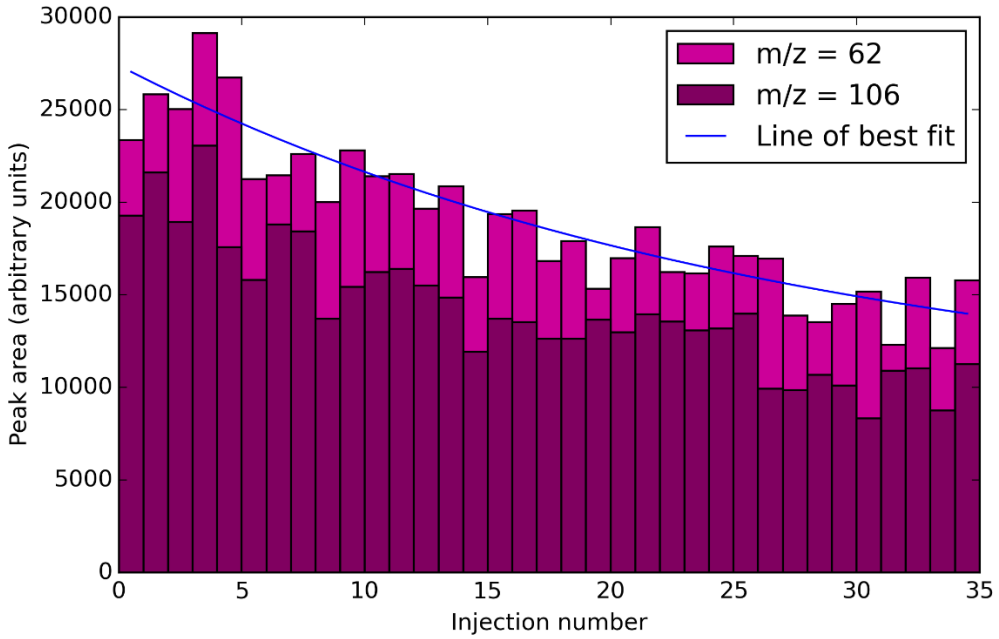


Figure 2-4. Paired bars showing peak areas of product ions from repeated 50 µl injection of 10 µg/L metaldehyde standards. The line of best fit shows the results of an exponential decay equation.

There is a decrease in response with each subsequent injection. Fitting the data to an exponential decay equation (below) resulted in the fitted parameters; A = 18257±435, K = -0.036±0.019, C = 877±502.

$$y = Ae^{-Kx} + C$$

Based on this observation it was decided that when testing different LCMS conditions, 20 repeated injections using a single condition would be used to condition the column so that response decreasing in this way would be clearly visible in a set of injections and be distinguishable from differences in quantification arising from experimentally altered conditions.

2.3.2 No obvious sources of noise in parts of default protocol that take place before MS when analysing metaldehyde

2.3.2.1 Injection volume test using default method conditions

To achieve the LOQ of 0.1 µg/L required for regulatory work carried out by FERA, injections of up to 100 µl were sometimes required – larger than the usual 1-5 µl injections used for other applications. A hypothesised reason for the unreliability of metaldehyde quantification was inconsistent injection volumes by the injection system at these unusually large volumes. An experiment testing the impact of different injection volumes was conducted. 500 pg of metaldehyde was injected each time, using metaldehyde concentrations given in Table 2-2. Standards were injected in sets of 20 for each, in order of increasing volume, and then 20 times in decreasing volume. This palindromic pattern allows detection of long term changes in response by direct comparison of sets of the same standard at different parts of the run. The first and second sets of 20 are referred to as sets A and B in the results below.

Standard conc. (µg/L)	Injection vol. (µl)
100	5
50	10
25	20
10	50

Table 2-2. Metaldehyde concentrations of standards and injection volumes used in this set of experiments.

Mean retention time was 1.59±0.01 min. Figure 2-5 shows the peak area of each product ion for each injection. The first 40 injections appear to show a drop in response similar to that seen in Figure 2-4. Beyond this no other general trends are visible. The peak areas are smaller with 10 µl injections than other injection volumes, shown by the mean values in Figure 2-6. This is also seen in all subsequent experiments in this section. As the main objective of this experiment was to identify quantification errors resulting from imprecise injection volumes, all injections at a particular volume were made using a single standard preparation. It is possible but unlikely that the injection system is consistently inaccurate in this way, but more likely that the 50 µg/L standard used in the 10 µl injections has a concentration lower than intended. The means and standard deviations of peak areas of the first two sets (5 µl and 10 µl) can be seen to be affected by decreasing response with increasing injection number in Figure 2-5 and Figure 2-6, and so only set B of both standards was considered when evaluating the precision of different injection volumes.

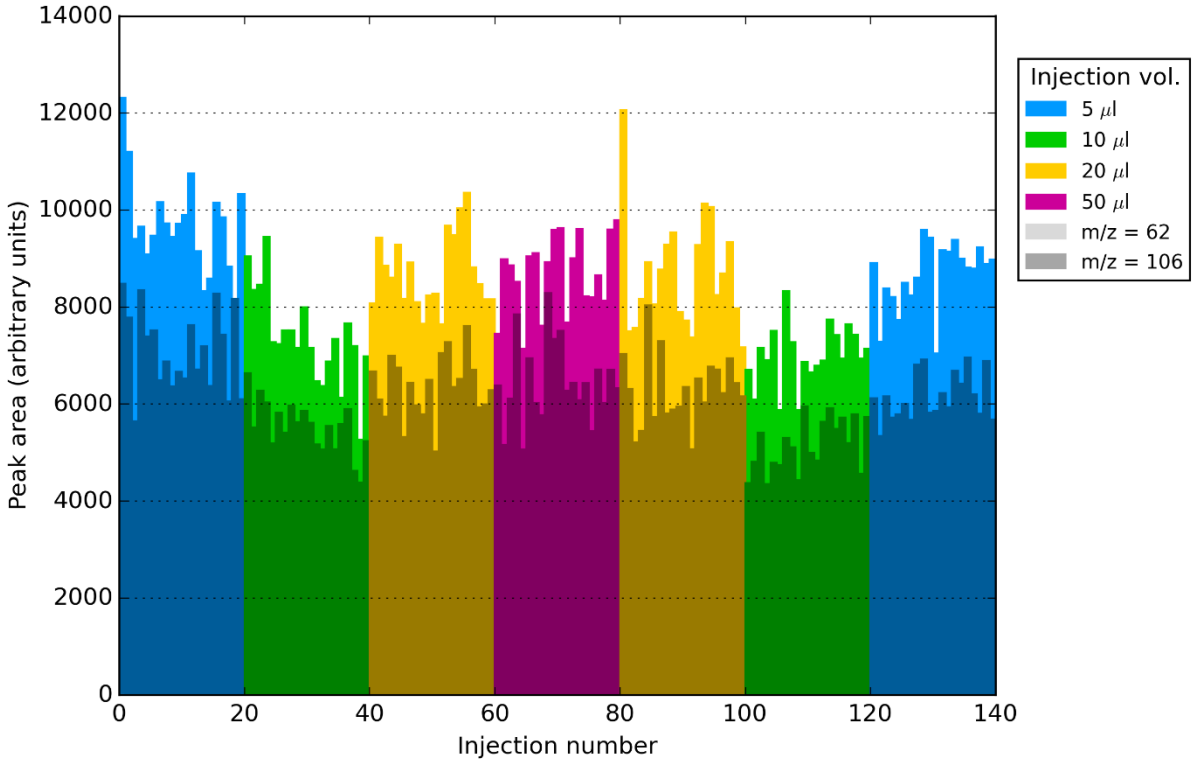


Figure 2-5. Peak areas resulting from the measurement of 500 pg injections of metaldehyde using LCMS with default ionisation conditions. Injection volumes and product ions are indicated by colour and shade respectively.

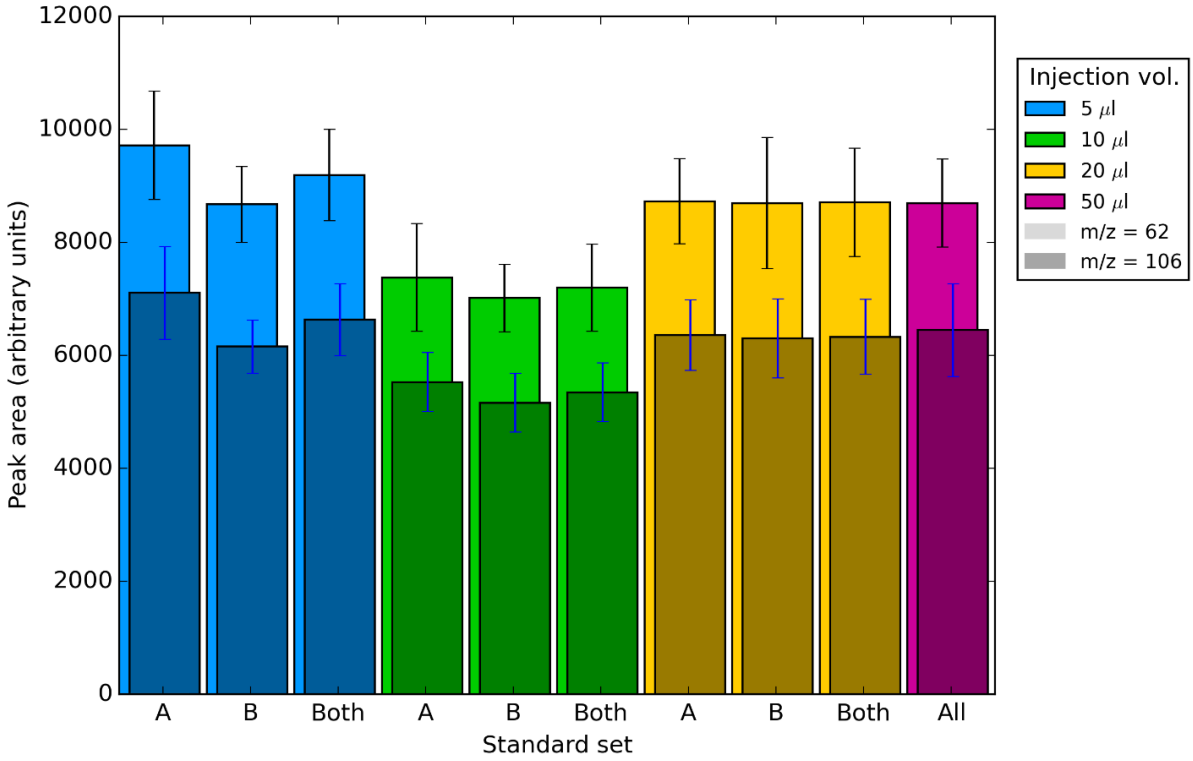


Figure 2-6. Mean response and standard deviation of peak areas of sets of 500 pg of metaldehyde measured by LCMS using default ionisation conditions (n = 20 per set).

The RSD of the different sets are given in Table 2-3. The high RSD of 20 µl, set B, *m/z* = 62 is largely a result of injection number 81 (Figure 2-5) which has a large peak area of 12066. This is an outlier as defined by Tukey (1977, pp.43–44). The interquartile distance (IQD) is the difference between the first and third quartiles (Q1 and Q3). A value that is greater than Q3, or less than Q1, by >1.5 times the IQD is defined as an outlier. The peak area of the 81st injection in Figure 2-5 is 2.25 times the IQD from Q3. This measurement does not appear to be part of a general trend of significantly decreased precision at higher volumes.

µl	Set	RSD (%) m/z = 62	m/z = 106
5	A	9.8%	11.5%
5	B	7.8%	7.6%
5	Both	8.8%	9.6%
10	A	12.9%	9.5%
10	B	8.5%	10.1%
10	Both	10.7%	9.8%
20	A	8.6%	9.9%
20	B	13.3%	11.2%
20	Both	11.0%	10.5%
50	N/A	8.9%	12.7%

Table 2-3. RSD of peak areas of sets of 500 pg injections of metaldehyde measured by LCMS using default ionisation conditions (n = 20 per set). The redness of each cell is proportional to the value of the cell.

To determine the significance of the differences in RSD, KS tests were employed to assess each distribution all others. Set A of the 5 and 10 µl standards was excluded due to the instability of response in the first 40 injections. The different distributions are not significantly different (Table 2-4).

µl/µl		5	10	20
10	*D*	0.1		
	p	0.98		
20	*D*	0.17	0.12	
	p	0.35	0.77	
50	*D*	0.12	0.1	0.07
	p	0.89	0.98	0.99

Table 2-4. Results of Kolmogorov-Smirnov tests of pairs of normalised distributions of peak areas from 500 pg injections of metaldehyde using default conditions. The statistic D is the greatest difference in the cumulative fraction in the distributions, p is the probability of this difference being present by chance given the degrees of freedom present in the data.

The same series of injections were repeated in subsequent experiments investigating LC conditions. It was not thought that there would be any interaction between injection volume and the other parameters of the chromatography but this allowed comparison to the default conditions without an additional run using a single injection volume.

2.3.2.2 Standards assayed with Atlantis T3 column

As a comparison the same sets of standards described in the previous section were quantified as described in methods using the Atlantis T3 column with a flow rate of 0.3 ml/min and double length gradient (Table 2-1). This column and protocol was previously used by FERA but was abandoned as the Kinetex column requires a third of the run time per sample. Peak areas for each injection and means for sets of standards are shown in Figure 2-7 and Figure 2-8. Mean retention time was 4.09±0.02 min. The Atlantis took 40 injections for response to stabilise. The most likely explanation for the increasing response is analyte being carried over (not eluting fully) between injections.

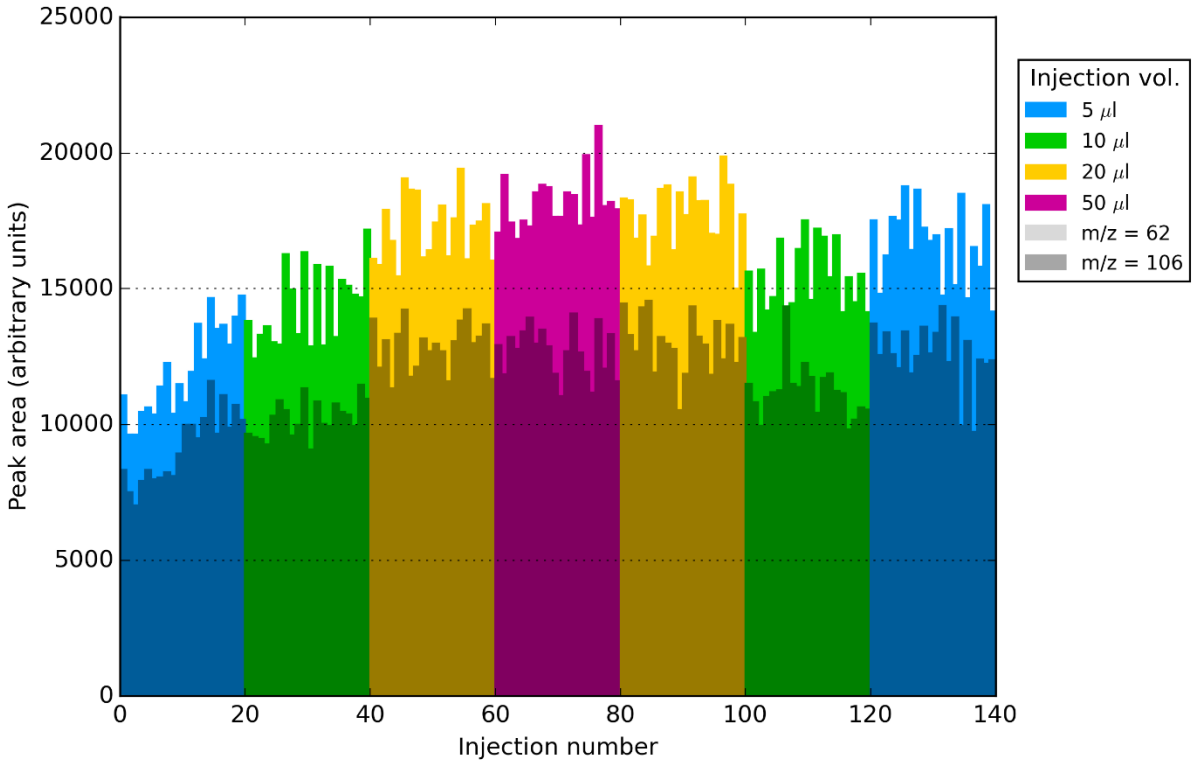


Figure 2-7. Peak areas of both product ions resulting from the measurement by LCMS of 500 pg injections of metaldehyde, of different volumes. An Atlantis T3 column was used.

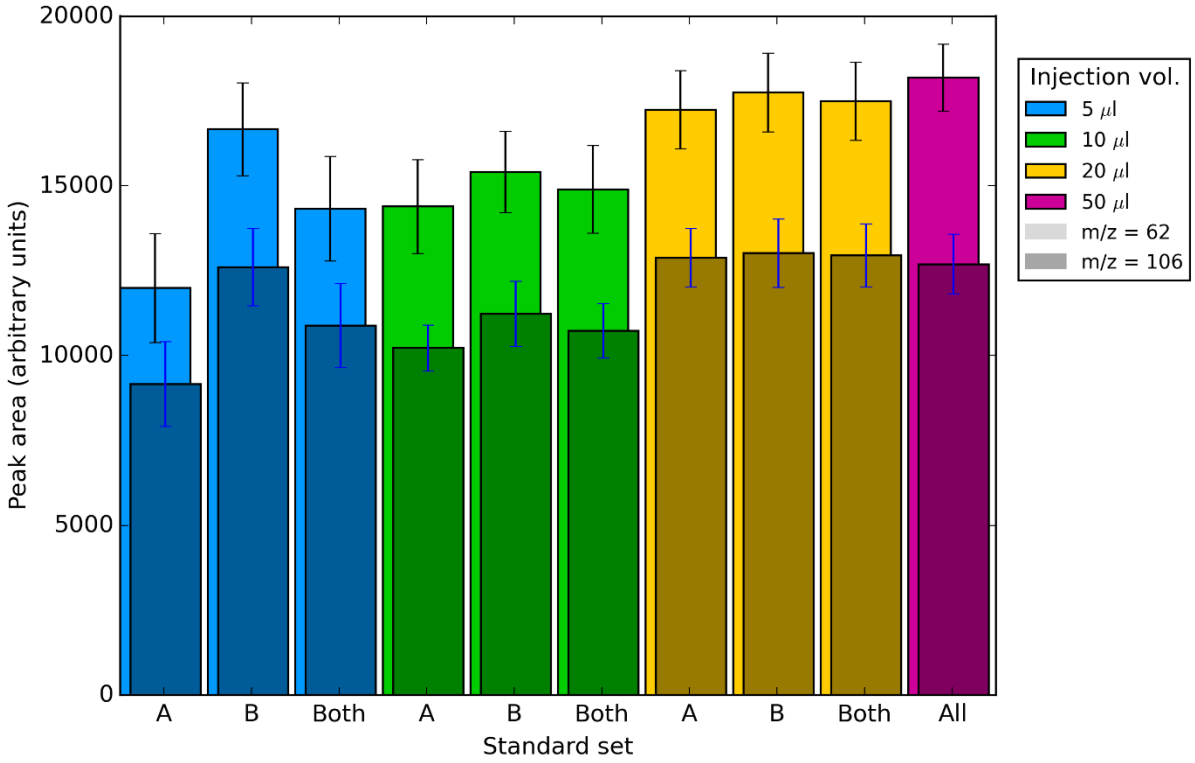


Figure 2-8. Mean and standard deviation (n = 20) of peak areas of both product ions resulting from the measurement by LCMS of 500 pg injections of metaldehyde, of different volumes. An Atlantis T3 column was used to perform the chromatography.

µl	Set	RSD (%) *m/z* = 62	RSD (%) *m/z* = 106
5	A	13.3%	13.7%
5	B	8.2%	9.0%
5	Both	10.8%	11.3%
10	A	9.6%	6.5%
10	B	7.8%	8.5%
10	Both	8.7%	7.5%
20	A	6.7%	6.7%
20	B	6.5%	7.8%
20	Both	6.6%	7.2%
50	N/A	5.4%	6.9%

Table 2-5. RSD (n = 20) of peak areas of both product ions resulting from the measurement by LCMS of 500 pg injections of metaldehyde, of different volumes. An Atlantis T3 column was used to perform the chromatography. The redness of each cell is proportional to the value of the cell.

A general trend of decreasing variance with increasing injection volume was observed

Table 2-5), however paired KS tests for the normalised values of each standard combination

(excluding set A for 5 and 10 µl standards) do not show significant differences between them (Table 2-6).

µl/µl		5	10	20
10	*D*	0.15		
	p	0.72		
20	*D*	0.1	0.1	
	p	0.94	0.94	
50	*D*	0.15	0.15	0.12
	p	0.72	0.72	0.77

Table 2-6. Results of Kolmogorov-Smirnov tests of pairs of normalised distributions of peak areas resulting from the measurement by LCMS of 500 pg injections of metaldehyde, of different volumes. An Atlantis T3 column was used to perform the chromatography.

The overall RSD of peak areas normalised to the mean for that standard is 10.0% for standards run on the Kinetex column and 7.5% for those run on the Atlantis column, though the difference in distributions is not significant (KS test; $D = 0.09$, $p = 0.3766$).

2.3.2.3 Reduced flow rate

Though the difference in variance between the Atlantis and Kinetex columns was not significant, a variant of the default Kinetex protocol, using a double length gradient (Table 2-1), was tested using the Kinetex column. Other aspects of the analysis used the default conditions. Results are shown below in Figure 2-9, Figure 2-10 and Table 2-7.

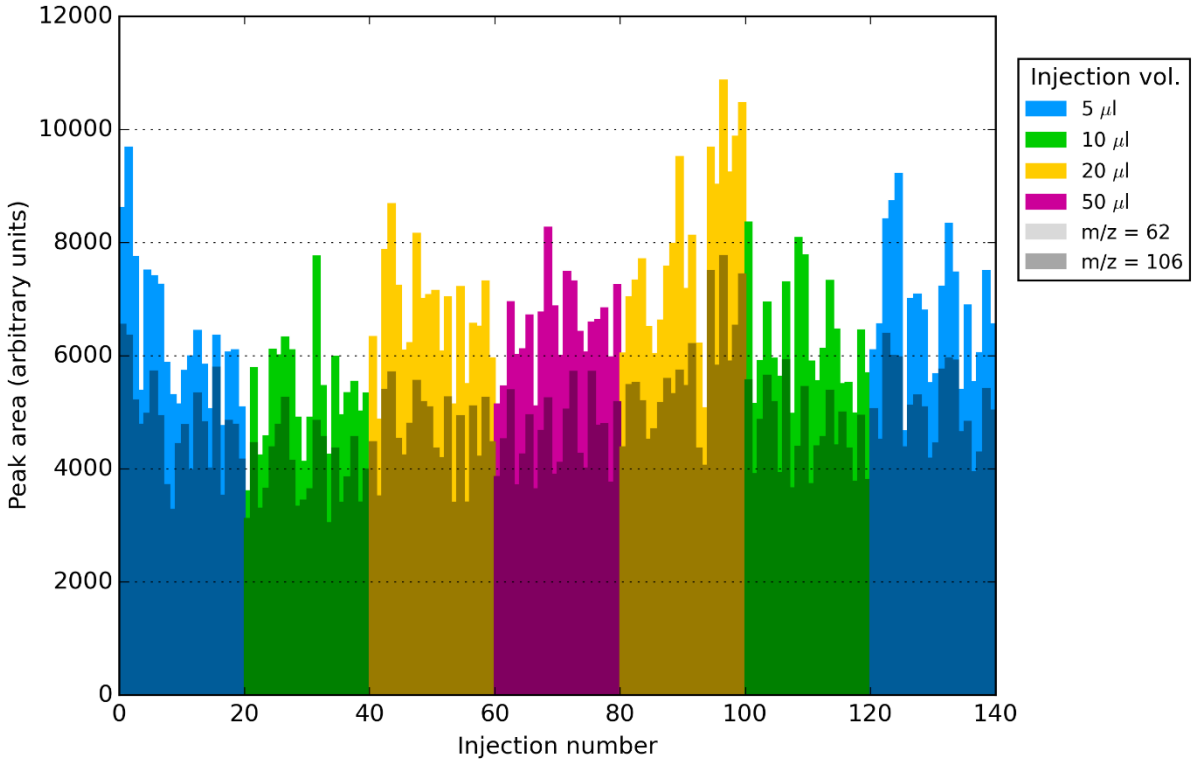


Figure 2-9. Peak areas of both product ions resulting from the measurement by LCMS of 500 pg injections of metaldehyde, of different volumes. A double length time gradient was used during the chromatography.

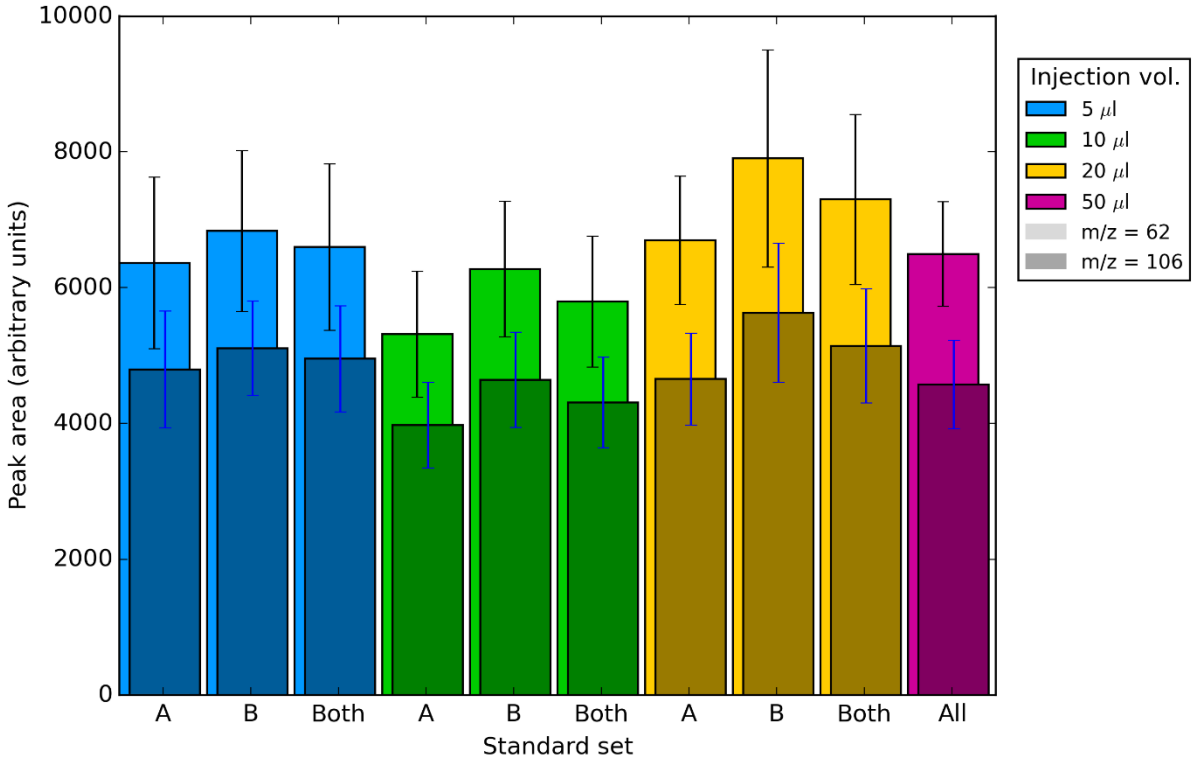


Figure 2-10. Mean and standard deviation (n = 20) of peak areas of both product ions resulting from the measurement by LCMS of 500 pg injections of metaldehyde, of different volumes. A double length time gradient was used during the chromatography.

µl	Set	RSD (%) *m/z* = 62	*m/z* = 106
5	A	19.8%	17.9%
5	B	17.3%	13.6%
5	Both	18.6%	15.8%
10	A	17.5%	15.9%
10	B	15.9%	15.1%
10	Both	16.7%	15.5%
20	A	14.1%	14.5%
20	B	20.3%	18.2%
20	Both	17.2%	16.3%
50	N/A	11.9%	14.3%

Table 2-7. RSD of the peak areas of both product ions resulting from the measurement by LCMS of 500 pg injections of metaldehyde, of different volumes. A double length time gradient was used during the chromatography. The redness of each cell is proportional to the value of the cell.

No overall trends are visible in the peak areas (Figure 2-9). This experiment was conducted using a column that had been used to analyse metaldehyde samples immediately prior and so had already stabilised in response. The mean peak areas shown in Figure 2-10 are lower than those shown for the default conditions in Figure 2-6 and the RSD (Table 2-7) are significantly greater than those shown in Table 2-3 (KS test of the distribution of peak areas normalised to standards' means: $D = 0.25$; $p = 1.1 \times 10^{-6}$).

The mean retention time was 1.79±0.01 min. This is only 6 seconds longer than the default conditions, meaning the elution of metaldehyde occurs at a different point in the gradient curve of aqueous:methanol.

2.3.2.4 0.5 mM ammonium acetate

In the development of the method at FERA no lower concentrations than 1 mM ammonium acetate in the mobile phase had been attempted (though higher concentrations had) and so a run using the default conditions and standards previously described was used replacing 1 mM with 0.5 mM ammonium acetate as the aqueous solvent in the mobile phase. Mean retention time in this experiment was 1.59±0.01 min

Figure 2-11, Figure 2-12 and Table 2-8 show the peak areas of each injection, the mean and standard deviation of sets of these injections, and the RSD of sets. The alteration of aqueous buffer resulted in a large increase in RSD when compared to default conditions where the median RSD was 9.35% (Table 2-3).

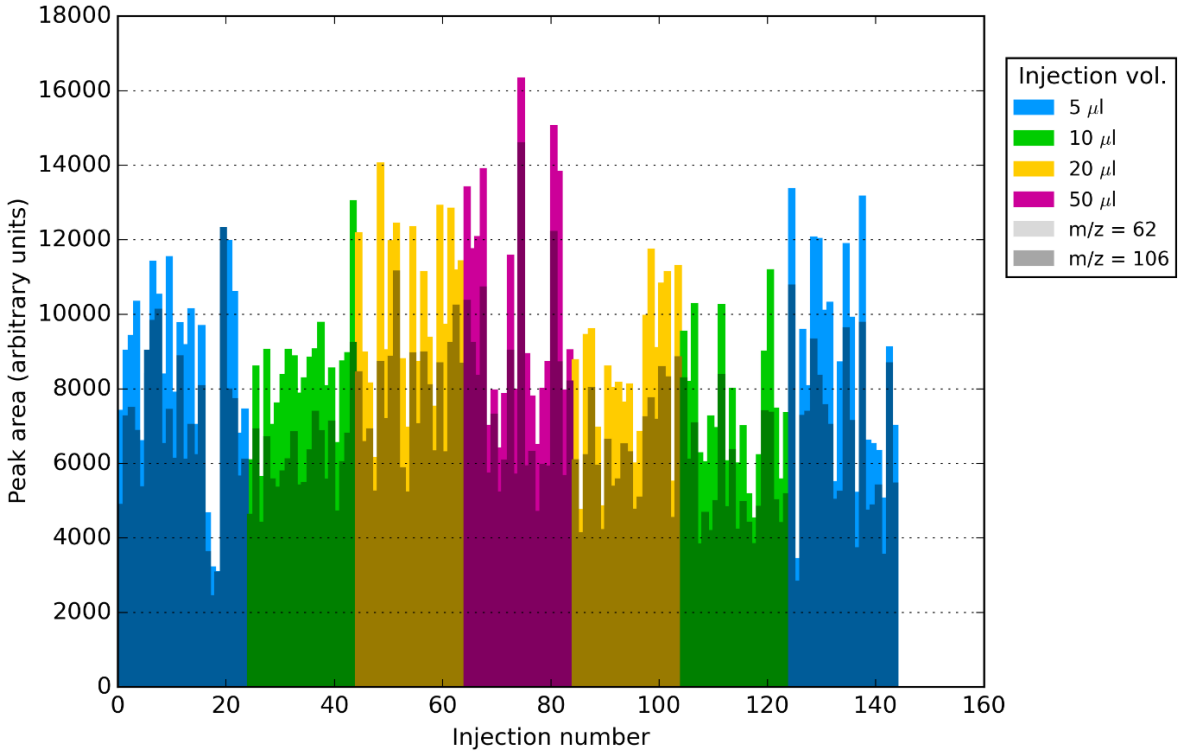


Figure 2-11. Peak areas of both product ions resulting from the measurement by LCMS of 500 pg injections of metaldehyde, of different volumes. The mobile phase used for chromatography was 0.5 mM ammonium acetate.

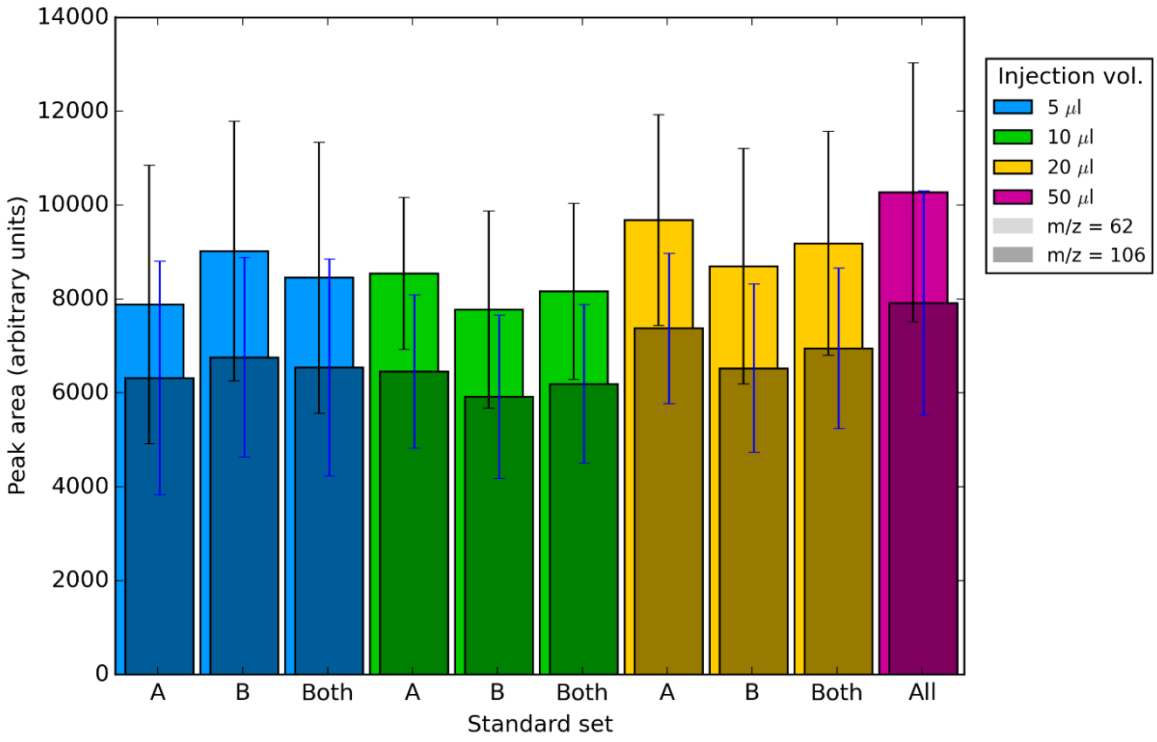


Figure 2-12 Mean and standard deviation (n = 20) of peak areas of both product ions resulting from the measurement by LCMS of 500 pg injections of metaldehyde, of different volumes. The mobile phase used for chromatography was 0.5 mM ammonium acetate.

µl	Set	RSD (%) m/z = 62	RSD (%) m/z = 106
5	A	29.1%	31.2%
5	B	30.7%	31.4%
5	Both	29.9%	31.3%
10	A	18.9%	25.3%
10	B	27.0%	29.4%
10	Both	23.0%	27.4%
20	A	23.2%	21.7%
20	B	28.8%	27.5%
20	Both	26.0%	24.6%
50	N/A	26.9%	30.1%

Table 2-8. RSD of the peak areas of both product ions resulting from the measurement by LCMS of 500 pg injections of metaldehyde, of different volumes. The mobile phase used for chromatography was 0.5 mM ammonium acetate. The redness of each cell is proportional to the value of the cell.

2.3.3 The desolvation temperature significantly affects response and capillary voltage affects precision during the quantification of metaldehyde

A good candidate for the stage at which noise is produced during MS is the ionisation of the sample. For this reason, assays of the ESI capillary voltage and desolvation temperature were conducted. A range of voltages between 4.0 kV and 6.0 kV were tested. Lower temperatures were investigated as it was hypothesised that heat induced decomposition of metaldehyde may affect the measurement.

The first two experiments described here used the repeated quantification of 50 µl injections of 10 µg/L metaldehyde standards. Conditioning injections were used prior to experimental injections to stabilise the signal before measurement began. In the first experiment a range of capillary voltages were tested at 500°C, followed by some voltages at lower desolvation temperatures (Figure 2-13A). It was clear that lower ionisation temperatures resulted in higher peak areas and so, in the second experiment, a similar range of voltages was tested at 425°C (Figure 2-13B).

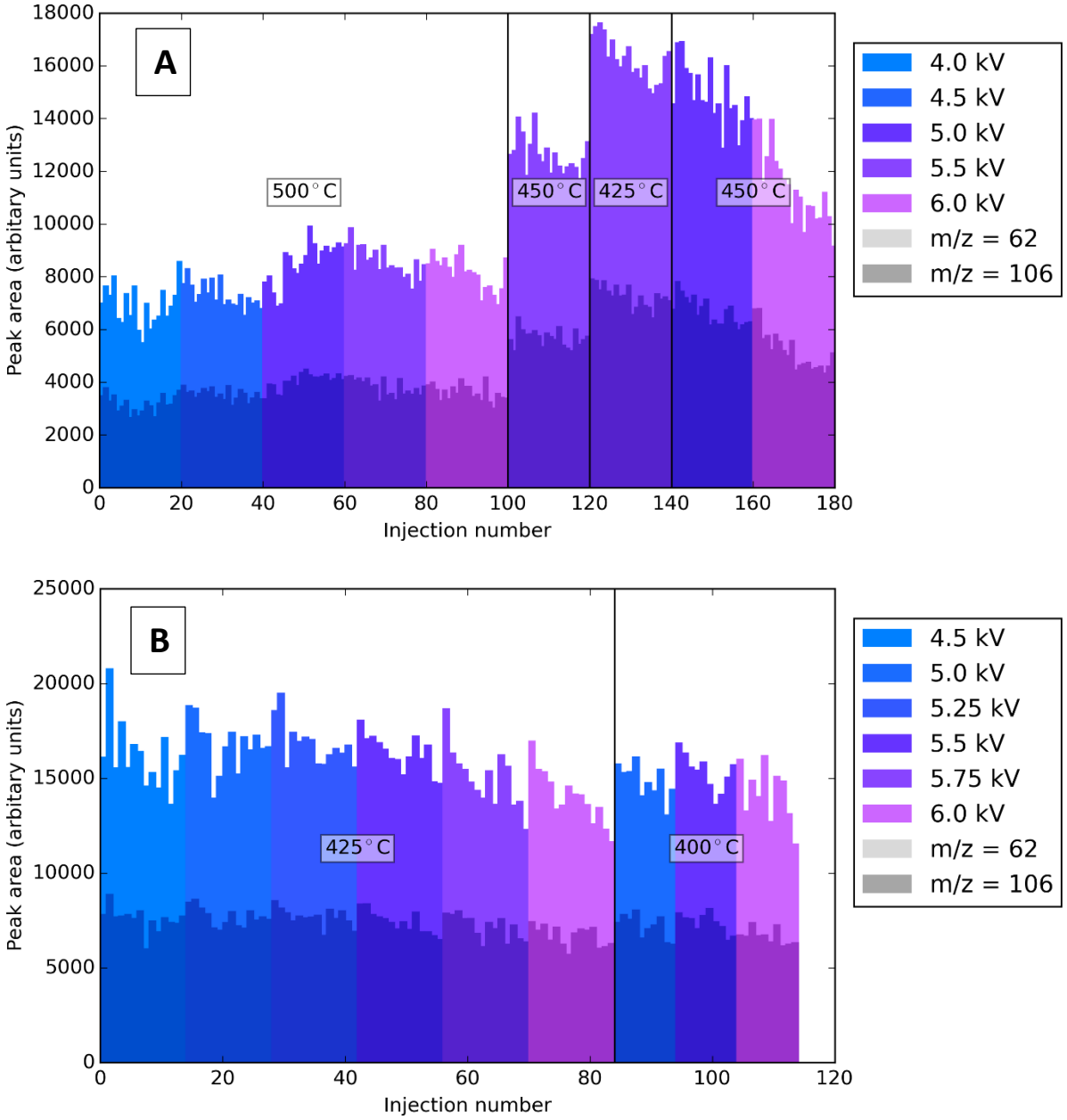


Figure 2-13. Peak areas of both product ions resulting from the measurement by LCMS of 500 pg injections of metaldehyde, using different capillary voltage and temperatures. Black lines indicate where the desolvation temperature was changed to that given on the labels. Different experimental runs are shown in A and B.

Ionisation condition 5.5 kV and 425°C was used in both experimental runs shown in Figure 2-13. The mean peak area for this condition in the second run was 0.4% higher for *m/z* = 62 and 1.4% higher for *m/z* = 106. This similarity gives confidence that the data produced in both experiments can be directly compared, and so the mean peak areas from both experiments are shown in Figure 2-14. Mean retention time was 1.58±0.01 min for both experiments. The number of measurements made at each condition was limited by equipment availability in the second experiment. The number of

observations for different conditions vary; n = 20 where ionisation temperature is 450°C or 400°C, n = 14 where temperature is 425°C with the exception of 5.5 kV and 425°C where n = 34, n = 10 where temperature is 400°C.

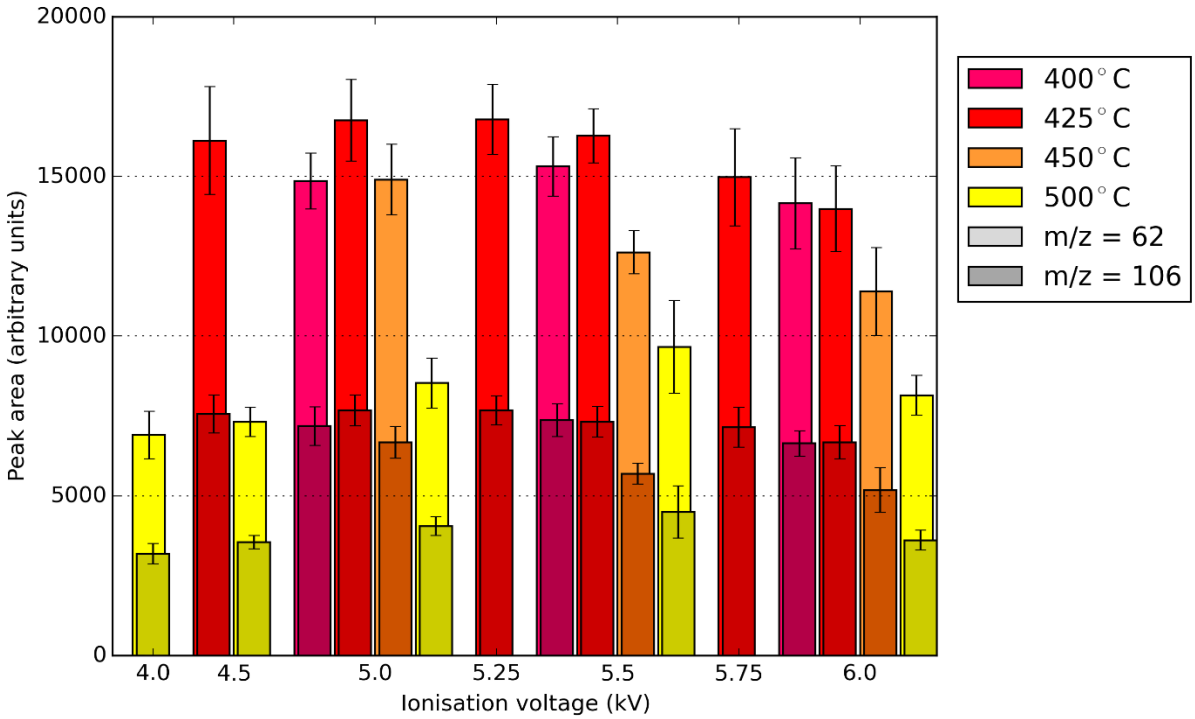


Figure 2-14. Means and standard deviations of peak areas for both product ions resulting from repeated assays of 500 pg of metaldehyde by LCMS, using different capillary voltages and temperatures. Groups of bars separated by white space give measurements made at the voltage indicated on the X-axis. The number of measurements for each condition are given in the text.

The RSD of different voltage and temperature combinations is given in Table 2-9.

Temp. (°C)	Voltage (kV)	RSD (%) *m/z* = 62	RSD (%) *m/z* = 106
400	5.0	6.2%	8.8%
	5.5	6.4%	7.3%
	6.0	10.6%	6.2%
425	4.5	10.8%	8.3%
	5.0	7.9%	6.6%
	5.25	6.8%	6.2%
	5.5	5.3%	6.7%
	5.75	10.5%	9.0%
	6.0	9.9%	8.1%
450	5.0	7.6%	7.8%
	5.5	5.5%	5.8%
	6.0	12.4%	13.9%
500	4.0	10.40%	11.10%
	4.5	6.5%	6.1%
	5.0	9.4%	7.4%
	5.5	15.2%	18.5%
	6.0	7.8%	8.7%

Table 2-9. RSD of peak areas resulting from the repeated measurement of 500 pg of metaldehyde by LCMS, using different ionisation temperature and voltage combinations. Redder shades of pink indicate higher RSD. The number of measurements for each condition are given in the text.

2.3.3.1 Interaction of temperature and voltage

There may be an interaction between ionisation temperature and voltage. The interaction plot (Figure 2-15) shows the change in response at different voltages does not appear to be entirely consistent at different temperatures. The nature of this possible interaction was not investigated further. The effect of each parameter is discussed individually in the following sections, taking into consideration the possibility of an interaction between them.

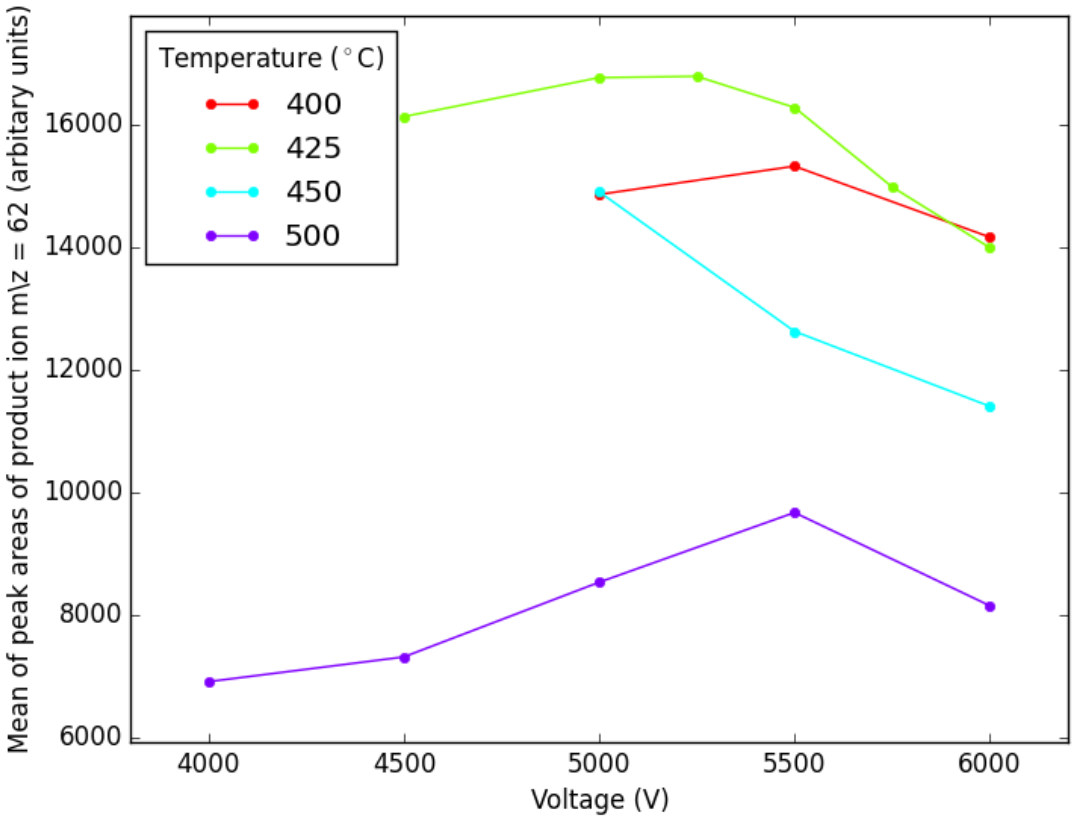


Figure 2-15. Interaction plot showing the mean peak areas observed for the m/z *= 62 product ion, when quantifying 500 pg of metaldehyde by LCMS using different combinations of capillary voltage and temperature.*

2.3.3.2 Effect of voltage

The RSD resulting from default conditions in this experiment (Table 2-9; 500°C and 5.5 kV) is 15.2% and 18.5% for *m/z* = 62 and *m/z* = 106. This is higher than was seen in in earlier tests using the same conditions, where the highest RSD from a set of 20 measurements was 13.3% and the mean RSD in a set was 9.97% and 10.36% for *m/z* = 62 and *m/z* = 106 (Table 2-3). There are no outliers in the measurements. The peak area of these injections (Figure 2-13A; 161-180), as well as the following set of injections (161-180), tend to decrease with each injection. Gradually decreasing response also occurs following changes to the capillary voltage (Figure 2-13B; injections 43-84). Whether these trends have some underlying cause related to changing the ionisation parameters, or are stochastic, cannot be determined from these data.

Considering all measurements made at a specified temperature across a range of capillary voltages, one-way analysis of variance (ANOVA) tests show there are significant differences in mean response

in assays done at all temperatures (data not shown). To establish which voltages produced significantly different responses, Tukey's honestly significant difference (HSD) test was conducted. The tests were done for all voltage pairs at each temperature and for each product ion and for family-wise error rates (FWER) ranging between 0.5 and 0.001. Table 2-10 gives the lowest FWER that could be obtained for each pair of conditions.

		400°C		425°C		450°C		500°C	
1st kV	2nd kV	m/z = 62	m/z = 106	m/z = 62	m/z = 106	m/z = 62	m/z = 106	m/z = 62	m/z = 106
4.0	4.5							>0.5	0.245
4.0	5.0							**0.001**	**0.001**
4.0	5.5							**0.001**	**0.001**
4.0	6.0							**0.005**	0.105
4.5	5.5			>0.5	>0.5			**0.001**	**0.001**
4.5	5.75			0.18	0.335				
4.5	6.0			**0.001**	**0.001**			0.16	>0.5
5.0	5.5	>0.5	>0.5	>0.5	>0.5	**0.001**	**0.001**	**0.005**	0.07
5.0	5.75			**0.01**	0.31				
5.0	6.0	0.39	0.08	**0.001**	0.115	**0.001**	**0.001**	>0.5	0.185
5.25	5.75			**0.005**	**0.001**				
5.25	6.0			**0.001**	0.3				
5.5	5.75			**0.025**	0.11				
5.5	6.0	0.085	**0.015**	**0.001**	**0.001**	**0.005**	**0.015**	**0.001**	**0.001**
5.75	6.0			0.33	**0.005**				

Table 2-10. Lowest FWER for pairs of capillary voltages obtained by Tukey's HSD tests, using data obtained by the measurement of 500 pg of metaldehyde by LCMS using different capillary voltages and ionisation temperatures. FWER between 0.5 and 0.001 were tested for. Sets of peaks areas from assaying 500 pg of metaldehyde at different temperatures were tested. FWER below 0.05 are in bold. Voltage pairs not shown did not result in FWER < 0.5 at any temperature tested. These pairs are; 4.00, 5.25; 4.50, 5.25; 5.00, 5.25; and 5.25, 5.50 kV.

At 400°C the only significant difference is of product ion *m/z* = 106, 5.5 versus 6.0 kV. At 425°C all voltage combinations of product ion *m/z* = 62 involving 5.75 and 6.0 kV – excepting 4.5 versus 5.75 kV – show the higher voltages have significantly lower peak areas. All of the significant differences for the *m/z* = 106 ion also include the same voltages but with many more exceptions. At 450°C the peak areas decrease significantly with increasing voltage at the three voltages tested. At 500°C the high response at 5.5 kV and low response at 4.0 kV are shown to be significantly different in most cases.

2.3.3.3 Effect of ionisation temperature

The differences in RSD given in Table 2-9 for measurements made using the same voltage at different temperatures are not large enough for KS tests to be able to show significant differences with the

exception of comparisons against condition 5.5 kV, 500°C. This distribution is significantly different to distributions produced using the same voltage at 425°C (D = 0.31, p = 0.001) and 450°C (D = 0.35, p = 0.002), but not 400°C (D = 0.31, p = 0.07).

Peak areas of measurements made using <500°C are higher at all voltages tested. Determining the conditions that give the highest response is part of the purpose of these experiments and so Welch's T-tests were used to determine if 425°C produces a significantly higher response to 400°C at 5 kV and 5.5 kV, as well as 450°C at 5.5 kV (Table 2-11). The majority of tests (4 of 6) confirm that 425°C gives significantly greater response, though the result p = 0.06 given in Table 2-11 is not strongly indicative given the number of tests done here. At 5.5 kV 425°C is significantly higher for product ion m/z = 62, the means of the product ions m/z =106 do not differ substantially (Figure 2-14).

Voltage (mV)	Temp. 1 (°C)	Temp. 2 (°C)	m/z = 62	m/z = 106
5.0	425	400	T = 4.16 p = 4.1E-04	T = 2.02 p = 0.060
5.0	425	450	T = 4.27 p = 2.4E-04	T = 5.59 p = 5.3E-06
5.5	425	400	T = 2.79 p = 0.015	T = 0.26 p = 0.80

Table 2-11. Results of Welch's T-tests comparing the repeated measurement of 500 pg of metaldehyde by LCMS using the given ionisation conditions.

2.3.3.4 Reduced flow rate improves response at lower ionisation temperatures

An issue with lower ionisation temperatures is that it can lead to insufficient evaporation of the mobile phase solvent. Lowering the mobile phase flow rate can improve evaporation. Ionisation temperatures of 400°C, 350°C and 300°C and flow rates of 0.6, 0.5 and 0.4 ml/min were investigated. To avoid possible issues of gradually diminishing response after changing ionisation conditions, discussed in Section 2.3.3.2, each ionisation temperature was tested in multiple pairs of injections, rather than all measurements at particular conditions being made in a single set of injections. 50 µl injections of 5 µg/L metaldehyde were used and conditioning injections were applied before quantification took place. Peak areas for each injection and means for each condition are given in Figure 2-16 and Figure 2-17. Injection number 31 was excluded from all statistical analysis.

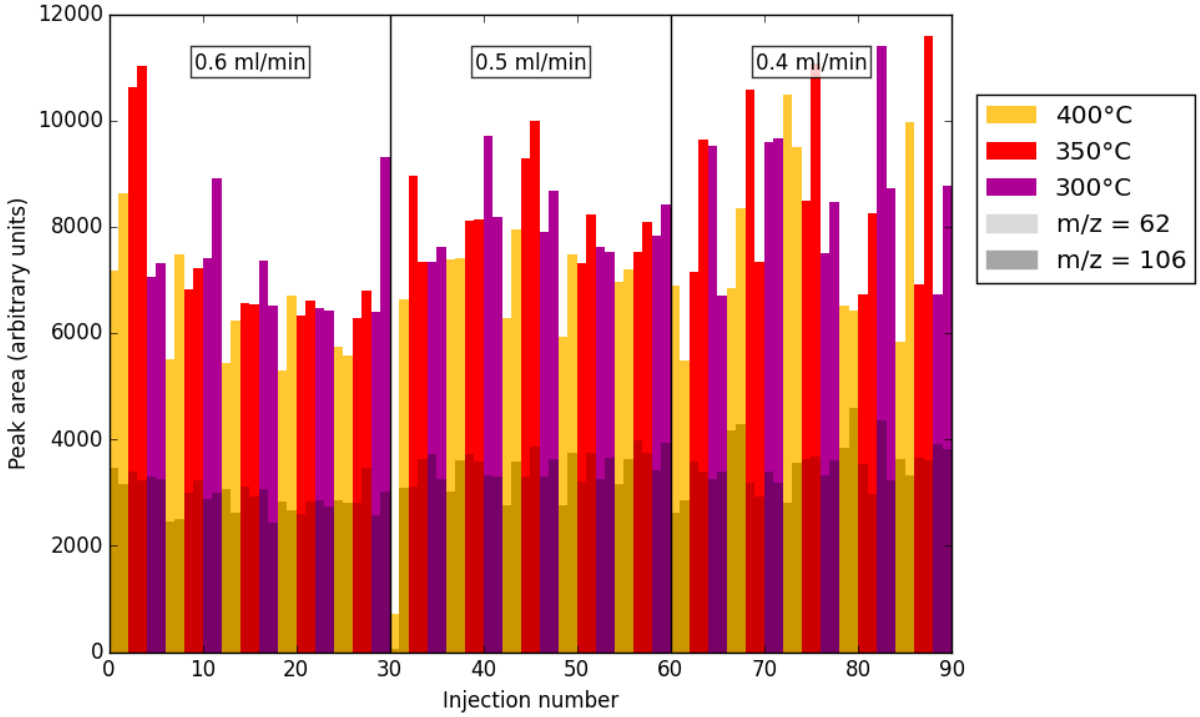


Figure 2-16. Peak areas for both product ions resulting from the repeated measurement of 250 pg of metaldehyde by LCMS, at different ionisation temperatures and flow rates.

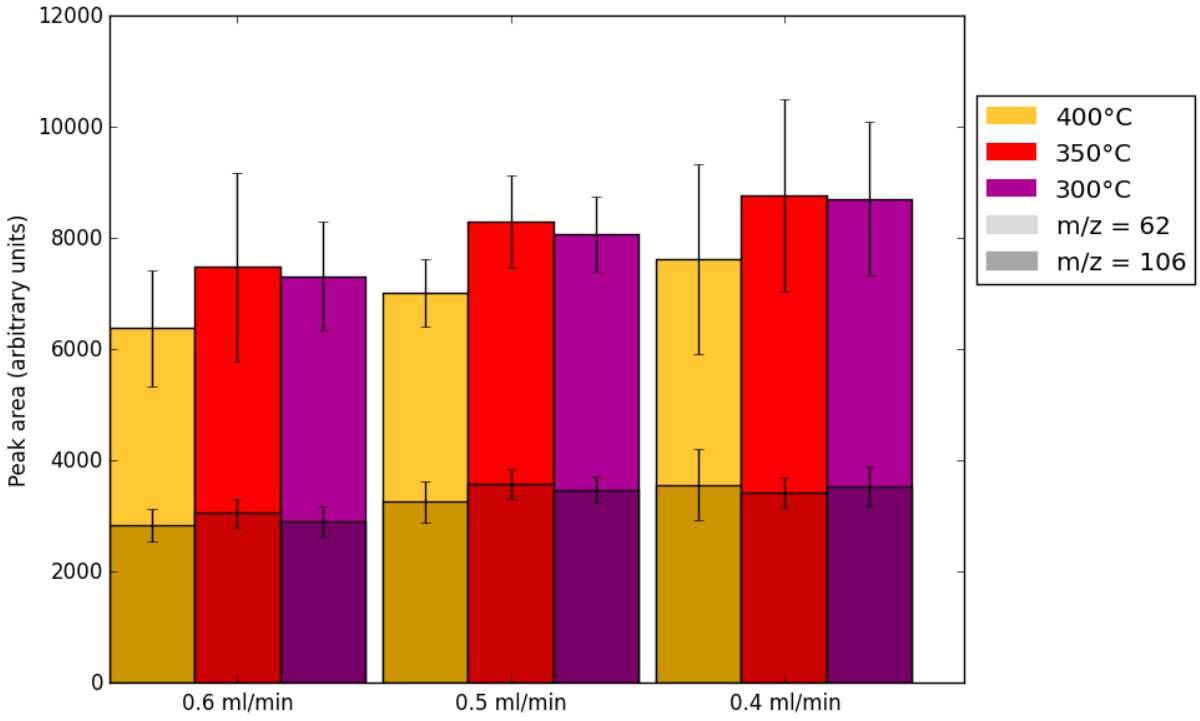


Figure 2-17. Means and standard deviations of repeated assays of 250 pg of metaldehyde by LCMS, at different ionisation temperatures and mobile phase flow rates (flow rates given below the x-axis). n = 10, with the exception of 0.5 ml/min 400°C where n = 9.

One way ANOVAs showed there were significant differences between the 9 groups shown in Figure 2-17 (F = 3.5, p = 0.0015), the different flow rates (F = 6.8, p = 0.0019) and the different temperature groups (F = 6.1, p = 0.0033). The application of Tukey's HSD test, using a family wide error rate of 0.05, shows that desolvation temperatures of 300°C and 350°C result in significantly higher response compared to 400°C, and that a flow rate of 0.4 ml/min results in significantly higher response than 0.5 or 0.6 ml/min.

		RSD (%)	
Flow (ml/min)	**Temp. (°C)**	***m/z* = 62**	***m/z* = 106**
0.6ml	400	16.4	10.5
	350	22.6	8.6
	300	13.4	9.0
	All temp.	19.4	9.9
0.5ml	400	8.6	11.2
	350	10.0	7.7
	300	8.4	6.6
	All temp.	11.5	9.3
0.4ml	400	22.4	17.8
	350	19.7	7.9
	300	15.9	10.1
	All temp.	20.3	13.0

Table 2-12. RSD of peak areas from the repeated measurement of 250 pg of metaldehyde by LCMS, using different solvent flow rates and desolvation temperatures. Redder shades of pink indicate higher RSD.

The RSD of quantification when using 0.5 ml/min solvent flow rate is notably lower for the m/z = 62 ion, but this is not shown to be significant by KS tests.

2.3.4 Quantification of metaldehyde shows good precision across a range of concentrations but assessment of accuracy was weakened by standard preparations

Previous experiments in this chapter showed that higher responses were achieved by using a flow rate of 0.4 ml/min, ionisation temperatures 300-350°C, and voltages 5.0-5.5 kV. A flow rate of 0.4 ml/min added too much time to the length of a sample's acquisition to be practical and so the values of 0.5 ml/min for flow rate, 325°C and 5.25 kV for ionisation temperature and voltage were used in the quantification of a range of metaldehyde concentrations. All other aspects of the quantification were as described in the methods and conditioning injections were performed. The assessment mimicked the method of quantifying samples of unknown concentration with "calibration brackets"; 50 µl injections of known standards containing 100, 50, 5, 1 and 0.5 µg/L metaldehyde that were

made between sets of 15 samples. In this experiment the "samples" were also standards of known concentration, including the given concentrations of the calibrants, as well as 0.25, 0.125 µg/L metaldehyde and a blank standard using only 0.1% culture matrix. Each standard was prepared once and aliquoted to multiple vials. The standards treated as samples were quantified in sets of 15 injections; 2 sets of the full range of standards, then a set of high (only 50 and 100 µg/L) concentration standards, low (0-1 µg/L) standards, blank standards and high concentration standards. The number of measurements made for each standard is given in Table 2-13.

MA (µg/L)	100	50	5	2.5	1	0.5	0.25	1.25
n	29	23	11	5	16	13	5	2

Table 2-13. Total number of measurements made of each metaldehyde standard at the given concentrations.

Normalised response (peak area/expected metaldehyde concentration) for each injection are given in Figure 2-18. Higher responses are seen in the first 4 injections. No other anomalous trends are apparent. The mean retention time was 1.91±0.01 min.

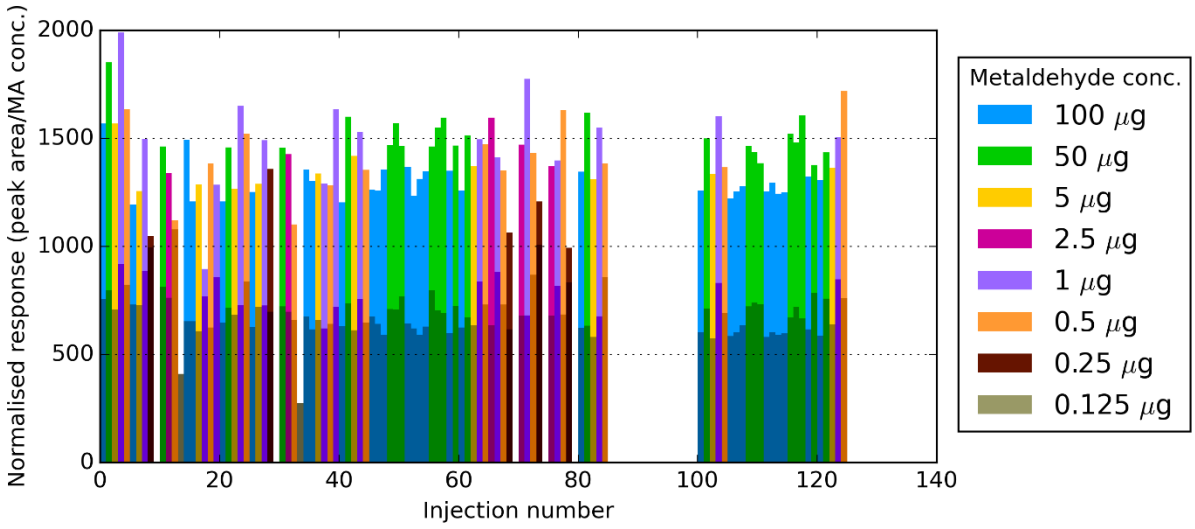


Figure 2-18. Normalised response (peak area/standard concentration) for both product ions resulting from the measurement of standards containing different concentrations of metaldehyde by LCMS. Product ion m/z = *62 are lighter shades,* m/z = 106 *are darker shades.*

The mean normalised response for each standard, including those used as calibrants and those treated as samples, is shown in Figure 2-19.

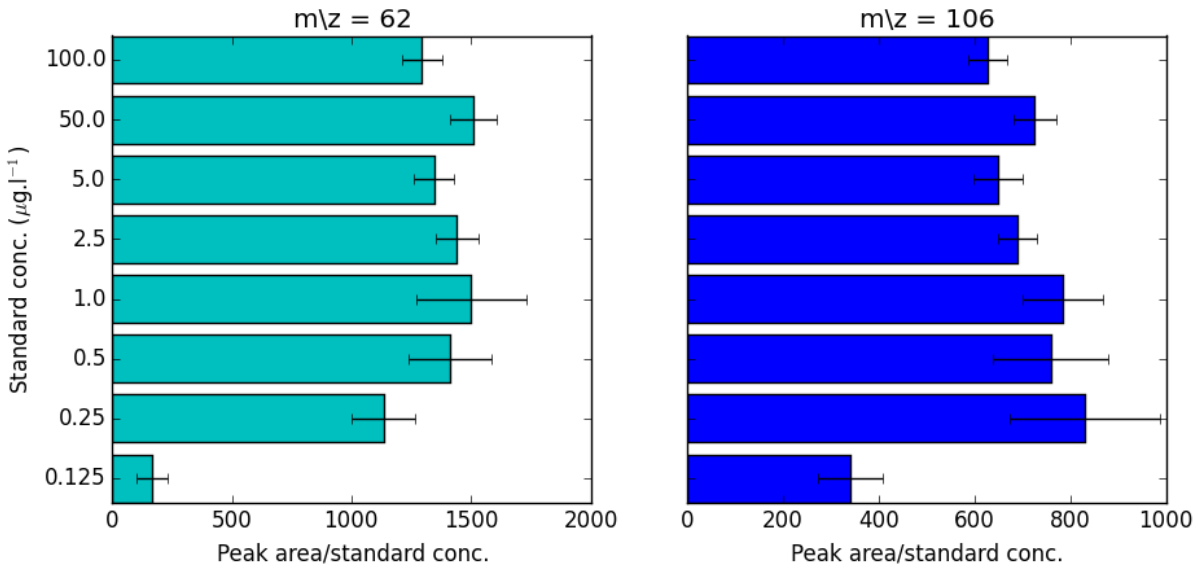


Figure 2-19. Mean normalised response (peak area/standard concentration) for the repeated measurement of metaldehyde standards of different concentrations by LCMS. Error bars give one standard deviation. Left and right charts give the results for the indicated product ion. Number of measurements for each standard are given in Table 2-13.

The RSD (Table 2-14) for all standards >1 µg/L are similar. Higher RSD are seen at lower concentrations as the signal:noise ratio becomes lower.

	RSD (%)	
Metaldehyde (µg/L)	***m/z* = 62**	***m/z* = 106**
100	6.4	6.4
50	6.5	6.1
5	6.3	8.0
2.5	6.2	6.0
1	15.3	10.8
0.5	12.2	15.8
0.25	11.7	18.9
0.125	38.3	19.9

Table 2-14. RSD of peak areas obtained for the repeated measurement, by LCMS, of metaldehyde standards of different concentrations. Number of measurements for each standard are given in Table 2-13. Redder shades indicate greater RSD.

Quantification of metaldehyde was achieved using weighted (1/X) least-squares linear regression of *m/z* = 62 calibrant peak areas (Figure 2-20). The regression yields the following equation with a coefficient of determination of 0.974:

[peak area units] = -56.64 + 1434*[standard µg/L]

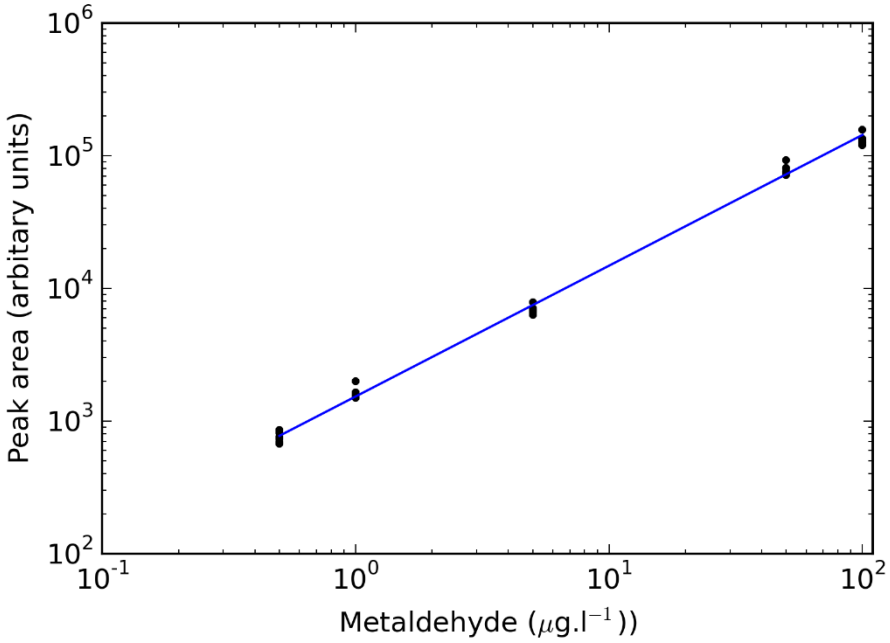


Figure 2-20. Calibration curve of metaldehyde concentration plotted against peak areas observed for the m/z *= 62 product ion, produced by LCMS. Blue line represents a linear regression of the points.*

The regression was used to calculate the concentration of the standard injections treated as samples. The mean calculated concentrations of these are presented as percentages of known concentrations in Figure 2-21. The 2 0.125 µg/L injections were below the limit of detection.

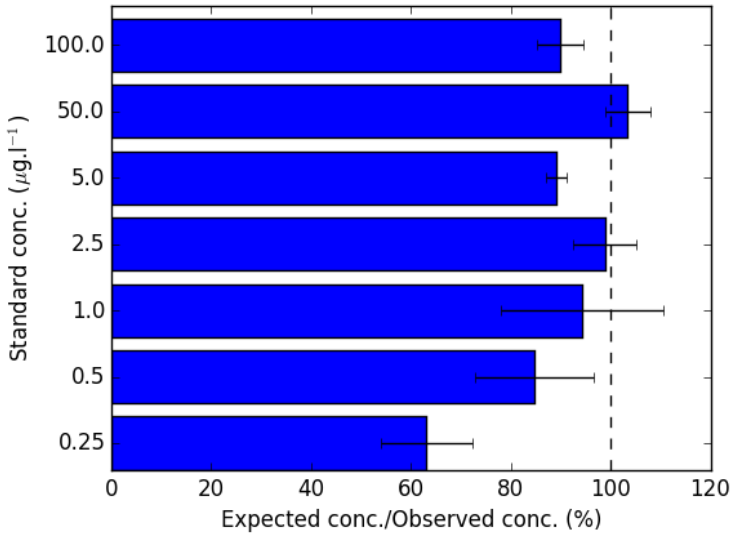


Figure 2-21. Mean calculated metaldehyde concentration of samples quantified using LCMS, as a percentage of their expected concentrations. Error bars show 1 SD.

All but one of the samples were calculated to be less than their expected value on average. The mean calculated concentrations of standard injections used in the calibration are shown in Figure 2-22. The calibrants are calculated to be around 100% as expected, which shows that the inaccuracy seen in Figure 2-21 results from the measurements, rather than the calibration being performed incorrectly.

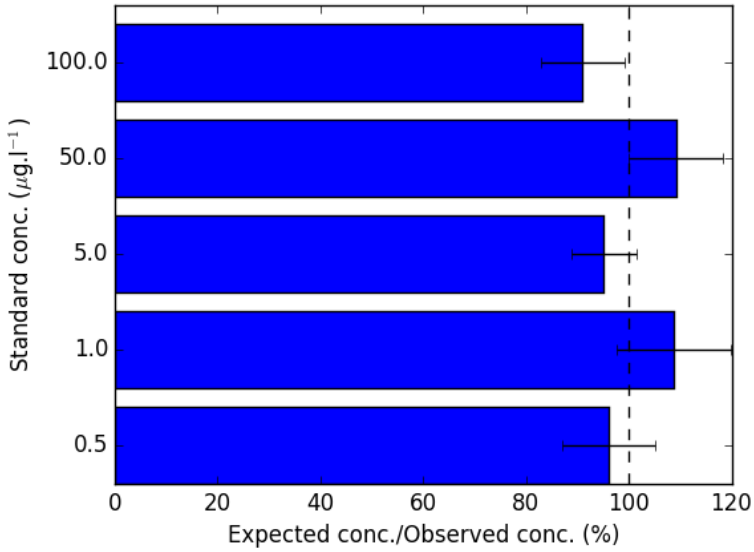


Figure 2-22. Mean calculated metaldehyde concentration of calibrant injections, quantified by LCMS, as a percentage of their expected concentrations. Error bars show 1 SD.

Given the larger RSD of calibrants 1 and 0.5 µg/L (mean RSD 13.7% versus 6.35% for calibrants > 1 µg/L) excluding them from the calibration was trialled, starting with 0.5 µg/L. The equation derived from the regression is shown below. The R^2 was 0.964.

[peak area units] = 297.4 + 1318*[standard µg/L]

Using the results of the second regression the concentration of all standards was repeated. The mean recalculated concentration of standard injections that were used as calibrants is shown in Figure 2-23 and those used as samples in Figure 2-24.

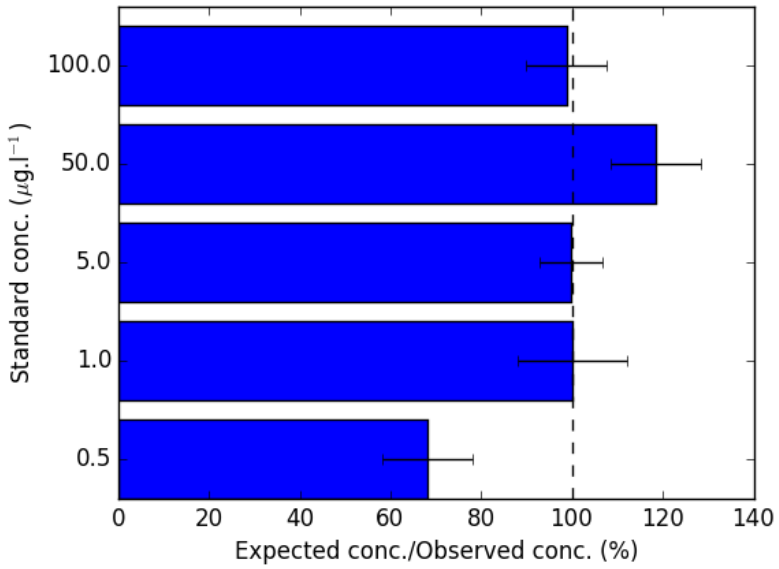


Figure 2-23. Mean calculated metaldehyde concentration of calibrants, quantified by LCMS, as a percentage of their expected concentrations. Calculation derived from calibration that excluded standards with 0.5 µg/L metaldehyde. Error bars show 1 SD.

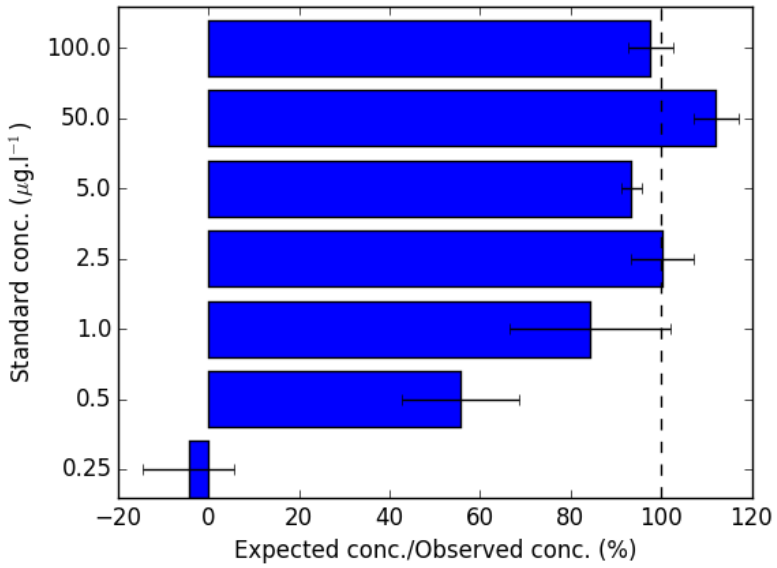


Figure 2-24. Mean calculated metaldehyde concentration of samples, quantified by LCMS, as a percentage of their expected concentrations. Calculation derived from calibration that excluded standards with 0.5 µg/L metaldehyde. Error bars show 1 SD.

The larger value for 50 µg/L standards across all figures in this section are best explained as an inaccuracy in the preparation of that standard. Inaccuracies in the preparation of other standards is likely the cause of mean values that deviate from the expected concentration. Decreased response caused by non-linearity of the quantification at high analyte concentrations also cannot be ruled out as affecting the quantification of the higher concentration standards. These inaccuracies affect the ability of the calibration curve to be able to calculate the metaldehyde concentration of unknown samples.

2.3.5 Comparison of precision of default protocol to optimised protocol

To compare the original, default, LCMS protocol to the optimised one, data from assays of 500 pg of metaldehyde using the default protocol, presented in Section 2.3.2.1, and assays of standards with a metaldehyde concentration ≥ 2.5 µg/L (125+ pg) or more from Section 2.3.4 were used.

The RSD of peak areas normalised to the mean value for each standard used in both sets of data, including both product ions, were 10.0% (n = 200) for the default protocol and 6.47% (n = 136) for the optimised protocol. A KS test comparing these distributions found them to be significantly different (D = 0.17, p = 0.017). Comparing peak areas of the product ion used in quantification, m/z = 62, in the same way does not show a significant difference (D = 0.19, p = 0.10).

2.4 Discussion

Alterations were made to an already established LCMS method for the quantification of metaldehyde in an attempt to; (i) increase precision of measurements at all concentrations, (ii) reliably quantify metaldehyde at low concentrations, and (iii) eliminate declining response seen at the beginning of an experimental run. An optimised protocol that uses 0.5 ml/min flow rate, 5.25 kV capillary voltage and 325°C solvation temperature (compared to the original values of 0.6 ml/min, 5.5 kV and 500°C) was developed that has greater response and improved precision. Conditioning injections were added to the protocol to ameliorate the worst of the declining response and allow more accurate quantification.

A comparison of the default and optimised protocol was performed (Section 2.3.5). By normalising the peak areas obtained to the mean for each standard used it was possible to compare RSD, and significantly greater precision in the optimised protocol was found when quantifying 125 pg of metaldehyde or more. A direct comparison, using the same standards, was not performed and so changes in precision at lower metaldehyde masses could not be assessed.

Kolmogorov-Smirnov tests were used to assess changes to the precision of measurements. While this method can result in p-values < 0.05 it is insensitive to differences in standard deviation between sets of measurements with the same mean where the number of measurements is low. A better test for precision was not found.

The vaporisation conditions in the ionisation chamber were found to have a significant effect on the response. High temperatures are required to produce the ion stream for MS. High temperatures also cause the depolymerisation of metaldehyde. An ionisation temperature of 425°C gave the highest response at 0.6 ml/min flow rate. Reducing the flow rate to 0.5 ml/min improved the response at lower temperatures with the highest responses seen at 350°C and 300°C. A value halfway between those two temperatures, 325°C, was chosen as the ionisation temperature for all subsequent analyses, replacing 500°C used in the default protocol. Using this lower temperature would be expected to improve the quantification of low levels of metaldehyde by increasing the signal seen above the baseline noise.

Evident in experiments presented was decreasing response with each subsequent injection which gradually stabilised. This may result from compounds in the sample annealing to the matrix of the LC column reducing the overall retaining surface area inside the column and therefore reducing the mass of analyte retained, though this is not demonstrated in the data shown here – no specific cause was determined in these experiments. A method to avoid the quantification problems caused by this

was adopted. Conditioning the column with injections of high concentration samples or standards causes the response to drop, and be more stable level before quantification begins.

The mobile phase ammonium acetate concentration of 1 mM, and capillary voltage of 5.5 kV used in the default protocol were found to be approximately optimal. It was shown that reducing the ammonium acetate concentration to 0.5 mM significantly reduced accuracy and response. At voltages less than 5 kV and greater than 5.5 kV accuracy and response appears to be diminished and so 5.25 kV was adopted as the voltage for subsequent investigations.

Injection volume did not impact quantification in any respect.

The poor results seen using a reduced flow rate of 0.3 ml/min may be the result of changes to the chromatography, or from increasing the negative impact of the high ionisation chamber temperature, or both.

The analysis of different concentrations of metaldehyde using the ionisation conditions of 5.25 kV and 325°C with a flow rate of 0.5 ml/min was complicated by standard concentrations that varied from the intended concentration. The accuracy of samples quantified using bracketed calibration sets was low, giving values between 89.5-75.4% of those expected. Removing the lowest concentration standard from the calibration increased the accuracy of samples >1 µg/L but raised the effective limit of quantification to 2.5 µg/L. The updated method was shown to be more precise in a comparison of the >1 µg/L standards quantified using it. The increased response would be expected to improve the precision of lower concentration standards by a greater degree, but this was not tested in the data presented here.

2.4.1 Comparison to other published methods

The published method by Li et al. (2010) uses an isocratic flow of 20 mM ammonium acetate and methanol in 15:85 at 0.2 mL min^{-1} using a column with the same dimensions as the Kinetex column used in this work. This results in a retention time of 0.89 min, less than the 1.5 min seen in the default protocol, likely due to the high methanol content of the mobile phase. Typically the advantage in using a gradient elution is that it can reduce the retention time of analytes (Kromidas, 2008, p.93). Optimisation of retention time was not among the objectives in this set of experiments and so isocratic flows were not investigated.

The MS conditions used by Li et al. (2010) were similar to the default protocol. The capillary voltage used there is 3 kV, lower than was investigated in this chapter. A general trend of greater noise away from an optimum value of ~5.25 kV was observed in the experiments described here, and so values further from this were not investigated. The difference in signal response and precision between

measurements made with 4-6 kV was not large so it may be that this parameter would not affect the quantification significantly if lowered to 3 kV, or it may yield a significant improvement. They use 500°C as the desolvation temperature along with a 0.2 ml/min flow rate. Based on the results described in this chapter it would be expected that reducing this temperature would improve sensitivity in the quantification of metaldehyde. The sensitivity of the method described by Li et al. cannot be directly compared as samples were pre-processed by solid phase extraction in their work.

Schumacher et al. (2016) have developed a method that significantly improves upon the sensitivity of the previously published method, with a LOQ of 4 ng/L in tap water. The significant innovation described in their work is the use of methylamine as the mobile phase additive. This has a relatively high affinity for metaldehyde, which suppresses the formation of other, unquantifiable, metaldehyde adducts, and requires a lower collision energy to obtain the methylamine-acetaldehyde adduct which is quantified. The method uses online enrichment and a 5-minute gradient. This degree of sensitivity is not required for the purposes of this project, but the use of methylamine as the mobile phase additive without online enrichment and some optimisation of the gradient to improve throughput could improve on the method developed in this chapter. The capillary voltage used is 3 kV (Schumacher et al., 2016), the same as Li et al (2010). It may be that Schumacher et al. based their protocol on Li et al. and capillary voltage is unimportant. It would be of interest to compare 3 kV to the 5.25 kV used in this chapter. The desolvation gas temperature used by Schumacher et al. is 300°C, close to the 325°C used in our optimised protocol.

2.4.2 Summary

In the work described here it was found that lowering the desolvation temperature resulted in significant increases in the response seen when quantifying metaldehyde by LCMS. This may be expected to improve the quantification of metaldehyde at low concentrations. The precision of the method may also have been improved. The method was shown to be able to quantify standards to at least 1 µg/L. The improved method, using a desolvation gas temperature of 325°C and capillary voltage of 5250 V, rather than 500°C and 5500 V, was used to quantify metaldehyde in the work that will be described hereafter.

Chapter 3: Enrichment and characterisation of metaldehyde degrading organisms

3.1 Introduction

3.1.1 The isolation of microorganisms with specific metabolic capabilities

As discussed in Chapter 1, genes within bacterial genomes are quickly evolving and encode a diverse repertoire of proteins, some of which have evolved to catalyse the metabolism of xenobiotics. The identification of organisms that metabolise xenobiotics and the enzymes they use can be achieved in several ways that take advantage of the catabolism itself.

The inoculation of a selective media containing the compound of interest as the sole source of an essential nutrient with an environmental sample thought to contain microorganisms with catabolic activity against the compound of interest, can be used to enrich for and isolate those microorganisms. First, inoculation of a liquid media causes the proliferation of catabolising organisms, then this culture (that is enriched for with catabolising organisms) can be plated onto a selective solid media containing the compound of interest. Organisms that are capable of clonal growth under the conditions used can then be easily isolated for further study. As relatively recent examples, this general method was used by Snellinx et al. (2003) in the isolation of 2 strains that can use 2,4-dinitrotoluene as a sole source of nitrogen, by Dejonghe et al. (2003) in the isolation of a linuron degrading organism and by Takenaka et al. (2013) in the isolation of several 4-aminopyridine degrading strains. Pure cultures of a single strain that does not grow on solid media can be obtained by repeated subculturing of liquid cultures (Sutherland et al., 2002). In the work presented here, the enrichment and isolation of metaldehyde catabolising organisms from a variety of substrates is described.

3.1.2 Bacterial phylogeny and the identification of species

The phylogenetic classification of strains isolated in this study is described in this chapter. The classification of bacterial species is fraught with ambiguity, but can aid in the discussion and understanding of isolates. As bacteria reproduce by binary fission, and do not share a pool of genes, genetic compatibility cannot be used as a defining characteristic of a species. In the past, a range of morphological, chemical and physiological traits were assessed to determine the species membership of a bacterium, as can be seen in Bergey's Manual of Determinative Bacteriology (Holt, 1994), for example.

It was recognised that a potentially excellent resource for systematics is the genetic information of bacterial cells, even before the nature of this information was well understood (Ravin, 1963). Genetic

information is passed down to daughter cells and is the primary cause of the phenotypic characteristics that have been used for classification. Prior to the development of sequencing technologies, quantification of the degree of DNA-DNA hybridisation (DDH) was used to assign strains to "genospecies". The standard method for this involved immobilising fragmented genomic DNA from one strain to a nitrocellulose membrane, and probing the membrane with radioactively labelled DNA from another strain (Gillespie & Spiegelman, 1965). As a guideline, a minimum DDH value of 70% was adopted as the cut-off for assignation of two strains to the same genospecies. This was based on observed DDH values for groups of related strains declared to be of the same species or not by other methods (Tindall et al., 2010).

The advent of nucleotide sequencing technologies has contributed greatly to the understanding of bacterial systematics. The primary sequence of the 16S ribosomal RNA (rRNA) was recognised as a particularly useful sequence for higher order classification of bacteria due to its low rate of mutation and universality (Fox et al., 1977; Woese & Fox, 1977). However, the low rate of mutation and limited depth of information available from the 16S sequence limits its usefulness in making finer phylogenetic distinctions. As the mass of genomic information that is available has increased, thanks to increasingly inexpensive sequencing techniques – bacterial genome sequencing can be purchased for £70 (MicrobesNG, Birmingham; non-commercial rate, at time of writing) – there has been a shift over to the use of genomic data to assign a species name to an isolate. Average nucleotide identity (ANI) is a method that attempts to quantify the overall degree of similarity between the homologous regions of two genomes. The genomes of interest are fragmented *in silico*, the basic local alignment search tool (BLAST) algorithm is used to identify fragments with a minimum level of homology, and the average nucleotide identity of these homologous fragments is calculated to give the final statistic (Konstantinidis & Tiedje, 2005). The vast majority of named species can be distinguished by having ANI values of < 95% (Konstantinidis & Tiedje, 2005; Kim et al., 2014; Rodriguez-R & Konstantinidis, 2014).

3.1.3 Aims

Metaldehyde can be quickly degraded in soils (Agriculture & Environment Research Unit, 2012; Zhang et al., 2011a) and is oxidised to carbon dioxide under aerobic conditions in unsterilised soils (European Food Safety Authority, 2006), in comparison to its long half-life in sterile conditions. This strongly suggests the involvement of microbial life in its degradation. The aim of the work described here was to isolate and identify those organisms capable of degrading metaldehyde. The genus of, and identity of species closely related to, the isolated organisms was established first using 16S rRNA gene sequences. This information is easily obtained and was used to guide which whole genome

comparisons should be made, limiting the number of comparisons that were required. ANI was used to assign species names to the sequenced isolates.

3.2 Materials and methods

3.2.1 Minimal media

Minimal salts media (MSM) were prepared using the concentration of salts given in Table 3-1 dissolved in ultrapure water. The salts solution was autoclaved. For solid plates, 1.5 g agarose would be added to 200 ml MSM prior to autoclaving. Agarose was used in place of agar to eliminate carbohydrates from the plates that could be used for growth by organisms. 2 ml of a trace elements solution (Vishniac & Santer, 1957 Table 3-2) was added for each 1 litre of salts solution.

	Concentration (mM)
Na_2HPO_4	55
KH_2PO_4	11
NH_4Cl	6
$MgSO_4$	0.4

Table 3-1. Salts concentrations for minimal media. All purchased from Sigma-Aldrich (Dorset, UK)

	Concentration (mM)
Na_2EDTA	140
$ZnSO_4$	7.6
$CaCl_2$	37
$MnCl_2$	25
$FeSO_4$	18
$(NH_4)Mo_7O_{24}$	0.9
$CuSO_4$	6.4
$CoCl_2$	6.7

Table 3-2. Concentration of compounds in trace elements solution. EDTA is ethylenediaminetetraacetic acid.

Up to 150 mg/l (0.851 mM) metaldehyde was added to liquid media in powder form and dissolved overnight on a magnetic stirrer. When used in solid media molten MSM-agarose would be cooled in a water bath to 50°C and 100 mg metaldehyde would be added to 200 ml molten agarose to form a suspension. Agarose plates with metaldehyde added will be referred to as "metaldehyde plates" those with no added carbon substrate will be referred to as "no-carbon plates".

Sodium acetate, when used, was dissolved in deionised water at 1 M concentration, filter sterilised (0.2 um pore size; Merck Millipore, Darmstadt, Germany), and added to MSM to achieve the desired concentration, typically 10 mM.

3.2.2 Enrichments and isolations of metaldehyde degrading organisms

In general, enrichments were performed using solid and liquid minimal media containing metaldehyde. Environmental substrates were added to liquid media cultures in flasks, which were then incubated at 30°C. Samples of enrichment cultures, or environmental samples, were spread across plates which were incubated at 30°C. The specifics of enrichments for the different samples are given below.

3.2.2.1 Domestic soils

Samples of domestic soils were obtained from two gardens in York. One from a large plant pot that had been treated with slug pellets that contained metaldehyde as the active ingredient for at least 2 years. The other from a flower bed in a separate garden that had not been treated with metaldehyde.

Enrichment cultures for metaldehyde degrading organisms from two domestic soil samples were made using 100 ml MSM with either 100 µl of SlugClear (SlugClear, UK; contains 228 g/l metaldehyde as the active ingredient), or no additional carbon, and 1 g of soil. 250 ml flasks containing the enrichment mixtures were shaken on an orbital shaker at 150 rpm, 500 ml culture flasks were not shaken. Cultures were incubated at 30°C.

After 3 days of incubation, 1 ml of each culture was used to inoculate equivalent 100 ml cultures which were incubated in the same way as the source culture.

The optical density at 600 nm (OD_{600}) of cultures was taken periodically.

Dilutions of cultures were made and spread onto metaldehyde and no-carbon plates and incubated at 30°C.

3.2.2.2 Agricultural soils

Agricultural soils were collected from farms by Paul Harrington of FERA, with the permission of the land owners. These were from fields that are used to grow oil seed rape, which is often treated with metaldehyde. Collected soils were stored in plastic bags at 4°C for approximately 2 weeks. Before the enrichment began the soils were incubated at 30°C for 4 days. The 3 soils used in enrichment experiments were from 2 fields in Coleby Hall Farm, Scunthorpe and 1 field from Grange Farm, Thorney.

One gram of each soil was transferred to 100 ml of MSM containing 100 mg/L, 10 mg/L or 0 mg/L metaldehyde in duplicate flasks, all were incubated at 30°C. One of each duplicate flask was shaken at 150 rpm in an orbital shaker, the other not shaken. Fresh liquid media was inoculated with 1 ml of

the first generation of culture after 7 days. Samples of the liquid cultures of volumes 100, 10 and 1 µl were used to inoculate metaldehyde and no-carbon plates using 2 and 10 day old cultures of both generations. Plates were incubated at 30°C for up to 2 weeks. OD_{600} measurements of the second generation liquid were made at least every 2 days.

3.2.2.3 Slow sand filter

Samples from two lab scale experimental slow sand filters were supplied by C. A. Rolph (Cranfield University, personal communication). One had been exposed to raw water spiked with metaldehyde at 50 µg/l for some unspecified time, the other had been exposed to raw water only (Rolph et al., 2014). Sand and water from a slow sand filter at Ravensthorpe Water Treatment Works, Northamptonshire was also supplied by C. A. Rolph. Five grams of each sand sample and approximately 10 ml of water that was present in the samples were used to inoculate 100 ml of three culture media; MSM + 50 mg/l metaldehyde, a ten-fold dilution of MSM containing 100 µg/l metaldehyde (final concentration) and MSM only. Controls of uninoculated media containing 50 mg/l and 100 µg/l metaldehyde were also made. Cultures and controls were incubated at 30°C without shaking for 2 weeks. OD_{600} was taken at least every 2 days. Two spread plates for each volume were made using 100, 10 and 1 µl of enrichment cultures after 2 and 5 days of incubation. Spread plates were incubated at 30°C for at least 7 days. Spread plates were also made using LB agar and 0.1 µl of the original slow sand filter water, incubated for 2 days at 30°C.

3.2.3 Substrate utilisation assays

To determine if an isolated bacterial strain could grow on substrates as sole carbon and energy source, 1 M stocks of those substrates were made and added to MSM to achieve the final concentration used. Triplicate 10 ml cultures in sterile sample bottles (Sterilin) were inoculated by pipette tip transference of colony material. The following substrates were tested in this manner, their final concentration given in brackets: sodium acetate (6 mM; Sigma-Aldrich), L-malic acid (3 mM; Sigma-Aldrich), glucose (2 mM; Fisher Scientific), fructose (2 mM; Sigma-Aldrich), arabinose (2.5 mM; Sigma-Aldrich) and acetaldehyde (1 and 2 mM; Fluka brand, Sigma-Aldrich). In addition, 4 µl/10 ml glycerol/MSM cultures were made. Growth was assessed by OD_{600} daily for up to 6 days.

Utilisation of ethanol was tested by supplementation of 6 mM sodium acetate MSM with 0, 0.1, 0.2 and 1.0% (V/V) ethanol (≥ 99.8%; Sigma-Aldrich). The effect on growth of acetaldehyde was assayed using 10 mM sodium acetate and 1 mM acetaldehyde.

3.2.4 Extraction of genomic DNA

Isolated bacterial strains were cultured using 10 ml MSM with 150 mg/l metaldehyde. Cells were harvested when they reached OD_{600} of ~0.2 and ~0.16 for the *Acinetobacter* and *Variovorax* respectively. Cells were harvested by centrifugation at 4000 × g for 10 minutes. Genomic DNA was extracted using a Genomic DNA 20/G kit, with proteinase K (Qiagen) and a fresh preparation of lysozyme (Affymetrix, CA, USA), following the manufacturer's instructions, with the following exceptions. Lysis of cells was done using double volumes of all reagents. Lysed solutions were split and loaded onto 2 Genomic DNA 20/G columns each. Ethanol precipitation was used to recover the DNA as described in the manufacturer's instructions and 2 DNA extractions per strain were resuspended in 200 µl of 9 mM tris(hydroxymethyl)aminomethane, 0.1 mM EDTA (TE) 8.0 pH buffer.

3.2.5 16S polymerase chain reaction (PCR) and Sanger sequencing

PCR of ribosomal 16S sequences was performed using universal primers 8F and U785R or U1492R.

8F	AGA GTT TGA TCC TGG CTC AG
785R	GGA TTA GAT ACC CTG GTA GTC C
U1492R	GGT TAC CTT GTT ACG ACT T

Template material was added to PCR mixtures (Appendix 7.1.3) from colonies on streaked metaldehyde plates using a sterile tip. Thermocycler program: 98°C, 30 s; 30 × (98°C, 10 s; 50°C, 30 s; 72°C, 60 s), 72°C, 10 min; held at 10°C. PCR products were purified using the QiaPrep Spin Miniprep kit following manufacturer's instructions and resuspended in ultrapure water.

For RFLP analysis, 1 microgram of purified DNA from successful PCR was digested for 1 or 3 hours at 37°C using restriction enzyme *HhaI*.

Sanger sequencing of purified PCR products was performed by University of York Technology Facility staff using a 3130xl Genetic Analyzer (Applied Biosystems, Foster City). The peaks recorded by the Genetic Analyzer were interpreted using Applied Biosystems Sequence Scanner Software V1.0. Peaks were manually reviewed to find anomalies in general but specifically for evidence of multiple sequences present, and the DNA sequence determined using the software's default settings.

3.2.6 Genomic sequencing

Genomic sequencing was performed by the University of York Technology Facility using the Ion Torrent PGM system (ThermoFisher). Assembly of the short reads was performed using Newbler by Technology Facility staff. The reads in ACE format and assembled contigs in FastA format were provided.

Genomic sequencing was also provided by MicrobesNG (http://www.microbesng.uk) using MiSeq technology (Illumina) and assembled using their bioinformatics pipeline. Reads were also provided in ACE format along with contigs.

3.2.7 BLAST based method used to identify gnomically sequenced bacterial strains close to a strain isolated in this study

Predicted amino acid sequences of isolated and sequenced *Acinetobacter* strain E1 were obtained by the *in silico* translation of open reading frames of 100 bp or greater in length using Artemis V15.0.0 software (Rutherford et al., 2000). BLAST searches of predicted protein sequences against the NCBI non-redundant database were performed with default settings. The results were downloaded and a custom python script was used to count the frequency of occurrence of *Acinetobacter* strains as the top scoring hit. Where the top hit was a sequence present in multiple *Acinetobacter* strains, each strain would be counted once.

3.2.8 Average nucleotide identity (ANI)

Average nucleotide identity (Goris et al., 2007) between genomes was determined using a calculation server (http://enve-omics.ce.gatech.edu/ani/). The ANI calculation is performed by this server as follows: the query and subject genomes are fragmented *in silico* every 200 bp into overlapping 1000 bp fragments. Locally collinear blocks (LCB) between the genomes are determined by the BLAST algorithm using a gap drop-off value of 150 (this is the maximum score penalty that is allowed by the introduction of a gap before alignments containing that gap are no longer explored by the algorithm; default value is 0), nucleotide mismatch penalty of -1 (default is -3) and low complexity filtering enabled. Other BLAST parameters are default. The highest scoring LCB for each query fragment that have >70% identity and >=700 bp length are used to calculate ANI. Percentage identity is recalculated across the whole of matching fragments and the mean percentage identity across all matching blocks is reported as ANI. ANI values >95% are used to assign a species name to a particular strain (Rodriguez-R & Konstantinidis, 2014).

3.3 Results

3.3.1 *Acinetobacter* strain E1 and *Variovorax* strain E3 were isolated in enrichments for metaldehyde catabolism

Metaldehyde enrichment cultures were made with two domestic soils as described in the methods. OD_{600} were taken after allowing the cultures that contained soil particles to settle for 2 hours (Table 3-3).

Label	MA history	Shaken	MA added	Day 0	2	3	4	5	6
A	+	+	+	0.165	0.071	0.074	0.056	0.042	0.044
B	-	+	+	0.173	0.003	0.041	0.003	0.032	0.013
C	+	-	+	0.09	0.131	0.121	0.058	0.027	0.033
D	-	-	+	0.108	0.181	0.043	0.026	0.032	0.031
E	+	+	-	0.009	-0.008	0.021	-0.003	-0.005	0.01
F	-	+	-	0.022	0.022	0.036	0.01	0.036	0.031
G	+	-	-	0.004	-0.012	0.02	0.015	0.019	0.017
H	-	-	-	0	-0.002	0.003	0.014	0.03	0.021

Table 3-3. OD_{600} of 100 ml liquid enrichment cultures of domestic soil. MA history indicates whether the soil inoculum had been treated with metaldehyde in the past, shaken flasks were shaken at 150 rpm, and metaldehyde was added as 100 µl of SlugClear to those cultures indicated. Darker green shades indicate higher OD_{600}.

The metaldehyde substrate used in enrichment cultures A-H contained substances that clouded the cultures and did not settle, interfering with the optical density of cultures A, C and D. After day zero the contents of culture B aggregated, leaving the medium clear after settling.

Metaldehyde and no-carbon plates were inoculated with spreads of 100 µl of 100-fold and 1000-fold dilutions of cultures A-H at 3 days after inoculation. Plates were incubated at 30°C for 2 days. Minimal growth was seen on no-carbon plates. Metaldehyde plates inoculated with 100 µl of 1000-fold dilutions of cultures A, C, D contained 50-200 colonies. The cultures used to inoculate these plates all contained metaldehyde. The morphology of all the colonies was white, round and glossy.

Suspensions of the 2 soils were made – 1 g soil in 10 ml DI water shaken vigorously for 30 seconds and used for inoculation immediately. Ten and 100 µl of these suspensions were used to inoculate 2 metaldehyde and 2 no-carbon plates, which were incubated for 2 days. Colonies were observed on metaldehyde plates, some with different morphology from those described previously. These were less thickly growing, patches of white growth.

One colony from each metaldehyde plate inoculated with culture A, C, D and 8 colonies from soil suspension plates were picked and streaked onto metaldehyde and no-carbon plates and incubated at 30°C for 2 days. Streaks from a no-metaldehyde-history-soil suspension plate grew equally well on both metaldehyde and no-carbon plates and during visual inspection of the colonies it was observed that they had sunk into the plate. It was surmised that this isolate was utilising agarose as a carbon and energy source. All colonies picked from culture inoculated plates grew on metaldehyde plates only. Four of the 8 colonies picked from soil suspension plates grew on metaldehyde plates only. Two originated from soil without metaldehyde history and had the appearance of colonies originating from the enrichment cultures and two originated from soil with metaldehyde history, and exhibited the second, less thickly growing, morphology.

7 subcultures in total were analysed further, their identities are given in Table 3-4 along with their isolation method (enrichment culture or soil suspension). A letter indicating the soil of origin is combined with a number to give a label for each isolate.

Label	MA history	Enrichment culture
E1	+	A
E2	+	C
E3	+	None
E4	+	None
M1	-	D
M2	-	None
M3	-	None

Table 3-4. Identifying labels given to metaldehyde catabolising isolates from 2 soils (E and M); with and without a history of metaldehyde treatment. Culture name refers to cultures in Table 3-3, or none when direct plating of soils suspensions was used.

Amplification of ribosomal 16S subunit genes of the 7 isolates given in Table 3-4 were amplified using 16S primers U8F and U785R. The PCR was successful for isolates E1, E2, M1, M2 and M3. Half of the reactions were loaded onto a 1.2% (w/v) agarose gel which was run for one hour at 100V (Figure 3-1).

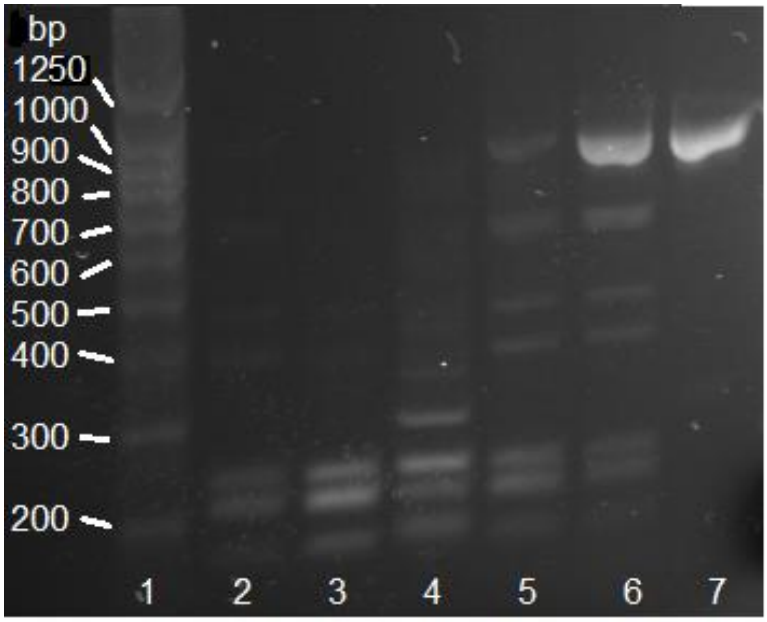


Figure 3-1. Gel electrophoresis of 16S amplicons of different isolates digested by HhaI*.* ***Lane 1,*** *Ladder (Q-step 4; York Bioscience, York);* ***2,*** *E1;* ***3,*** *E2;* ***4,*** *M1;* ***5,*** *M2;* ***6,*** *M3;* ***7,*** *E1 undigested.*

The shortest 3 bands add up to approximately 650bp and are consistent with completely digested fragments of the PCR product. The larger bands are most likely incompletely digested PCR product. The same pattern of bands can be seen in lanes 2, 3, 5 and 6. Lane 4 has a greater number of more indistinct partially digested bands which may indicate multiple 16S sequences were amplified by the PCR.

As many of the PCR failed in the previous experiment, colony PCR of the same isolates was performed using primers U8F and U1492R, rather than U785R, to give more distinct RFLP patterns. Two colonies were used of isolate E4 as there were 2 morphologies present on the streak plate, some colonies appearing to spread more. A PCR was also attempted using liquid culture samples to provide templates. A small pipette tip was dipped into liquid enrichment cultures that had been inoculated from cultures I and L and then transferred to the PCR tube. All PCR were successful and sufficiently pure DNA was obtained using a clean-up kit. One microgram of amplicons from isolates E1, E3, E4 (a and b), M1 and the enrichment cultures were digested using restriction enzyme *HhaI.* The reactions were incubated for 3 hours at 37°C. Gel electrophoresis using 1.2% (w/v) agarose was performed at 100V for 1 hour.

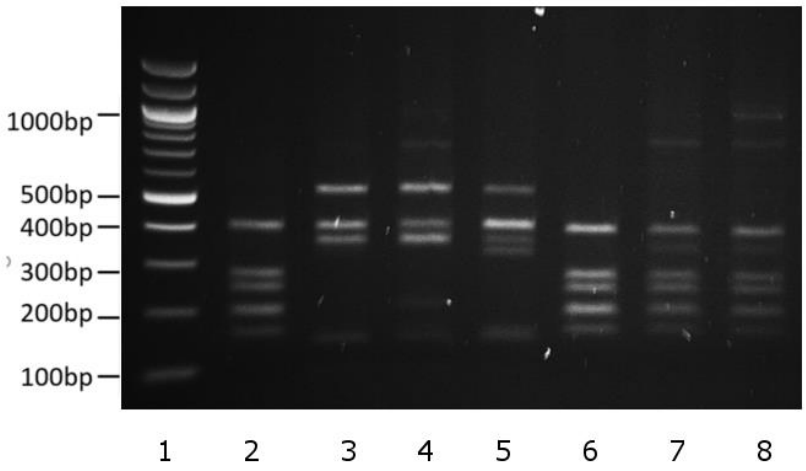


Figure 3-2. Gel electrophoresis of 16S amplicons of different isolates digested by HhaI. ***Lane 1,*** *NEB 100 bp ladder;* ***2,*** *E1;* ***3,*** *E3;* ***4,*** *E4a;* ***5,*** *E4b;* ***6,*** *M1;* ***7,*** *enrichment culture first inoculated with soil with metaldehyde history;* ***8,*** *enrichment culture of no-metaldehyde history soil.*

Isolates E1 and M1 were included in both RFLP analyses. In the second gel the fragment pattern of M1 (Figure 3-2, lane 6) is different to what was seen for the same isolate in the first gel (Figure 3-1, lane 4). The second gel appears to show the digest of a single sequence and is taken to represent the ribotype of M1. The inclusion of E1 in both analyses allows us to conclude that the ribotype of E1 is the same as E2, M1, M2, M3, and the predominant ribotype seen from samples of the enrichment cultures. A second, distinct ribotype is seen from isolates E3, E4a. E4b has an additional band ~350 bp, but this most likely results from star activity by the restriction enzyme.

Sanger sequencing was used to obtain the nucleotide sequences of the U8F-U1492R amplicons of E1, M1, E3 and E4 – two examples of each ribotype – using the U8F primer. Sequences of length 958, 964, 958 and 407 bp were obtained for each of the isolates respectively. The sequence of E1 was identical to that of M1, and the sequence of E3 was identical to E4.

The sequences of E1 and E3 were used as BLAST queries against the NCBI non-redundant nucleotide database using default settings. Thirteen sequences in the database had 100% identity and coverage to the E1 query sequence. Where species names were available they were given as *Acinetobacter calcoaceticus.* The vast majority of named species in the top 100 scoring alignments were *A. calcoaceticus* with some named as *A. rhizosphaerae.* The top scoring alignments for the E3 sequence were from *Variovorax paradoxus* strain BS064 and *Variovorax* sp. LZA10. These sequences were identical and the alignments contained one gap out of 958 identities. The top 100 hits were all *Variovorax* or an unspecified genus and the majority of named species were *V. paradoxus* with a

minority of *V. boronicumulans*. These data were taken as sufficient to assign genera to the isolates; they will be referred to as *Acinetobacter* E1 and *Variovorax* E3.

3.3.2 Metaldehyde degrading organisms could not be isolated from agricultural soils or slow sand filter samples

The enrichment and isolation of metaldehyde degrading bacteria from three agricultural soils was attempted as described in the methods. No changes in OD_{600} were observed in liquid media cultures over the 9 days they were monitored for. Bacterial growth on both metaldehyde and no-carbon plates was light and did not visibly increase after one day of incubation. Four to six colonies were picked from each metaldehyde plate after 2 days of incubation. In total 89 colonies were picked and streaked onto both metaldehyde and no-carbon plates. No streaked colony showed greater growth on metaldehyde plates.

Enrichment and isolation of metaldehyde degrading organisms was also attempted using samples of slow sand filter (SSF) sands and water from 2 experimental SSF (one that had been treated with metaldehyde and one that had not) and Raventhorpe commercial SSF, as described in the methods. No changes in the OD_{600} of enrichment cultures were seen over the course of the experiment. As in previous experiments spread plates of enrichment cultures contained small, thin colonies which ceased to grow after one day. A total of 30 colonies were picked from spread plates of enrichment cultures and streaked onto metaldehyde and no-carbon plates. None of these colonies showed greater growth on metaldehyde plates.

Spread plates of 0.1 µl of the SSF water samples onto LBA plates and incubated at 30°C. The growth after 26 and 53 hours is shown in Figure 3-3. A range of morphologies can be seen on all plates. The plates inoculated with water from the Raventhorpe SSF and the SSF treated with metaldehyde spiked water contain more colonies than the experimental SSF not exposed to metaldehyde.

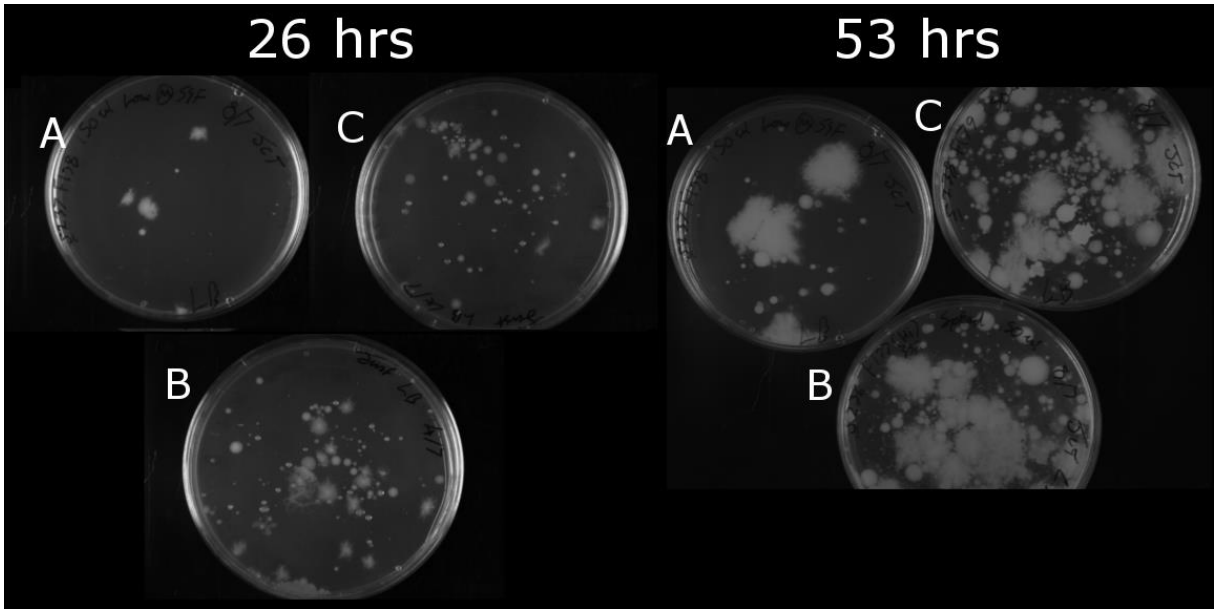


*Figure 3-3. Plates of LBA that had been inoculated with 0.1 µl of water from SSF samples and incubated at 30°C for 26 or 53 hours. Inoculum came from experimental SSF water that was not treated with metaldehyde (**A**), experimental SSF that had been treated with metaldehyde (**B**) and from Raventhorpe water treatment SSF (**C**).*

3.3.3 The disappearance of metaldehyde from minimal media is proportional to the growth of *Acinetobacter* E1 and *Variovorax* E3 in pure cultures

Three each of 2 L flasks containing 300 ml MSM with 100 mg/l metaldehyde were inoculated with a picked colony of *Acinetobacter* E1 or *Variovorax* E3 from overnight metaldehyde spread plates. An additional 3 flasks of media were not inoculated. Periodic samples were taken from each, OD_{600} was measured. At the same time 1.5 ml was centrifuged at 13 krpm and 1 ml of the supernatant transferred to a clean 1.5 ml tube avoiding the cell pellet. These samples were stored at -20°C for later analysis by LCMS.

Growth curves are shown in Figure 3-4. Log_2 plots of the growth phase of each species are shown in Figure 3-5.

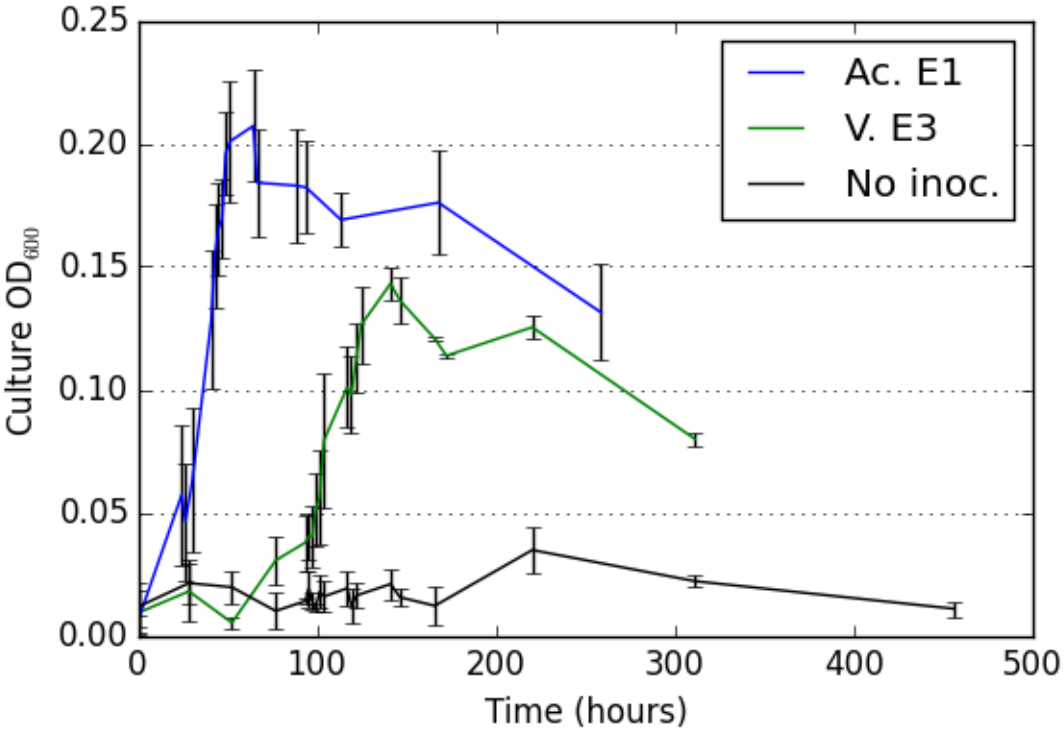


Figure 3-4. Mean (n=3) OD_{600} of cultures of MSM + 100 mg/l metaldehyde inoculated with single colonies of A. calcoaceticus *E1,* Variovorax *E3, or not inoculated. Error bars give SD of biological replicates.*

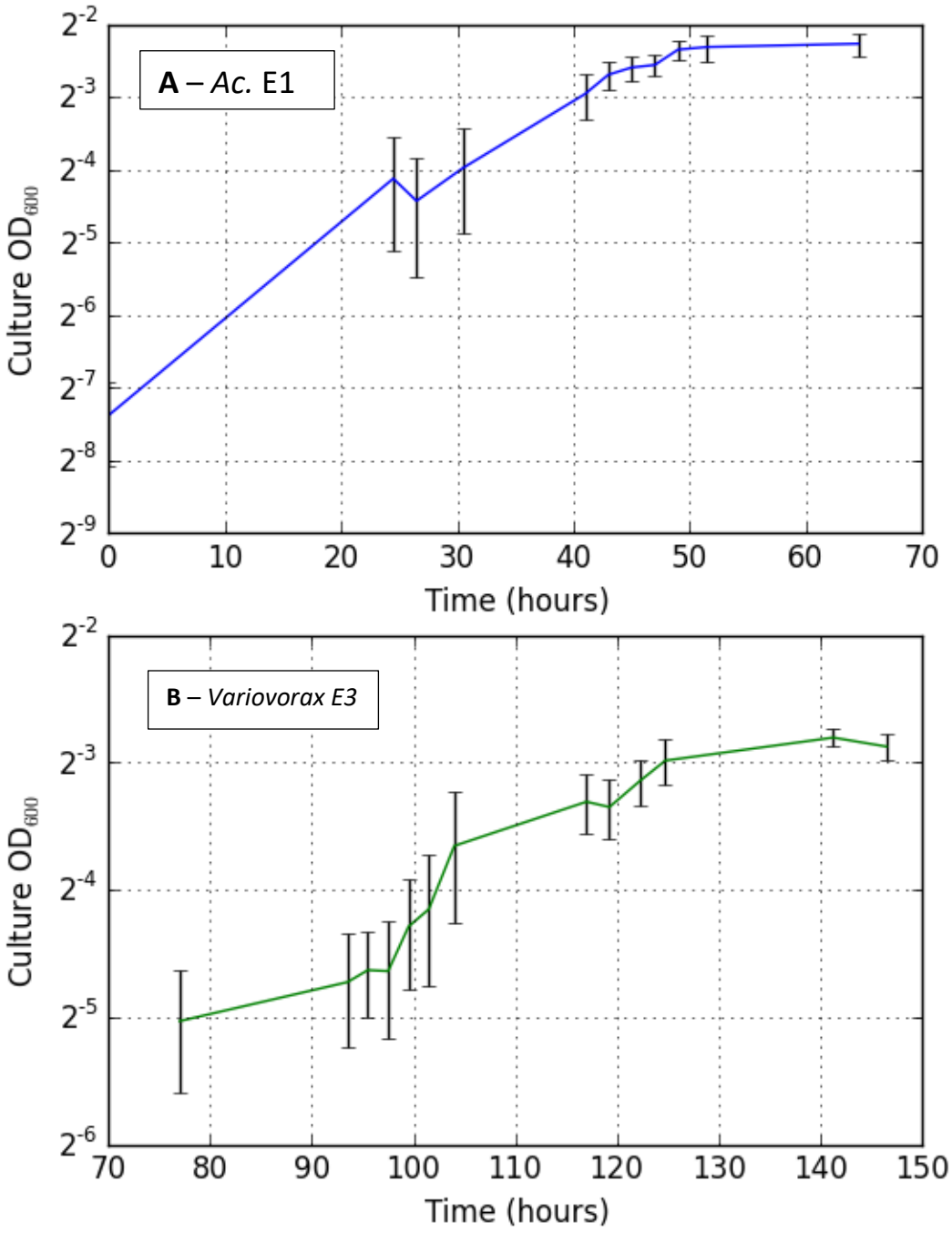


Figure 3-5. Log plots of mean (n = 3) OD_{600} of cultures of MSM + 100 mg/l metaldehyde inoculated with single colonies of the indicated strains, or not inoculated. Error bars give SD.

Doubling time was calculated for the species by taking the $\log_2$ of OD measurements during exponential growth (24.5 to 49 hours for the *Acinetobacter* and 97.5 to 122.25 hours for the *Variovorax*) and performing a linear regression of these data. Table 3-5 shows growth statistics for the two isolates. The *Variovorax* catabolises metaldehyde less efficiently; it has a longer lag time, greater doubling time and higher final concentration of remaining metaldehyde. It also peaks at lower optical density. The ratio of OD_{600} to dry cellular mass is not necessarily the same across

different bacterial species and the proportion of carbon being utilised for mass rather than being oxidised for energy may be different. *Variovorax* are motile and so may expend energy in this way.

	Acinetobacter E1	*Variovorax* E3
Doubling time (hrs.)	10.9	18.7
Lag time (hrs.)	< 24	90
Length of growth phase (hrs.)	41.5	48
MA conc. at start of stationary phase (mg/l)	< 0.1	7.08±1.41
Peak OD_{600}	0.21±0.023	0.14±0.0069

Table 3-5. Growth statistics of metaldehyde catabolising isolates grown in minimal media with 150 mg/l metaldehyde.

Metaldehyde concentration of samples taken at time points were quantified by LCMS. Samples were diluted 1000-fold with ultrapure water. Standard brackets were made up of 150, 50, 5, 1 and 0.5 µg/l metaldehyde standards measured between every 15 samples. The 0.5 µg/l standards had an RSD of 40% and were excluded from the calibration giving an LOQ of 1 µg/l for the experiment.

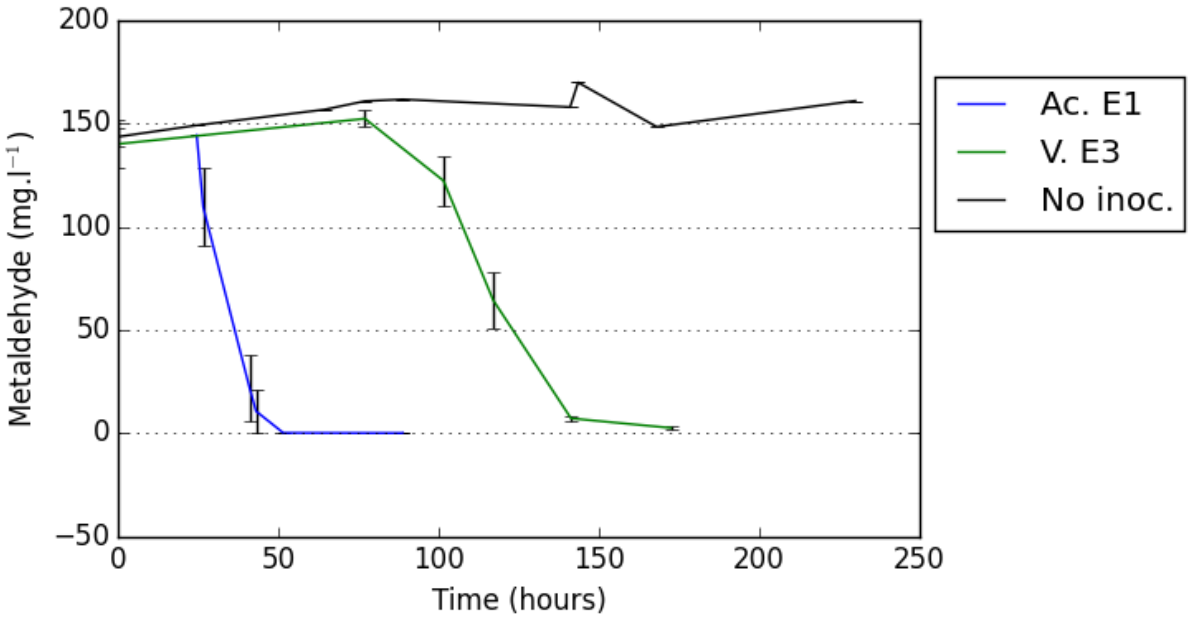


Figure 3-6. Mean (n = 3) metaldehyde in samples of cultures of MSM + 100 mg/l metaldehyde inoculated with single colonies of the indicated strains, or not inoculated. Error bars give SD.

The disappearance of metaldehyde from the cultures was correlated with the growth of the isolates (Figure 3-7). As the sole carbon and energy source present in the culture medium it can be concluded that the strains were catabolising metaldehyde for growth.

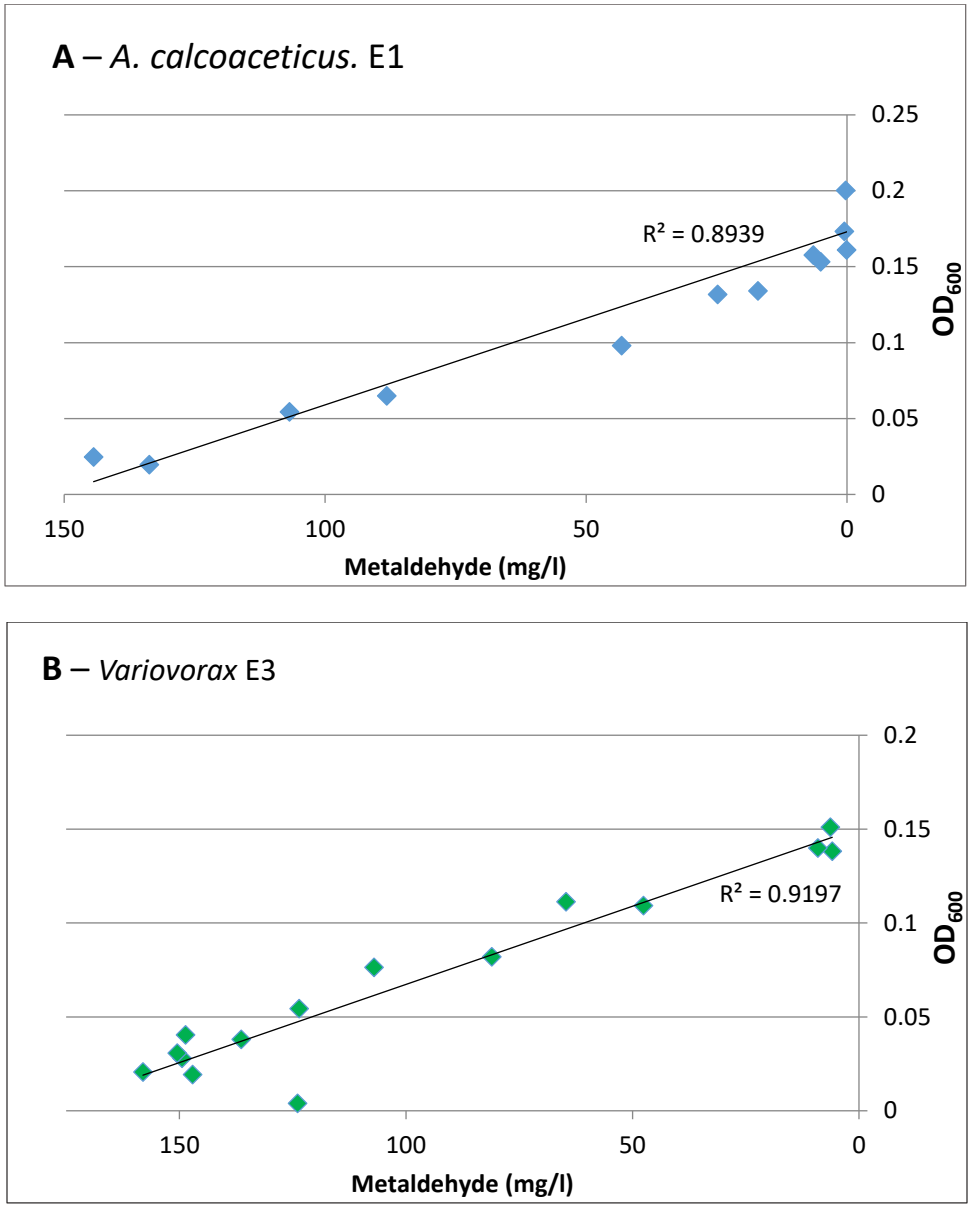


Figure 3-7. Metaldehyde concentration and optical density of samples taken from cultures of (A) Acinetobacter *and (B)* Variovorax *grown in minimal media containing metaldehyde as the sole carbon source. The lines show unweighted linear regressions of the data using the least-squares method. The R^2 value of the regression is given on the charts.*

3.3.4 Genomic sequence of isolates

The genomes of *Acinetobacter* E1 and *Variovorax* E3 were sequenced *de novo.* This was done to allow the identification of any proteins extracted from the isolates, for comparative genomic work and to further characterise the isolates

The genome of isolate *Variovorax* E3 was sequenced and assembled using Ion Torrent as described in the methods. Isolate *Acinetobacter* E1 was initially sequenced using the same method, in a separate reaction, but the assembled sequence was found to contain frequent errors and so, after attempts to

fix these errors failed (see Section 3.3.4.1, below), the isolate was resequenced by MicrobesNG. Statistics for the Ion Torrent *Variovorax* E3 and MicrobesNG *Acinetobacter* E1 assemblies are given in Table 3-6. Genes were predicted by GeneMarkS using default settings for bacteria (Besemer et al., 2001). The lengths of the draft genomes are larger than those of sequenced related species. The NCBI genome database records *Acinetobacter calcoaceticus* genomes in the range of 3.8 - 4.1 Mbp and *Variovorax paradoxus* genomes range between 6.6 – 9.7 Mbp.

	Acinetobacter E1	***Variovorax* E3**
Number of contigs >100 bp	362	459
Mean contig length (bp)	12,139	18,060
Median contig length (bp)	4559.5	8348
Mean read depth	23.38	55.31
Total length (bp)	4,394,323	8,289,841
GC content (%)	38.8%	67.4%
Predicted genes	4395	9193

Table 3-6. Assembly statistics from the sequencing of Acinetobacter *E1 and* Variovorax *E3. Read depth was calculated using data obtained by the software package Tablet V1.15.09.01 (Milne et al., 2013).*

3.3.4.1 Sequencing errors

It was discovered that the *Acinetobacter* E1 sequence contained numerous errors as part of a protein mass fingerprinting experiment that will be described fully in Chapter 5. Briefly, this experiment used amino acid sequences inferred from the sequenced genome to identify proteins from a trypsinised protein extracted from *Acinetobacter* E1 cell lysate. The 3 most abundant predicted amino acid sequences were from non-overlapping open reading frames (ORF) that were within 1500 bp of each other (Figure 3-8; ORFs labelled 2, 4 and 6). A BLAST search of the largest of the 3 predicted amino acid sequences (260 amino acids in length) against the NCBI non-redundant database returned an aldehyde dehydrogenase from *Acinetobacter baumannii* 1451147, GenBank identifier KCX52113.1, as the top hit. Manual alignment of this sequence to translated amino acid sequences in the *Acinetobacter* E1 genome lead to the construction of a sequence that is mostly identical to KCX52113.1 with some single amino acid substitutions. The arrangement of the reconstructed sequence on the *Acinetobacter* E1 genome is shown in Figure 3-8. Reanalysis of the protein mass fingerprinting data resulted in a stronger signal for the presence of the entire reconstructed protein, compared to the fragments initially identified, confirming this to be a sequencing problem.

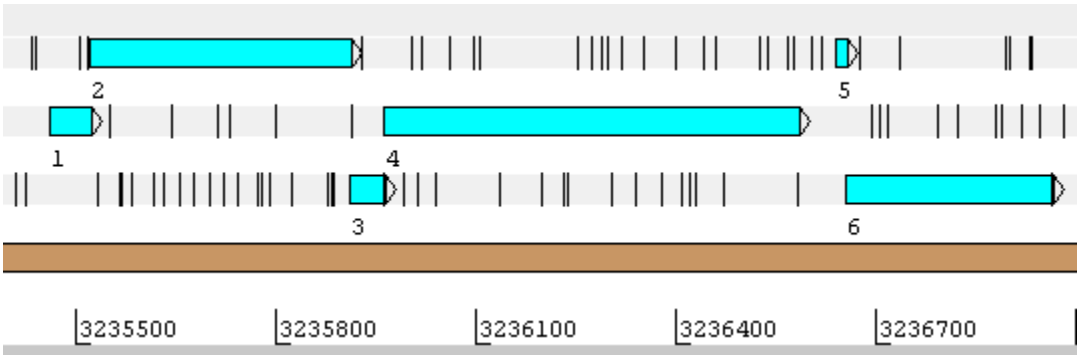


Figure 3-8. The genomic region of the putative aldehyde dehydrogenase gene in A. calcoaceticus *E1 visualised using Artemis (Rutherford et al., 2000). The scale gives nucleotide positions. The three forward reading frames are grey horizontal bars. Blue horizontal bars indicate the coding sequence for ALDH. Black vertical lines indicating stop codons.*

Examination of the of the NA sequence at the 5 apparent frame shifts showed that in each case it could be explained by an under-read of the length of homopolymeric bases; so that, for example, the sequence `ATGGGCT` was read as `ATGGCT`.

To determine the extent of these under-reads, manual alignment of housekeeping gene sequences; denylate kinase (*adk*), shikimate dehydrogenase (*aroE*) and glucose-6-phosphate dehydrogenase (*gdh*) from *A. calcoaceticus* PHEA-2 was performed all of which were found on separate contigs in the *Acinetobacter* E1 genome. *Adk* contained no homopolymer under-reads, *aroE* contained one, and *gdh* contained two. The same method was applied to the *Variovorax* E3 genome; no errors were found.

Examination of the assembly using Tablet (Milne et al., 2013) showed that the majority of reads at the locus of an error were under-reads, Figure 3-9 shows an example. Thirteen of the reads in the column denoted by the star have trimer-thymine and 7 have 4mer-thymine. This was erroneously decided to be the former in the final assembled sequence. Variation in the homopolymer lengths between 1 and 4 adenosines can also be seen in Figure 3-9.

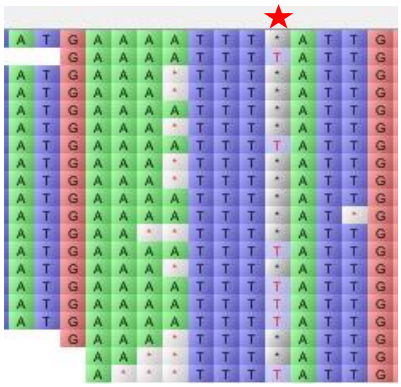


Figure 3-9. A portion of the A. calcoaceticus *E1 assembly visualised in Tablet. Individual reads are orientated horizontally and aligned. A star indicates an incorrectly called homopolymer run.*

As many of the reads did give the correct homopolymer lengths, reassembly using a different program – MIRA rather than Newbler – was attempted. It was thought that the reads selected for the final assembly may vary between the programs and that this may fix the problem. The MIRA assembly increased the overall length of ORF in the genome (Figure 3-10) which is consistent with fewer frame-shifting errors, but examination of the same genes as were previously examined in the Newbler assembly showed that 4 of the 8 previously identified errors still existed in the MIRA assembly.

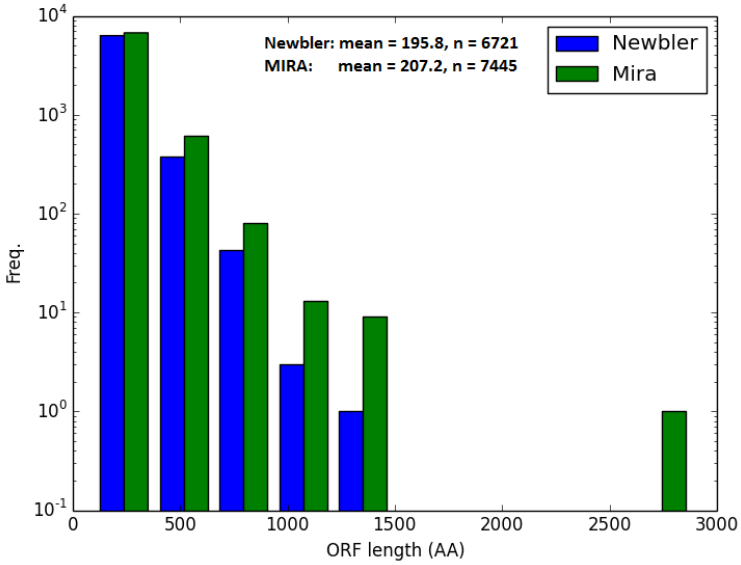


Figure 3-10. Histogram showing size of ORF in two different assemblies of A. calcoaceticus *E1 sequencing data. Mean ORF length and number of ORF are given at the top of the figure.*

Homopolymer length errors are the most frequent type produced by Ion Torrent, however errors typically include both under- and over-reads (Bragg et al., 2013). The consistent under-reads seen

here could result from an error in the sequencing reaction mixture resulting in smaller than expected pH changes when multiple bases were sequenced at once.

The genome was resequenced by MicrobesNG. The returned sequence did not contain any read-length errors resulting in frameshifts in *ALDH*, *adk, aroE* or *gdh* in the final assembly.

3.3.5 Identification of *A. calcoaceticus* RUH2202 as a strain closely related to *Acinetobacter* E1 that does not degrade metaldehyde

For future work it was desirable to identify a strain closely related to *Acinetobacter* E1 for which living samples could be obtained. BLAST of 6622 amino acid sequences from open reading frames >100 bp in *Acinetobacter*. The sum of all strain frequencies counted was 9715 due to multiple strains having the same sequence. The results were reviewed to correct for inconsistently formatted strain names, for example in the use of spaces. The top ten most frequently occurring strains and species are given in Table 3-7.

Acinetobacter strain or species	Frequency
calcoaceticus [strain not specified]	3464
calcoaceticus ANC 3680	1261
calcoaceticus RUH 2202	1074
calcoaceticus NIPH 13	1043
baumannii [strain not specified]	900
oleivorans DR1	720
sp. NIPH 542	562
sp. NIPH 817	374
calcoaceticus ANC 3811	308
calcoaceticus PHEA-2	205

Table 3-7. Frequency of Acinetobacter *species and strains returned as the top scoring hit in a BLAST search of predicted amino acid sequences from* Acinetobacter *E1 against the NCBI non-redundant database.*

Acinetobacter calcoaceticus RUH 2202 was found to be available in a culture collection, and a freeze-dried sample was purchased from the Belgium Coordinated Collection of Microorganisms. The sample was revived by being resuspended in LB, streaked onto an LB agar plate and incubated at 30°C for 24 hours. Glycerol stocks of the strain were prepared.

The lack of metaldehyde degrading capability of the strain was ascertained by streaking colonies from an LB plate onto a metaldehyde plate and the inoculation of liquid media containing 150 mg/l metaldehyde. There were no signs of growth in either media after 4 days' incubation at 30°C.

3.3.6 Phylogeny of isolates E1 and E3

The DNA sequence of the 16S ribosomal RNA genes in *Acinetobacter* E1 and *Variovorax* E3 were obtained by Sanger sequencing of purified PCR products using primers U8F and U1492R for both the PCR and the sequencing. The 2 sequences produced by the 2 primers were combined. 16S gene sequences for the type strains for species of the genera *Acinetobacter* and *Variovorax* were obtained from The List of Prokaryotic Names with Standing in Nomenclature (Parte, 2014) excepting *A. grimontii*, which has since been shown to be a synonym of *A. junii* (Vaneechoutte et al., 2008). Phylogeny for each isolate was determined separately using the type strains from its genus. Multiple alignments were performed using a MUSCLE server (Edgar, 2004) hosted by the European Bioinformatics Institute (Li et al., 2015). Preliminary alignments of all type strain and isolate 16S sequences in each genus were used to determine the common sequence regions for all species in that genus and sequences were trimmed so that only these portions were used for phylogeny. Trimmed sequences for *Acinetobacter* E1 and *Variovorax* E3 were 1261 and 1394 bp respectively. Multiple alignments were performed on the trimmed sequences, as before, and the results of this alignment used to construct a phylogenetic tree using bootstrapped Randomized Axelerated Maximum Likelihood algorithm (Stamatakis, 2006; Stamatakis et al., 2008). As outgroups, to root the trees, species were chosen from the same family as the isolates; *Polaromonas vacuolata*, a Comamonadaceae, for the *Variovorax* tree and *Pseudomonas putida*, a Pseudomonadales, for the *Acinetobacter* tree. Trees shown in Figure 3-11 visualised using Tree View, part of the software package ETE Toolkit (Huerta-Cepas et al., 2016).

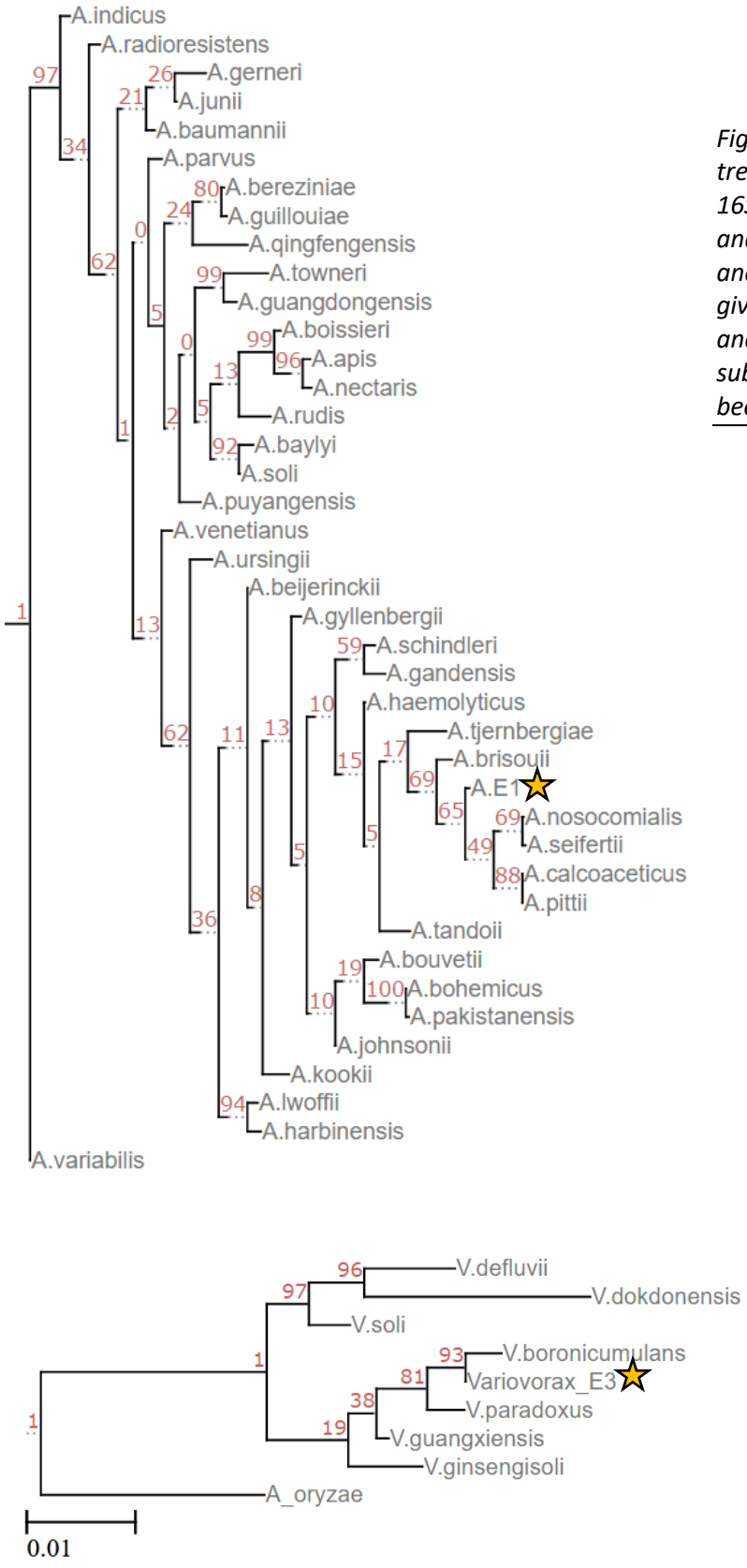


Figure 3-11. Maximum likelihood trees produced using alignments of 16S sequences of Acinetobacter *(top) and* Variovorax *(bottom) type strains and isolates. Numbers at branches give the bootstrap percentage values and scales give average number of substitutions per base. A star has been used to highlight the isolates*

Selected genomes of species that have closely related 16S sequences to the isolates had their ANI calculated against that isolate. When available, two strains for each species were selected that had at least a draft quality genome. No genome was available for *V. guangxiens or V. ginsengisoli. A. baumannii* and *A. seifertii* are included as genomic analysis of *Acinetobacter* by Touchon et al. (2014) demonstrates they form a clade with *A. nosocomialis. A. oleivorans* DR1 was also included in the analysis as it was the most frequently occurring *Acinetobacter* strain after *A. calcoaceticus* in the analysis of top scoring BLAST hits of *Acinetobacter* E1 open reading frames (Table 3-7). Results are summarised in Table 3-8.

Species	Genetic similarity between strains (%)	Strain	Mean ANI (%)	Median ANI (%)
Acinetobacter				
A. calcoaceticus	87	ANC 3680	**96.0**	**96.5**
		RUH 2202	**95.8**	**96.4**
A. oleivorans	N/A	DR1	**91.4**	**91.6**
A. pittii	84	PHEA-2	**89.2**	**89.3**
		ANC 3678	**89.2**	**89.2**
A. seifertii	N/A	NIPH 973	**86.8**	**86.8**
A. nosocomialis	85	6411	**86.1**	**86.2**
		NIPH 2119	**86.0**	**86.1**
A. baumannii	89	ACICU	**86.1**	**86.3**
		AB031	**86.1**	**86.3**
Variovorax				
V. boronicumulans	N/A	NBRC 103145	**87.7**	**88.2**
V. paradoxus	58	EPS	**87.1**	**87.5**
		S110	**86.9**	**87.2**
V. soli	N/A	NBRC 10624	**82.1**	**82.3**

Table 3-8. ANI of strains to metaldehyde degrading isolate of the same genus. Mean and median ANI values between listed strains and Acinetobacter *E1 or* Variovorax *E3. Genetic similarity gives the percentage similarity, as determined by genomic BLAST and reported by the NCBI, between the two strains of the same species used for the ANI comparison.*

The species identity of *A. calcoaceticus* can be assigned to isolate *Acinetobacter* E1 based on >95% ANI to other strains of this species (henceforth abbreviated to *A. calcoaceticus* E1). No species can be assigned to *Variovorax* E3 based on these data.

3.3.7 Carbon sources utilised by *A. calcoaceticus* E1

A. calcoaceticus E1 was found to be unable to utilise sodium acetate, glucose, fructose arabinose, ethanol, acetaldehyde and glycerol as sole carbon and energy sources. It was found to grow using sodium acetate and L-malic acid with peak OD_{600} of 0.35 and 0.13 respectively, observed after 19 hours of incubation. It is also capable of growth on MSM agarose plates containing 200 µl/L phenol, forming small colonies. *A. calcoaceticus* RUH 2202 could not utilise phenol under the same conditions.

Ninety-one per cent of *A. calcoaceticus* strains are able to utilise ethanol as a carbon and energy source (Nemec et al., 2011) and so the ability of *A. calcoaceticus* strains E1 and RUH 2202 to utilise ethanol was further investigated by supplementing 6 mM acetate with ethanol and inoculating the media with either strain E1 or RUH 2202. The peak OD_{600} after 20 hours (Figure 3-12) shows that no additional growth occurs when *A. calcoaceticus* E1 cultures are supplemented with ethanol, and it therefore cannot utilise ethanol for growth.

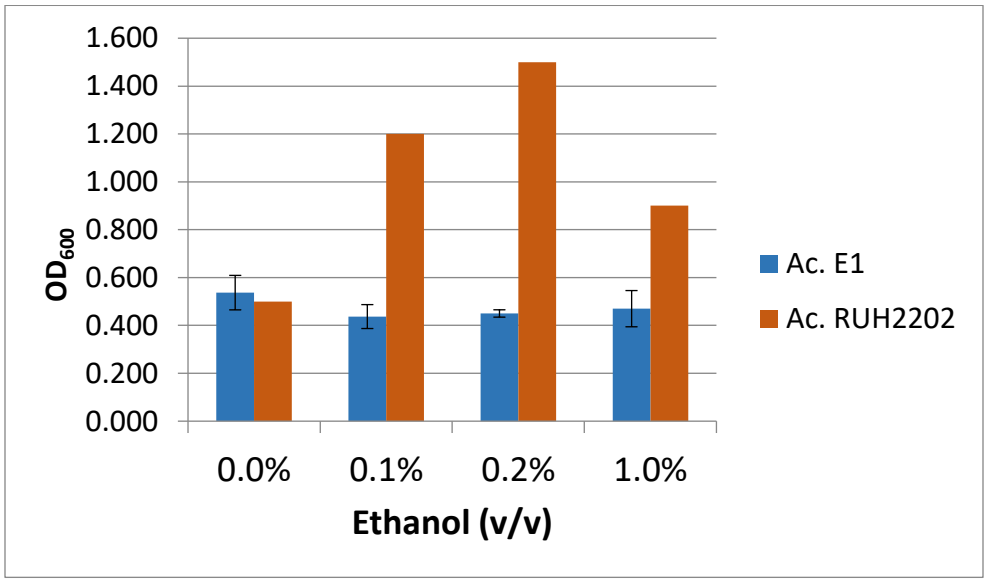


Figure 3-12. OD_{600} of 10 ml cultures containing acetate and the given volume of ethanol inoculated with one of two strains of A. calcoaceticus, *n = 3 for* Ac. *E1 and n = 1 for* Ac. *RUH 2202. Error bars represent biological repeats. All samples were diluted by half before measurements were made, the given OD values are double the observed values.*

As acetaldehyde is a likely product of metaldehyde degradation, 10 mM sodium acetate was supplemented with 1 mM acetaldehyde to observe the effect on *A. calcoaceticus* E1 growth (Figure

3-13). Concentrations of acetaldehyde that would produce measurable growth are inhibitory to the organism.

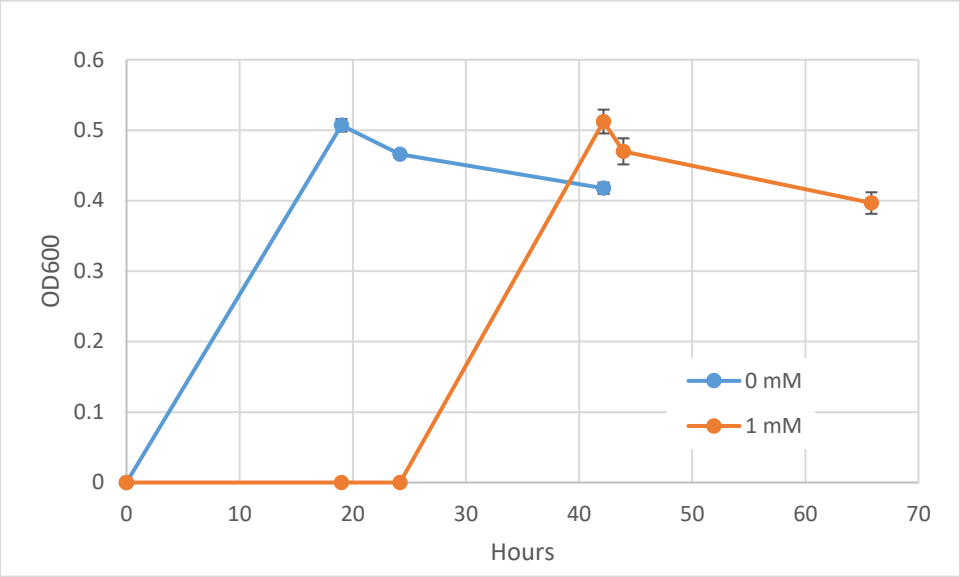


Figure 3-13. Mean OD_{600} of 3 biological replicate A. calcoaceticus *E1 cultures with 10 mM sodium acetate as the carbon source with or without 1 mM acetaldehyde supplement.*

3.4 Discussion

3.4.1 Enrichments

Prior to this work no metaldehyde degrading bacteria have been identified. In this study metaldehyde catabolising bacteria were isolated from domestic soils that were able to grow as colonies on metaldehyde plates. Amplified and digested 16S rRNA genes showed the existence of two ribotypes among the isolates. The sequence of 16S rRNA genes was obtained for 2 examples of each ribotype which were found to be identical for the ribotypes. One isolate of each ribotype from soil that had metaldehyde history were used for further experiments. Their genomes were sequenced and their species identity was investigated using average nucleotide identity. One was shown to have >95% ANI to *Acinetobacter calcoaceticus* strains and was assigned the strain name of E1. The other, a *Variovorax* with <95% ANI to known species, was assigned the strain name E3. Ribotypes consistent with *A. calcoaceticus* strains were present in both domestic soils tested, whereas *Variovorax* were only found in soil that had a history of metaldehyde treatment. *Variovorax* was not isolated from enrichment cultures, probably due to be being outcompeted by the faster growing *A. calcoaceticus.* The isolates were shown to be able to grow utilising metaldehyde as a sole source of carbon and energy. Similar enrichments and attempts at isolations did not isolate metaldehyde catabolising bacteria from agricultural soils believed to have been treated with metaldehyde in the previous 5 years, or from experimental and commercial SSF samples.

The method of isolation used could only be successful in identifying organisms that were capable of visible growth, on agarose plates or in liquid media, at 30°C using a particular minimal salts media with metaldehyde as the sole carbon and energy source. The vast majority of microorganisms do not typically grow under such conditions, even ignoring the requirement for metaldehyde catabolism (Rappé & Giovannoni, 2003). This was not considered a problem as the purpose of the isolations was to acquire strains that would be amenable for use in further experiments to identify the molecular basis for metaldehyde degradation, but it does mean that absence of cultivatable organisms from most of the enrichments is not indicative of an absence of metaldehyde catabolism. Indeed, the slow sand filters were shown in a separate experiment to degrade metaldehyde, particularly after being exposed to higher concentrations of metaldehyde for some time (Rolph et al., 2014).

A greater diversity of organisms may have been cultivated under different conditions. Pesticides would not be expected to be present at high concentrations in soils for significant lengths of time, and most bacteria would not be adept at growing in high concentrations. (Gözdereliler et al., 2013) found that different populations of bacteria were produced by liquid enrichment cultures using 25 mg/l or 0.1 mg/l of the herbicide 2-methyl-4-chlorophenoxyacetic acid. The use of gellan gum as a

solidifying agent for media improves the number of colony forming units seen when isolating bacteria as many organisms appear to be inhibited by agar (Kamagata & Tamaki, 2005). When cultivating organisms from soil, the use of a standard nutrient broth diluted to 1/100 of its usual concentration, solidified with gellan, enabled Janssen et al. (2002) to isolate 7.5% of viable cells in the sample after 12 weeks' incubation at 25°C. It is not clear to what extent the diversity of isolates cultivated is improved by this method, though several of the 30 isolates that had their 16S genes sequenced were members of subclasses or other subdivisions previously detected only in metagenomic studies. Cultivation methods using media that mimics soil, by the addition of supernatant of soil treated with a NaOH solution (Hamaki et al., 2005) or humic acids to the media (Stevenson et al., 2004) were successful in isolating novel bacteria. An issue with the application of the methods described above to the identification of metaldehyde degrading organisms is that they do not select for a particular catabolism. The addition of other potential carbon sources such as humic acids would cause obvious problems, and at low concentrations of substrate, organisms that are more efficient at scavenging other carbon substrates would also be enriched. Some aspects of the methods discussed, such as the use of gellan rather than agarose, along with lower incubation temperatures and longer incubation times could be easily implemented and may have revealed more culturable organisms. The first enrichment conducted was successful in isolating *A. calcoaceticus* E1 which seemed a very amenable organism for further studies, and so the possibility of isolating other, slow growing and fastidious organisms was not a priority. Had the inexpensive genome sequencing offered by MicrobesNG – approximately 1/10 of the price paid for sequencing of the *Variovorax* E3 genome – been available at the earlier stages of this work, then the acquisition of more metaldehyde degrading isolates that could be used for comparative genomic studies may have been more vigorously pursued.

Other methods were considered and dismissed due to difficulty and expense. Ashida et al. (2010) describes the utilisation of a micromanipulator to sort cells that are elongating in the presence of a substrate. Their elongation is made visible to a detector by their successful staining with 5-carboxyfluorescein diacetate acetoxymethyl ester which binds to esterases present in cells during elongation. Stable isotope probing is a method that can be employed to identify the genomic material of cells that catabolise, directly or indirectly, a ^{13}C labelled substrate (Radajewski et al., 2000). An environmental sample is treated with the labelled substrate for some time. Cells that utilise the substrate for growth will have daughter cells that contain DNA made of the labelled carbon. Genomic material is then extracted from the sample and the denser, ^{13}C labelled DNA can be separated by ultracentrifugation.

Metaldehyde degrading organisms that were not cultured in this study could be investigated using culture independent methods. Where enrichment methods fail to isolate individual strains that can grow clonally on the substrate of interest then functional metagenomic approaches may be used (Ferrer et al., 2005). Extraction of a bacterial population's genomic information directly from a substrate and its transformation into, for example, a lab strain of *E. coli,* allows for interrogation of this genomic information without the need for cultivating environmental strains in minimal media. Enrichment of xylanase degrading bacteria using rice straw, extraction of DNA from the culture, cloning and expression of this DNA into *E. coli* which was then assayed for xylanase activity was used to identify xylanase genes by Mo et al. (2010).

3.4.2 Isolated bacteria

3.4.2.1 Phylogenetics

In this work, ANI was used to assign taxonomic names to the isolates. *Variovorax* E3 could not be assigned a species name based on ANI. The isolate was only compared to strains with a defined species, and may be similar to other *Variovorax* isolates that do not have a species.

A different approach can be employed to determine phylogenetic relationships among a group of related organisms that uses the inferred proteome of an organism. The general method is to take the predicted amino acid sequences from a genome, determine those which are orthologous in the group, concatenate the orthologous protein sequences into one long sequence for each strain, perform multiple sequence alignment on the orthologous sequences and construct a phylogenetic tree based on evolutionarily plausible sequences of mutations. Many different algorithms and software packages can be used to perform the different steps of such an analysis. Two recent examples of the construction of phylogenetic trees of *Acinetobacter* strains are given by (Fanelli et al., 2015; Fondi et al., 2016). Using concatenated amino acid sequences allows each organism's proteome to be treated as a single sequence. However, each protein has its own evolutionary history subject to different selection pressures on the proteins and by horizontal gene transfer. For these and other reasons, the evolutionary history revealed by phylogenetic analysis of concatenated orthologous proteins is often not congruent with the analysis of individual proteins. This casts doubt on the interpretation of aligned concatenated sequences, especially at deeper phylogenetic levels (Thiergart et al., 2014).

Due to the nature of bacterial genetics in particular, there is no definitive way to determine the taxonomy of an isolate. The purpose of assigning taxonomic names here is to aid in the discussion and research of the isolates. For this, ANI is sufficiently robust.

3.4.2.2 General description of Acinetobacter and Variovorax from the literature

The genus *Acinetobacter* (Brisou & Prévot, 1954) are gram-negative γ-proteobacteria of the order Pseudomonadales. They are described as rod-shaped, obligately aerobic, non-motile (the latin *acinetus* means non-moving), though twitching motility has been observed in at least 3 species including *A. calcoaceticus.* (Gohl et al., 2006; Henrichsen, 1975; Skiebe et al., 2012). *Acinetobacter calcoaceticus* was designated the type species for *Acinetobacter* by Baumann et al. (1968) when reclassifying oxidase negative, catalase positive *Moraxella* strains based on their utilisation of a wide range of carbon substrates.

The genus *Variovorax* was defined by Willems et al. (1991). Its latin meaning is given as "devouring a variety [of substrates]". It is described as oxidase positive, catalase positive, aerobic and motile, with yellow colonies. Chemoorganotrophy is the primary trophic classification, but *Variovorax* strains are also capable of lithoautotrophy utilising hydrogen as an energy source, leading Willems et al. to assign the name *V. paradoxus* to the type strain for the genus. Motility on agarose plates and yellow colour in liquid cultures using a clear medium were observed for *Variovorax* E3 (data not shown).

3.4.2.3 *Other* Variovorax *isolations*

The species most closely related to *Variovorax* E3 with a sequenced genome is *V. boronicumulans* (highest observed mean ANI of 87.7%) which was first isolated from soil and identified in a screen as having the ability to accumulate boric acid to a higher concentration that closely related species (Miwa et al., 2008). The second most closely related species *V. paradoxus* (highest observed mean ANI of 87.1%).

Both species are frequently isolated in enrichments using synthetic compounds as the substrate. Linuron (a pesticide) mineralising *Variovorax* strains have been isolated in at least 4 separate studies. Enrichments using Belgian agricultural soil with 10 years' history of linuron application and linuron as the sole carbon, nitrogen and energy lead to the isolation of 5 strains on solid media. Only *Variovorax* strain WDL1 could grow in pure culture (Dejonghe et al., 2003). A study using Danish agricultural soil as inoculum found two strains that could mineralise linuron, one of which was *Variovorax* strain SRS16 which had a 16S rRNA sequence 99% identical to *Variovorax* WDL1 and grew more quickly on linuron in pure culture (Sørensen et al., 2005). Breugelmans et al. (2007) repeated the work of Sørensen et al. using a sample from the same field with the addition of samples from two other Belgian fields, one of which had no history of linuron treatment. From the replicated field a *Variovorax* with a 99% identical 16S sequence to *Variovorax* SRS16 was found as the only linuron degrader; from the other field with linuron usage history 2 strains with 99% and 98% 16S sequence

identity to *Variovorax* WDL1 were isolated, along with 2 non-*Variovorax* strains; from the field with no linuron history, enrichment cultures did not produce a linuron degrading consortium. Satsuma (2010) isolated linuron degrading bacteria from Japanese river sediment using microcosms. Four unique 16S sequences were found in 8 isolates tested, which had least 98.6% identity. In these 4 studies, *Variovorax* strains have been repeatedly isolated as the primary degraders of the pesticide linuron. Only one environmental sample was seen to contain cultivatable primary linuron degraders that were not *Variovorax.* A culture independent study has shown the involvement of other taxa that have not been isolated by enrichments (Dealtry et al., 2016).

A *V. boronicumulans* strain able to utilise ibuprofen as sole carbon and energy source has been described (Murdoch & Hay, 2015). Reference is made, by Murdoch and Hay, to the (unpublished) isolation of *V. paradoxus* able to degrade polybutylene sebacate (a new generation plastic) and the aromatic herbicide bromoxynil octanoate. From soil that showed a high level of degradation of the neonicotinoid pesticide thiacloprid, a *V. boronicumulans* strain was identified as catalysing the first reaction in its degradation (Zhang et al., 2012).

The industrially useful chemical intermediate mercaptosuccinate was found to be degraded by 15 isolates present in water or soil samples near to a chemical plant that produced the compound. One isolate was found to be a *V. paradoxus* and given the strain name B4, and 14 belonged to two genera that were found to be potential opportunistic human pathogens and so were not studied further (Carbajal-Rodríguez et al., 2011).

Enrichment of soil from an industrial area lead to the isolation of 3 bacterial strains, 2 of which were *V. paradoxus* (strains TBEA6 and SWFT), that could utilise the antioxidant thioester 3,3'-thiodipropionic acid as sole carbon source (Bruland et al., 2009).

From deep sea sediment, *V. paradoxus* T4 was isolated by enrichment cultures and shown to mineralise dimethyl terephthalate in minimal media (Wang & Gu, 2006).

Of two 2,4-dinitrotoluene (DNT; a precursor to the explosive 2,4,6-trinitrotoluene) degrading consortia enriched from soil and water samples from near a former ammunition production site, one consortium contained a strain of *V. paradoxus* that catalysed the first steps of DNT degradation but required the other two members of the consortium for sustained growth (Snellinx et al., 2003).

Three isolates were obtained from enrichment cultures using sulfolane, a xenobiotic used to remove hydrogen sulphides from natural gas, as the sole carbon and sulphur source. The isolate determined to have the highest sulfolane to sulphate turnover rate was *Variovorax* WP1. (Greene et al., 2000)

The biopolymer, poly(3-hydroxybutyrate) was found to be utilised as sole carbon and energy by 9 bacterial and 10 fungal species in microcosms. *Variovorax paradoxus* strains were enriched from 4 of the 5 soils tested at 28°C (Mergaert et al., 1993)

In these enrichments and isolations, it is seen that strains related to *Variovorax* E3 are either isolated alone, are the primary degrader or are the only isolated strain capable of full mineralisation of a substrate. Typically, the enrichments are from soils, but also from a river bed and deep-sea sediment. *Variovorax* are metabolically diverse, widely distributed and amenable to cultivation under typical enrichment conditions, the enrichment of a *Variovorax* strain from soil using a xenobiotic substrate is not a surprise.

3.4.2.4 Other A. calcoaceticus *isolations*

Acinetobacter calcoaceticus strains are able to utilise a wide range of carbon sources as growth substrates, but few carbohydrates (Fewson, 1967). A review of the literature shows that *A. calcoaceticus* strains have been isolated from environmental samples in enrichments and shown to be able to utilise as a substrate for growth; the herbicide atrazine (Mirgain et al., 1993), nylon precursor caprolactam (Rajoo et al., 2013), crude oil (Reisfeld et al., 1972), diesel (Singh & Lin, 2008), ethanol (Abbott et al., 1973), malonate (Kim & Kim, 1985), oxhratoxin A (Hwanga & Draughon, 1994), palm oil (Koh et al., 1985), phenol (Yuquan et al., 1999), swainsonine (a plant produced toxin; Zhao et al., 2008), and volatile fatty acids from "swine wastes" (Jolicoeur & Morin, 1987). In many of these isolations only *A. calcoaceticus* strains were identified. Atrazine and caprolactam are, in common with metaldehyde, xenobiotic compounds, but it does not appear that any further work with the isolates produced in those studies has been conducted. Though *A. calcoaceticus* strains seem to be less metabolically versatile then *Variovorax* strains, it is also a species that can easily be enriched using a variety of carbon substrates.

Chapter 4: Dynamics and constitution of the MA degrading pathway in *A. calcoaceticus* E1

4.1 Introduction

From the perspective of a bacterium, regulation of the expression of proteins is desirable to prevent resources being wasted on the production of unnecessary proteins, to avoid a protein performing an action that has a negative effect, and to ensure the production of appropriate proteins in response to a constantly changing environment. Regulation can be achieved by control exercised over any stage in the production of a mature protein; transcription, post-transcriptional modification of ribonucleic acid (RNA), translation and post-translational modification of the polypeptide. Typically, a recently evolved catabolic process will be constitutively active at a low level. Basic regulation of its expression will, over time, be superimposed at the transcriptional level by the opportunistic recruitment of transcriptional regulators that happen to have some small degree of influence over the gene or operon (Cases & Lorenzo, 2001).

Quantification of protein and gene expression is possible. Differences in the quantity of expressed proteins can be determined by the separation of the expressed proteome by polyacrylamide gel electrophoresis of proteins denatured by sodium dodecyl sulphate (SDS-PAGE) and visualisation of the protein bands produced. A protein visualised on a gel can then be extracted, treated with the proteolytic enzyme trypsin – which fragments the polypeptide in a specific and predictable way – and its mass spectrum obtained. This mass spectrum can be compared to the predicted mass spectra for predicted proteins trypsinised *in silico,* and the identity of the extracted protein determined. This process is known as protein mass-fingerprinting. The quantification of specific transcribed RNA is possible using quantitative polymerase chain reactions (qPCR), and total transcribed RNA can be quantified by whole transcriptome RNA sequencing. Where a gene is induced in a specific way by a substrate, the application of these methods can lead to the identification of that gene.

Reconstitution of a metabolic activity of interest *ex vivo* can lead to the identification of the enzyme responsible following the fractionation of the proteome and testing those fractions for activity. For example, Wang et al. (2015) identified the enzyme complex responsible for degrading the herbicide acetochlor (2-chloro-N-(ethoxymethyl)-N-(2-ethyl-6-methylphenyl)-acetamide) from the lysate of a *Rhodococcus* species, fractionating its proteome first using ammonium sulphate precipitation and then further separating the active fraction by hydrophobic interaction chromatography, identifying the peptides present in the final active fraction by protein mass fingerprinting.

The metabolism of a substrate by a bacterium proceeds through a series of chemical transformations catalysed by enzymes. The product of one reaction becomes the substrate for the next. The non-fermentative oxidation of a substrate for the ultimate production of adenosine triphosphate (ATP) results in the reduction of oxygen to water, and the oxidation of the carbon source to carbon dioxide. Some quantity of the carbon can also be diverted to anabolic processes or stored in some form to satisfy future anabolic or energetic requirements. Exploration of the kinetics of metaldehyde degradation were hindered by the lack of a real-time assay that could be repeated in a reasonable time-frame. The number of samples that could be quantified by LCMS was limited by the time available on the equipment. For these reasons, an assay that uses real-time measurements of oxygen utilisation in washed cell suspensions treated with metaldehyde and other carbon substrates was developed and a novel step-wise metabolic model was employed in an attempt to gain insights into the kinetics of the metabolism of metaldehyde.

Experiments described below were undertaken to determine if genes in the metaldehyde degrading pathway were induced by the presence of metaldehyde, and the expressed proteome was examined to identify those involved in metaldehyde metabolism. Attempts were made to reconstitute metaldehyde degradation using cell lysates. The kinetics of metaldehyde metabolism were investigated by quantifying metaldehyde degradation in washed cells directly and several approaches were attempted to infer information about the pathway from observations of the oxygen utilisation of cells treated with various carbon sources.

4.2 Materials methods

4.2.1 Transposon library generation and validation

To induce genetic lesions in the genome of *A. calcoaceticus* E1, plasmids with transposable selective markers were transferred by mating with *E. coli* hosting the plasmid.

4.2.1.1 E. coli SM10 λ pir-pGCS62

The transposon-containing plasmid used was pGCS62 (Cornells et al., 1991). It features the transposon Tn5 containing tetracycline resistance gene Tc1 (Tn5-Tc1), a β-lactamase gene and the R6K origin of replication. The host strain is SM10 λ *pir* that produces the pir protein required by the R6K origin of replication (Simon et al., 1983).

4.2.1.2 Mating

Rifampicin resistant *Ac.* E1 and *E. coli* SM10 were grown overnight in LB with 50 µg/ml rifampicin and 20 µg/ml tetracycline, respectively. Aliquots of the cultures were span down at 13 krpm and washed twice with LB before being resuspended in 1/5 volume of LB. Typically 100 µl of each culture were mixed in a sterile tube (negative control *Acinetobacter* were placed in a sterile tube and mixed with sterile water) and dispensed onto a sterile nitrocellulose membrane placed on an LBA plate with no antibiotic. The plate with membrane was then incubated at 30°C or 37°C for 1-24 hours to allow mating to occur. The membrane was placed in a petri dish, the cells washed off and spread onto LBA plates with 50 µg/ml rifampicin and 20 µg/ml tetracycline before being incubated overnight at 30°C. Colonies growing on the plates could then be assayed for loss of metaldehyde metabolism.

4.2.1.3 Polymerase chain reaction (PCR) assay to determine plasmid maintenance

As low-level innate resistance to ampicillin in *A. calcoaceticus* E1 meant that acquired ampicillin resistance could not be used, primers (Table 4-1) were designed to determine if the plasmids transferred by mating were being maintained in the strain. To give an amplicon (1029bp) from the IS50R insertion sequence of Tn5 (Auerswald et al., 1981) and an amplicon (708bp) of part of the vector pGCS62 (specifically a fragment of pGP704 used in its construction; sequence from Christie-Oleza et al. (2013)). Colonies to be assayed were streaked onto fresh plates and incubated overnight at 30°C to avoid plasmid contamination from *E. coli* on the transformant plate. Colony PCR was performed using the following concentrations of reagents: 1X GoTaq Green buffer, 200 µM dNTPs, 200 µM primers, 0.5 U GoTaq DNA polymerase (Promega) in 20 µl total volume. The following thermocycle was used: 2m 95°C, 30 * (60s 95°C, 45s 60°C, 65s 72°C), 5m 72°C. Amplicons were separated by electrophoresis 1.2% agarose gel and visualised with SYBR safe.

Primer name	Oligo sequence
Tn5 IS50R F	CCA TTG AGG ACA CCA CCT CT
Tn5 IS50R R	CCA TAA AAC CGC CCA GTC TA
pGP704 F	CCT TCC AGA CGA ACG AAG AG
pGP704 R	CGA TAT GGA TGT GCA GGT TG

Table 4-1. Sequences of primers for colony PCR used to amplify regions of the plasmid pGCS62 to determine if it is present in the colony.

4.2.2 Washing and concentration of cells for *in vivo* assays, and the preparation of cell lysate for *in vitro* assays

Oxygen and metaldehyde utilisation assays were performed with washed live cells or cell lysate. The process to obtain washed cells is described here.

A colony of *A. calcoaceticus* E1 grown on a metaldehyde plate, or *A. calcoaceticus* RUH 2202 grown on LBA with 50 µg/L rifampicin, was used to inoculate a 5 ml LB starter culture with 50 µg/L rifampicin and incubated for 17 hours at 30°C. From the starter culture a flask of minimal media was inoculated from the starter culture; 500 ml minimal media containing 0.85 mM metaldehyde was inoculated with 500 µl *A. calcoaceticus* E1 starter culture, and 150 ml minimal media containing 20 mM sodium acetate was inoculated with cultures of *A. calcoaceticus* E1 or RUH 2202. Cultures were then incubated for 20 hours at 30°C and shaking at 140 rpm. Cells grown with 0.85 mM metaldehyde were harvested when their OD_{600} reached 0.18 and cells grown on 20 mM sodium acetate were harvested at OD_{600} of 0.80.

Sodium acetate cultures were diluted to 500 ml using pH 7.0 100 mM phosphate buffer (PB). Cultures were centrifuged at 5000 × g and 4°C for 15 minutes (Sorvall Evolution RC with SLC-6000 rotor, Thermo Fisher Scientific), the supernatant removed and the cell pellet resuspended in 500 ml 100 mM PB, and this step repeated with the pellet resuspended in 100 ml 100 mM PB. This was split between two 50 ml tubes and centrifuged again at 4000 × g (Allegra X-22R; Beckman Coulter, Brea, USA) and 4°C for 15 minutes. The supernatant was removed, and the cells were resuspended in the final buffer to be used, or ultrapure water.

Where *in vivo* oxygen assays alone were to be performed, the cell pellets were resuspended to OD_{600} 1.0 ± 0.1 in PB or 20 mM phosphate buffer (pH 7.0). Cell suspensions were kept on ice for 1 hour before being used in experiments.

Where cells were lysed they were resuspended in PB or (where other buffers were being used) in ultrapure water and lysed by sonication using 3 minutes of total process time with 1.5 seconds on and 7.0 seconds off, and sonication energy set to give a maximum reading of 50 watts. Samples were

kept on ice to prevent the build-up of heat. Lysate in ultrapure water was mixed with buffers at twice the desired final concentration in 1:1 ratios.

4.2.3 Oxygen utilisation in live cells or lysate

4.2.3.1 Oxygen electrode

The quantification of oxygen was achieved using a Clark electrode. This instrument measures the current at an electrode that results from the following reaction taking place in the saturated potassium chloride medium through which the current passes: $O_2 + 2H_2O + 4e^- \rightarrow 4OH^-$. Oxygen from the sample chamber passes through an oxygen permeable membrane at a rate proportional to the oxygen concentration in the chamber, and so changes in the concentration of oxygen in the sample result in changes in the current produced.

Perspex Dissolved Oxygen Electrodes (Rank Brothers, Bottisham, UK) controlled using a Dual Digital Model 20 controller (Rank Brothers) were used here. The electrodes were prepared with a reservoir of saturated potassium chloride next to the electrode, separated from the sample chamber with a 13 µm thick polytetrafluoroethylene sheet. The electrode was set to -0.61 polarising volts. Analogue electrical signals produced by the oxygen electrode were converted to digital by a PicoLog High Resolution Data Logger ADC 20 (Pico Technology, St Neots, UK) attached to a personal computer running PicoLog Recorder V5.24.8. PicoLog was set to record at a voltage range ±1250 mV and conversion time 180 ms.

4.2.3.2 Experimental runs

From cell suspensions or lysate on ice, 3.5 ml would be taken and transferred to a 15 ml tube immersed in room temperature water for at least 10 minutes to equilibrate the sample's temperature. The temperature of the sample would be taken and incubation would continue until it was at most 2°C below room temperature. 3 ml of cells or lysate were transferred to the oxygen electrode chamber and the oxygen level monitored for at least 2 minutes to ensure it was stable before proceeding. Injections of substrates were made into the oxygen electrode using gel loading tips (Alpha Laboratories, Eastleigh, UK). The apparatus includes magnetic stirrers to maintain homogenous oxygen concentration in the sample chamber, these were set to 5.5 units, greater speeds were not used to avoid lysing the cells.

4.2.3.3 Quantification of data

The voltage generated by oxygen-equilibrated deionised water was determined each day for each electrode before any experiments were run. This was taken to represent 290, 287 or 278.75 µM dissolved O_2 for room temperatures of 19°C, 20°C or 21°C respectively (U. S. Geological Survey,

2011). Voltage readings taken during experiments were converted to µM O_2 assuming linear proportionality between observed voltage and dissolved oxygen.

4.2.4 Oxidation of acetaldehyde by *Saccharomyces cerevisiae* aldehyde dehydrogenase

Aldehyde dehydrogenase (ALDH) from *S. cerevisiae* (Sigma-Aldrich) catalyses the reaction acetaldehyde + β-NAD^+ + H_2O → acetate + β-NADH + H^+. The lyophilised enzyme was dissolved into the buffer being tested plus 0.02% (w/v) bovine serum albumin, at a concentration of 0.3 mg/ml (0.24 U/ml). The enzyme preparations were portioned into 1 ml single use aliquots and stored at -20°C. β-NAD^+ was prepared in deionised water made to pH 4 with HCl at concentrations of 5 mM and stored in single use aliquots at -80°C. Dithiothreitol (DTT; PlusOne) stocks were prepared at 100 mM, stored at -20°C and defrosted a maximum of 3 times before being disposed of. Assays used final concentrations of 200 µM DTT, 20 mM KCl and 200-400 µM β-NAD^+. Different buffers, reaction conditions and substrates were used, as described in the results.

Real time assays of the reduction of β-NAD^+ to β-NADH + H^+ was performed using a UV-1601 spectrophotometer running UV Probe v1.01 (Shimadzu, Kyoto, Japan).

4.2.5 Approximate steady state kinetics for carbon substrate utilisation

The fastest rates of oxygen reduction were determined from oxygen utilisation data by performing multiple linear regressions of 30 second regions, excluding the first 30 seconds where the treatment injection sometimes caused a transient peak. Regressions were done over 30 seconds to balance the effect of noise and the fact that most treatments do not result in a steady rate. Values for K_M and V_{max} were calculated by fitting the parameters of the Michaelis-Menten equation, shown below, to observed fastest oxygen utilisation rates (v) and the initial metaldehyde treatment concentration (S).

$$v = \frac{V_{max}\, S}{K_M + S}$$

Michaelis-Menten equation

4.2.6 Simulation of the metaldehyde oxidation pathway

Step-wise models of metabolic pathways were constructed using the Python programing language. The general model used to construct curves fitted to oxygen utilisation data is described here. With each tick of the model, the concentration of 4 simulated substrates changes according to the equations shown below, starting with the transformation of Substrate 1 (S1) to Substrate 2 (S2) as a function of Reaction A (Rxn A):

$$[S1]_n = [S1]_{n-1} - [Rxn\ A]([S1]_{n-1})$$

$$[S2]_n = [S2]_{n-1} - [Rxn\ B]([S2]_{n-1}) + [Stoich] * ([S1]_n - [S1]_{n-1})$$

$$[S3]_n = [S2]_{n-1} - [S2]_n$$

$$[Oxy]_n = [Oxy]_{n-1} - [S3]_n * OxyF$$

The stoichiometry of S1 to S2 (Stoich) is 4 when the experimental data being fitted resulted from metaldehyde treatment, representing the 4 acetaldehydes or equivalent yielded from metaldehyde. The final equation depletes oxygen proportional to the product of Rxn B (i.e. S3) multiplied by the oxidation factor (OxyF). Note that S3 does not accumulate. Reactions A and B behave according to Michaelis-Menten kinetics or convert substrate at a rate linearly proportional to its concentration:

$$MM\ kinetics = \frac{[S] * V_{max}}{[S] + K_M}$$

$$Linear\ kinetics = [S] * Rate$$

Values fitted to experimental data for OxyF, Rate or the K_M and V_{max} for each modelled reaction were obtained using the Levenberg–Marquardt damped least-squares algorithm, using the Python package Lmfit v0.9.5 (http://lmfit.github.io/lmfit-py).

Other simulated pathways are discussed that are constructed from chains of these reactions.

4.3 Results

4.3.1 Transformation of *A. calcoaceticus* E1 with pGCS62

To successfully transform *A. calcoaceticus* E1 by mating with *E. coli* SM10 different incubation times and temperatures were investigated. The key condition was found to be temperature. Incubation of the membrane with mixed cells at 37°C produced transformants while incubation at 30°C did not. Incubation times of 6 hours or more lead to visible growth of the bacteria on the membrane and so 3 hours incubation was used. Colonies that grew on the rifampicin and tetracycline plates were mostly *A. calcoaceticus* E1 transformants. *E. coli* SM10 that had gained rifampicin resistance could be distinguished by morphology, *E. coli* colonies are darker than *A. calcoaceticus* E1 when growing on LBA.

4.3.1.1 Assay of plasmid maintenance

Colony PCR of transformants produced amplicons of the correct size for the vector (Figure 4-1) showing that the vector was maintained in all tested colonies. This result means tetracycline resistance could not be used to isolate transposon mutants and the method was abandoned.

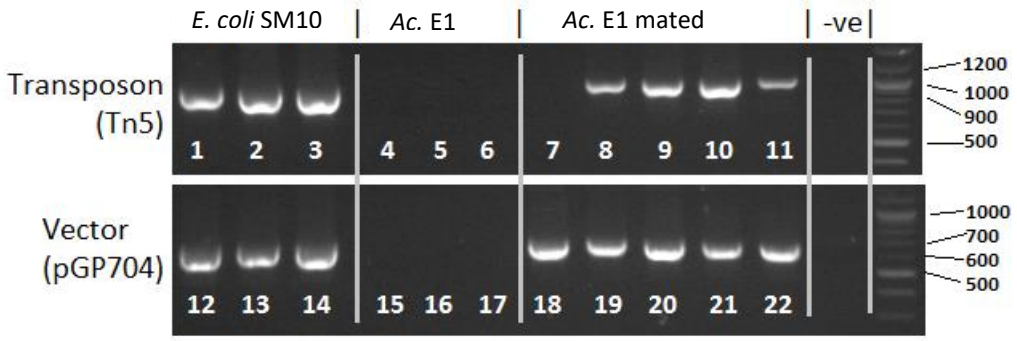


Figure 4-1. Agarose (1.2% w/v) gel electrophoresis of the products of colony PCR using primers specific for the Tn5 or pGP704 regions of pGCS62. E. coli *SM10-pGCS62 untransformed* A. calcoaceticus *E1 and E1 that acquired tetracycline resistance after mating with the SM10 were used to provide template DNA in the reactions.*

4.3.2 Metaldehyde-dependent oxygen utilization is induced by metaldehyde

To determine the feasibility of isolating enzymes in the metaldehyde degrading pathway by analysis of the proteome, it was important to establish whether or not enzymes involved in the pathway were induced by the presence of metaldehyde. Measurements of the metaldehyde-dependent oxygen

utilisation of cell suspensions of *A. calcoaceticus* E1 grown on different substrates were made to determine if oxidation pathway was induced by metaldehyde.

The assays were performed on *A. calcoaceticus* E1 cells that had been grown using a 10-fold dilution of LB, or MSM with 10 mM acetate or 0.85 mM metaldehyde as the carbon source that had been concentrated to OD_{600} 0.25 suspended in 100 mM PB. The cells were treated with 85 µM metaldehyde. Figure 4-2 shows that oxygen was utilised at a faster rate when the cells were grown with metaldehyde, and the pathway is therefore induced by metaldehyde.

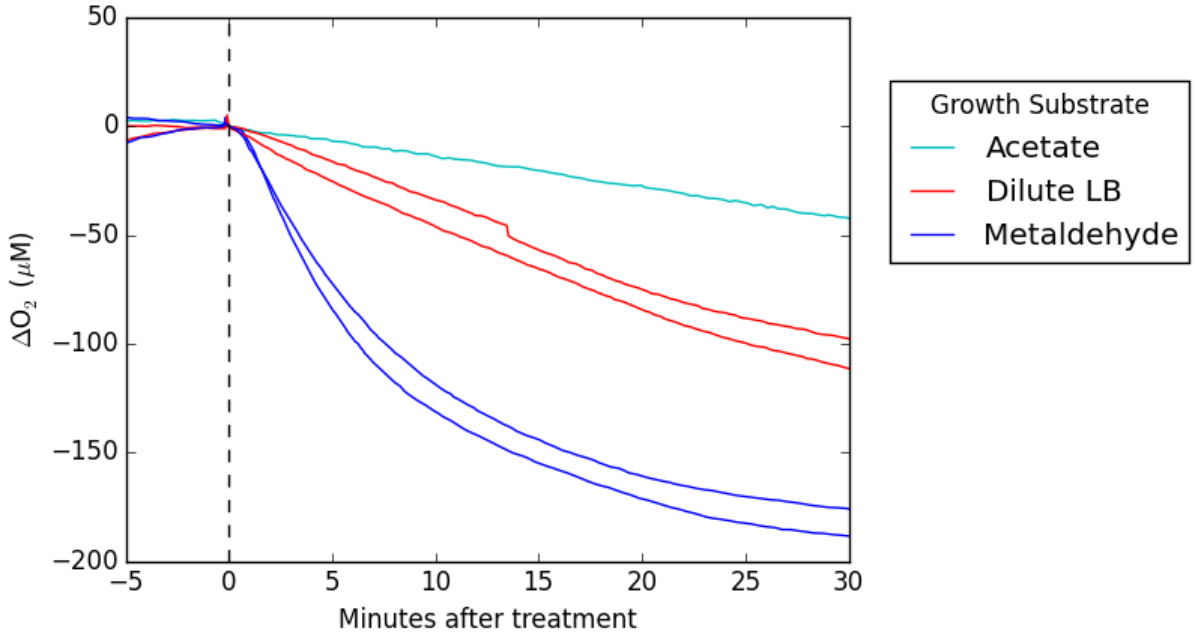


Figure 4-2. Oxygen utilisation of A. calcoaceticus *E1 grown using minimal media with acetate or metaldehyde as the carbon source, or 10X dilute LB media. Each line represents a single assay.*

4.3.3 Metaldehyde degradation by *A. calcoaceticus* E1 treated with metaldehyde is not dependent on culture substrate

4.3.3.1 Oxygen utilisation and metaldehyde degradation of A. calcoaceticus *E1 cells grown with metaldehyde or acetate*

To determine if the degradation of metaldehyde (not only the utilisation of oxygen) was induced by growth on metaldehyde, washed cells grown on different substrates were treated with metaldehyde. Cultures of *A. calcoaceticus* E1 grown in minimal media with 0.85 mM metaldehyde or 20 mM sodium acetate, and *A. calcoaceticus* RUH 2202 grown with 20 mM acetate were harvested in the late growth phase. The cells were washed and concentrated to OD_{600} 1.0 in PB. The oxygen and metaldehyde utilisation of cell samples treated with 53 µM metaldehyde was measured in cell suspensions (Figure 4-3). Cells that were used to assay metaldehyde directly were placed into oxygen

electrodes without the stopper being affixed and stirred at the same rate as the oxygen assayed suspensions. Samples were taken for LCMS analysis and filtered before being stored at -20°C. Biological triplicates were used in each case. The data files for the oxygen utilisation of the third *A. calcoaceticus* E1 and RUH 2202 grown using acetate (Figure 4-3D) were corrupted and could not be recovered, so only two of each are shown.

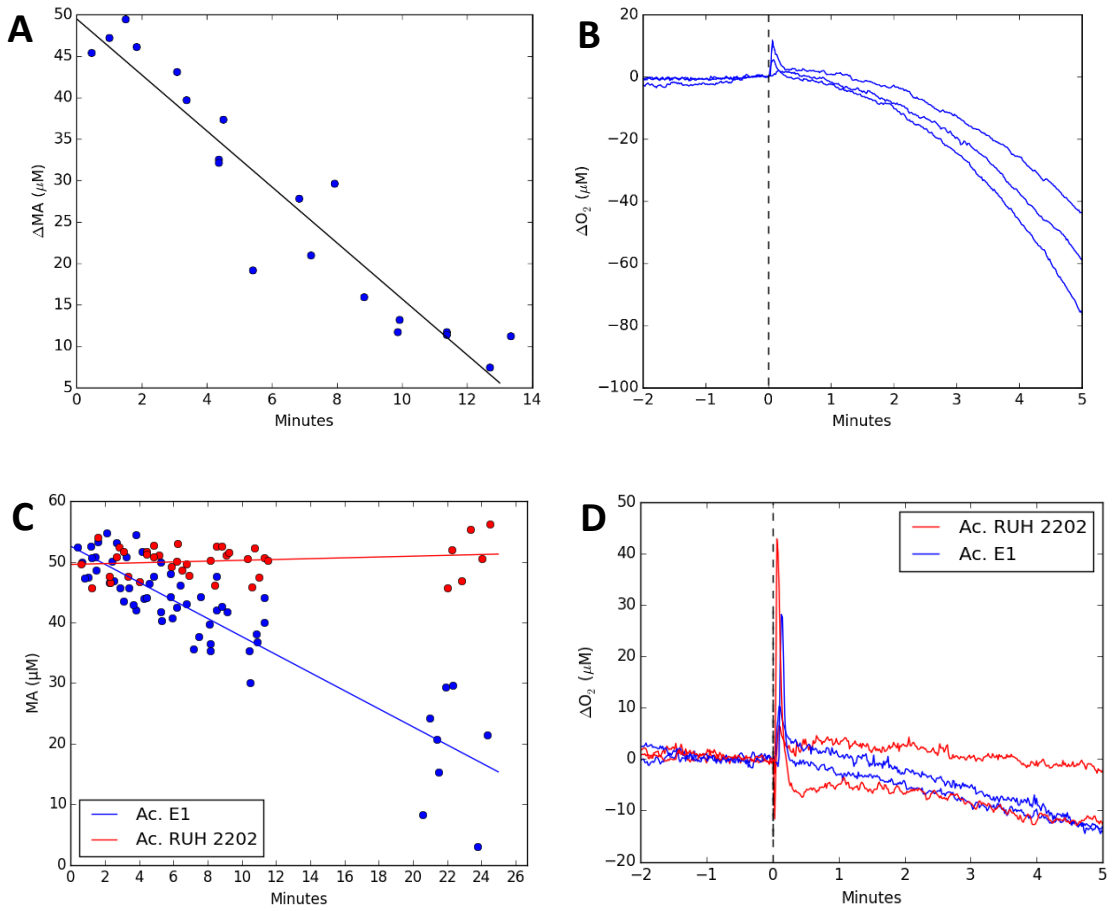


*Figure 4-3. Oxygen and metaldehyde utilisation in samples of washed cells at OD_{600} 1.0 treated with 53 μM metaldehyde. The X-axis gives minutes after treatment. Straight lines show the results of linear regressions. Charts **A** and **B** show results from* A. calcoaceticus *E1 grown with metaldehyde. Charts **C** and **D** show* A. calcoaceticus *E1 and RUH 2202 grown with acetate.*

Linear regressions were performed of the data shown in Figure 4-3. Oxygen measurements between 3-5 minutes were used to determine the rate of oxygen utilisation, while all metaldehyde measurements were used in regressions (Table 4-2).

Growth substrate	O_2 (μM/min)	O_2 R^2	MA (μM/min)	MA R^2
Metaldehyde	-27.8 -20.8 -16.9	>0.99 for all	-3.38	0.904
Acetate	-4.11 -2.49	0.925 0.907	-1.49	0.645

Table 4-2. Linear regressions of metaldehyde and oxygen utilisation data from A. calcoaceticus *E1 cells grown in either acetate or metaldehyde and treated with metaldehyde. Oxygen regressions were performed on measurements made between 3-5 minutes after the cells were treated with metaldehyde.*

The mean rate of oxygen utilisation by cells that had been grown on metaldehyde was 6.6-fold greater than cells grown with acetate, while the observed metaldehyde utilisation rate was 2.3-fold greater. These data show that the metaldehyde degrading step itself is less tightly regulated than subsequent steps which are required for the oxygen dependent catabolism of metaldehyde.

To obtain a more precise picture of the kinetics of metaldehyde degradation, the direct quantification of metaldehyde was repeated with some refinements. More frequent samples were taken from each biological repeat and the cells were suspended in 20 mM PB (rather than 100 mM) so that the samples taken need only be diluted 2-fold prior to LCMS analysis. This produced larger absolute readings less affected by background noise. This method required larger initial starting volumes, and so 100 ml of cell suspension was added to a 1 L beaker, treated with metaldehyde and manually shaken for 10 seconds before the first sample was taken and briefly after each sample. The R^2 for the calibration was 0.988 and an LOQ of 0.057 µM. The resulting metaldehyde quantification data were sufficiently noise free that individual curves for each culture were distinct (Figure 4-4). Oxygen utilisation by cells treated with metaldehyde (Figure 4-5) was not obtained under the same conditions as the metaldehyde data, and so is not directly compared.

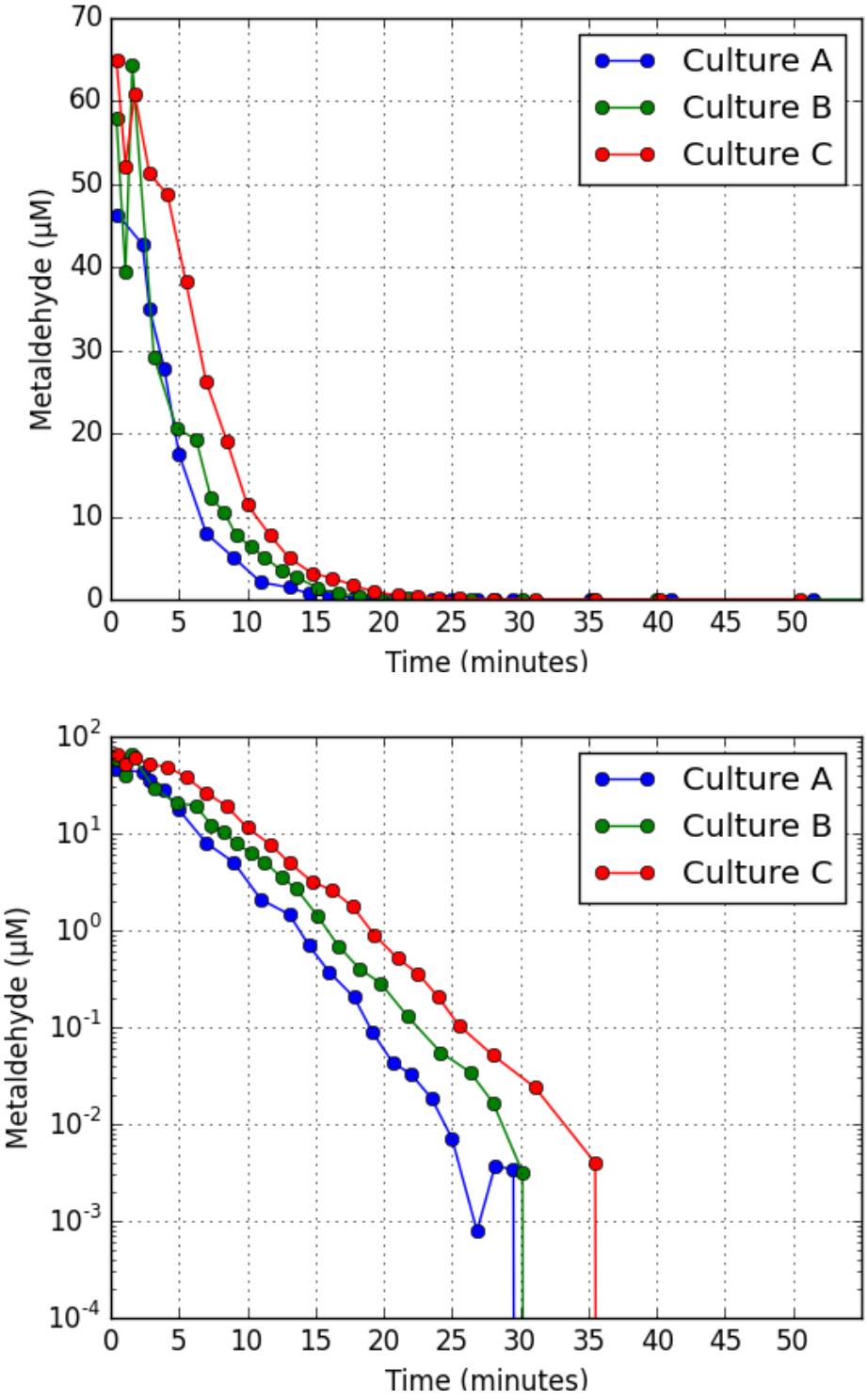


Figure 4-4. Degradation of 53 µM metaldehyde by washed A. calcoaceticus *E1 cell suspensions cultured with metaldehyde as the substrate. The same data are shown in both charts with linear (top) and logarithmic (bottom) scales.*

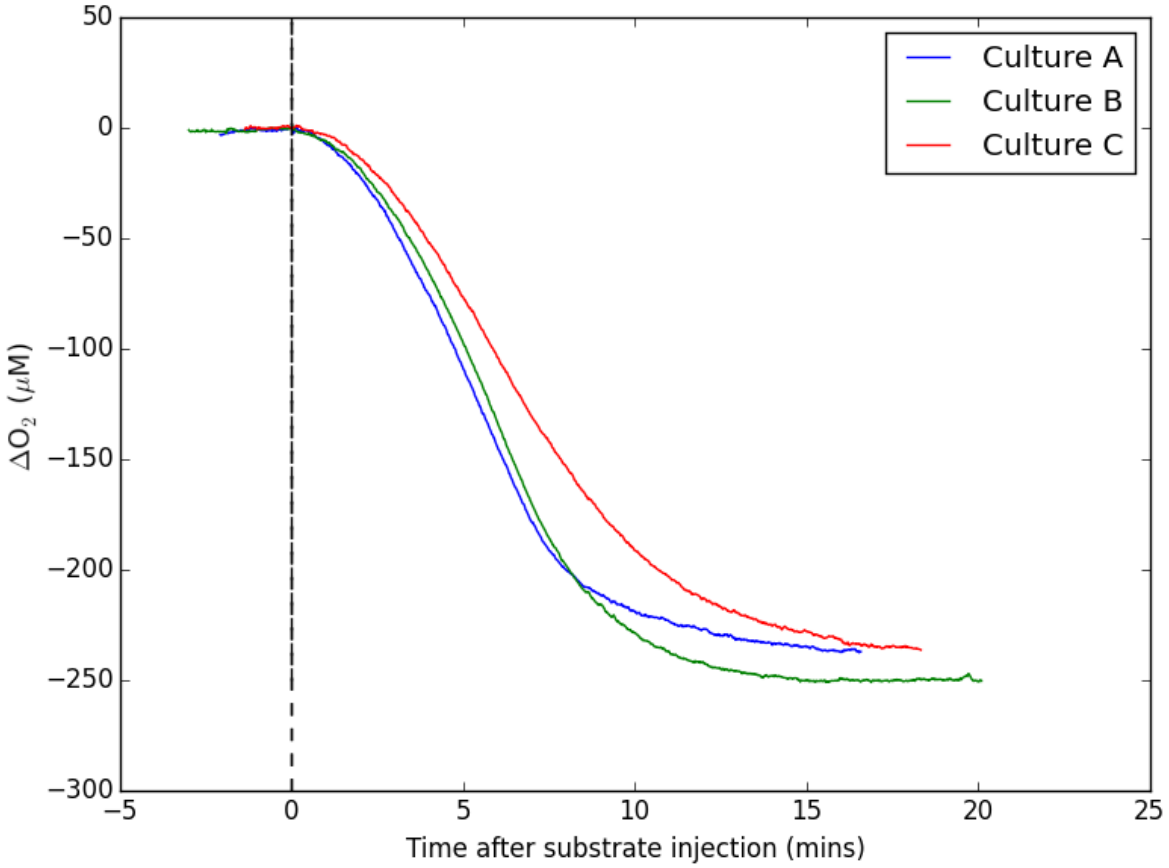


Figure 4-5. Oxygen utilisation of washed A. calcoaceticus *E1 cell suspensions cultured with metaldehyde as the substrate treated with 53 µM metaldehyde.*

The earliest metaldehyde measurements (< 3 minutes) deviate from the expected values of 53 µM, especially in cultures B and C, possibly due to insufficient mixing, but after this the curves look smooth.

Values for Michaelis-Menten kinetics for this process were obtained by fitting the metaldehyde utilisation data shown in Figure 4-4. The first measurement for culture A, and the first 3 measurements of cultures B and C were excluded from the fits.

The velocity between each measurement was calculated using the following equations:

$$v_n = \frac{S_n - S_{n+1}}{t_{n+1} - t_n}$$

These values were fitted to the Michaelis-Menten equation using $S_n + (S_n - S_{n+1})/2$ as the value for *S.* The weight for v_n used when calculating the fit was determined by the length of its interval $(t_{n+1} - t_n)$. The observed velocities and fitted lines are shown in Figure 4-6; fitted values for K_M and V_{max} are given in Table 4-3.

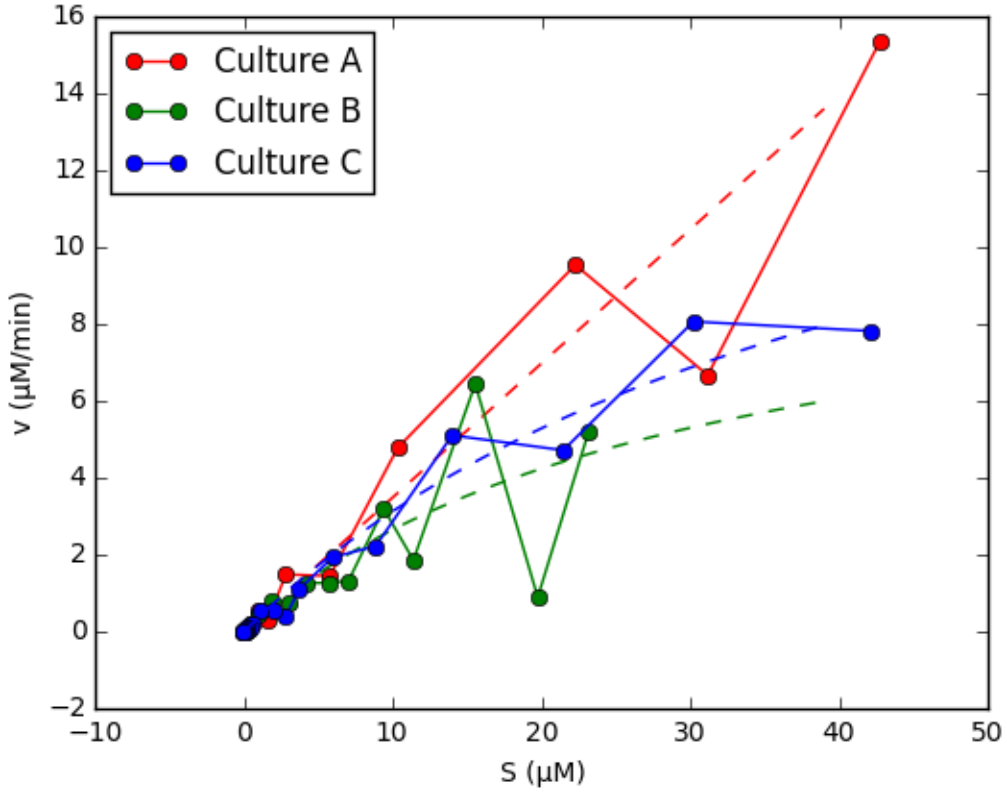


Figure 4-6. Calculated velocities versus substrate concentration from metaldehyde degradation data. Dashed lines show fitted Michaelis-Menten kinetics.

Culture	K_M ± SE (µM)	V_{max} ± SE (µM/min)	K_M / V_{max}
A	4.05e+8 ± inf	1.41e+8 ± inf	2.87
B	30.1 ± 39.1	10.6 ± 9.65	0.768
C	44.5 ± 14.5	17.0 ± 3.41	2.62

Table 4-3. Fitted values for K_M and V_{max} from metaldehyde degradation data transformed into velocity versus substrate concentration.

The value for K_M for culture A exceeds the concentrations used in the experiment by 10^7-fold. This results in an effectively straight line at the concentrations seen, and as many solutions to the Michaelis-Menten equation result in a straight line the error of the fit is infinite.

4.3.4 Acetaldehyde is the primary product of metaldehyde degradation

4.3.4.1 An aldehyde dehydrogenase is highly expressed in A. calcoaceticus *E1 grown with metaldehyde*

As the metaldehyde degrading pathway in *A. calcoaceticus* E1 was shown to be regulated to some degree, an examination of its proteome was undertaken. *Acinetobacter calcoaceticus* E1 was

cultured in MSM using 0.85 mM metaldehyde or 10 mM sodium acetate as the carbon source. An SDS-PAGE was performed. Two gels are shown in Figure 4-7. An additional band is visible in the gels at around 55 kDa only in lysate from cells grown using metaldehyde. The location of the band or its absence is indicated by arrows in Figure 4-7A.

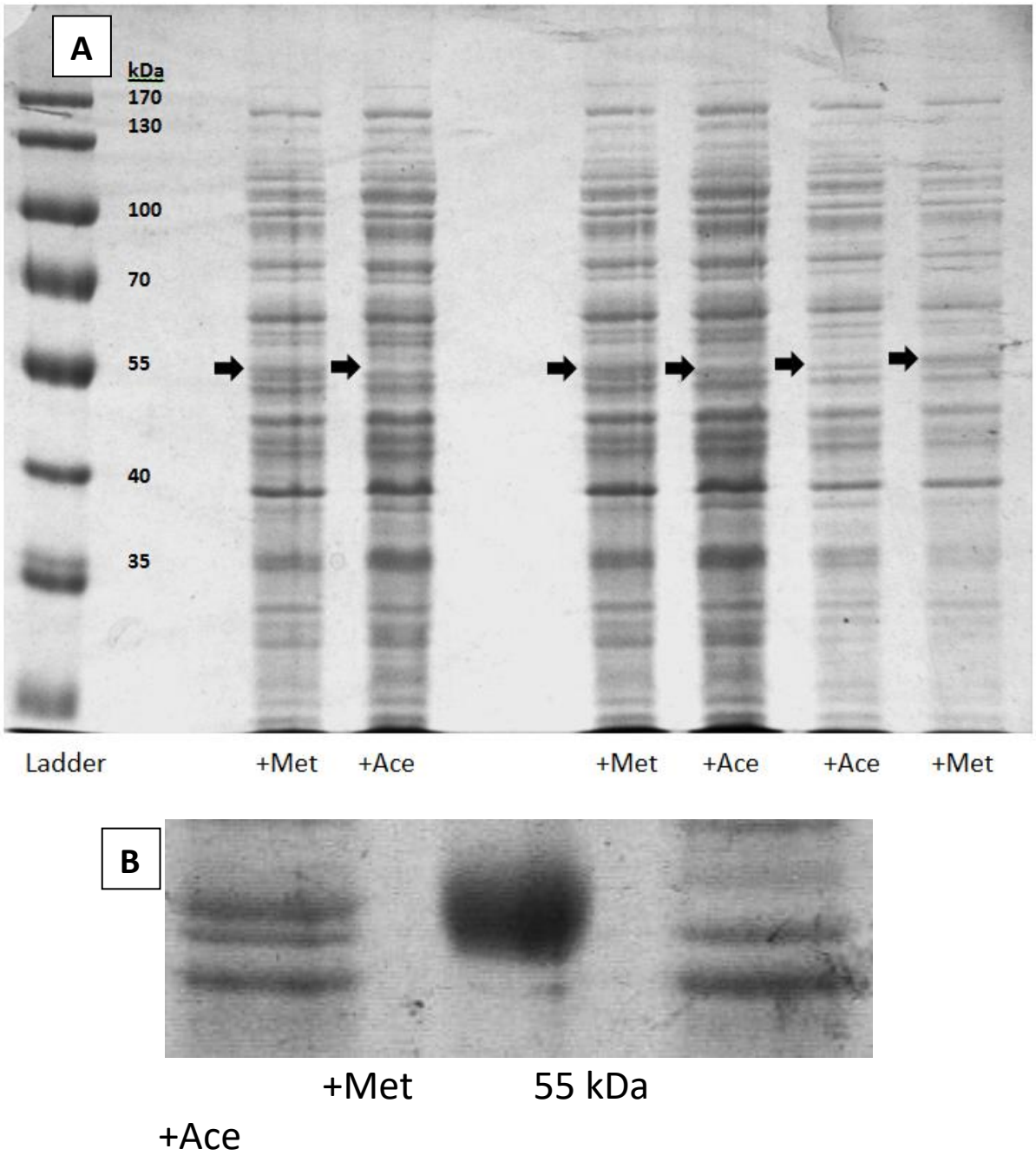


Figure 4-7. Images of 2 SDS-PAGE protein gels (12% acrylamide) of denatured cell lysate of A. calcoaceticus *E1 cultured using metaldehyde (+Met) or sodium acetate (+Ace) as the carbon source. Image **A** shows lysate from 6 cultures, arrows indicate the location of an additional band only seen in lysate from metaldehyde grown cells. Image **B** shows a zoomed in section of a different gel where the additional band is clearly visible in the +Met lane.*

The additional band present in the gel shown in Figure 4-7, and the equivalent region from the +Ace lane were excised and the identity and abundance of proteins present in the excised sections were determined. One protein with a calculated mass of 55.3 kDa was found to be enriched from 1.0% to 33.7% of the total molar mass of proteins in the sections. The next largest enrichment was from 0.9% to 3.1%. InterPro (Hunter et al., 2012) was used to detect signatures of protein family or domains in the amino acid sequence (Table 4-4).

Database	Detected signature	Database reference
SUPERFAMILY	Aldehyde dehydrogenase-like superfamily	(Gough et al., 2001)
CATH	Aldehyde Dehydrogenase; Chain A, domain 1 (3.40.605.10) & Aldehyde Dehydrogenase; Chain A, domain 2 (3.40.309.10)	(Sillitoe et al., 2015)
PFAM	Aldehyde dehydrogenase family (PF00171)	(Finn et al., 2016)
Prosite	Aldehyde dehydrogenase glutamic acid and cysteine active site residues conserved	(Sigrist et al., 2013)

Table 4-4. Functional classification of a protein expressed in A. calcoaceticus *E1 grown using metaldehyde as a carbon source. InterProScan was used to query multiple databases and the results from each database are shown here.*

The enriched protein was predicted to be an aldehyde dehydrogenase (ALDH). The predicted gene ontology term was GO:0016620, "oxidoreductase activity, acting on the aldehyde or oxo group of donors, β-NAD or β-NADP as acceptor". Such proteins can have a wide range of aldehyde substrates (Riveros-Rosas et al., 2013). The two active site residues (Joernvall & Persson, 2006) were found to be conserved. Three amino acids (Met124, Cys301 and Cys303; numbering as human ALDH2) shown to be involved in aldehyde binding (Riveros-Rosas et al., 2013) were found to be conserved in the induced ALDH based on the examination of an alignment of the induced protein with the human ALDH2.

The oxidation of aldehydes by ALDH occurs when the aldehyde enters a catalytic tunnel, indicated in Figure 4-8 by a white ring. Catalytic residues are located on the peptide loop inside the ring in Figure 4-8; it does not seem feasible that metaldehyde, given its "puckered crown" structure, could enter the catalytic tunnel to be oxidised. Given this, and the fact that an enzyme identical to the *A. calcoaceticus* E1 ALDH is present in *A. calcoaceticus* RUH 2202 (as determined by a BLAST search of the sequence against the *A. calcoaceticus* RUH 2202 predicted proteome), it is very unlikely that metaldehyde could be acted on directly by this enzyme and so it is theorised that metaldehyde is depolymerised by an unknown enzyme to acetaldehyde which is then oxidised by the ALDH.

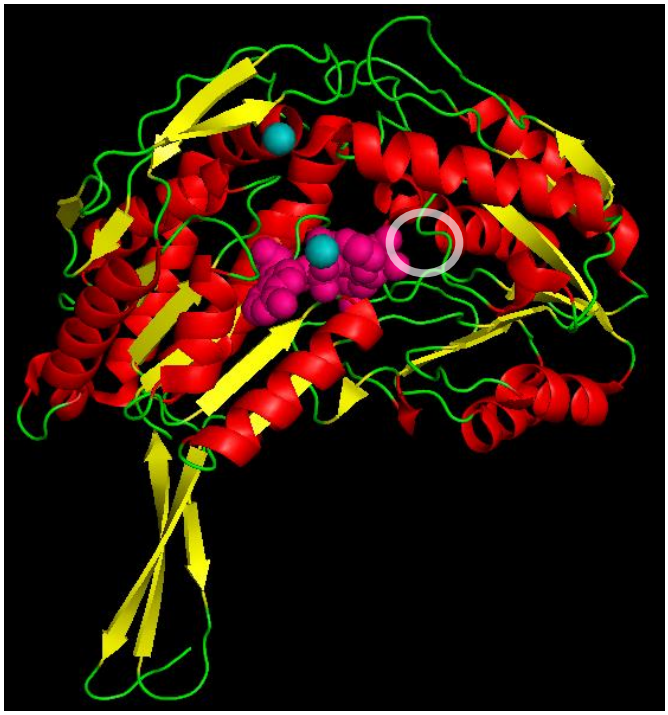


Figure 4-8. Structure of one subunit of the homo-tetramer human mitochondrial ALDH2 (Larson et al., 2007); Protein Database ID 2ONP). The catalytic tunnel where aldehydes are oxidised is indicated by a translucent white ring. Secondary structures are indicated in red and yellow, a molecule of nicotinamide-adenine-dinucleotide is shown in pink and bound sodium ions are shown in blue.

4.3.4.2 Acetaldehyde is oxidised by cells grown on metaldehyde more quickly than acetate or metaldehyde

Oxygen assays were performed using *A. calcoaceticus* E1 cultured using metaldehyde as the carbon source. The washed cells were treated with metaldehyde, acetaldehyde or sodium acetate. The different treatments contained similar concentrations of carbon; the carbon concentration of the metaldehyde treatment used was 94.0% of the acetaldehyde and 92.5% of the acetate treatment. The results of the assays are shown in Figure 4-9.

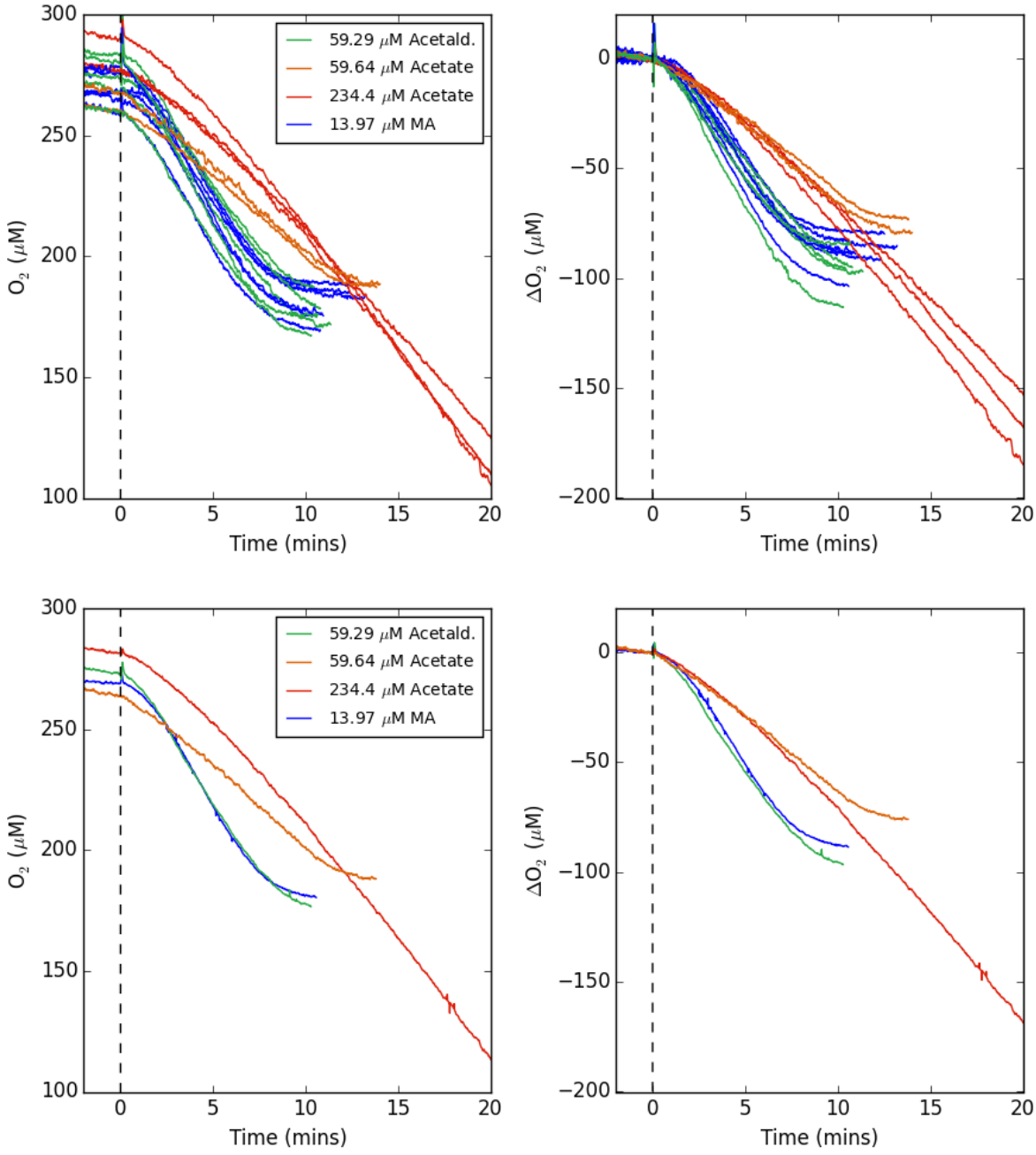


Figure 4-9. Oxygen utilisation of A. calcoaceticus *E1 cells grown with metaldehyde treated with the acetaldehyde, metaldehyde or acetate at 0 minutes. Upper charts show every assay; lower charts show mean readings. Left charts show absolute oxygen measurements, and right charts show oxygen deviation from the start of the treatment.*

Metaldehyde and acetaldehyde are oxidised at similar rates in *A. calcoaceticus* E1 grown with metaldehyde, supporting the notion that the two substrates share some of the same oxidisation pathway. The maximum rate of acetate oxidisation is less than that of acetaldehyde or metaldehyde, demonstrating that acetate cannot be the majority metabolite of metaldehyde, though acetate could be produced by, for example, the oxidation of one acetaldehyde monomer of the cyclic tetramer that is metaldehyde.

4.3.5 Kinetics of the metaldehyde degrading pathway in vivo

4.3.5.1 Oxygen utilisation of A. calcoaceticus *E1 treated with different substrates*

Oxygen assays of *A. calcoaceticus* E1 treated with different concentrations of metaldehyde, acetaldehyde or acetate were performed. The results are shown in the following pages.

Oxygen utilisation of A. calcoaceticus *E1 treated with 5 concentrations of metaldehyde*

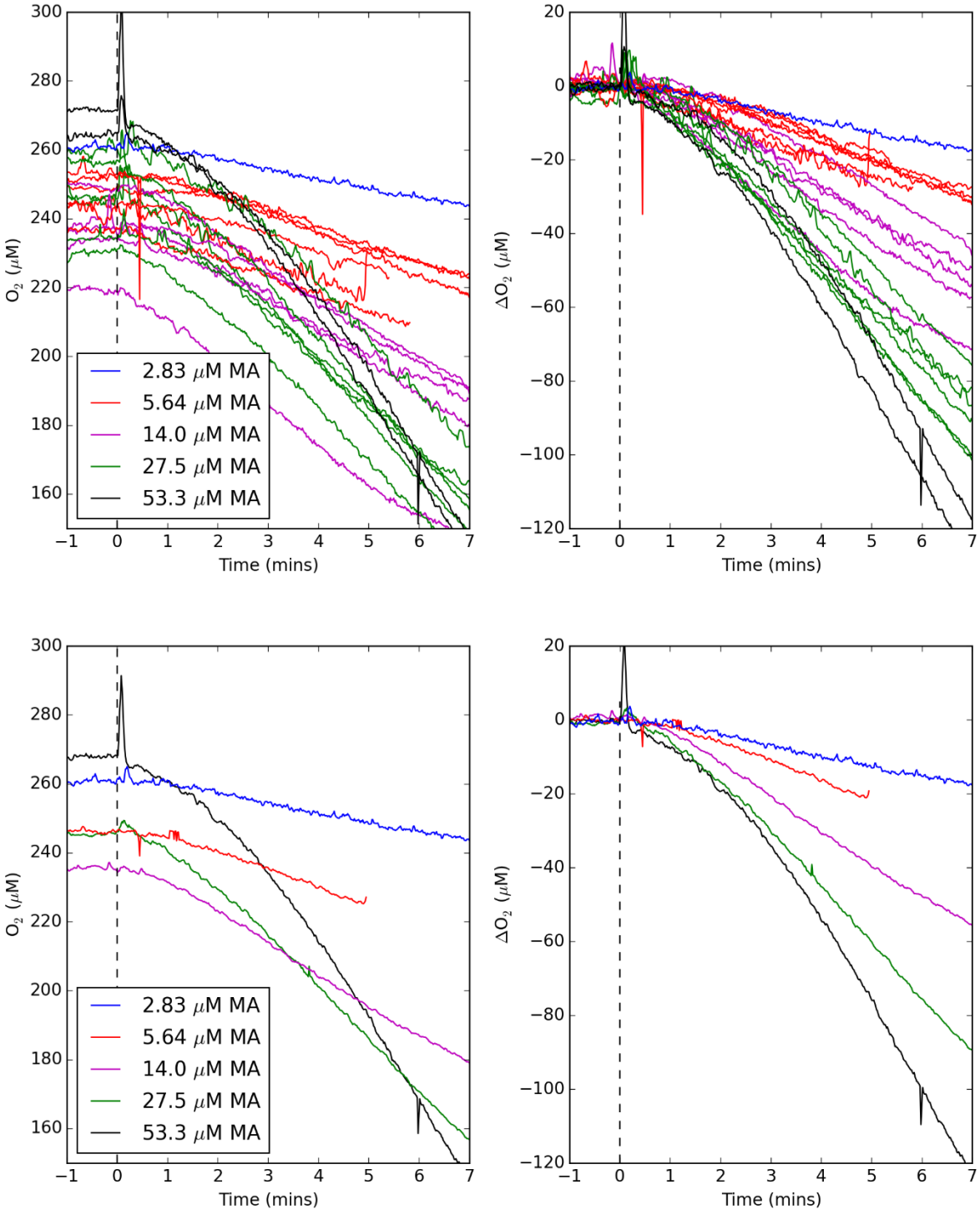


Figure 4-10. Measured oxygen levels of washed samples of A. calcoaceticus *E1 grown in minimal media using metaldehyde treated with different concentrations of metaldehyde. Upper charts show every assay; lower charts show mean readings. Left charts show absolute oxygen measurements, and right charts show oxygen change from the start of the treatment.*

Oxygen utilisation of A. calcoaceticus *E1 treated with 3 concentrations of metaldehyde*

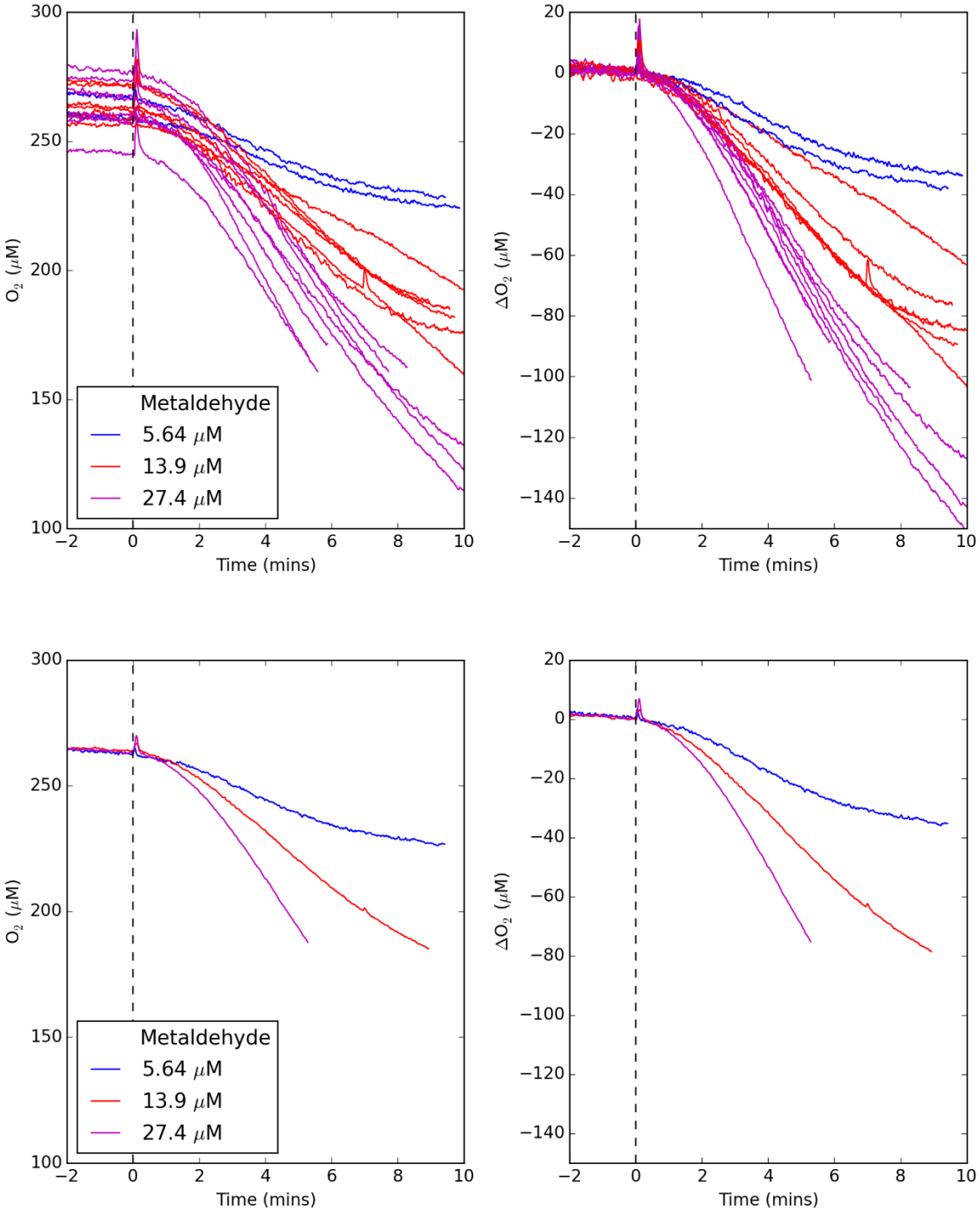


Figure 4-11. Measured oxygen levels of washed samples of A. calcoaceticus *E1 grown in minimal media using metaldehyde treated with different concentrations of metaldehyde. Upper charts show every assay; lower charts show mean readings. Left charts show absolute oxygen measurements, and right charts show oxygen change from the start of the treatment.*

Oxygen utilisation of A. calcoaceticus *E1 treated with 4 concentrations of metaldehyde*

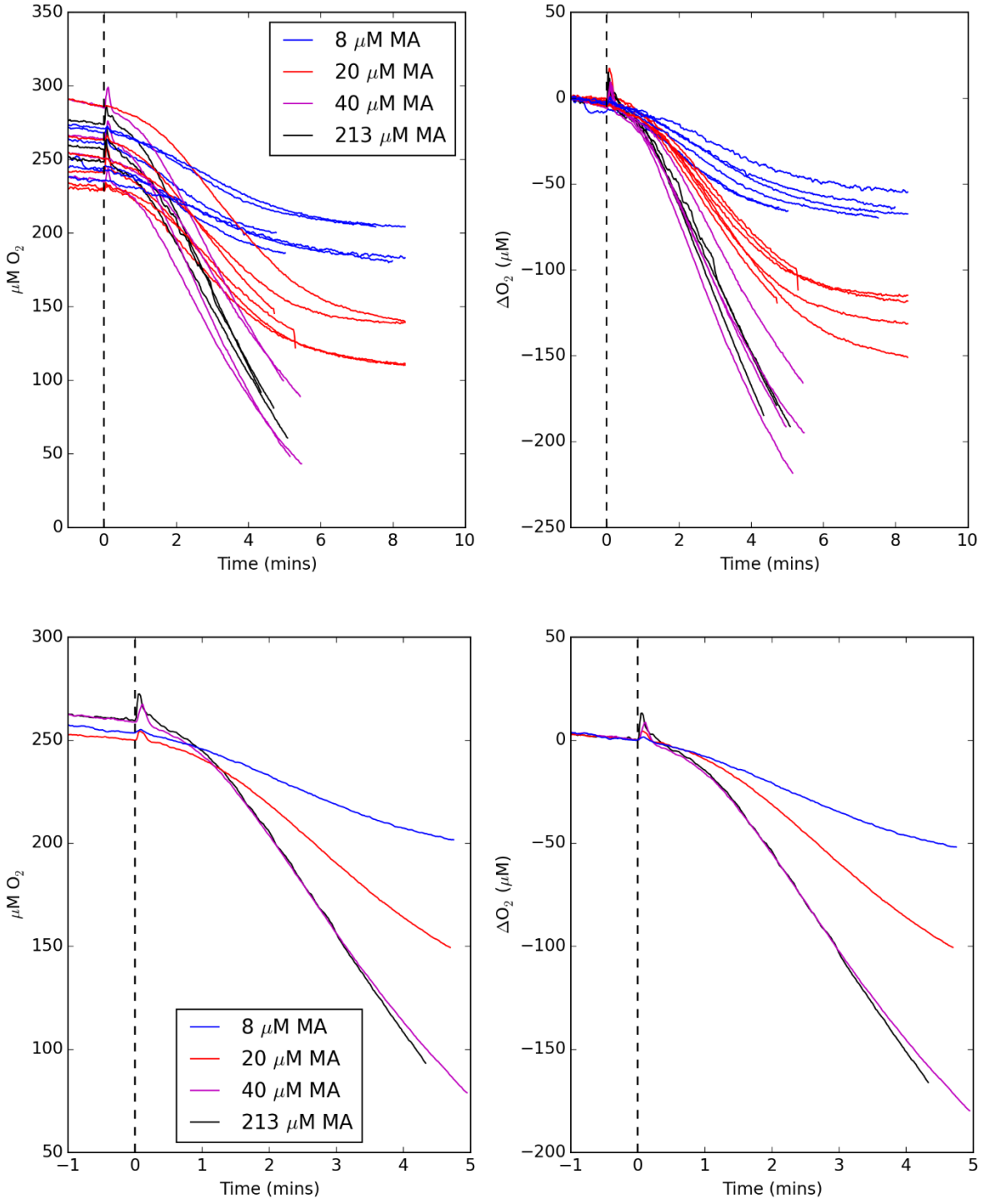


Figure 4-12. Measured oxygen levels of washed samples of A. calcoaceticus *E1 grown in minimal media using metaldehyde treated with different concentrations of metaldehyde. Upper charts show every assay; lower charts show mean readings. Left charts show absolute oxygen measurements, and right charts show oxygen change from the start of the treatment.*

Oxygen utilisation of* A. calcoaceticus *E1 treated with 4 concentrations of acetaldehyde or 27.5 µM metaldehyde

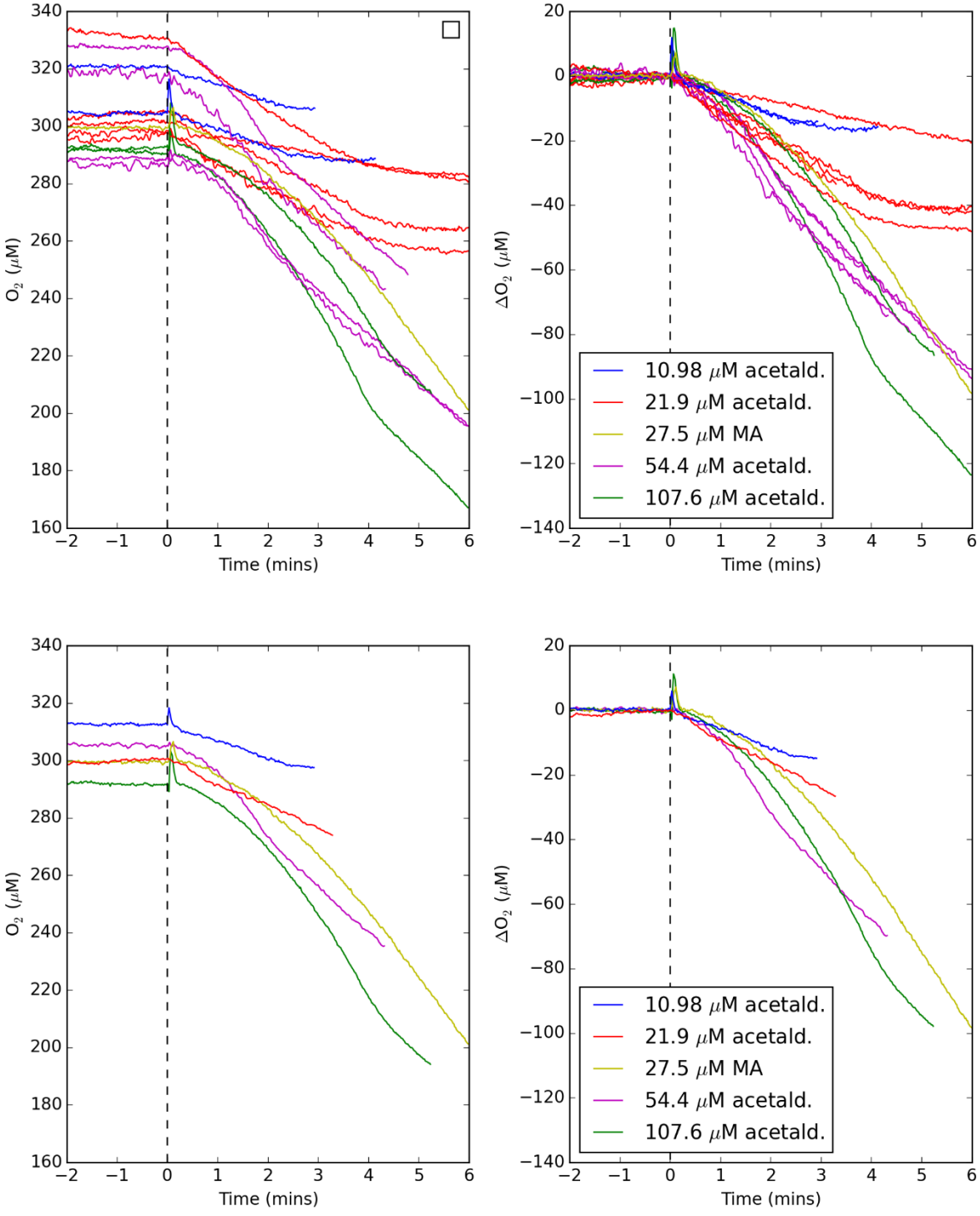


Figure 4-13. Measured oxygen levels of washed samples of A. calcoaceticus *E1 grown in minimal media using metaldehyde treated with different concentrations of acetaldehyde, and one concentration of metaldehyde. Upper charts show every assay; lower charts show mean readings. Left charts show absolute oxygen measurements, and right charts show oxygen change from the start of the treatment.*

Oxygen utilisation of* A. calcoaceticus *E1 treated with 5 concentrations of acetate or 27.4 µM metaldehyde

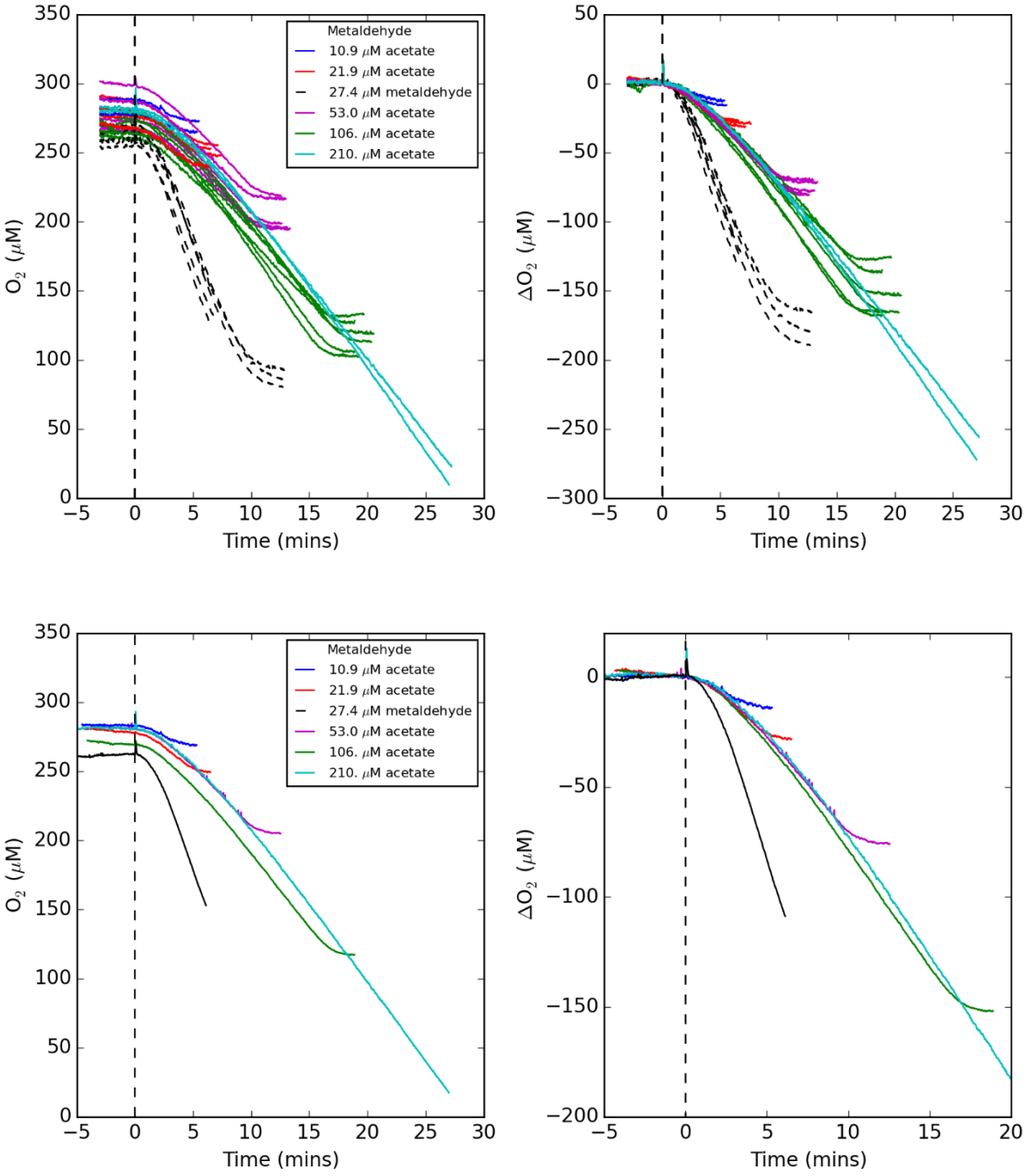


Figure 4-14. Measured oxygen levels of washed samples of A. calcoaceticus *E1 grown in minimal media using metaldehyde treated with different concentrations of acetate. Upper charts show every assay; lower charts show mean readings. Left charts show absolute oxygen measurements, and right charts show oxygen change from the start of the treatment.*

In all the assays presented above, a general pattern of oxygen utilisation can be seen where the rate of oxygen consumption increases with time, sometimes reaching a steady state, and then can be seen to decrease where measurements were continued for sufficient time. This can be explained as resulting from the fact that oxygen reduction is preceded by several metabolic processes, its rate is dependent on the concentration of its substrates, which each take time to accumulate.

The acetaldehyde data (Figure 4-13) appears to show some substrate inhibition. In the region of 1.5 minutes the rate of oxygen utilisation is at its peak in the cells treated with 54.4 µM acetaldehyde and exceeds the rate seen in cells treated with 107.6 µM acetaldehyde. The higher acetaldehyde treatment does not reach peak oxygen utilisation until after 3 minutes. This could be explained as the result of inhibition caused by acetaldehyde that diminishes as the acetaldehyde is removed from the reaction. Acetaldehyde at 107 µM was oxidised faster than the equivalent metaldehyde treatment of 27.5 µM despite the apparent inhibition. This is a different result than was seen in Figure 4-9 where 14 µM metaldehyde and 60 µM acetaldehyde produced similar rates of oxygen utilisation in the cells.

4.3.5.2 Apparent K_M of oxygen utilisation

The peak oxygen utilisation values over 30 second periods were determined for the assays shown in Section 4.3.5.1. The treatment concentration and rates of oxygen utilisation were plotted against each other for each experiment and Michaelis–Menten kinetics were determined by fitting the equation to the oxygen utilisation data (Figure 4-15). The velocities were corrected to take into account the diluting effect of the substrate treatment. This was particularly important for the higher concentrations of metaldehyde; to achieve 213 µM, 1 ml of 150 mg/L metaldehyde was added to a 3 ml cell sample. The derived apparent K_M and V_{max} values are given in Table 4-5.

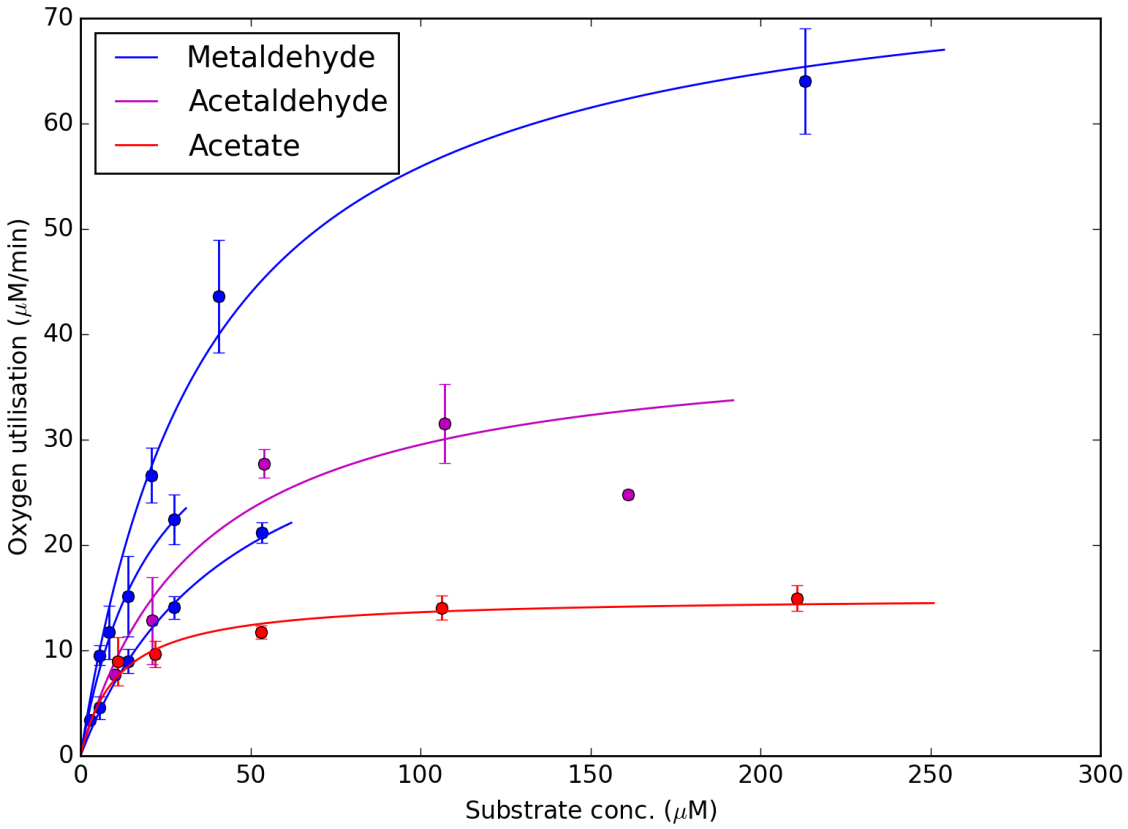


Figure 4-15. Mean maximum oxygen utilisation vs initial substrate concentration (given on the x-axis) of cell samples treated with different concentrations of the substrates metaldehyde, acetaldehyde or sodium acetate. Error bars give one standard deviation from the mean. Solid curves are plotted using the fitted parameters of the Michaelis-Menten equation.

Substrate	V_{max} (μM/min)	Apparent K_M (μM)
Metaldehyde 1 (Figure 4-10)	38.3 ± 4.52	45.4 ± 8.69
Metaldehyde 2 (Figure 4-11)	40.0 ± 9.69	21.8 ± 10.5
Metaldehyde 3 (Figure 4-12)	76.9 ± 4.52	45.4± 8.69
Acetaldehyde (Figure 4-13)	39.9 ± 5.82	35.3 ± 12.8
Acetate (Figure 4-14)	15.1 ± 0.746	11.0 ± 2.78

Table 4-5. Derived V_{max} and apparent-K_M of the reduction of oxygen utilisation resulting from the oxidation of different substrates by A. calcoaceticus *E1.*

The response of the cells to metaldehyde treatment is not very consistent. Care was taken to be consistent in the treatment of the cells but minor variations in the culture conditions and in the washing procedure could result in changes to the activity and relative quantity of the different enzymes and cofactors involved in the oxidation of metaldehyde.

There are several caveats to the interpretation of the results presented above. Scanning a noisy line for the steepest slope using many regressions will most likely result in the selection of a region that has a higher than average steepness by chance. In the other direction, taking rates from a dynamic process will underestimate the fastest rate as some sub-fastest velocity will be included in the regression. The most significant caveat is that Michaelis-Menten kinetics describe a very simple case of enzyme catalysed reactions, while what is being measured here is the end result of a large pathway.

4.3.5.3 The apparent K_M of a simulated pathway made of multiple reactions

To investigate how measuring the end product of multiple reactions can affect apparent K_M and V_{max}, a simulated pathway was constructed using versions of the general step-wise model described in the methods. The fastest rate of product formation was calculated across a 100 time-unit region using linear regression for simulated pathways containing up to 5 reactions. All the simulated reactions had the same K_M and V_{max} values of 20 and 0.1 and the apparent K_M and V_{max} values of the pathways was calculated in the same way as was done using the data obtained *in vivo* (Figure 4-15). The number of reactions in each simulated pathway was varied. The maximum velocity of final product formation was determined for initial substrate values of 500, 100, 50, 20 and 10. The results are shown in Figure 4-16 and Table 4-2.

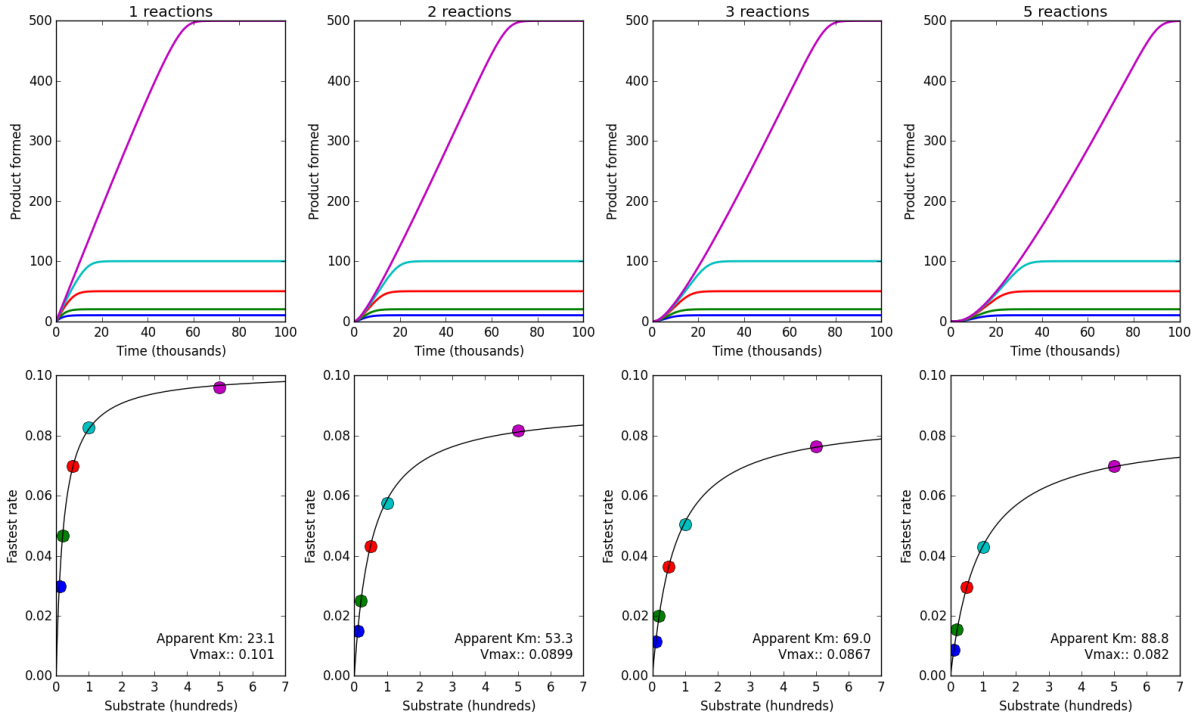


Figure 4-16. The results of simulated pathways containing the number of reactions given at the top of each vertical pair of charts. The quantity of final product formed over time for each starting concentration of substrate is shown in the top row of charts, and starting substrate versus fastest rate of final product formation is shown in the bottom row. A line fitted to the data using the Michaelis-Menten equation is shown on each lower graph along with the derived K_M and V_{max} parameters. Colours of lines and circles match for identical starting substrate values of 500 (purple), 100 (teal), 50 (red), 20 (green) and 10 (blue).

No. of reactions	1	2	3	5
Derived K_M	23.1	53.3	69.0	88.8
Derived V_{max}	0.101	0.0899	0.0867	0.0820

Table 4-6. K_M and V_{max} derived by fitting the fastest rate of final product formation from a series of modelled reactions, each with a K_M and V_{max} of 20 and 0.1, to the Michaelis-Menten equation.

The calculated K_M increases and V_{max} decreases with each additional reaction. These data demonstrate that analysis of the rate of final product formation cannot be interpreted to give details of individual components of a pathway without understanding the other components in that pathway.

4.3.5.4 Step-wise metabolic models fit to oxygen utilisation data

It was hoped that more information about the individual components involved in the oxidation of metaldehyde could be determined by fitting the observed oxygen utilisation curves of *A. calcoaceticus* E1 treated with metaldehyde to a model that simulates multiple reactions. To explore the idea, models with two reactions were fitted to some oxygen utilisation data. The two reactions had a stoichiometry of 1:4 to represent the hypothesised depolymerisation of metaldehyde to acetaldehyde and values for the parameters K_M and V_{max} were derived for each reaction. Preliminary experiments demonstrated that the ratio of initial substrate concentration to oxygen reduced should be included as a parameter in the fit; this allows for the fact that metaldehyde was not completely oxidised and also allows the model to correct for errors in the treatment concentration. This is referred to as the oxidation factor or "OxyF". The models discussed in this section are described in Table 4-7.

Model	Reaction A	Reaction B	Reaction C
A	MM	MM	OxyF
B	MM	Linear	OxyF
C	Linear	MM	OxyF

Table 4-7. Structure of models fit to oxygen utilisation data. Reactions are described in the methods.

Oxygen utilisation data of *A. calcoaceticus* E1 cells treated with metaldehyde, shown in Section 4.3.5.1 (page 121), was fitted using Model A (all fits are shown in Appendix 7.2). It was observed that most fits resulted in very high derived K_M and V_{max} values with high standard deviations for one of the reactions, usually reaction B (as can be seen in Figure 4-17D). As observed above, values of K_M far above the substrate concentration results in a linear relationship between substrate concentration and reaction velocity. For this reason, the data were fitted to Models B & C that use linear reaction kinetics (Figure 4-17)

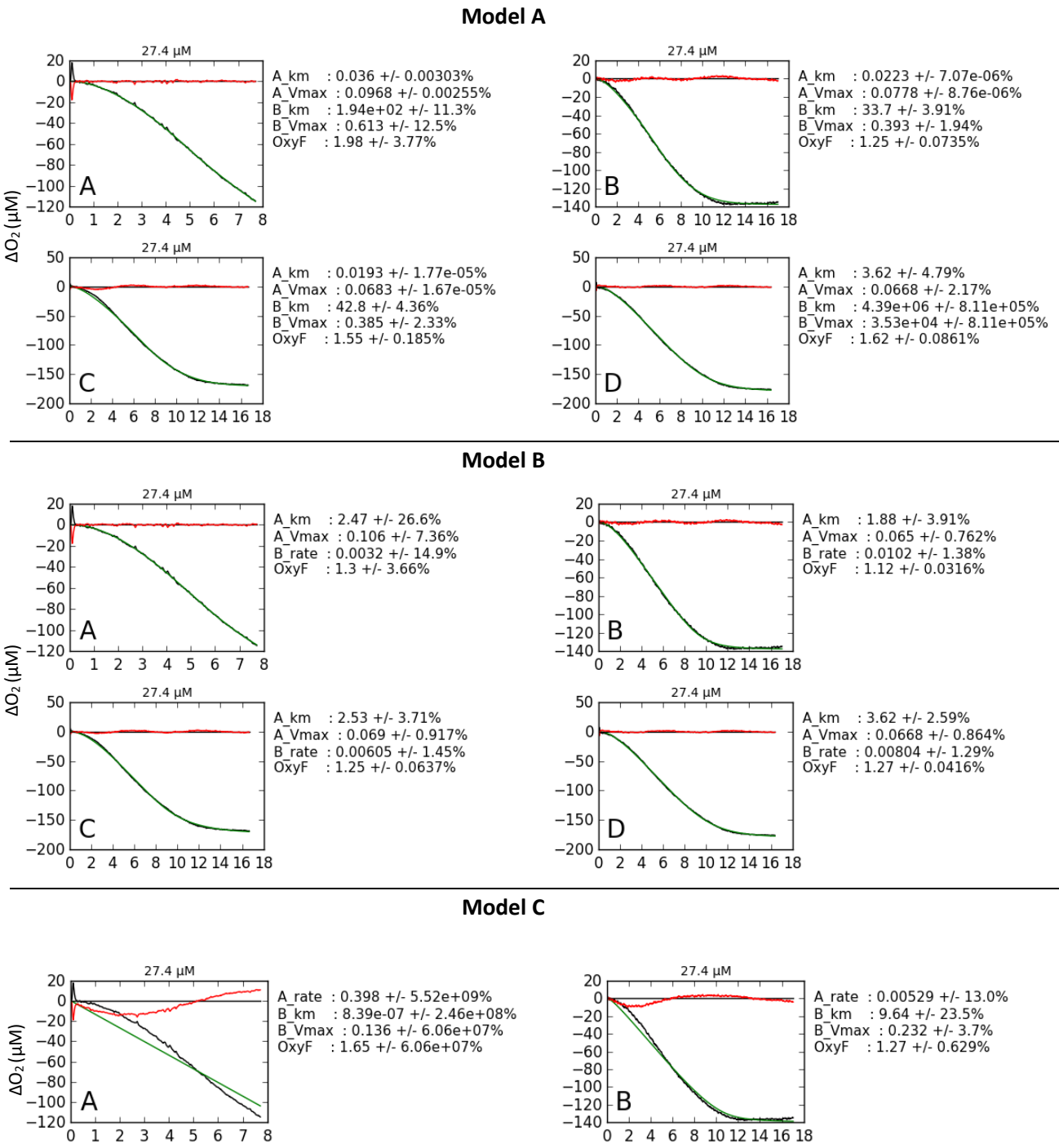


Figure 4-17. Oxygen utilisation data (black curves), curves fitted to the oxygen data (green) and residuals of the fit (red) are shown for selected oxygen utilisation curves first shown in Figure 4-11.

The exemplar fits shown in Figure 4-17 demonstrate that Models A & B are able to produce a curve that fit the data. Model C performed much worse, curves could not be constructed for most data. This supports the notion that the oxygen utilisation curves seen result from the accumulation of intermediates in the metabolic pathway. With a linear first reaction, the flux through the first

reaction drops more quickly at the beginning of the process, meaning that less of the substrate for the second reaction can accumulate.

To investigate to what degree, if at all, the parameters fitted to the data directly reflect the properties of particular reactions in the real pathway. To investigate this question, models with between 2-4 reactions that follow Michaelis-Menten kinetics (KM = 10, V_{max} = 0.2) were used to generate data. These data were fitted using Model B (Figure 4-18). As the number of simulated reactions increases, the poorer the fit becomes, showing that the reduction of oxygen in response to metaldehyde treatment is not the result of a process similar to a series of Michaelis-Menten reactions.

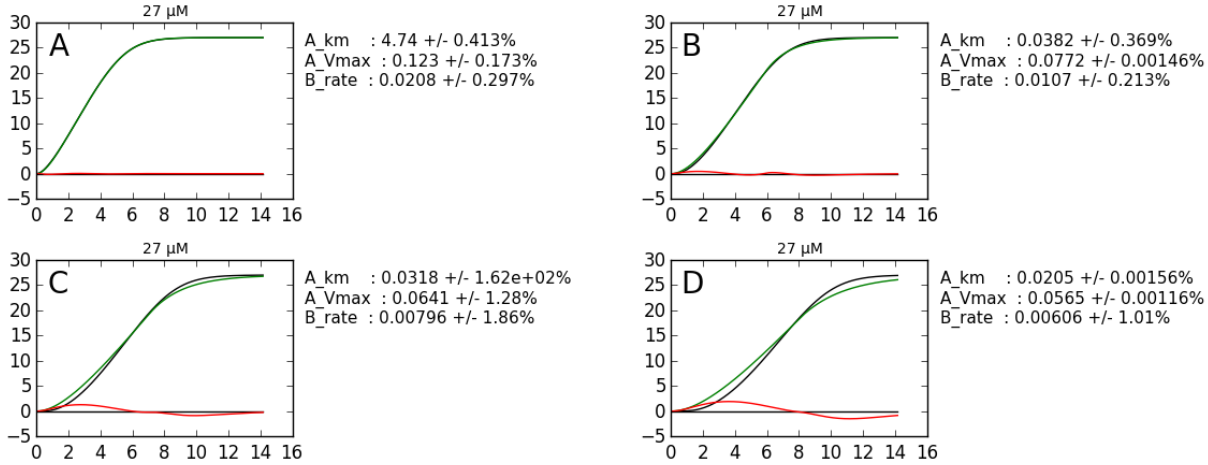


*Figure 4-18. Simulated final substrate accumulation produced using series of Michaelis-Menten reactions with K_M of 10 and V_{max} of 0.2 (black curves), fitted curves using Model B (green), residuals of the fit (red). The number of simulated reactions for each panel: **A** = 2, **B** = 3, **C** = 4 and **D** = 5.*

4.3.6 Metaldehyde degradation could not be reconstructed *in vitro*

The reconstitution of metaldehyde degradation in lysate from *A. calcoaceticus* E1 grown with metaldehyde was attempted using various buffers and cofactors.

Spectrophotometric assays of the reduction of β-NAD^+ by *S. cerevisiae* ALDH in 100 mM Tris-HCl buffer (pH 8.0), are shown in Figure 4-19. Reaction mixtures were treated with acetaldehyde or metaldehyde, with or without the addition of lysate from *A. calcoaceticus* E1 grown with metaldehyde and concentrated to OD_{600} 1.02 in 100 mM PB. Oxygen assays showed that the live cells were capable of metaldehyde dependent oxygen utilisation before being lysed (data not shown).

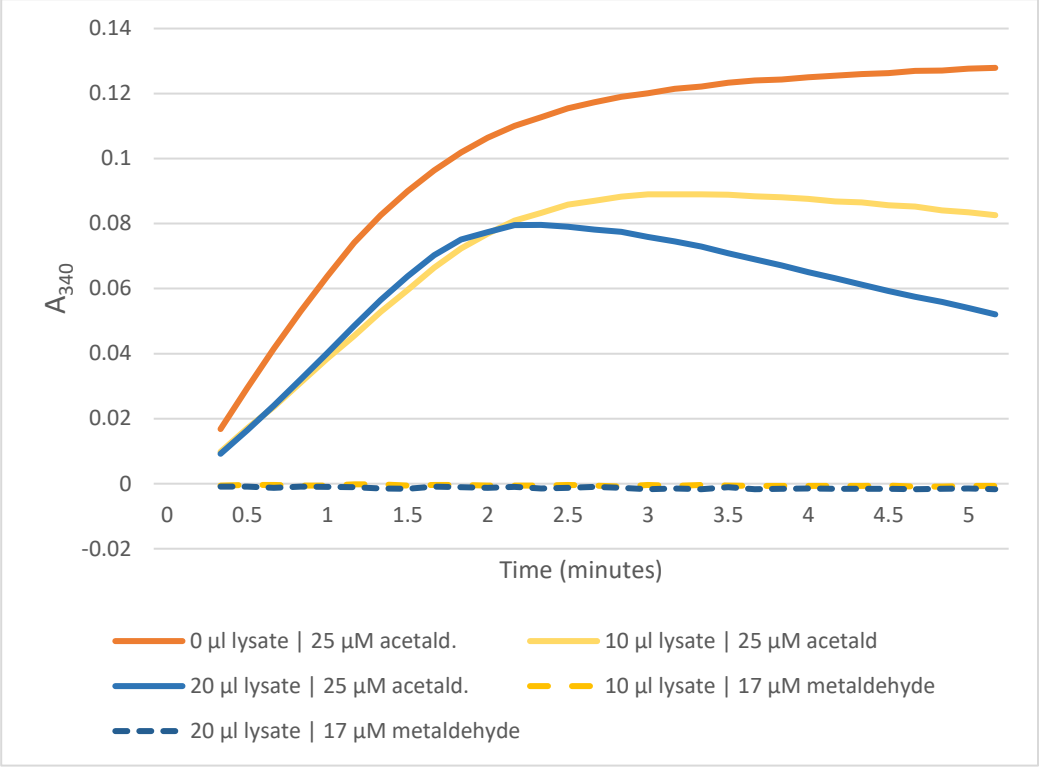


Figure 4-19. Absorbance at 340 nm measured over time in reactions mixtures containing S. cerevisiae *ALDH and treated with acetaldehyde or metaldehyde, with or without the addition metaldehyde grown* A. calcoaceticus *E1 cell lysate.*

No quantifiable yield of β-NADH was produced when metaldehyde was added to the reaction mixture. β-NAD^+ was reduced upon the addition of acetaldehyde to reaction mixtures. The addition of lysate reversed the reduction of β-NAD^+. It was possible that metaldehyde was being depolymerised and acetaldehyde being oxidised at a rate insufficient to produce a measurable signal as the β-NADH was being reoxidised or destroyed by the lysate, and so the direct quantification of metaldehyde degradation by lysate was undertaken.

Triplicate cultures of *A. calcoaceticus* E1 were grown using MSM with 0.85 mM metaldehyde as the carbon source. The cells were washed and concentrated to OD_{600} 1.99 in PB. An aliquot from each was tested for metaldehyde dependent oxygen utilisation, each was found to be positive (data not shown). The cells were lysed and split into two aliquots. Half were treated with 56.7 µM metaldehyde, 200 µM each β-NADH and β-NAD^+ only. Half were additionally treated with of 200 µM coenzyme A. Samples were taken at 2, 7, 15, 40, 120 and 260 minutes after metaldehyde addition and diluted 10-fold with 80% (v/v) chilled ethanol, to inhibit further enzymatic reactions, and stored at -20°C. Untreated samples of the lysates were also taken and stored at -20°C. The relative metaldehyde concentration of all samples were determined by LCMS (Figure 4-20). No degradation of metaldehyde was observed in any of the samples.

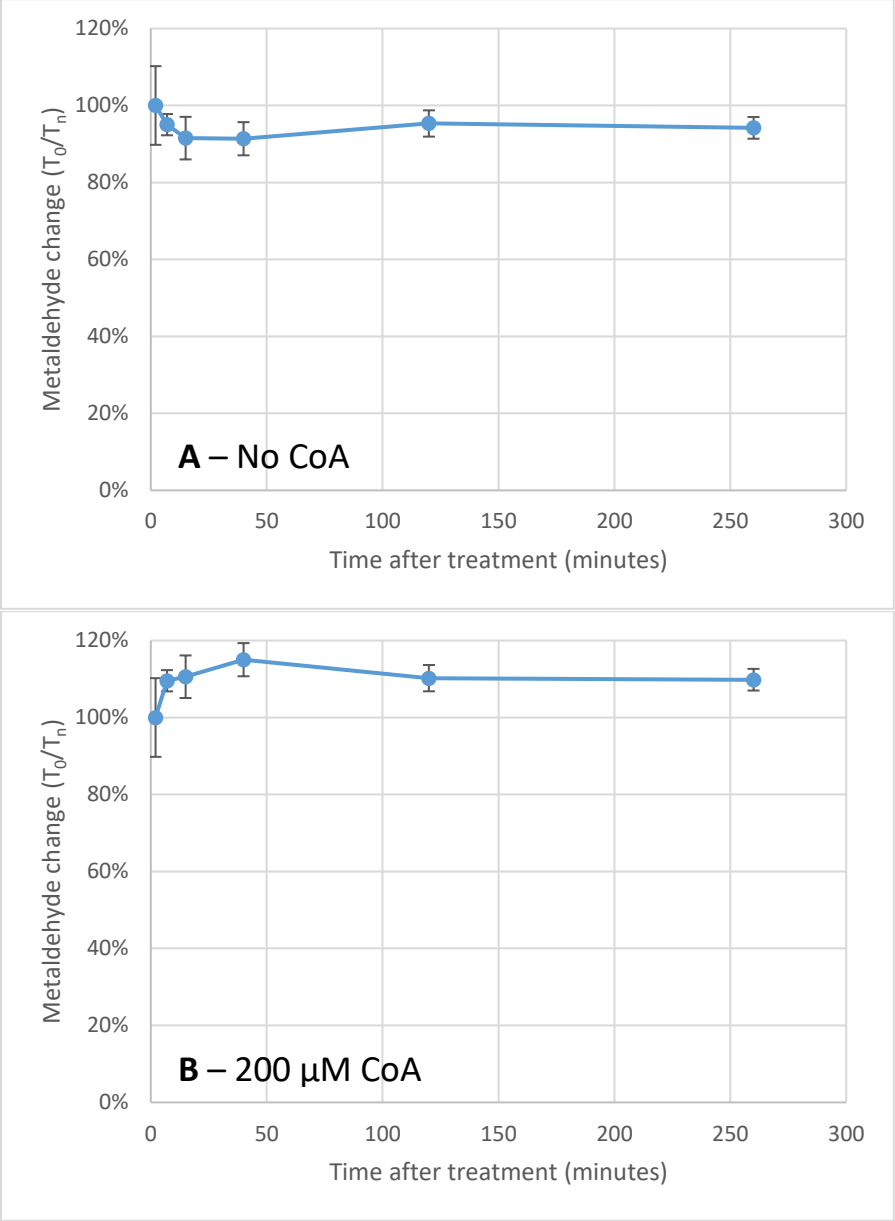


Figure 4-20. Change in measured metaldehyde concentration over time in cell lysate of A. calcoaceticus *E1 in 100 mM PB pH 7.0, treated with 85 µM metaldehyde. Chart **A** shows the addition of 200 µM each β-NAD$^+$ and β-NADH, chart **B** shows 200 µM CoA along with the 200 µM β-NAD(P)$^+$. Error bars show standard deviation from the mean of biological replicates.*

It was hypothesised that acetaldehyde was being produced by the depolymerisation of metaldehyde and inhibiting the progress of the reaction. To test this *Saccharomyces cerevisiae* aldehyde dehydrogenase and cofactors required for the oxidation of acetaldehyde were included in some experiments. The effect of the composition of buffers on the *in vitro* degradation of metaldehyde was also tested. The buffers used were 100 mM phosphate (Na_2HPO_4-NaH_2PO_4) at pH 6, 7 and 8; 100 mM citric acid-trisodium citrate at pH 5 and 6; and 100 Tris-HCl at pH 8.

Buffered lysates were prepared from a culture of *A. calcoaceticus* E1 grown with metaldehyde concentrated to OD_{600} 0.82. Duplicate samples for each condition were treated with 85 µM metaldehyde or 400 µM β-NAD^+ only or *S. cerevisiae* ALDH, with required cofactors given in the methods, as well as 85 µM metaldehyde. The treated buffered lysates were incubated at 30°C. Samples were taken from the treated lysate immediately after metaldehyde treatment, 30 minutes later and the following day. Samples were immediately diluted 10-fold with 80% ethanol. Metaldehyde in each sample was quantified using LCMS and the relative metaldehyde concentration of each sample is shown in Figure 4-21. Metaldehyde was not found to be degraded by the lysate.

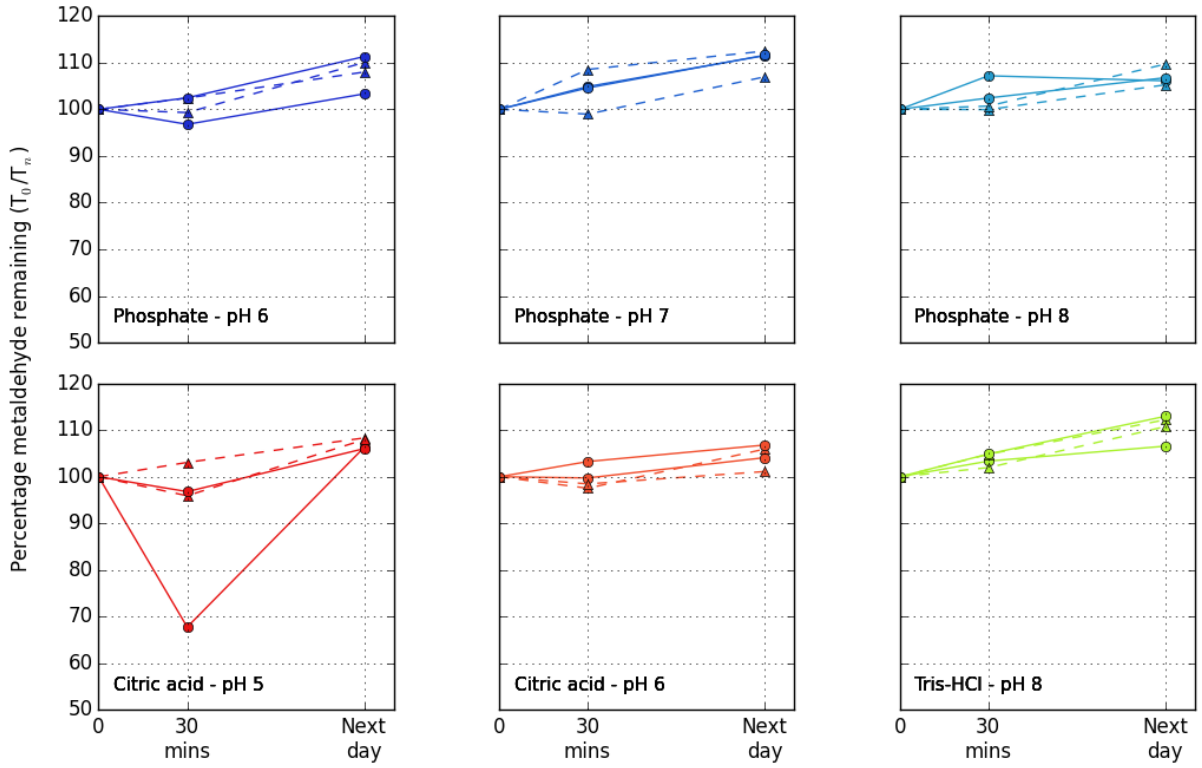


Figure 4-21. Change in metaldehyde concentration in A. calcoaceticus *E1 lysate samples in the given buffers, treated with 85 µM metaldehyde. Solid lines and circular markers show experiments using metaldehyde and 400 µM β-NAD^+ alone, dashed lines and triangular markers show experiments using 85 µM metaldehyde and* S. cerevisiae *ALDH with required cofactors.*

4.4 Discussion

4.4.1 Genetic screens

The ability of *A. calcoaceticus* E1 to maintain suicide plasmids meant that a practical method of producing random genetic lesions was not available for this strain. The production of a genetic library using genomic material from the isolates expressed in *E. coli* was begun, but concern regarding the likelihood of success using this method meant that it was not completed to a sufficient degree to present here. Such a method may have been successful if attempted.

4.4.2 Acetaldehyde is a product of metaldehyde degradation

Evidence that acetaldehyde is an intermediate in the metabolism of metaldehyde was given in the work presented here. Acetaldehyde is oxidised more rapidly than metaldehyde in cells grown using metaldehyde. Significant enrichment of an aldehyde dehydrogenase was observed. Together these facts show that the capacity to quickly oxidise aldehydes is specifically induced by metaldehyde. The most labile type of bond in the metaldehyde molecule would be expected to be the carbon-oxygen bonds. Breaking any of these bonds would result in unstable linear polymers of acetaldehyde. Acetaldehyde is generally toxic, forming adducts on proteins. The elevated production of ALDH to minimise the concentration of acetaldehyde present in the cells would explain why there is a larger capacity for acetaldehyde flux than could be produced by the depolymerisation of metaldehyde.

The first step may still involve the oxidation of an acetaldehyde monomer, producing an acetate, or some other reaction. The predicted function of the ALDH is the production of acetate, rather than the acetylation of coenzyme-A. However, if the relatively slow oxidation of acetate by *A. calcoaceticus* E1 grown on metaldehyde results from a bottleneck in the maximum flux of acetate that can be oxidised then this demonstrates that acetate cannot be a major intermediate in the oxidation of metaldehyde. The hypothesised intermediates, and their relationships are shown in Figure 4-22.

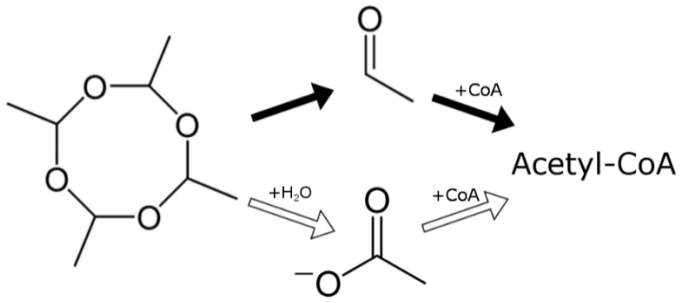


Figure 4-22. Primary intermediates in the proposed pathway for the metabolism of metaldehyde.

4.4.3 Regulation of primary metaldehyde degrading enzyme(s) compared to the rest of the metaldehyde oxidising pathway

Cell suspensions of *A. calcoaceticus* E1 grown with metaldehyde were found to utilise oxygen at a 6.6-fold increased rate and remove metaldehyde at 2.3-fold increased rate compared to cells grown using acetate and treated with metaldehyde. The cells for these assays were grown on different substrates, using metaldehyde or acetate as the carbon source. The cells grow more quickly using acetate and so higher concentrations of it were used to ensure the cells did not reach the stationary phase when grown over night. However, it may be the case that the acetate grown cell suspensions contained a greater proportion of cells that were stationary and therefore metabolically inactive, partially explaining the reduced rates of oxygen and metaldehyde utilisation. Proportionally, the rate of metaldehyde degradation was less affected by growth substrate than oxygen utilisation, and there was a 34-fold increase in the quantity of ALDH present in cells grown on metaldehyde. These facts suggest that the metaldehyde specific part of the pathway is either constitutive or only partly down-regulated in the absence of metaldehyde, while the enzymes that oxidise acetaldehyde are more tightly regulated. These scenarios are consistent with the evolution of regulatory control of a recently evolved metabolic activity.

4.4.4 Reconstitution of metaldehyde degradation *in vitro*

Attempts to reconstitute metaldehyde degradation *in vitro* using cell lysate failed. Buffers with different pH and constituents were tested. The cofactors β-NAD^+ and CoA may be required for catabolism of released acetaldehyde, depending on the nature of the ALDH produced by the organism, and so these cofactors were included in some experiments. β-NADH was included in the reaction mixture in one experiment as it is a potential cofactor. The *S. cerevisiae* ALDH and required cofactors for the oxidation of acetaldehyde to acetate were also included in some experiments to avoid product inhibition. It may be that the primary enzyme in metaldehyde degradation requires some other cofactor, or that it does not survive the lysing process.

4.4.5 Models of oxygen utilisation

A general model, one that involves the accumulation of intermediate substrates in a pathway, for the utilisation of oxygen in response to metaldehyde treatment was suggested by the shape of the oxygen utilisation curve. A specific model with two reactions, the first using Michaelis-Menten kinetics and the second using linear kinetics, could be fit to the oxygen utilisation data. This model could not be fit to data from simulated pathways made up of a series of Michealis-Menten reactions. In the cells of *A. calcoaceticus* E1 treated with metaldehyde, the reduction of oxygen is the final step in a complex series of reactions involving the citric acid cycle and the electron transport chain. The final oxygen curve seen in the experimental data will be dependent on all the nature of all these reactions, and how they interact. The success of this model in fitting real data may yield some insight into the pathway – the cumulative behaviour of a large part of this pathway may be represented by first order linear kinetics. It is, however, hard to establish that the values of parameters fit to the data directly represent real processes. To find out what the parametrised values of the model actually represent would require a more detailed understanding of the events that occur between the addition of metaldehyde to the oxygen chamber and the eventual reduction of oxygen.

Chapter 5: Identification of enzymes for primary metaldehyde catabolism by comparative genomics

5.1 Introduction

As has previously been discussed (Chapter 3), two isolates that are capable of degrading metaldehyde have been isolated, both from the same soil sample: *Acinetobacter calcoaceticus* E1 and *Variovorax* E3. Their genomes were *de novo* sequenced. Species related to the isolates have been found to degrade a wide range of synthetic and naturally occurring compounds. While it is possible that both strains independently evolved the ability to catabolise metaldehyde it is more likely that they acquired the ability by horizontal gene transfer.

This chapter is concerned with the use of genomic sequence information to identify candidates for the primary metaldehyde degrading protein (MDP).

5.1.1 Horizontal gene transfer (HGT)

There are several mechanisms by which HGT can occur. Canonically these are plasmid transfer by conjugation, transduction of genomic material by bacterial viruses (phages) and transformation by DNA acquired from the environment.

Plasmids are autonomously replicating circular DNA molecules that carry genes promoting the conjugation of the host to other bacterial cells, allowing the transfer of the plasmid from cell to cell. Plasmids frequently carry genes that may be beneficial to the host in certain environments; genes such as those for antibiotic resistance, heavy metal resistance (Silver & Misra, 1988) and the metabolism of xenobiotics (Sayler et al., 1990; Dunon et al., 2013) . Bacteria related to the *A. calcoaceticus* E1 and *Variovorax* E3 have been identified that acquired metabolic abilities through plasmids. *Acinetobacter* strains have acquired oxygenases that allow the catabolism of diesel (Mengoni et al., 2007) and aniline (Fujii et al., 1997), and a *Variovorax paradoxus* strain acquired the ability to degrade the xenobiotic pesticide 2,4-dichlorophenoxyacetic acid through the acquisition of a plasmid (Vallaeys et al., 1998). Some plasmids contain toxin-antitoxin systems. These produce a toxin, and a more short lived antidote to that toxin, so that daughter cells that do not inherit the plasmid are inhibited by the longer lasting toxin (Melderen & Bast, 2009).

Transposons are mobile genetic elements that can be associated with plasmids. These can carry genetic information that is located between specific inverted repeats. The excision or duplication of the genetic material that lies between the inverted repeats is performed by transposases, the genes for which are typically within the transposon – though not always (De Palmenaer et al., 2004). This translocation of genetic information can result in the transfer of genes to or from a plasmid – some

transposases preferentially transfer DNA to plasmids (Finn et al., 2007) – ultimately enabling the transfer of genomic material from one bacterial cell to another.

The inactivated genomes of phages (prophages) can be found in the sequenced genomes of bacteria (Canchaya et al., 2003). The activity of phages can result in the transduction of genomic material when some becomes packaged in a viral particle instead of the phage genome. This can result in the transfer of metabolic activities, shown experimentally in the transduction of a lactose degrading phenotype between *E. coli* and *Shigella dysenteriae* strains (Luria et al., 1960) and inferred in the case of experimentally mated *A. baumannii*, where the donor strain lacked any plasmids or secretion system, but a transposon containing an antibiotic resistance gene was found to have been translocated into the recipient strain's genome (Krahn et al., 2016). In an exceptional case, it has been observed *ex vitro* that cyanobacterial photosynthesis genes have become part of the functional genome of a phage that infects them, resulting in the back and forth translocation of these genes (Lindell et al., 2004).

Some bacteria produce gene transfer agents: phage like particles that contain random genetic material from the host's genome. These are released upon lysis of the cell (though they do not induce the death of the cell) and are capable of transferring DNA to a wide range of recipient species (Lang et al., 2012).

Transformation occurs when bacteria take up free DNA from the environment through specialised transporters for this purpose, as a source of potentially valuable genetic material (Chen et al., 2005; Takeuchi et al., 2014). The genetic material can then be integrated into the genome by DNA repair mechanisms. *Acinetobacter calcoaceticus* has been identified as being a very naturally competent species (Juni & Janik, 1969) and natural transformation of the species has been demonstrated to occur in soil microcosms (Nemec et al., 2011).

5.1.2 Evolution of protein function

Enzymes function by direct interaction with compounds. The final shape of a protein and the relative location in space of the different chemical moieties that form it determines the mode of interaction that the enzyme has with other molecules it encounters. This structure is itself determined by the linear chain of amino acids that makes up the protein, encoded by an organism's genome. The genome of an organism is subject to many sources of mutation; for example point mutations caused by chemical modification of the bases (Stolarski et al., 1987) or ionising radiation; and inversions, deletions, duplications or translocations caused by the action of DNA replication machinery or mobile genetic elements (Darmon & Leach, 2014). These mutations have the potential to alter an enzyme's

function, or change the way in which expression of the enzyme is regulated by altering regulatory sequences.

New functions for an enzyme can occur as a result of mutations to DNA that alter its amino acid sequence and hence structure. Enzymes may be capable of catalysing reactions using substrates chemically similar to their primary substrate, with lesser efficiency. Duplication of the responsible gene can then allow a divergence in function of the duplicated genes; with one gene evolving greater efficiency and specificity for a different substrate (O'Brien & Herschlag, 1999).

5.1.3 The possible nature of enzymes ancestral to MDP

Acetaldehyde is the primary product of metaldehyde degradation by *A. calcoaceticus* E1 (Section 4.4.2). The production of acetaldehyde from metaldehyde is most likely to be the result of the molecule depolymerising following the disruption of an ether bond. Enzymes that have the ability to lyse ether bonds, either by hydrolysis or by other means may evolve the ability act in the degradation of metaldehyde. This depends entirely on the mechanism by which the enzyme functions, however. For example, based only on the reaction catalysed, one might expect an isochorismate-pyruvate lyase to be a good candidate for a proto-MDP. However, the initial step of the reaction catalysed by these enzymes involves a pericyclic transition state, where a hydrogen bound to C2 (Figure 5-1) is abstracted to form a methyl (C9), causing the formation of a conjugated pi system and scission of the ethyl bond. This mechanism would not be supported by the structure of metaldehyde.

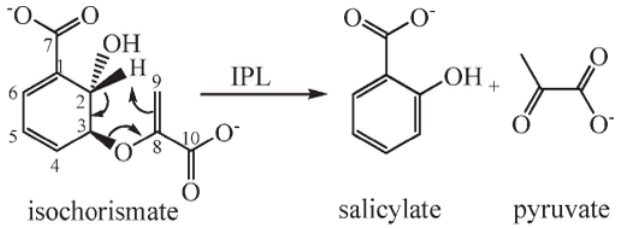


Figure 5-1. The proposed reaction mechanism of isochorismate-pyruvate lyase, catalysed by Pseudomonas aeruginosa *PchB. Taken from Luo et al. (2009).*

There are some chemical similarities between an ether and an amine that is made up of an NH moiety bound to 2 carbons (C-NH-C). Both are polar, having greater electronegativity at the O or N. The nitrogen in such an amine has a single lone pair of electrons that could be used in hydrogen bonding or other interactions with an enzyme, while the ethereal oxygen has two lone pairs. It is plausible that an enzyme could evolve from acting only on an amine bond, to acting on an ether bond.

5.1.4 Function prediction of uncharacterised proteins

In the work presented here, the nature of reactions predicted to be catalysed by proteins are evaluated to identify candidates for MDP. The function of uncharacterised proteins can be predicted based on the similarity of their amino acid sequence to proteins of known function. A tool used for this purpose in the work presented here is InterPro (Hunter et al., 2012). This service integrates multiple databases that look for signatures present in a given amino acid sequence that are indicative of enzyme function (such as the conservation of residues directly involved in a catalysis), regulatory domains, intermembrane domains, and others.

The structure of proteins is more conserved than the primary amino acid sequence. (Arnold et al., 2006). SWISS-MODEL is a tool for structure prediction that uses solved structures as templates. The amino acid sequence is aligned to sequences of proteins for which structural determinations have been performed. A model is computed using the template, attempting to align the predicted protein's backbone to the template where there is homology between the two. Low homology loops are predicted *de novo* or, failing this, the database is searched for homologous loops to use as a template. The probable conformation of side chains is calculated and the structure adjusted to minimise its energy (Schwede et al., 2003). The final model is evaluated (Benkert et al., 2009).

Genes that function in the same metabolic pathway are often collocated in the genome and may be co-transcribed together as an operon (Yan & Moult, 2006). The predicted function of genes close to those of interest may also be considered when evaluating the function of a gene.

5.1.5 The molecular basis for the biotic degradation of atrazine

The herbicide atrazine is an example of a xenobiotic compound that is degraded by a specifically evolved pathway, the molecular and genetic basis for which is well studied. It is presented here as a case study for how pathway can evolve and spread. Many of the concepts discussed above are demonstrated in this case study.

Field studies that took place between 1969 and 1985 examined by (Krutz et al., 2010) showed that the half-life of atrazine was between 28 and 178 days. Studies since 1993 show a half-life between 1 and 12 days, with the primary first step being the dechlorination of the compound. Pre-1993 studies showed that dealkylation of branches from the S-triazine were the main first step. This change in the fate of atrazine has occurred due to the evolution and spread of genes that catalyse the de-halogenation and de-alkylation of the compound.

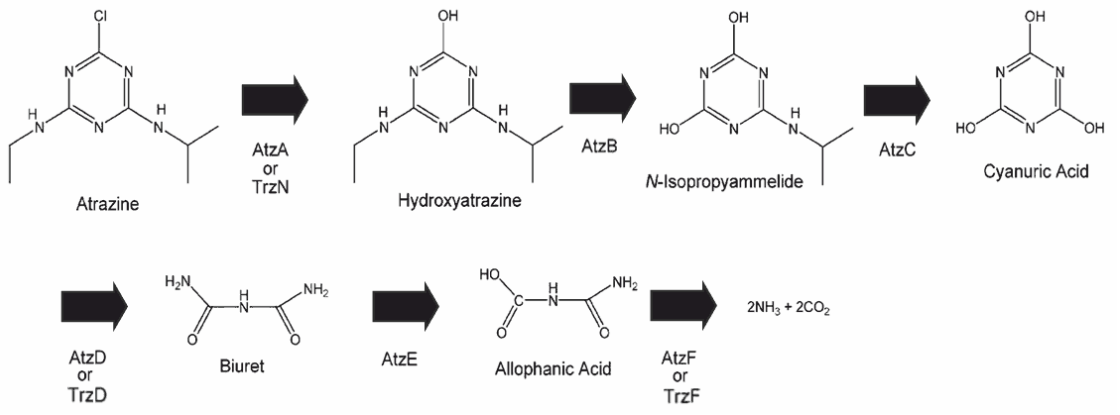


Figure 5-2. The biotic mineralisation pathway of atrazine and the proteins identified as catalysing each step. Adapted from (Shapir et al., 2007).

The genes involved in the mineralisation pathway for atrazine (Figure 5-2) are found spread among various organisms. It is common to see organisms that do not appear to have some of the genes required for complete mineralisation. Smith et al. (2005) enriched samples from soil that had been treated with atrazine for 15 years using a liquid media spiked with atrazine or one of the metabolic intermediates found in the pathway. They isolated 8 bacterial species and assayed them for known atrazine catabolising genes using PCR with primers for *atzA, -B, -C, -D, -E, -F*, *trzD* and *-N* the results of which are shown in Table 5-1. *AtzA, -E* or *-F* were not found in their study.

Identification	Genes[a]			
	trzN (100%)	*atzB* (100%)	*atzC* (100%)	*trzD* (100%)
Sphingomonas yanokoikuyae	–	–	99	100
Variovorax paradoxus	–	–	100	100
Caulobacter crescentus	–	–	99	–
Pseudomonas putida	–	–	99	100%
Nocardia sp.	100	–	100	–
Rhizobium sp.	–	100	100	100
Flavobacterium oryzihabitans	–	–	100	–
Agrobacterium tumefaciens	–	–	100	–

Table 5-1. Atrazine catabolising genes found in bacterial species from enrichment cultures by Smith et al. (2005). The numbers represent the sequence similarity of the amplified gene to reference genes.

Only *Nocardia* sp. was able to dechlorinate atrazine. This species was also able to remove the alkyl-amine ($(CH_3)_2CHNH_2$) from hydroxyatrazine using an unknown enzyme. An earlier enrichment culture study of a sample from the same soil found *Clavibacter michiganense* ATZ1 that contained *atzA* and *Pseudomonas* sp. CN1 that contained *atzB* and *atzC* (de Souza et al., 1998).

Pseudomonas sp. ADP was the first organism found capable of fully mineralising atrazine (Mandelbaum et al., 1995). It is able to do so due to the plasmid pADP-1 that contains all the *atz* genes (A-F). This plasmid was sequenced by Martinez et al. (2001). A map is shown in Figure 5-3.

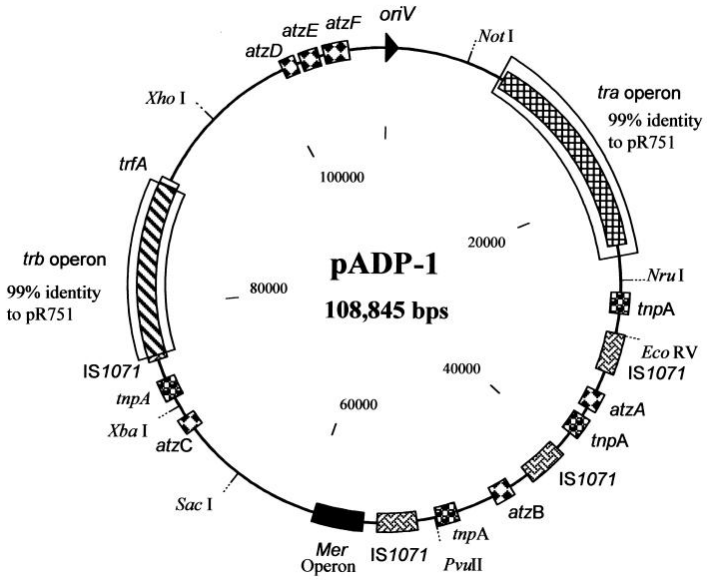


Figure 5-3. The plasmid pADP-1 found in Pseudomonas sp. ADP. Taken from Martinez et al. (2001).

AtzA, -B and *-C* are individually flanked by transposases suggesting they were recruited to the plasmid in separate events (Martinez et al., 2001). In a study of 17 atrazine degraders from 3 different sites, Devers et al. (2007) found that it was common for *atzA* to be on a separate plasmid from other genes in the pathway, and 3 of the 17 strains isolated had an additional copy of *atzC* on another plasmid.

Expression of *atzA, -B* and *C* is unregulated (Govantes et al., 2009).

Genes for the mineralisation of atrazine are not always found in the same organism, meaning that organism does not fully benefit from the catabolism of the compound. Genes in sequential steps of atrazine hydroxylation to cyanuric acid are found on separate plasmids and only on plasmids. Their expression is unregulated – unlike the expression of AtzD, -E and -F that are tightly regulated by the presence of cyanuric acid and the availability of other nitrogen sources (Govantes et al., 2009). All this is best explained as being due to atrazine specific genes only recently having evolved these activities, AtzA having evolved from a melamine deaminase (Seffernick et al., 2001), whereas cyanuric acid is a naturally occurring compound (Seffernick & Wackett, 2016). Consolidation, as has occurred in the case of pADP-1, and appropriate regulation can only be selected for after the individual genes are extant and functioning to the benefit of their host organism. It is worth noting

that the full diversity of atrazine catabolising genes is unlikely to have been described and most studies rely on testing for known genes without looking for new genes.

Practically, this means that when searching for genes that degrade a synthetic compound one cannot rely on complete operons being available for transformation as a functional unit into a host, complicating functional genomics approaches using gene libraries. The fact that the evolution of a new function for an enzyme precedes the evolution of its regulation is illustrated in the case of AtzA. Expression of MDP is also apparently unregulated, meaning that functional genomics approaches that rely on substrate induced expression (Uchiyama & Watanabe, 2008) would not succeed in identifying it. Another thing to take away from the biotic degradation of atrazine is that once a useful metabolic pathway exists, it will spread from species to species and place to place much more quickly than similar metabolic activities can evolve in those places – a point also made by Janssen et al. (2005) discussing dehalogenases.

5.1.6 Rationale

If we postulate that (i) MDP has evolved from an enzyme with a similar function and (ii) *A. calcoaceticus* E1 and *Variovorax* E3 acquired MDP by horizontal gene transfer, then the identification of horizontally transferred genes and the examination of their predicted function could lead to the identification of MDP.

Two HGT scenarios are possible: (i) The same enzyme is responsible metaldehyde degradation in both *A. calcoaceticus* E1 or *Variovorax* E3 (being acquired from the other isolate, or independently from another source), or (ii) different horizontally transferred sequences are responsible for the degradation. In the first case it can be hypothesised that sequences that are more similar in the degrading strains than in related strains known to lack metaldehyde catabolism are shared horizontally transferred sequences, and therefore candidates for MDP. In the second case it can be hypothesised that enzymes similar to those present in strains that lack metaldehyde catabolism can be ruled out as candidates for MDP, and so the comparison of a degrading strain's proteome with non-degrading strains' may reveal good candidates for MDP.

An argument for preferring the first scenario is that both strains were found to have the same capacity and exist in the same soil; as previously stated it is easier for a metabolic ability to be shared between bacteria than it is to evolve multiple times. An argument for preferring the second scenario, where metaldehyde degrading proteins evolved multiple times, is that fact that the two strains show different affinities for metaldehyde.

Comparative genomics methods were applied here, using amino acid similarity scores, that tested for both of the HGT scenarios described.

5.2 Methods

5.2.1 BLAST score ratio

This method is based on that described by Rasko et al. (2005).

BLAST score ratios (BSR) of amino acid sequences were obtained for a subject sequence using the following method. All BLASTp searches were performed with default settings.

A reference score for each subject sequence was obtained by performing a BLASTp search using the sequence as both the query and the subject. BLASTp searches were then performed using each sequence in a query strain's predicted proteome against the subject sequence. The value of the highest scoring alignment was divided by the reference score for the subject sequence yielding the BSR.

In the work described here, predicted amino acid sequences of *A. calcoaceticus* E1 were the subject sequences. BSR were obtained using various query proteomes.

5.2.2 InterPro

Amino acid sequences were submitted to the InterPro v5 service (Hunter et al., 2012). A submitted amino acid sequence is queried using multiple databases and algorithms. Signatures of the functional domains, protein family membership and other features identified by these other databases have been manually integrated and described by the developers of InterPro. An example of an integrated InterPro record is shown in Figure 5-4. The contributing signatures box on the right of Figure 5-4 indicates which external database signatures are associated with this record.

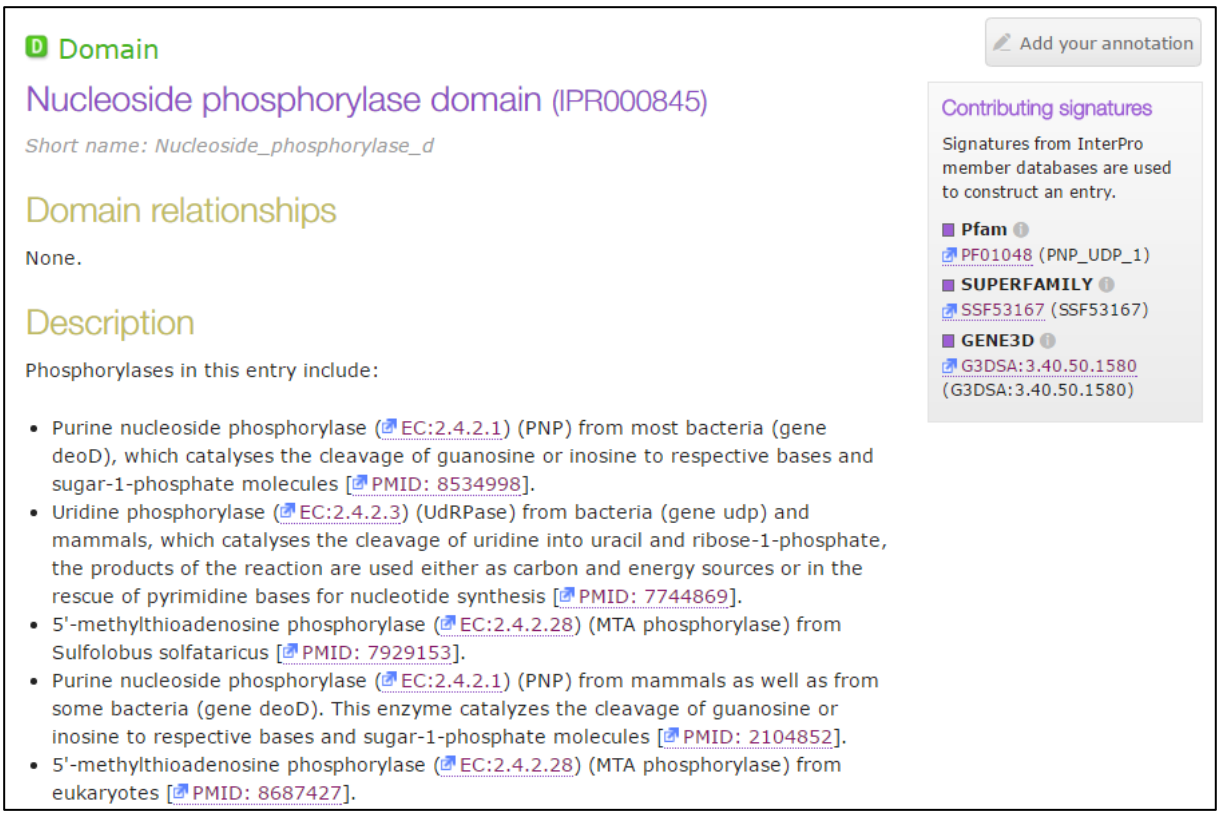


Figure 5-4. An example of an integrated InterPro record. This record can be accessed using the URL https://www.ebi.ac.uk/interpro/entry/IPR000845

The domain and protein family classifications were extracted from the XML files returned. InterPro accession numbers, in the format "IPR######", will be given for the classifications, these can be retrieved using the URL https://www.ebi.ac.uk/interpro/entry/IPR###### with the appropriate InterPro accession number as the last part of the URL.

5.2.3 SWISS-MODEL

Computationally predicted structures for amino acid sequences were produced using SWISS-MODEL workspace (Arnold et al., 2006) in automated mode. A QMEAN4 value is reported for models presented in this work. This score is based on 4 parameters: Distance dependent interaction potential of Cβ atoms and all atoms, desolvation potential and torsion over 3 residues. Scores are normalised by solved structures; the mean score for solved structures of a particular size is used as zero, and the QMEAN4 score for computed structures are relative to this.

Additional alignments were sometimes selected manually, but the models produced from these selections never had QMEAN4 scores higher than the automatically constructed models.

5.2.4 Operon detection

Genes that are encoded in the same direction on the genome and are 25 or fewer base pairs separated are believed to form an operon (Yan & Moult, 2006). This was evaluated manually using Artemis to visualise the location of the genes (Rutherford et al., 2000).

5.2.5 Identification and evaluation of MDP candidates

The above described methods are applied to identify and evaluate candidates for MDP, using the *de novo* sequenced genomes of *A. calcoaceticus* E1 and *Variovorax* E3 (Section 3.2.6) and other genomes acquired from the NCBI database.

To identify MDP candidates assuming that *MDP* is a recently acquired gene shared between *A. calcoaceticus* E1 and *Variovorax* E3 (the first scenario described in the Rationale for this chapter, Section 5.1.6), predicted proteins from the *A. calcoaceticus* E1 genome that were more homologous (had greater BSR) to predicted proteins in the *Variovorax* E3 genome than in the *A. calcoaceticus* RUH 2202 genome were identified. A minimum BSR of 0.5 to *Variovorax* E3 was used as a safe cut-off. Any recently horizontally transferred enzyme would be expected to have a BSR much higher than this. Deviations from the mean guanine-cytosine content (GC%) found in either organisms' genome in the coding sequence for homologous genes will be considered evidence for the recent horizontal transfer of the genes. The mean *A. calcoaceticus* E1 GC% is 38.8%, and *Variovorax* E3 GC% is 67.4%. Enzymes that have these properties were examined in more detail.

The identification of MDP candidates assuming separate origins for the capacity in *A. calcoaceticus* E1 and *Variovorax* E3 was performed for *A. calcoaceticus* E1 only. As this strain has higher affinity for metaldehyde, it is the more interesting hypothetical MDP of the two for future applications. Other *Acinetobacter* strains that lack metaldehyde catabolism were acquired. The predicted proteome of *A. calcoaceticus* E1 was compared to the proteomes of all the other *Acinetobacter* and any predicted protein in *A. calcoaceticus* E1 that did not have a BSR > 0.4 to any other strain was described as being unique to that strain and examined in more detail.

The starting point for examining a protein in more detail was the InterPro records of detected domains and family signatures (e.g. Figure 5-4). If the reactions associated with the signatures included the scission of carbon-oxygen or carbon-nitrogen bonds, then the plausibility of the protein as a candidate for MDP was evaluated. This evaluation may include structure prediction by SWISS-MODEL, examination of the genomic context of the gene, or a review of relevant literature.

5.2.6 Expression vectors for MDP candidate protein

Two plasmids for heterologous expression of a protein identified as being a good candidate for MDP were purchased from GenScript (Piscataway, USA). The gene was synthesised *de novo* and used to construct the plasmids described here. The first was intended for transformation into a strain of *Acinetobacter* and used the vector pMAL-c4X (Figure 5-5). This vector was chosen as the tac promoter is expected to be compatible with *Acinetobacter* strains. The native gene sequence for the candidate (i.e. found in *A. calcoaceticus* E1) was ligated between restriction sites NdeI and EcoRI, excising the maltose binding protein (MBP) which was not required in this experiment. The second plasmid was constructed from pET22b(+) (Figure 5-5) and contained a gene sequence for the candidate protein optimised for expression in *E. coli*, ligated into the HindIII restriction site. This plasmid utilises the T7 promoter which requires the T7 polymerase to be expressed by the host. Both constructs contain the lac operon for controlled expression of the gene using induction by isopropyl β-D-1-thiogalactopyranoside (IPTG) and an ampicillin resistance gene for selection of transformed clones.

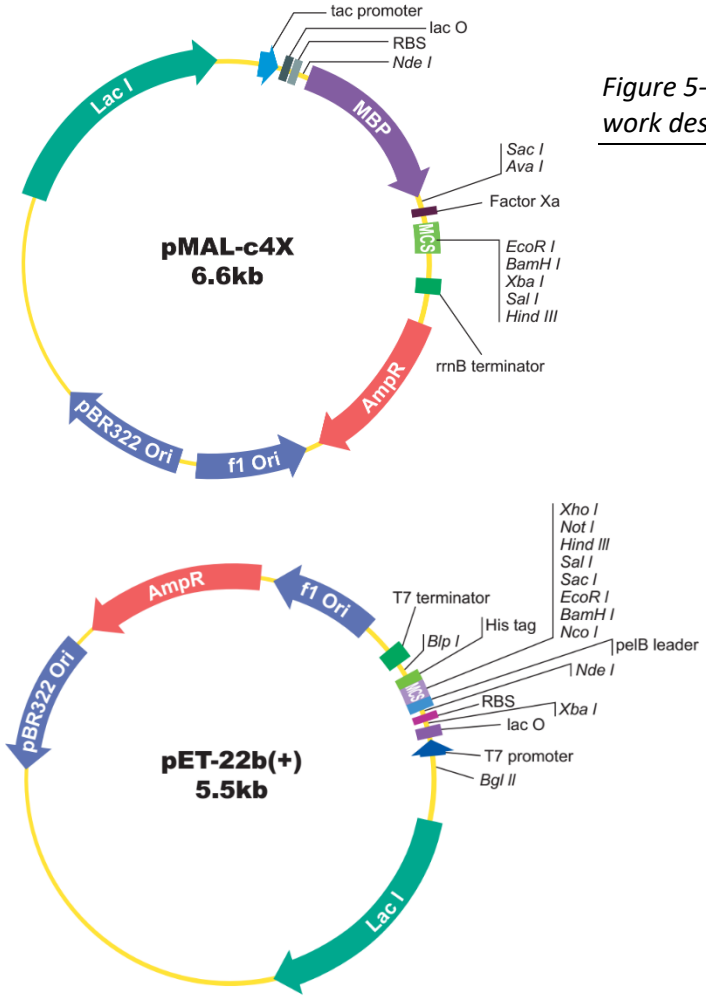


Figure 5-5. Maps of expression vectors used in the work described here. Maps supplied by GenScript.

5.2.7 Natural transformation

Natural transformation of naturally competent *Acinetobacter baylyli* strains can be achieved by growing them in liquid culture to the stationary phase, adding DNA and additional nutrients and incubating them for an hour or more, the switch from stationary phase to growth induces the uptake of DNA (Metzgar et al., 2004; Vaneechoutte et al., 2006). This method was used to transform overnight cultures of *A. baylyi* 107474T grown with sodium acetate by the addition of plasmid DNA and more sodium acetate.

5.3 Results

5.3.1 Identification of enzymes that have greater homology between *Variovorax* E3 and *A. calcoaceticus* E1, and their evaluation as MDP candidates

5.3.1.1 The BSR of A. calcoaceticus RUH 2202 and Variovorax E3 proteomes to A. calcoaceticus E1

Utilising predicted protein sequences from each strain, BSR were obtained of *Variovorax* E3 and *A. calcoaceticus* RUH 2202 proteomic query sequences versus the *A. calcoaceticus* E1 proteome. The query BSRs obtained for each *A. calcoaceticus* E1 subject sequence are given in Figure 5-6. As would be expected the majority of sequences scored highly (BSR > 0.8) in the intraspecies comparison. The group of sequences with BSR < 0.4 for both strains that cluster around the line is interesting; it may be hypothesised that these are proteins with highly conserved domains that are part of otherwise unconserved proteins. The sequences with BSR to *Variovorax* E3 that are greater than 0.5 and greater than the BSR of *A. calcoaceticus* RUH 2202 sequences (shown in blue in Figure 5-6) were investigated further.

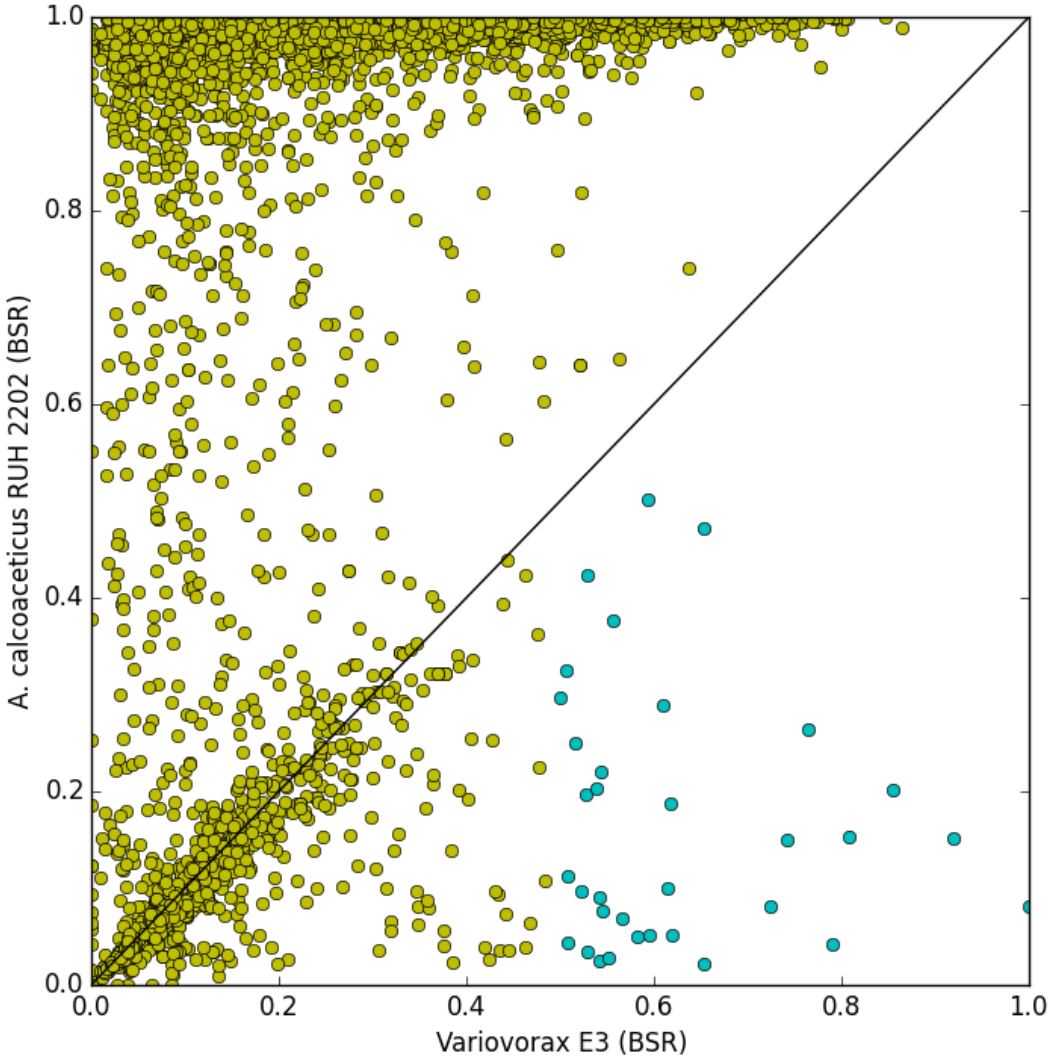


Figure 5-6. BSR values for two query proteomes versus the A. calcoaceticus *E1 proteome. Each circle represents an* A. calcoaceticus *E1 predicted protein. The circle's position on an axis gives the highest BSR obtained from that strain's proteome. The diagonal line shows where the two BSR are equal. Circles below the line therefore show* A. calcoaceticus *E1 sequences that are more similar to a* Variovorax *E3 sequence than any* A. calcoaceticus *RUH 2202 sequence. Blue circles show sequences with* Variovorax E3 *BSR > 0.5 and Variovorax E3 BSR > A. calcoaceticus RUH 2202 BSR.*

For the purposes of comparison, the process was repeated using the same *A. calcoaceticus* proteomes but with other *Variovorax* genomes: *V. boronicumulans, V. paradoxus B4, V. paradoxus EPS and V. paradoxus S110.* A chart showing BSR for the *V. boronicumulans and A. calcoaceticus* RUH 2202 proteomes are shown in Figure 5-7. The highest BSR to one of these *Variovorax* proteomes, that was higher than equivalent *A. calcoaceticus* RUH 2202 BSR, was 0.804.

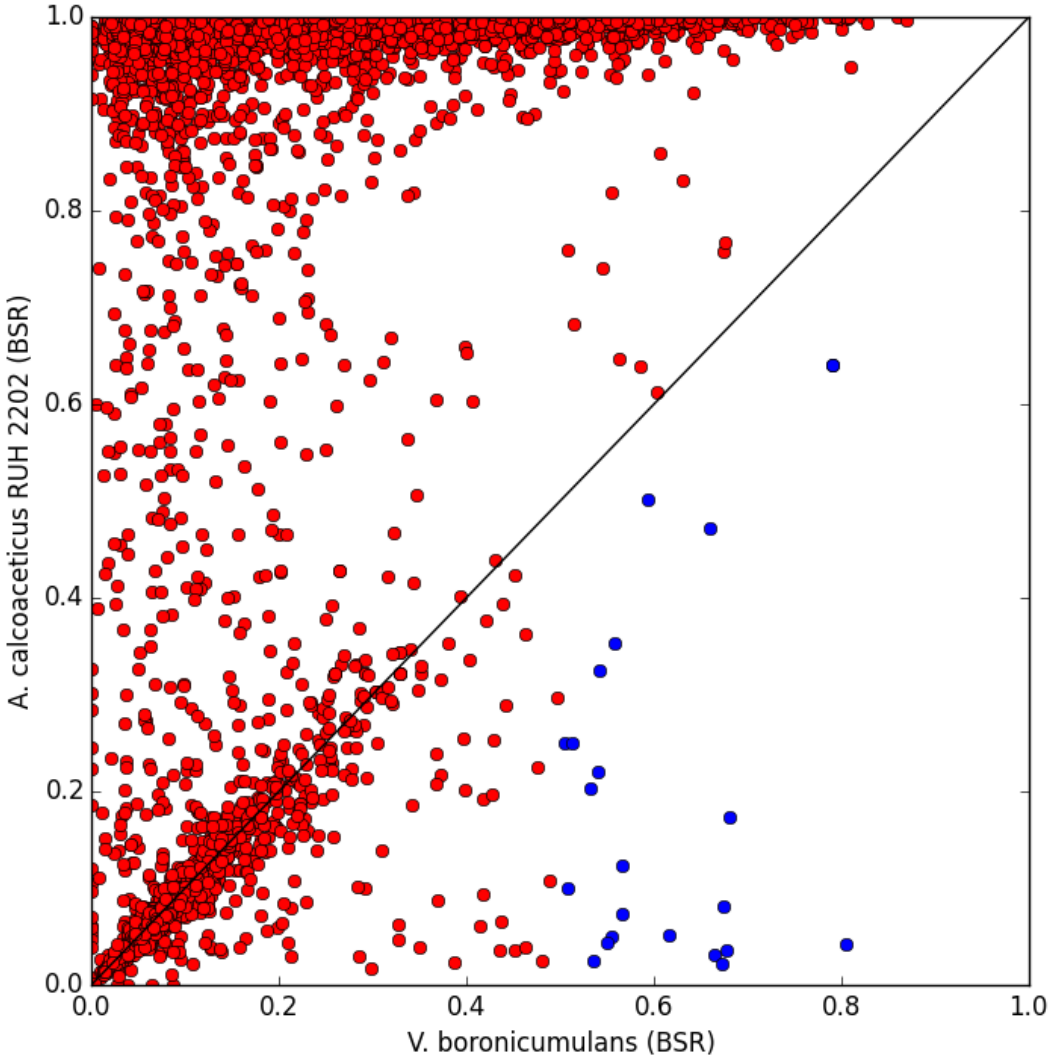


Figure 5-7. BSR values for two query proteomes versus the A. calcoaceticus *E1 proteome. Each circle represents an* A. calcoaceticus *E1 predicted protein. The circle's position on the axes gives the highest BSR obtained from that strain's proteome. The diagonal line shows where the two BSR are equal. Circles below the line therefore show* A. calcoaceticus *E1 sequences that are more similar to a* V. boronicumulans *E3 sequence than any* A. calcoaceticus *RUH 2202 sequence. The blue circles show sequences that have a BSR > 0.5 in* V. boronicumulans that are also below the line.

5.3.1.2 Characterisation of genes with higher homology to Variovorax E3

A. calcoaceticus E1 genes with greater homology to *Variovorax* E3 than *A. calcoaceticus* RUH 2202 and BSR > 0.5 are discussed here. Where these genes are closely collocated in the genome, they will be discussed together.

Mercury resistance operon

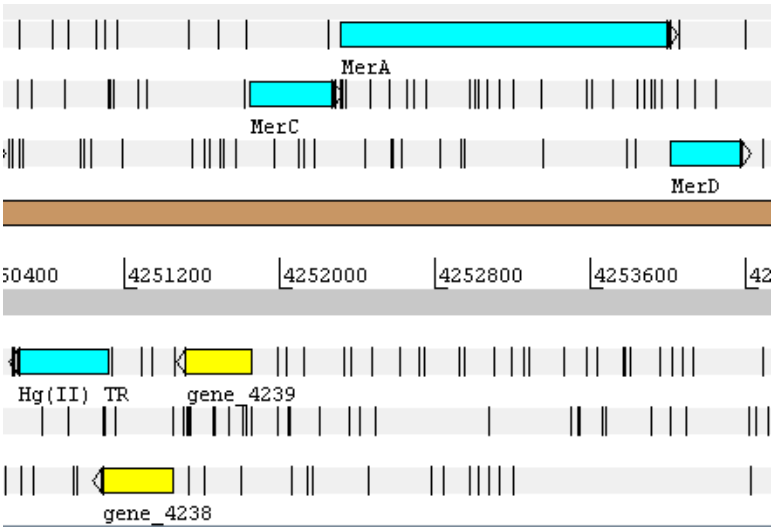


Figure 5-8. Relative position of mercury resistance genes in A. calcoaceticus *E1 genome. Genes with greater homology and BSR > 0.5 between the metaldehyde degrading strains shown in blue. Light grey lanes represent the 6 reading frames; vertical black lines indicate stop codons.*

	Predicted family	BSR rank	*V.* E3 BSR	*Ac.* E1 BSR
MerC	Mercury resistance protein (IPR004891)	1	1.00	0.08
MerA	Mercury(II) reductase (IPR021179)	2	0.92	0.15
Hg(II) TR	Hg(II) Responsive transcriptional regulator (IPR011794)	3	0.85	0.20
MerD	Mercuric resistance transcriptional repressor protein (IPR011797)	4	0.81	0.15

Table 5-2. Genes with greater homology and BSR > 0.5 between the metaldehyde degrading that are part of a mercury resistance operon. The BSR rank is the rank order of potentially horizontally transferred genes (shown in blue in Figure 5-6), from the highest Variovorax *E3 BSR to the lowest.*

The 4 genes with greatest homology between *A. calcoaceticus* E1 and *Variovorax* E3 are part of a mercury resistance operon. MerC transports Hg^{2+}. Genes 4238 and 4239 have no predicted domains or families and no significant homologues in the *Variovorax* E3 predicted proteome. A BLAST search of the *A. calcoaceticus* E1 contig containing these genes, using the megablast algorithm, against the *Variovorax* E3 genome identifies a syntenic locally collinear block that spans the MerC-MerA-MerD genes shown in Figure 5-8 and has 91.4% nucleotide identity with the *Variovorax* E3 block. The GC content for the *A. calcoaceticus* E1 contig is shown in Figure 5-9, with the region shown in Figure 5-8 highlighted.

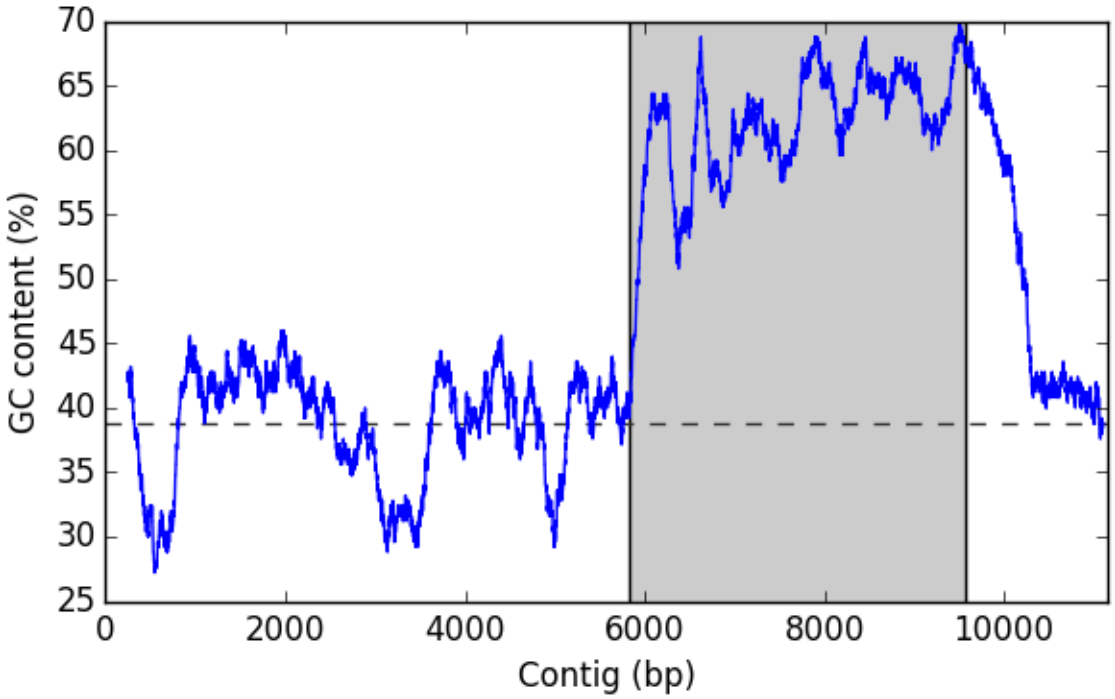


Figure 5-9. Mean GC content (%) of a 11.2 kbp contig using a 250 bp sliding window. The shaded region indicates the region inclusively spanning from HG(II) TR to MerD (Figure 5-8). The mean GC content for A. calcoaceticus E1, *38.8%, is indicated by the dashed line.*

A transposase is located shortly upstream of MerD. This contig may be part of a plasmid, or part of the genome. It is certain that MerC-MerA-MerD are horizontally transferred in *A. calcoaceticus* E1 and likely that this sequence has a recent shared origin with the *Variovorax* E3 genomic region.

MerA acts on mercury, and the other proteins are not similar to enzymes. These proteins are not good candidates for MDP.

Other annotated genes on the same contig are: a heavy metal resistance gene, a transposase and a HigB toxin gene.

MerA acts on mercury, and the other proteins are not similar to enzymes. These proteins are not good candidates for MDP.

The Rut pathway

Signatures for proteins in the Rut pathway, a pyrimidine degradative pathway, was found in proteins with greater homology between the metaldehyde degrading strains. The location of these genes in the *A. calcoaceticus* E1 genome are shown in Figure 5-10. Predicted family memberships and BSR are given in Table 5-3.

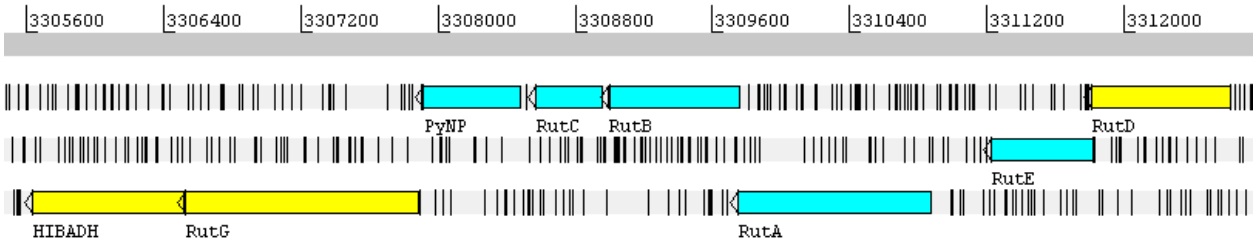


Figure 5-10. Position of putative Rut pathway genes in A. calcoaceticus *E1 genome. Genes with greater homology and BSR > 0.5 between the metaldehyde degrading strains shown in blue.. Lanes representing the 3 reading frames proceed from right to left, along which vertical black lines indicate stop codons. The scale bar units are numbers of base pairs.*

Abbr.	Predicted family	*V.* E3 BSR rank	*V.* E3 BSR	*Ac.* RUH 2202 BSR
RutC	Aminoacrylate peracid reductase (IPR019898)	6	0.76	0.26
RutA	Pyrimidine monooxygenase (IPR019914)	7	0.74	0.15
RutE	NADH dehydrogenase/NAD(P)H nitroreductase, putative (IPR023936)	13	0.61	0.10
RutB	Isochorismatase, amidohydrolase, RutB (IPR019916)	21	0.55	0.08
PyNP	Thymidine|pyrimidine-nucleoside phosphorylase (IPR000053)	29	0.54	0.22
RutG	Pyrimidine permease, putative (IPR019918)	-	0.25	0.55
RutD	Pyrimidine utilisation protein (IPR019913)	-	0.41	0.64
HIBADH	3-hydroxyisobutyrate dehydrogenase-related (IPR015815)	-	0.12	0.11

Table 5-3. Predicted protein family membership and BSR of genes shown in Figure 5-10.

The Rut operon enables the catabolism of pyrimidines (Loh et al., 2006). The functions of proteins RutA and RutB were demonstrated *in vitro* and other proteins in the pathway had their functions predicted based on sequence similarity to proteins of known function (Kim et al., 2010). The pathway proposed by Kim et al. (2010) is shown in Figure 5-11.

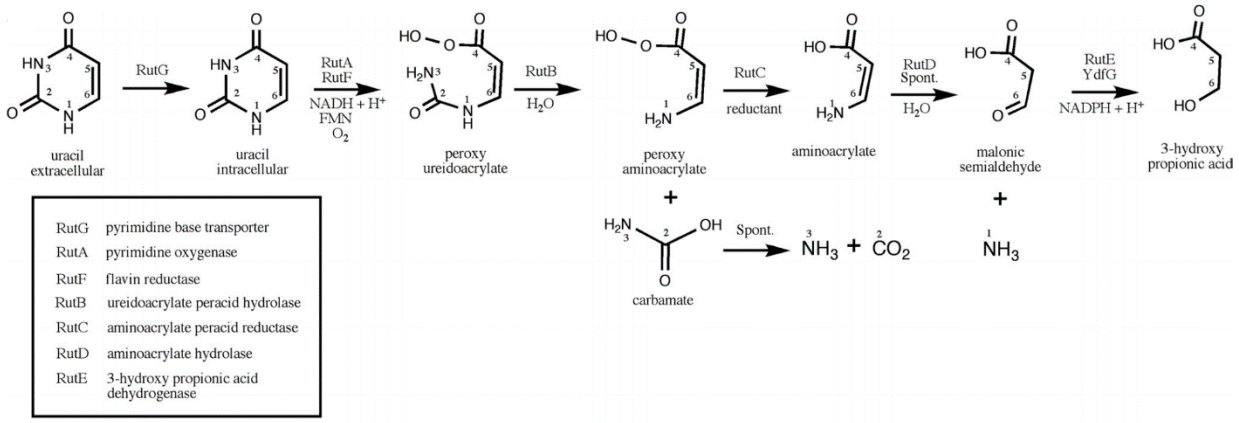


Figure 5-11. The proposed Rut pathway for uracil catabolism. Figure taken from Kim et al. (2010)

For the purposes of this study, the pyrimidine monooxygenase, RutA, and the amidohydrolase, RutB, are most interesting. Both cleave amines bound to 2 carbons (Figure 5-11). The results of an InterPro search using the *Variovorax* E3 sequence results in the same predicted function for the proteins. The GC% of the *Variovorax* E3 and *A. calcoaceticus* E1 genes found to be homologous are 64% and 43% respectively, which does not support the hypothesis that they were recently horizontally transferred.

Phenol degradation pathway

Figure 5-12 and Table 5-4 give details of genes found to be more similar in *Variovorax* E3. These are predicted to be part of a phenol degradation pathway.

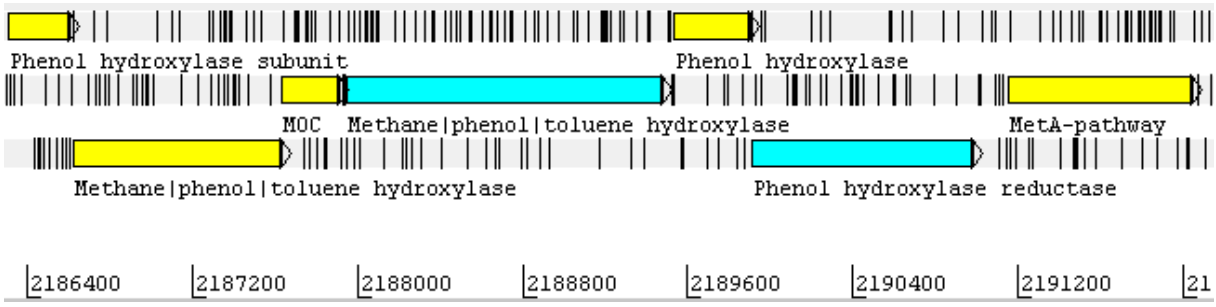


Figure 5-12. Position of putative phenol degradation pathway genes in A. calcoaceticus *E1 genome. Genes with greater homology and BSR > 0.5 between the metaldehyde degrading strains shown in blue. Lanes representing 3 reading frames proceed from left to right, along which vertical black lines indicate stop codons. The scale bar gives number of base pairs.*

Predicted family	BSR rank	*V.* E3 BSR	*Ac.* RUH 2202 *B*SR
Methane|phenol|toluene hydroxylase (PHOα; IPR003430)	10	0.65	0.02
Phenol hydroxylase reductase (IPR001221)	22	0.54	0.22
MOC (Monooxygenase component MmoB/DmpM; IPR003454)	-	0.48	0.11
Methane|phenol monooxygenase, hydroxylase component (IPR012078)	-	0.44	0.04
Phenol hydroxylase (IPR006756)	-	0.36	0.09
Phenol hydroxylase subunit (IPR010353)	-	0.31	0.10
MetA-pathway of phenol degradation, putative (IPR025737)	-	0.032	0.12

Table 5-4. Predicted protein family membership and BSR of genes shown in Figure 5-12

A. calcoaceticus E1, but not strain RUH 2202, was found to be able to utilise phenol (Section 3.3.7), confirming a possible function for the genes shown above. The ability of *Variovorax* E3 to catabolise phenol is currently undetermined. Other genes predicted to be involved in phenol degradation are present in the *A. calcoaceticus* E1 genome, but are not co-located. The oxidation of phenol requires a 3 component enzyme, made up of an oxidase, a hydroxylase and a reductase (Divari et al., 2003). The gene labelled methane|phenol|toluene hydroxylase has 93% amino acid identity to the phenol hydroxylase oxygenase alpha component (PHOα) identified by Divari et al. (2003). This subunit contains the active site for the 3 component enzyme complex. The *Variovorax* E3 gene with most homology to this was also annotated as a methane|phenol|toluene hydroxylase and has 66% identity to the *A. radioresistorans* PHOα.

Despite the evidence linking these genes to phenol oxidation, they were still considered to be of interest as a candidate for MDP due to the substrate promiscuity observed in enzymes of this family (Tinberg & Lippard, 2010).

Aldolase and ALDH

Three genes that are more similar in *Variovorax* E3 and appear to form an operon are presented here in a genomic context in Figure 5-13, and with BSR in Table 5-5.

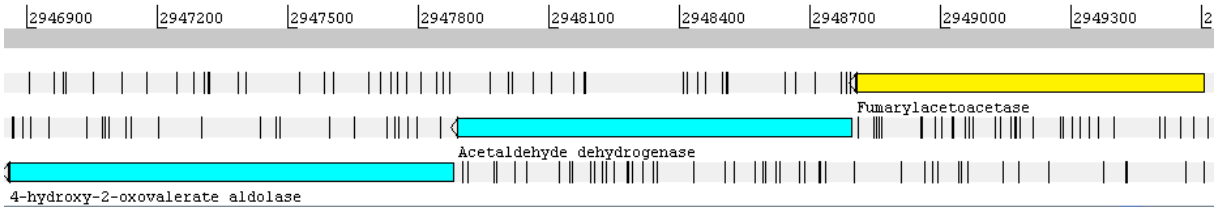
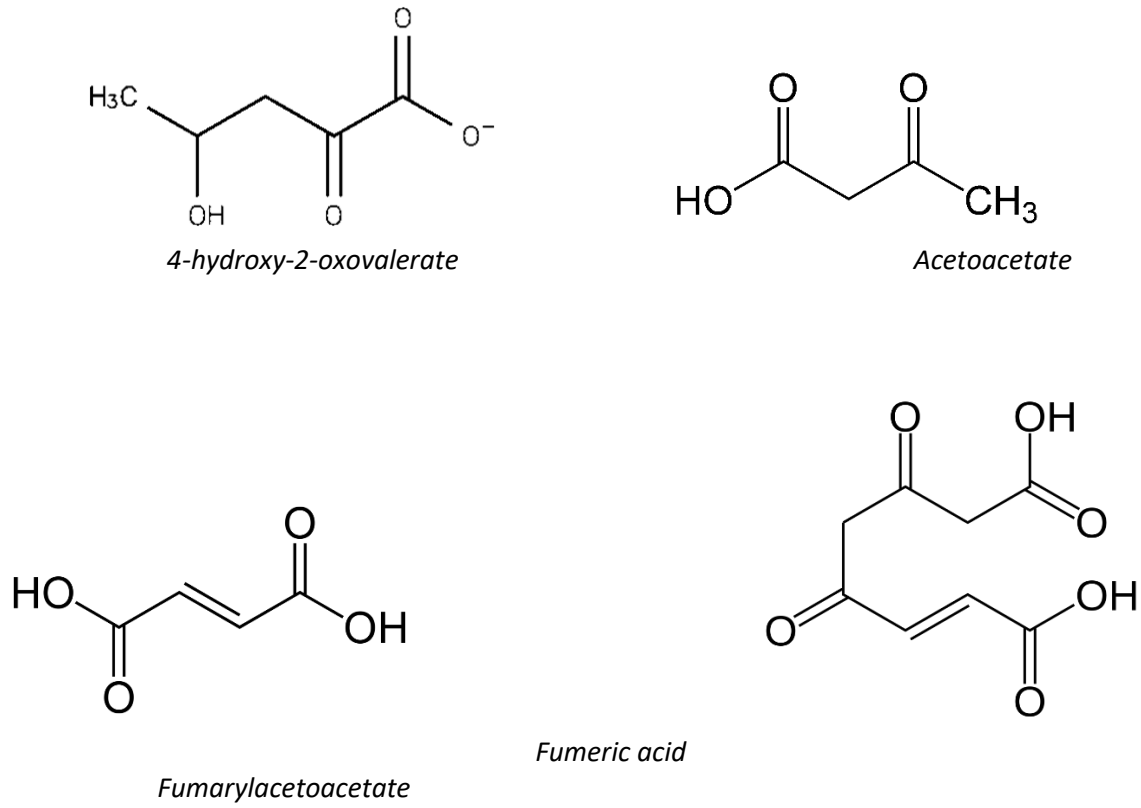


Figure 5-13. Position of genes in a putative operon in the A. calcoaceticus *E1 genome. Genes with greater homology and BSR > 0.5 between the metaldehyde degrading strains shown in blue.. Lanes representing 3 reading frames proceed from right to left, along which vertical black lines indicate stop codons. The scale bar units are numbers of base pairs*

Abbr.	Predicted family	BSR rank	*V.* E3 BSR	*Ac.* RUH 2202 *B*SR
ALDH	Acetaldehyde dehydrogenase (IPR003361)	5	0.79	0.04
4H2O aldolase	4-hydroxy-2-oxovalerate aldolase (IPR017629)	8	0.72	0.08
FAH	Fumarylacetoacetase (IPR005959)	-	0.42	0.04

Table 5-5. Predicted protein family membership and BSR of genes shown in Figure 5-13.



4-hydroxy-2-oxovalerate

Acetoacetate

Fumarylacetoacetate

Fumeric acid

Figure 5-14. Compounds that may be products or substrates of the enzymes given in Table 5-5.

The ALDH identified here is not the same one identified to be expressed by *A. calcoaceticus* E1 when growing on metaldehyde (see Section 4.3.4.1). If the hypothesis that this operon is responsible for the degradation of metaldehyde were correct, then this ALDH would also be expressed as it appears to be in an operon with the other two genes. However, this does not impact on the likelihood of the hypothesis, as the depolymerisation of 1 mole of metaldehyde would produce 3 or 4 moles of acetaldehyde, which could lead to an excess that induced the expression of an ALDH that was directly induced by acetaldehyde.

4H2O aldolases catalyse the lysis of 4-hydroxy-2-oxovalerate (Figure 5-14) to pyruvate and acetaldehyde (Platt et al., 1995). An ALDH is typically found associated with this gene. The carbon-carbon bond lysed by this type of enzyme is not similar to bonds in the metaldehyde ring and therefore the aldolase was not considered to be a good candidate for MDP.

The action of fumarylacetoacetases is to catalyse the hydrolysis of fumarylacetoacetate to fumarate and acetoacetate. It is possible that the FAH predicated here produces acetoacetate which is catabolised by the product of the predicted aldolase, plausible due to the structural similarity of acetoacetate to 4H2O (Figure 5-14).

The homologous predicted fumarylacetoacetase gene found in *Variovorax* E3 is not collocated with a 4H2O aldolase/ALDH pair. The GC content of the homologous regions are typical for each strains. These genes do not appear to be horizontally transferred between the species and the predicted reactions are not similar to those likely to occur in the primary steps of the degradation of metaldehyde.

Other enzymes with greater homology in Variovorax *E3*

The examination of the genomic context of the aldolase, discussed in the previous section, indicated that its function is likely not conserved between *A. calcoaceticus* E1 and *Variovorax* E3 based on the differences in genes present in the operons. With the exception of the mercury resistance operon, no genes discussed so far appear to have been recently horizontally transferred between the two strains. The classification of proteins with BSR > 0.5 were examined to determine if the reaction associated with the classification, or the substrate they act on indicated they may be candidates for MDP. A more detailed examination of them is not presented here.

The predicted enzymes given in Table 5-6 were rejected as candidates for MDP due to the nature of the reactions they catalyse, and substrates they act on. Proteins listed in Table 5-7 were identified as transporters and were therefore not considered as candidates. Proteins listed in Table 5-8 were predicted to be regulatory proteins or had no function predicted.

Predicted family	BSR rank	*V.* E3 BSR	*Ac.* RUH 2202 *B*SR
Coenzyme A transferase family I (IPR004165)	9	0.65	0.47
Aspartate 4-decarboxylase (IPR022518)	11	0.62	0.05
Glucose/ribitol dehydrogenase (IPR006003)	14	0.61	0.29
Kynureninase (IPR010111)	17	0.58	0.05
Haloacid dehydrogenase like domain (IPR023214)	20	0.55	0.03
FMN reductase (NADH) RutF (IPR019917)	23	0.54	0.09
None (tRNA nucleotidyltransferase domain, IPR015329)	27	0.53	0.03
Thymidine/pyrimidine-nucleoside phosphorylase (IPR000053)	29	0.53	0.42
Acrylyl-CoA reductase (IPR014188)	30	0.52	0.10
Glucose/ribitol dehydrogenase (IPR006003)	33	0.51	0.32
Lon protease, bacterial (IPR027543)	34	0.50	0.30

Table 5-6. Predicted protein family membership and BSR of genes for enzymes found in A. calcoaceticus *E1 that have BSR values to the* Variovorax *E3 that are greater than 0.5 and greater than the* A. calcoaceticus *RUH 2202 BSR.*

Predicted family or domain	BSR rank	*V.* E3 BSR	*Ac.* RUH 2202 *B*SR
Aldo/keto reductase/potassium channel subunit beta (IPR001395)	12	0.62	0.19
ABC-2 transporter (IPR025699)	18	0.57	0.07
ABC transporter ATPase like domain (CATH family 1.20.1580.10)	19	0.56	0.38
Aspartate-alanine antiporter (IPR022457)	24	0.54	0.02
ATPase/ABC transporter domain (IPR019195)	25	0.54	0.20
ABC transporter, permease (IPR001851)	26	0.53	0.04
ATPase domain of an ABC transporter (IPR027417)	31	0.52	0.25

Table 5-7. Predicted protein family membership and BSR of genes for transporters found in A. calcoaceticus *E1 that have BSR values to the* Variovorax *E3 that are greater than 0.5 and greater than the* A. calcoaceticus *RUH 2202 BSR.*

Predicted family or domain	BSR rank	*V.* E3 BSR	*Ac.* RUH 2202 *B*SR
(None predicted)	15	0.60	0.05
Transcriptional regulatory protein PcoR, heavy metal response (IPR006291)	16	0.59	0.50
(None predicted)	28	0.53	0.20
DNA-binding HTH domain, TetR-type (IPR001647)	32	0.51	0.11

Table 5-8. Predicted protein family membership and BSR of genes found in A. calcoaceticus *E1 that have BSR values that have BSR values to the* Variovorax *E3 that are greater than 0.5 and greater than the* A. calcoaceticus *RUH 2202 BSR.*

5.3.1.3 Candidates for MDP identified by homology between A. calcoaceticus *E1 and* Variovorax *E3*

Of the enzymes discussed above three stood out as having some potential for metaldehyde degrading activity. They were annotated as RutA, RutB and a phenol hydroxylating enzyme. None of these pathways appeared to be directly transferred between the two species, contradicting the hypothesis that MDP is shared between the two metaldehyde degrading isolates, but the reactions predicted to be performed by these enzymes made them interesting. A BLASTp search against the NCBI non-redundant database showed that proteins with high similarity to the *A. calcoaceticus* E1 PHOα and the Rut pathway proteins were present in other *Acinetobacter*.

5.3.1.4 Identification of Acinetobacter strains with RutA and phenol hydroxylase genes

Acinetobacter strains were identified that had, or were lacking, genes homologous to the genes annotated as encoding RutA, RutB and PHOα. This would allow a test of the hypothesis that one of these genes was responsible for metaldehyde degradation.

Strain names were extracted from the results of BLAST searches using the protein sequences for RutA, RutB and PHOα. A survey of culture collections from which *Acinetobacter* strains identified could be purchased was conducted. Several strains with genes homologous (amino acid identity ≥ 97.6%) to those for RutA, RutB and PHOα were found to be available from The Collection of Institut Pasteur (CIP; France). Strains were purchased, the details of which are given in Table 5-9. These strains were tested for metaldehyde catabolism using liquid cultures and solid media containing metaldehyde as the sole carbon source. None showed any signs of growth, indicating that these genes are not responsible for metaldehyde degradation.

Strain	CIP number	Other CC no.	AA identity to *Ac.* E1			Isolation details	Genome reference
			RutA	RutB	PHOα		
A. pittii	70.29		97.9%	97.6%	99.6%	Cerebrospinal fluid (Seifert et al., 1994)	NCBI bioproject 183266
A. calcoaceticus	110488	ANC 3680	98.7%	98.4%	3.0%	Beech forest soil, Czech Republic, 2008	Nemec et al. (2011)
A. pittii	110468	ANC 3678	98.4%	97.6%	99.8%	Pond, Czech Republic, 2008	Nemec et al. (2011)
A. calcoaceticus	110439	NIPH 13	25.0%	18.8%	3.0%	Burn, Czech Republic, 1991.	Nemec et al. (1999)
A. baylyi	107474T	DSM 14961	84.1%	78.8%	3.0%	Activated sludge, Australia (Carr et al., 2003)	NCBI bioproject 183290

Table 5-9. Strains purchased form CIP. Where the strain is primarily known by another culture collection number in the literature, this is given as ***Other CC no.*** *The highest amino acid identities found in each strain for 3* A. calcoaceticus *E1 protein sequences are given. The location and date of the initial isolation of the strain is given, along with a literature reference where available. Where the publication of the strain's genome is associated with a paper, this reference is given as the genome reference. Otherwise, the NCBI bioproject number is given; the project details can be retrieved by entering the number into the search field at www.ncbi.nlm.nih.gov/bioproject*

5.3.2 Identification and evaluation of MDP candidates by intrageneric comparison of *Acinetobacter* genomes

Intrageneric comparative genomics was performed using *A. calcoaceticus* E1, RUH 2202 and the strains given in Table 5-9. A survey of the genome *A. calcoaceticus* E1 was conducted to find proteins that were unique to the strain, a "unique" protein being defined as having BSR ≤ 0.4 when compared to any of the strains confirmed to lack metaldehyde catabolism.

A histogram of the distribution of BSR for each other strain compared to *A. calcoaceticus* E1 is shown in Figure 5-15. The *A. calcoaceticus* have the highest frequency of sequences with BSR ≥ 0.9, as would be expected. The predicted proteome of *A. baylyi* is not very similar to that of *A. calcoaceticus* E1.

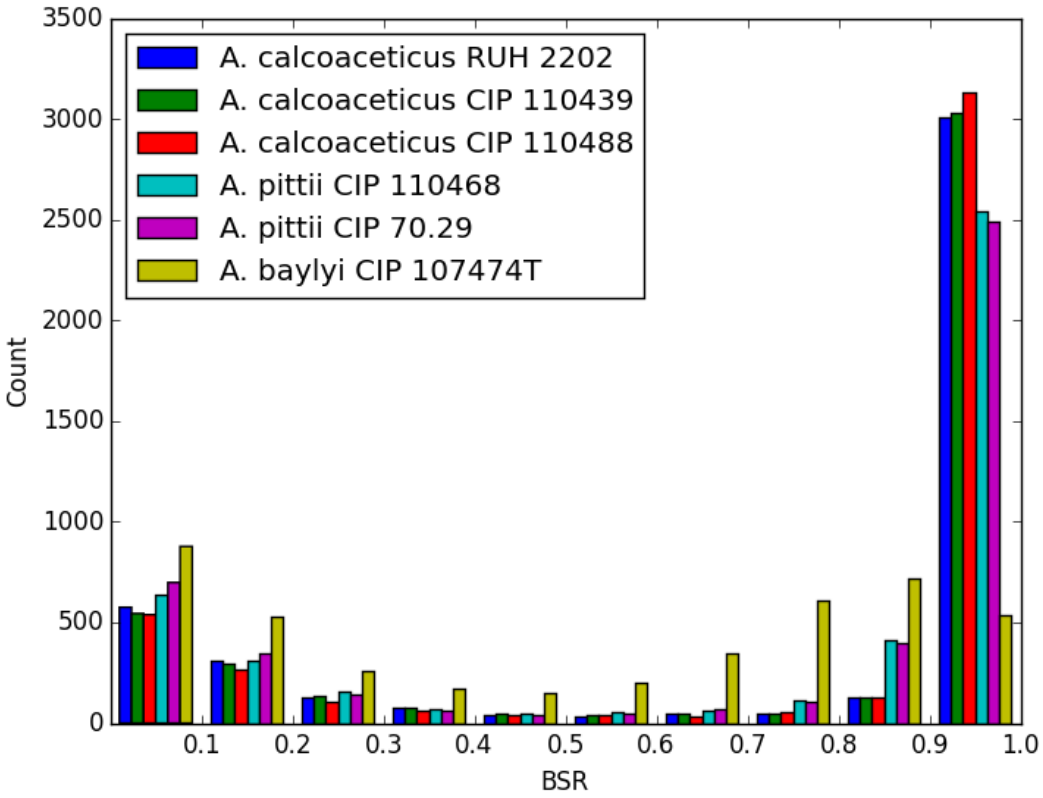


Figure 5-15. Frequency of BSR values of A. calcoaceticus *E1 putative gene sequences to different* Acinetobacter *strains.*

The detected protein family membership and domain signatures were retrieved for putative proteins from the *A. calcoaceticus* E1 genome that had BSR ≤ 0.4 against all *Acinetobacter* tested. Those with some annotation were manually reviewed, and those with a catabolic activity predicted were examined in more detail.

Table 5-10 gives the number of genes remaining at each stage of this process.

Gene filter	Remaining	
All genes	4395	100.0%
Unique	669	15.2%
Any annotation	273	6.2%
Catabolic activity	29	0.7%

Table 5-10. Number and percentage of total A. calcoaceticus *E1 genes remaining that are unique (have BSR ≤ 0.4 against tested* Acinetobacter *strains), had any annotation (domain or family signatures detected by InterPro) and had annotation that indicated some kind of catabolic activity.*

The 273 genes with some annotation were categorised based on their classification by InterPro scan (Figure 5-16). The second most common category was of phage related genes. This group only includes those explicitly annotated as being phage proteins; many of the DNA processing and genes/domains of unknown function may also originate from phages. A plasmid mobilisation gene and 5 toxin/antitoxin genes were observed. These observations are consistent with the hypothesis that the genes more unique in a bacterial genome are more likely to have been horizontally transferred.

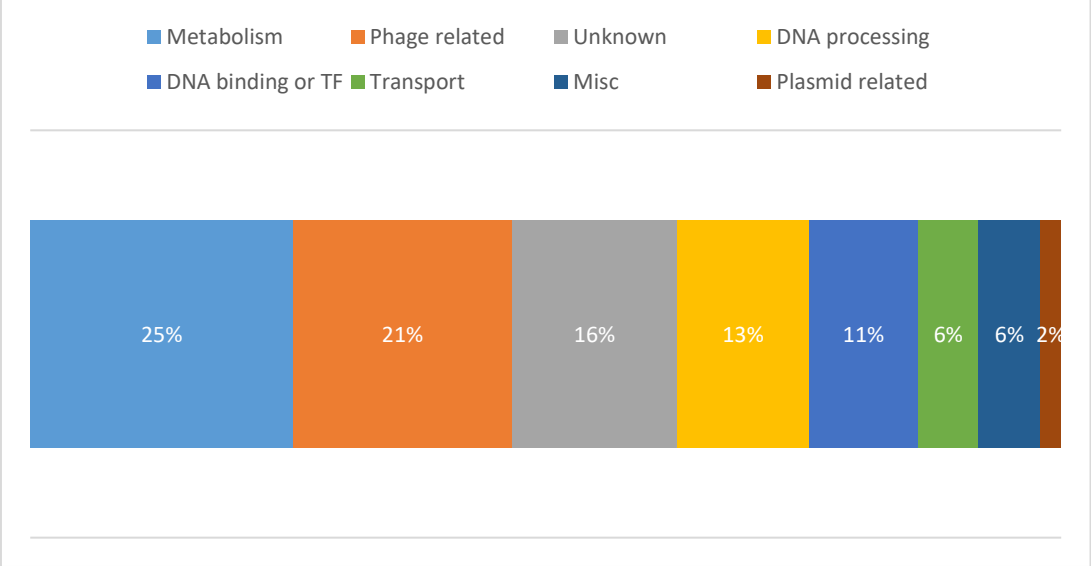


Figure 5-16. Percentage of annotated unique genes that were assigned to different groups. The group of unknown genes is made up of those that were annotated as having an unknown function.

5.3.2.1 Characterisation of A. calcoaceticus E1 proteins with low intrageneric homology

The 29 unique genes predicted to have a catabolic function are discussed here.

Protein 1137 – Isochorismatase-like

The *A. calcoaceticus* E1 predicted protein 1137 contains a isochorismatase-like domain (IPR000868). The isochorismate-like family of proteins is large and includes enzymes that act on substrates other than isochorismate such as RutB, discussed above. The reaction typically catalysed by enzymes with this domain signature is shown in Figure 5-17. This is the hydrolysis of an ether, making it of interest as a candidate for MDP.

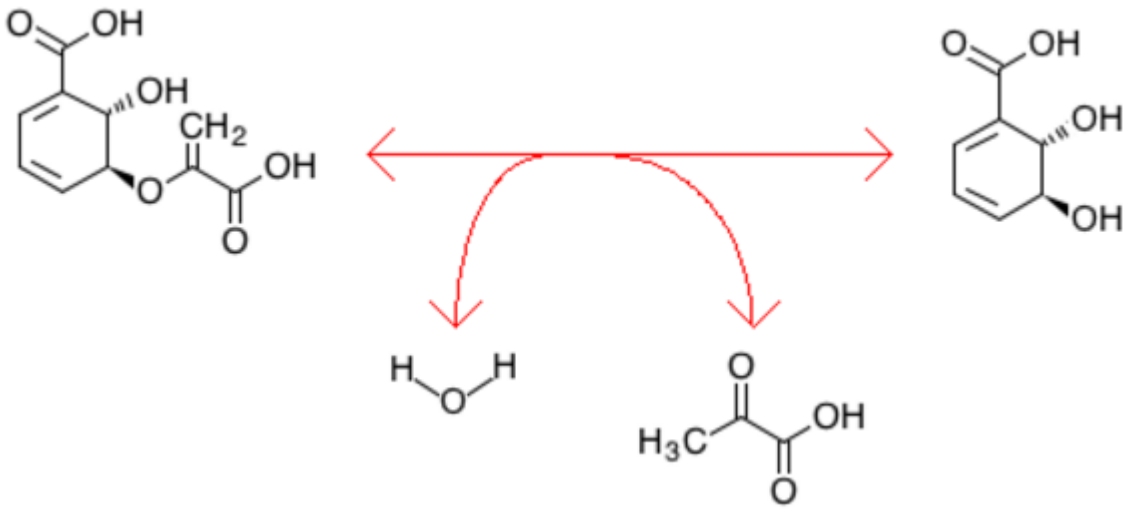


Figure 5-17. The reaction catalysed by isochorismatase. Isochorismate is hydrolysed to pyruvate and (2S,3S)-2,3-Dihydro-2,3-dihydroxybenzoate. Image is of R03037 taken from the KEGG website (Kanehisa et al., 2004)

To obtain more information on the likely substrate of, and reaction catalysed by, protein 1137, its structure was predicted using SWISS-MODEL. The computed structure with the highest QMEAN4 used the structure listed in the Protein Data Bank (PDB; Berman et al., 2000) as 4H17.1.b. This protein was isolated from *Pseudomonas putida*, and the computed structure had a QMEAN4 score of -3.75. Unfortunately, this protein had no published literature associated with it. The second best result was for the PDB structure 3LQY (Goral et al., 2012). The initial alignment resulted in 44% amino acid identity and the final structure had a QMEAN4 score of -4.12. This structure was of a protein from the Antarctic γ-proteobacterium *Oleispira antarctica*, a putative isochorismate hydrolase, referred to as OaIHL (Goral et al., 2012). Goral et al. performed *in silico* docking of OaIHL with isochorismate. In Figure 5-18 an image of the docking is presented (A) along with a ribbon model of the published structure for OaIHL (B), and the computed structure of protein-1137 (C).

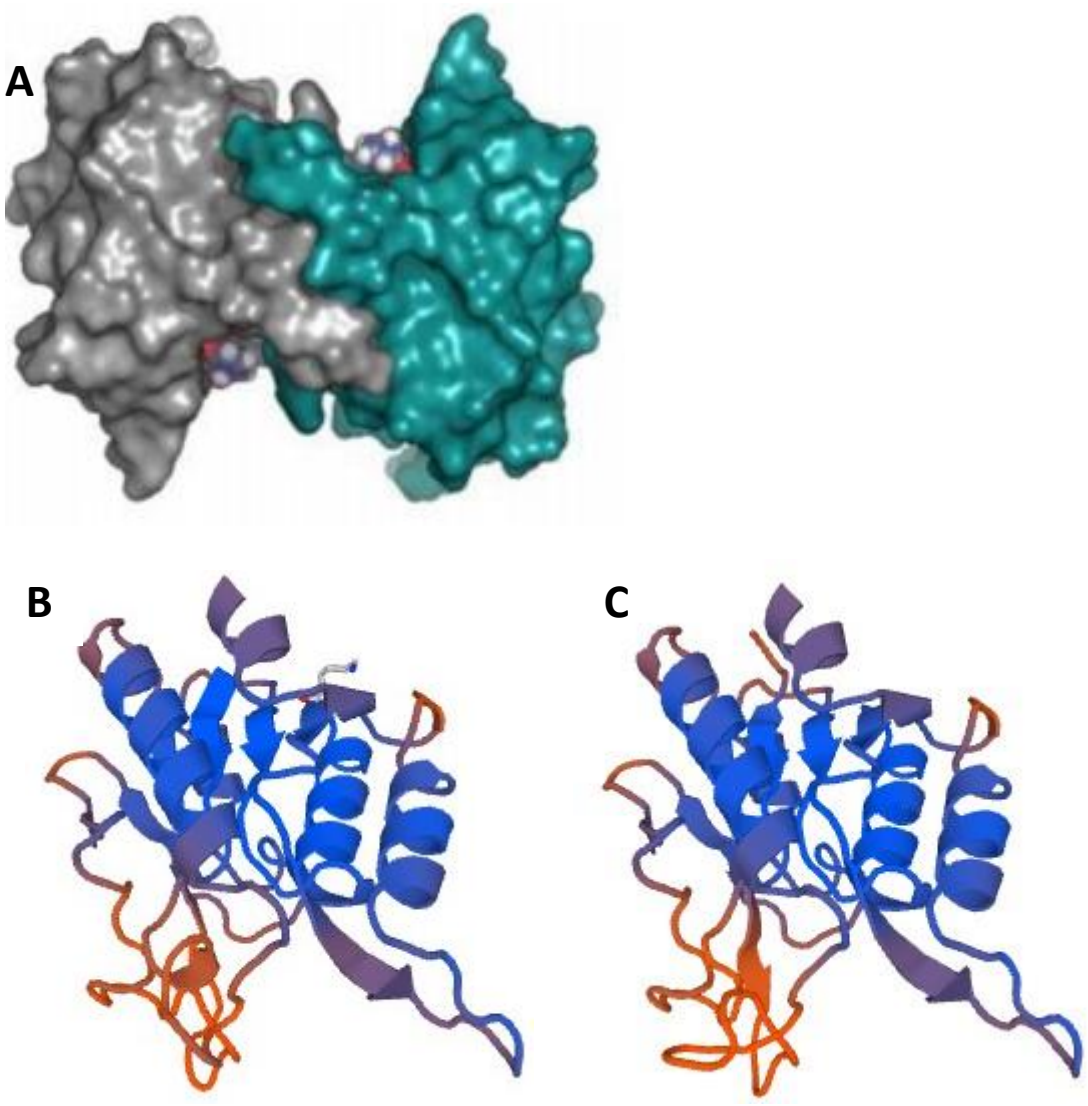


*Figure 5-18. **A**, Surface rendering of OaIHL with* in silico *docked isochorismate, taken from Goral et al. (2012). **B,** ribbon structure of OaIHL (3LQY). **C,** computed structure of protein 1137. The colours **B** and **C** indicate the degree of similarity between the two structures, blue is high and orange is low.*

The portion of the structures shown in orange in Figure 5-17B and C are not structurally conserved between the proteins. This unconserved region includes residues believed to interact with isochorismate in OaIHL and be involved in substrate selection (Goral et al., 2012).

A catalytic triad of Asp-Lys-Cys were identified to be involved in the hydrolysis of the ether in OaIHL, based on their conservation in the family of enzymes and the *in silico* docking they performed (Goral et al., 2012). The structure of OaIHL and protein 1137 were superimposed over one another and these residues visualised (Figure 5-19).

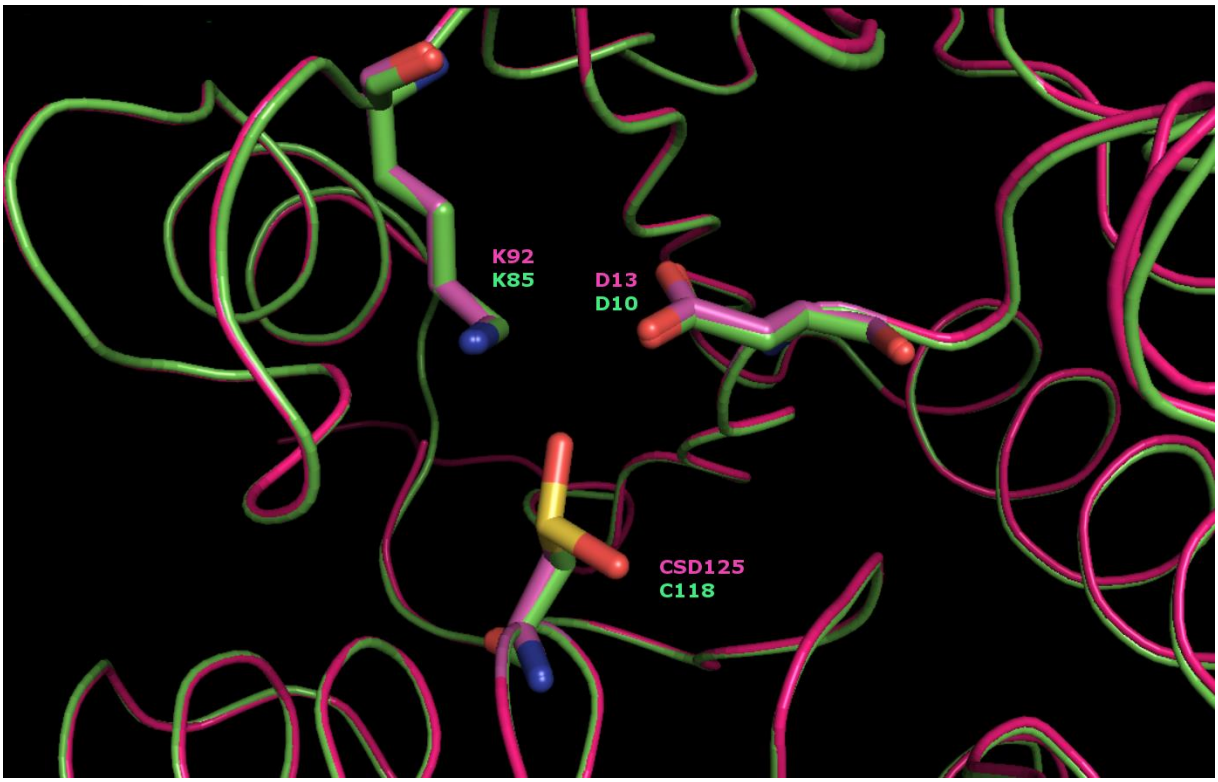


Figure 5-19. Ribbon structure of OaIHL (pink) and computed structure of protein 1137 (green). Residues of the catalytic triad of OaIHL, and residues that match these of protein 1137, are shown as sticks.

It can be seen that the positions of the catalytic residues are predicted to have conserved positions between the two enzymes (Figure 5-19), consistent with the possibility that the two enzymes perform the same reaction on the same chemical group. The lysine is predicted to coordinate the ethereal oxygen (Goral et al., 2012). The cysteine in 3LQY has been modified to a sulfino-alanine (CSD125). Protein 1137 C118 aligns in space with the unmodified portion of CSD125.

The initial step of the proposed reaction mechanism for OaIHL is protonation of the ethylene C3' in the pyruvyl sidechain leading to water binding to the pyruvyl C2' (Goral et al., 2012). For illustrative purposes a similar mechanism (though catalysed by a dissimilar protein) is shown for the hydrolysis of chorismate in Figure 5-20. Metaldehyde does not have an ethylene, but protonation of an ethereal oxygen could cause the depolymerisation of the molecule.

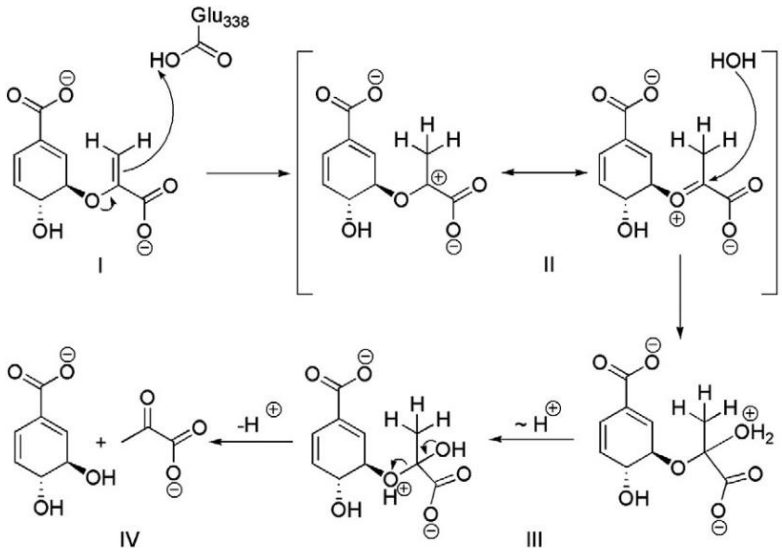


Figure 5-20. Reaction mechanism of chorismatase FkbO. Taken from Juneja et al. (2014).

Examination of the genomic region of protein 1137 reveals it to be plasmid borne. The contig containing this protein is shown in Figure 5-21. On this contig are proteins predicted to be an initiator of plasmid replication (labelled as Rep); a complete and a partial plasmid mobilisation protein (MobA_MobL and Mob (partial)); and toxin-antitoxin proteins, YafQ and RelB/DinJ.

A BLAST search of the protein 1137 amino acid sequence against the NCBI non-redundant database identifies the closest homologues as having 93% identity and 97% positive amino acids. The closest homologue in *Variovorax* E3 has 27.0% amino acid identity and 44.9% positives.

Protein 1137 is a horizontally transferred protein, the catalytic core of which is structurally similar to enzymes known to perform ether hydrolysis. The portion of the enzyme that is involved with substrate binding and selection is not conserved indicating that the substrate is something other than isochorismate. These properties make it a good candidate for MDP.

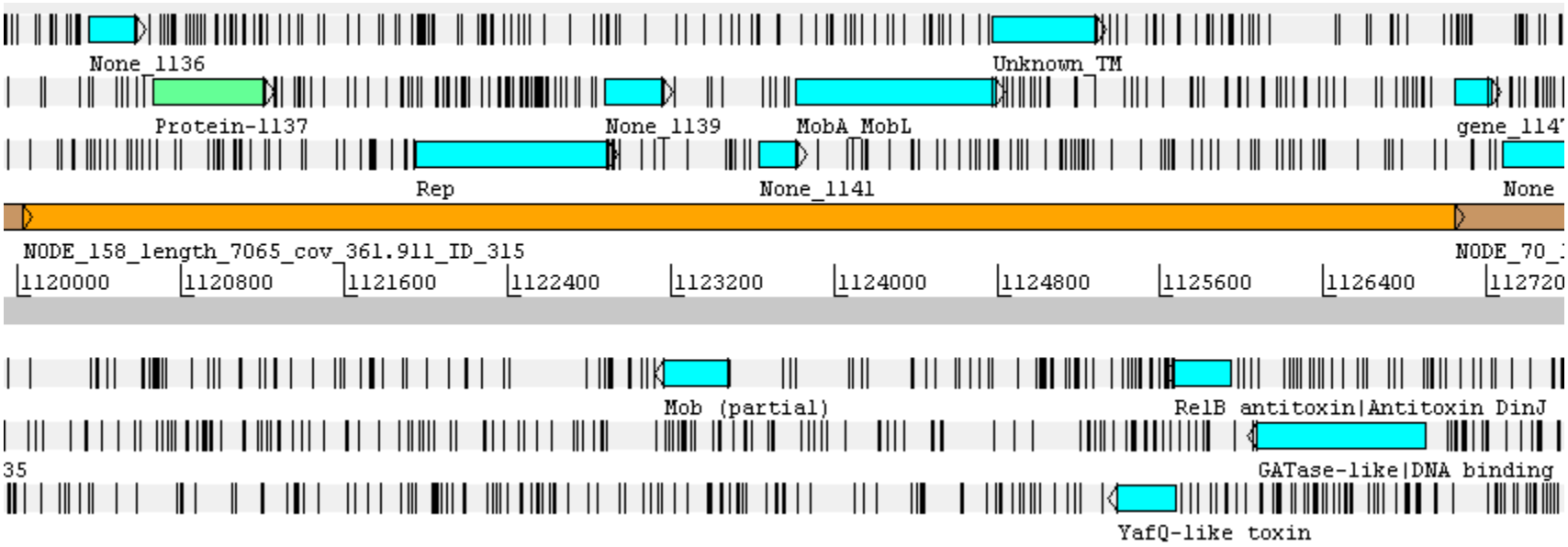


Figure 5-21. The full contig that contains protein 1137 in the A. calcoaceticus *E1 genome. Protein 1137 is shown in green. Genes identified by GeneMarkS are shown as boxes. 6 lanes show the 6 reading frames with vertical lines indicating stop codons. The scale gives base pairs.*

Proteins 1228, 2135 and 4392 – Esterases

Proteins 1228, 2135 and 4392 have SGNH hydrolase-type esterase (IPR013830) domain signatures, detected by InterPro. The same segment of protein 1228 is also identified as a GDSL lipase/esterase domain (IPR001087). These domains were identified in the C-terminal portion of proteins 1228 and 2135 (Figure 5-22). Only the N-terminal region of protein 4392 was available as it is truncated by the contig end.

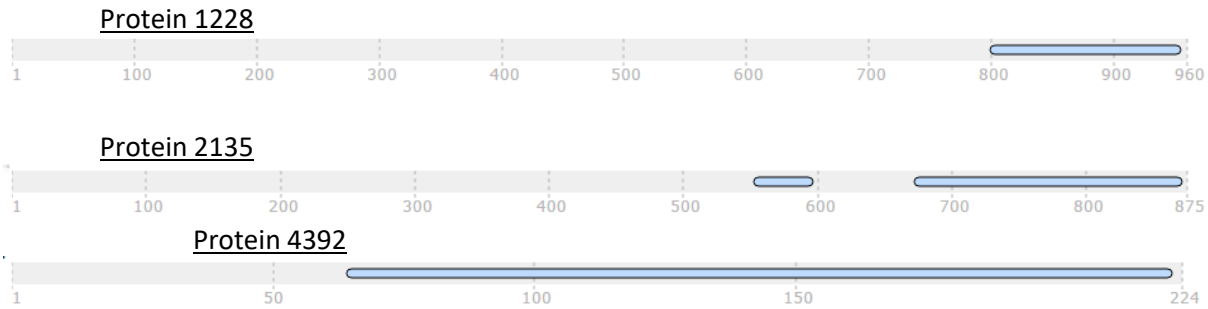


Figure 5-22. The location of domains detected by InterProScan in three proteins, shown as bars. The scale gives amino acid residues.

Esterases may hydrolyse an oxygen that is bound to two carbons, though the ester group is not found in metaldehyde. The catalytic residues of SGNH esterases are found in a groove in the protein (Wang et al., 2011; Baker et al., 2014; Wang et al., 2016) that may be accommodating to metaldehyde in overall shape (Figure 5-23). Related esterases have different patterns of electrostatic potential with similarly shaped catalytic cavities (Baker et al., 2014).

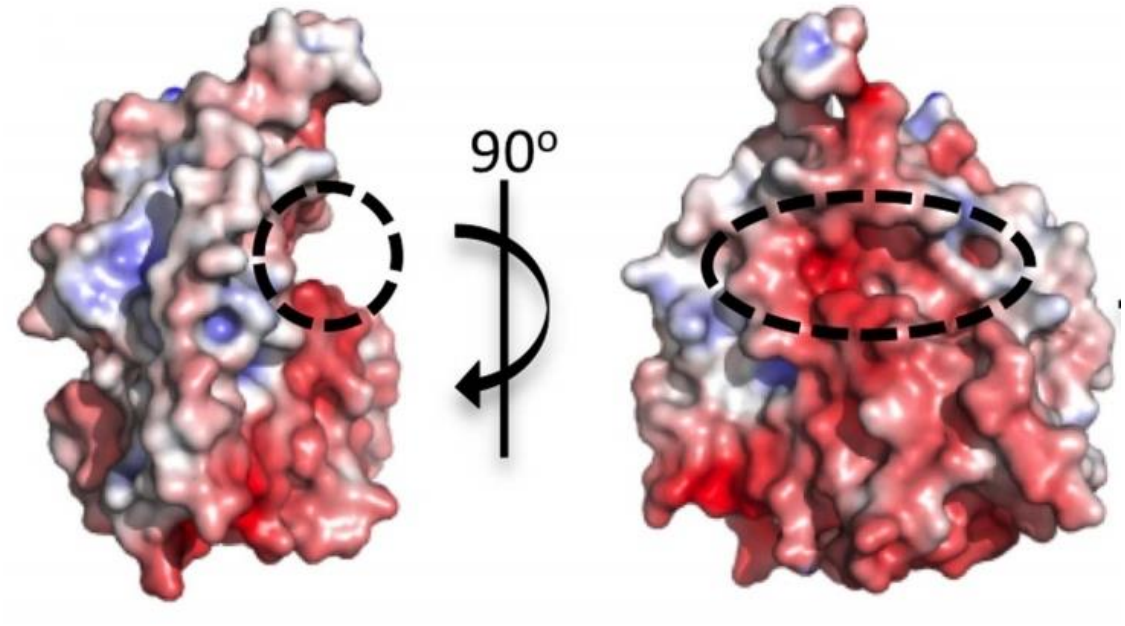


Figure 5-23. A rendering of the surface of the Pseudomonas putida *SGNH esterase AlgJ. The dashed lines indicate the substrate binding area. The surface is coloured red where the electrostatic potential is negative and blue where it is positive. Taken from Baker et al. (2014).*

Automated structure prediction using SWISS-MODEL did not result in a reliable prediction for any protein. The best QMEAN4 for computed structures were -12.07 for protein 1228 and -10.71 for protein 2135. As the sequence of protein 4392 is truncated by the contig end no model was attempted.

The gene for protein 1228 is found on a contig 74 kbp in length, in an apparent operon with 6 other genes, 5 of which are phage related. The genomic context of protein 2135 is similar; it is in an operon with 2 other genes, one of which is phage related. Genes immediately up and downstream of both 1228 and 2135 are also phage related. The contig containing gene 4392 is only 2.5 kbp long and contains 4 other genes, all phage related.

The probable function of these proteins is hard to predict. Most of the sequences are unannotated and the esterase domain predicted in protein 1228 is smaller than the smallest SGNH ester hydrolase domain in CathDB (Sillitoe et al., 2015), weakening the reliability of the prediction. Their structures are not very similar to any protein for which the structure is known. The fact that 3 proteins with similarly annotated domains occur in operons with several phage genes suggests that they are probably viral proteins of unknown function and not good candidates for MDP.

Protein 449 – Protease

Predicted families	Predicted domain
Clp protease proteolytic subunit /Translocation-enhancing protein TepA (IPR023562)	ClpP/crotonase-like domain (IPR029045)
ATP-dependent Clp protease proteolytic subunit (IPR001907)	

Table 5-11. Predicted domain and family membership of protein 449.

The sequence of protein 449 was classified as an endo peptidase (Table 5-11). As this would involve the lysis of a carbon-nitrogen bond, this enzyme was investigated further. Structure prediction using SWISS-MODEL selected PDB structure 1TG6 (Kang et al., 2004) as the template a structure was computed (Figure 5-24).

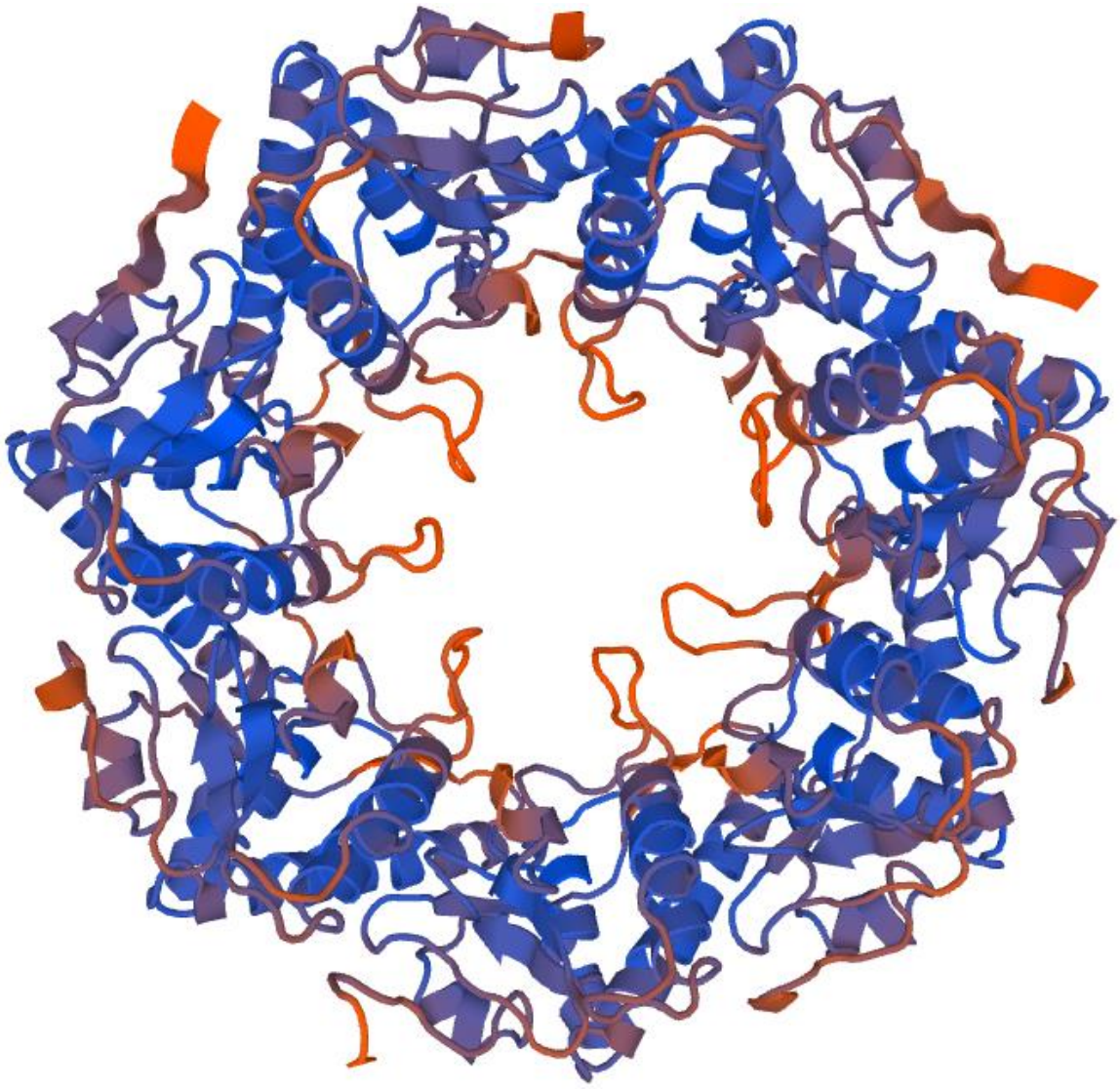


Figure 5-24. Computed structure of protein 449, as a heptamer. The colours show the degree of similarity between the computed structure and the template, blue is high and orange is low.

The protein for which 1TG6 is the structure is a human mitochondrial serine protease, ClpP. An ATPase domain forms a similar ring to that shown in Figure 5-24 and the two rings together form a barrel. The ATPase domain is responsible for unfolding a protein and feeding the polypeptide into the barrel, and proteolysis is catalysed by residues in the middle of the ring shown in Figure 5-24 (Kang et al., 2004).

The contig upon which protein 449 resides is 33 kbp long and exclusively contains phage genes, DNA manipulating genes and genes with no annotation. Serine proteases are known to be expressed by phages (Cheng et al., 2004).

The mechanism by which ClpP-like proteases operate, a linear peptide delivered to the catalytic region by a protein unfolding domain, seems an unlikely origin for a metaldehyde degrading enzyme.

This, and the fact that protein 449 appears to be a phage protein, make it an unlikely candidate for MDP.

Proteins 763, 3060 and 3973 – Cupins

Signatures associated with cupin domains were detected in 3 proteins (Table 5-12). No family membership was detected for these proteins.

Protein	Predicted domains
763	NAD(P)-binding domain (IPR016040) RmlC-like cupin domain (IPR011051) RmlC-like jelly roll fold (IPR014710) NAD-dependent epimerase/dehydratase, N-terminal domain (IPR001509)
3060	RmlC-like cupin domain (IPR011051) RmlC-like jelly roll fold (IPR014710)
3973	Cupin 2, conserved barrel (IPR013096) RmlC-like cupin domain (IPR011051) RmlC-like jelly roll fold (IPR014710)

Table 5-12. Predicted domains of some putative proteins from the A. calcoaceticus *E1 genome.*

Cupin domains are named after their barrel structure, which is formed from a series of β-sheets. The catalytic area of enzymes that have this domain is inside the cavity formed, and contains a metal ion. There is a great deal of structural and catalytic diversity in cupin enzymes, though most are dioxygenases that act on carbon-carbon bonds (Dunwell et al., 2004). Other reactions catalysed by cupins include the oxidation of various C-H groups, demethylation (Loenarz & Schofield, 2011), thiol dioxygenation (Stipanuk et al., 2011) and the lysing of a thioether (Todd et al., 2011). No cupin oxygen etherases were found in a review of the literature, but the domains are capable of a diverse set of reactions, and so proteins 763, 3060 and 3973 were examined in more detail.

The structures of the proteins were computed using SWISS-MODEL. The best QMEAN4 achieved for each protein is shown in Table 5-13 and the computed structures are shown in Figure 5-25.

Protein	QMEAN4	AA ident.	Template
3060	-6.03	0.28	1YHF.1.A
3973	-4.68	0.29	1YFX.1.A
764	-3.88	0.42	2ZKL.1.A

Table 5-13. QMEAN4 of computed structures for the numbered putative proteins from the A. calcoaceticus E1 genome, using the specified templates. The amino acid identity (AA ident.) is given of the protein aligned to the template.

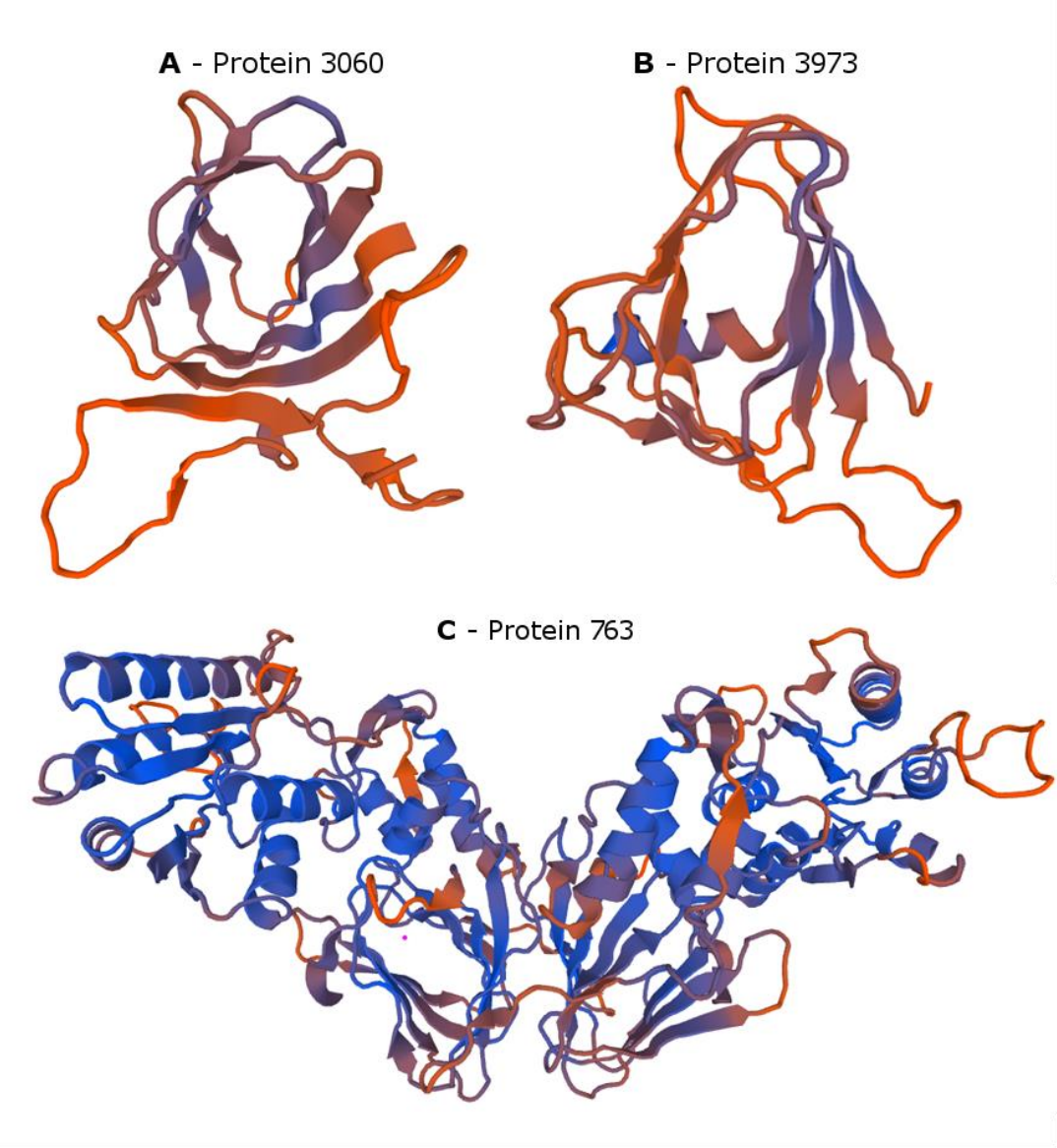


Figure 5-25. Computed ribbon models of the numbered putative proteins from the A. calcoaceticus *E1 genome. The colours indicate the degree of similarity between the two structures, blue is high and orange is low.*

The similarity of proteins 3060 and 3973 to their templates does not support a functional comparison to the template. Protein 3060 is located in an operon with 2 viral proteins. Protein 3973 is in an operon with two genes with no predicted function.

The template for protein 763 is *Staphylococcus aureus* CapF, which is involved in the biosynthesis of capsular polysaccharides (Miyafusa et al., 2012). The cupin domain of the *S. aureus* protein is an epimerase and the other domain acts as a reductase, catalysing the reactions shown in Figure 5-26.

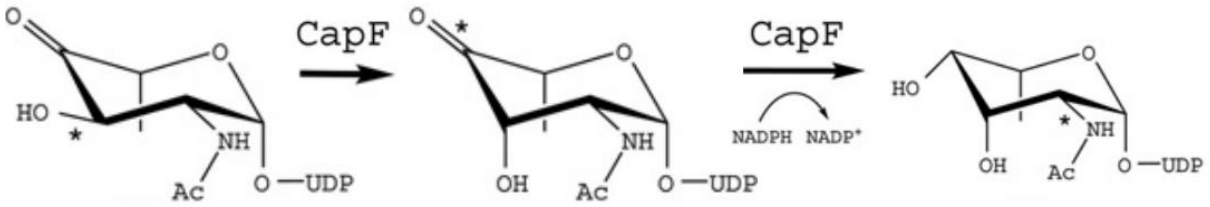


Figure 5-26. Reactions catalysed by the protein CapF. Adapted from Miyafusa et al. (2012).

The gene for protein 763 is in an operon with genes with classifications that are consistent with the biosynthesis of polysaccharides (Table 5-14).

Given the structural similarity of protein 763 (Figure 5-25C) to a protein of known function, the substrate of which is dissimilar to metaldehyde, and the gene's position in an operon, it is possible to rule out protein 763 as a good candidate for MDP. Proteins 3060 and 3973 are more difficult to judge, there is no particular reason to expect these proteins to act as metaldehyde degrading enzymes, but they also cannot be ruled out.

Gene No.	Predicted domains	Predicted family membership
758	Pyridoxal phosphate-dependent transferase, major region, subdomain 2	DegT/DnrJ/EryC1/StrS aminotransferase
	Pyridoxal phosphate-dependent transferase	
	Pyridoxal phosphate-dependent transferase, major region, subdomain 1	
759	Pre-ATP-grasp domain	Sialic acid O-acyltransferase NeuD-like
	Trimeric LpxA-like	
760	Bacterial sugar transferase	
761	Glycosyltransferase subfamily 4-like, N-terminal domain	
	Glycosyl transferase, family 1	
762	UDP-N-acetylglucosamine 2-epimerase domain	UDP-N-acetylglucosamine 2-epimerase WecB-like
763	RmlC-like jelly roll fold	
	RmlC-like cupin domain	
	NAD-dependent epimerase/dehydratase, N-terminal domain	
	NAD(P)-binding domain	
764	Polysaccharide biosynthesis protein, CapD-like domain	
	NAD(P)-binding domain	
	UDP-glucose 4-epimerase CapD, C-terminal domain	

Table 5-14. Predicted domains and protein family of genes in the same operon as gene 763.

Protein 3185 – Patatin/phospholipase

Protein 213 has a domain that was annotated as patatin/phospholipase A2-related (IPR016035) and acyl transferase/acyl hydrolase/lysophospholipase (IPR002641). Proteins with the IPR016035 domain includes lysophospholipases that act on ester bonds, but the substrate binding surface of these enzymes is large and the portion of the substrate bound is unlike metaldehyde (Figure 5-27). For this reason, protein 3185 was not judged a good candidate for MDP.

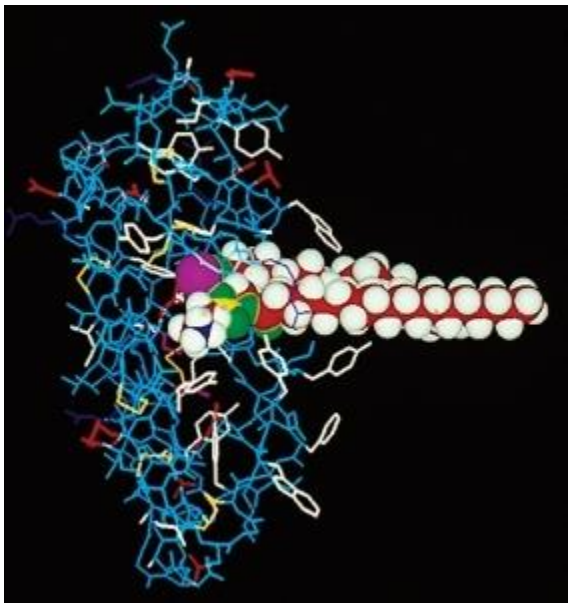


Figure 5-27. The phospholipase A2 shown as sticks with a phospholipid substrate shown as a space filling model in the active site. Taken from Burke & Dennis (2009).

Protein 3581 – leucyl aminopeptidase

The domain signature detected in protein 3581 was a "Peptidase M17, leucyl aminopeptidase, N-terminal" (IPR008283). Proteins with this domain typically cleave the N-terminal amino acid from polypeptides (Burley et al., 1990) but can also act on amino acid methyl esters (Arima et al., 2005). The best scoring SWISS-MODEL produced had a QMEAN4 of -5.69 using PDB 3H8G.1.A (Kale et al., 2010) as the template. The template protein was 497 amino acids long, and putative protein 3581 is 204 residues. Comparison of the computed structure to the template structure revealed that the portion of the enzyme that coordinates Zn^{+} ions, that are required for catalysis (Burley et al., 1990), were not present in protein 3581. The protein was not studied further.

Others proteins

A series of unique genes that are present in the *A. calcoaceticus* E1 genome are presented in Table 5-15. These genes were not considered likely candidates for MDP based on their predicted substrates

and/or the type of reactions they catalysed. However, these putative proteins have not been examined in detail and have not been ruled out as candidates for MDP.

No.	Domain	Family
244	Xylose isomerase-like, TIM barrel domain	
371	Domain of unknown function DUF4325	
	Histidine kinase-like ATPase, C-terminal domain	
648	FAD/NAD(P)-binding domain	
718	Fatty acid desaturase domain	
762	UDP-N-acetylglucosamine 2-epimerase domain	UDP-N-acetylglucosamine 2-epimerase WecB-like
861	Peptidase S24/S26A/S26B	
	Peptidase S24/S26A/S26B/S26C	
	Peptidase S24/S26, beta-ribbon domain	
2904	Reductase, C-terminal	
	Pyridine nucleotide-disulphide oxidoreductase, dimerisation domain	
	FAD/NAD(P)-binding domain	
	FAD/NAD-linked reductase, dimerisation domain	
2905	Extradiol ring-cleavage dioxygenase, class III enzyme, subunit B	
2906	Rieske [2Fe-2S] iron-sulphur domain	
2908	NTF2-like domain	Ring-hydroxylating dioxygenase beta subunit
2909	Rieske [2Fe-2S] iron-sulphur domain	
	Aromatic-ring-hydroxylating dioxygenase, alpha subunit	
	Aromatic-ring-hydroxylating dioxygenase, alpha subunit, C-terminal domain	
2911	Alpha/beta hydrolase fold-1	
	Alpha/Beta hydrolase fold	
2913	Pyruvate carboxyltransferase	4-hydroxy-2-oxovalerate aldolase
	Aldolase-type TIM barrel	
	DmpG-like communication	
2915	Fumarylacetoacetase, C-terminal-related	
3908	Luciferase-like domain	
3911	FAD/NAD(P)-binding domain	
	FAD-binding domain	
3976	NAD(P)-binding domain	Acrylyl-CoA reductase AcuI
	GroES-like 141 180 226	Alcohol dehydrogenase superfamily, zinc-type
	Polyketide synthase, enoylreductase domain	
	Alcohol dehydrogenase, N-terminal	
	Alcohol dehydrogenase, C-terminal	
4193	Cell wall hydrolase, SleB	
4288	Nucleoside phosphorylase domain	Nucleoside phosphorylase

Table 5-15. Unique genes in the A. calcoaceticus *E1 genome not examined in detail.*

The nearly sequential numbering of genes 2904-2915 is due to their proximity in the *A. calcoaceticus* E1 genome. The nucleotide sequence of the genomic region containing genes 2904-2915 was queried against the against the NCBI non-redundant database. The top results are shown in Table 5-16.

Description	Max score	Total score	Query cover	E value	Ident	Accession
Acinetobacter baumannii genome assembly R2091, chromosome : I	18119	18616	100%	0.0	93%	LN997846.1
Acinetobacter baumannii genome assembly CIP70.10, chromosome :	18119	18616	100%	0.0	93%	LN865143.1
Acinetobacter pittii strain IEC338SC, complete genome	16539	17057	99%	0.0	90%	CP015145.1
Acinetobacter sp. NCu2D-2, complete genome	3238	6670	91%	0.0	72%	CP015594.1
Acinetobacter equi strain 114, complete genome	3225	7303	86%	0.0	74%	CP012808.1
Pseudoalteromonas translucida KMM 520 chromosome I, complete se	913	1186	20%	0.0	71%	CP011034.1
Beggiatoa leptomitiformis strain D-402, complete genome	729	729	15%	0.0	69%	CP012373.1
Photorhabdus temperata subsp. thracensis strain DSM 15199, comp	574	796	26%	1e-158	66%	CP011104.1
Photorhabdus luminescens subsp. laumondii TTO1 complete genome	567	685	24%	2e-156	65%	BX571866.1
Acinetobacter sp. TTH0-4, complete genome	551	850	11%	2e-151	78%	CP012608.1
Vibrio natriegens NBRC 15636 = ATCC 14048 = DSM 759 chromoso	547	917	26%	2e-150	70%	CP016345.1
Photorhabdus asymbiotica ATCC43949 complete genome	547	947	26%	2e-150	71%	FM162591.1
Shigella sonnei strain FDAARGOS_90, complete genome	493	903	18%	4e-134	69%	CP014099.1
Shigella sonnei strain FORC_011, complete genome	493	903	18%	4e-134	69%	CP010829.1
Escherichia coli PCN033, complete genome	493	891	17%	4e-134	69%	CP006632.1

Table 5-16. Highest scoring BLAST hits from the querying of A. calcoaceticus E1 genomic region containing genes 2904-2915 against the NCBI non-redundant database.

The top results are essentially the same strain of *A. baumannii*; strain R2091 is the result of experimental mating of CIP 70.10 with another *A. baumannii* (Krahn et al., 2016). The last hit in Table 5-16 (to *E. coli* PCN033) is homologous to the region containing gene 2909 in *A. calcoaceticus* E1. This region was found in at least 40 other *E. coli* strains. This whole region is 12196 bp long and appears to have been horizontally transferred between several *Acinetobacter* species.

5.3.3 Transformation of *A. baylyi* 107474T with a plasmid containing gene 1137

Natural transformation was performed using two 1 ml cultures of *A. baylyi* grown overnight in MSM + 10 mM sodium acetate. Their OD_{600} was monitored over 2 hours to determine growth had stopped. Final OD_{600} values were between 0.42 and 0.46. 100 ng of plasmid DNA (pMAL-c4X + gene 1137) was added to one culture and the concentration of sodium acetate was increased by 10 mM in both cultures. The cultures were incubated for 3 hours at 30°C, OD_{600} of the plasmid treated culture had increased to 0.56 indicating growth had resumed. 100 µl from the both cultures were used to

inoculate LBA+100 µg/L ampicillin plates. These were incubated overnight at 30°C. Seven large colonies were present on the plate inoculated with plasmid treated culture along with ~200 barely visible colonies. No colonies were visible on the plasmid free plate. Both plates were incubated again overnight at 30°C. No colonies appeared on the plasmid free plate. All colonies continued to grow on the plasmid treated plates. The smaller colonies were uniformly distributed over the plate, and did not appear to be satellite colonies. Three of the large colonies were streaked onto LBA+amp plates and incubated overnight, then one colony from each of these streak plates was used to inoculate 1 ml MSM + 150 mg/L metaldehyde and MSMA+metaldehyde plates both with 100 µg/L ampicillin and 1 mM IPTG. No growth was seen on metaldehyde media. Three 1 ml MSM + metaldehyde + 100 µg/L ampicillin with 0, 0.5, 1 or 2 mM IPTG cultures were inoculated with an *A. baylyi* transformant colony. The cultures were incubated overnight at 30°C, and lysate in 1× SDS loading buffer was produced as described in the methods (Appendix 7.1.6). 10 µl samples of lysate from each culture was run on a 15% acrylamide SDS-PAGE gel and protein bands visualised using Coomassie stain (Figure 5-28). The predicted mass of protein 1137 is 19.5 kDa. No additional band of this mass is apparent in Figure 5-28 in the cultures treated with IPTG indicating that the protein is not being expressed.

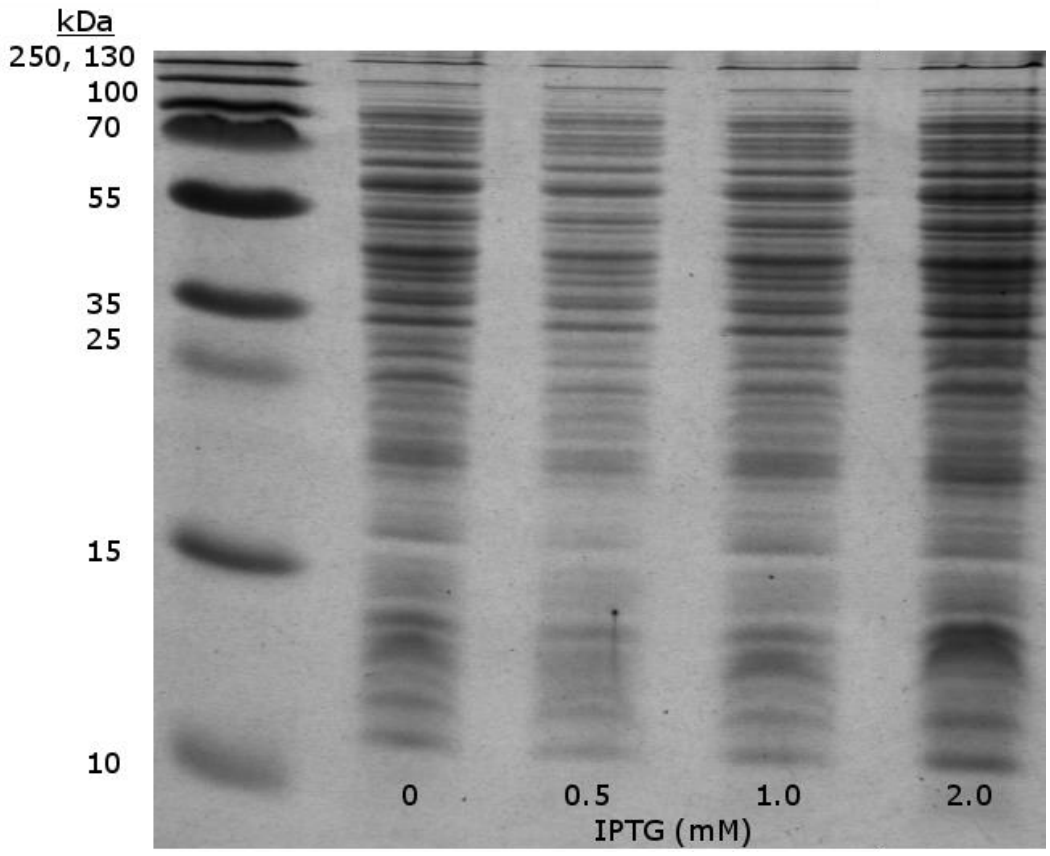


Figure 5-28 . 15% acrylamide SDS-PAGE gel of lysate of transformants grown with different concentrations of IPTG to induce the expression of protein 1137.

5.4 Discussion

In this chapter the genomes of the two metaldehyde degrading isolates, *A. calcoaceticus* E1 and *Variovorax* E3, were compared with each other and with *Acinetobacter* strains shown to not degrade metaldehyde in order to identify putative proteins that are good candidates for the primary metaldehyde degrading protein. Two experiments were conducted.

5.4.1 Detecting MDP candidates by intrageneric homology between the isolates

The first HGT scenario tested here was that the two isolates recently acquired the genetic material that enables growth on metaldehyde from the same source, or from one another. No candidate for MDP fit this scenario. Only the mercury resistance pathway presented in Section 5.3.1.2 appears to been recently shared between the two isolates. Other proteins and pathways discussed may represent homologous sequences that have a shared origin, however this shared origin may be far in the past and not the result of a recent shared environment.

It is not obvious how low a BSR can be and still be informative of a catabolic activity that is shared between the two proteins. The substitution of a single residue can result in a loss of function, a fact utilised in many protein mutagenesis experiments to identify catalytic residues, while residues that are not crucial to the protein's structure or directly involved in the interaction of the protein with other molecules may be more free to change.

Based on the reactions predicted to be catalysed by them, three enzymes (RutA, RutB and phenol hydroxylase) were thought to have some potential as candidates for MDP. *Acinetobacter* strains containing close homologues to these genes were not able to utilise metaldehyde as sole carbon source. This was not a definitive test for the *A. calcoaceticus* E1 proteins' hypothetical role as MDP as the equivalent genes may not be constitutive.

5.4.2 Detecting MDP candidates by lack of intrageneric homology in *Acinetobacter* strains

The second analysis (Section 5.3.2) tested the scenario that metaldehyde catabolism was performed by different enzymes in the two strains. This scenario is supported by the difference in affinity for metaldehyde between the two strains and the failure of the previous analysis to identify recently shared genes that are good candidates for MDP. Amino acid sequences that were unique to *A. calcoaceticus* E1 and were not found (BSR ≤ 0.4) in other *Acinetobacter* strains, incapable of metaldehyde catabolism, were examined. This successfully identified horizontally transferred material in the form of plasmid borne sequences, prophages and a genomically borne pathway that

has 93% nucleotide similarity to *A. baumannii* strain CIP 70.10, isolated in France, in 1970 (Bonnin et al., 2012).

A plasmid borne protein with a predicted function consistent with MDP was identified, the isochorismate hydrolase-like protein 1137. This protein is not present in *Variovorax* E3. The expression of the protein in *A. baylyi* CIP 10747T was attempted but does not appear to have been successful. Unfortunately, identifying the reason for this, or trying other hosts such as *A. calcoaceticus* CIP 110488, CIP 110439 or *E. coli* using the pET vector described in Section 5.2.6 was not possible in the time available. Natural transformation of *A. calcoaceticus* RUH 2202 was attempted, but it was found to not be a suitable host due to a low level of ampicillin resistance (data not shown).

When defining unique genes, the BSR value of ≤ 0.4 was chosen arbitrarily, following Rasko et al. (2005) who chose this value as the cut-off for conserved versus not conserved genes in their analysis. There is no *a priori* reason for assuming that horizontally transferred genes with novel functions would not have a BSR higher than this. The number of *A. calcoaceticus* E1 genes that have BSR less than different cut-off BSR values to all of the *Acinetobacter* strains given in Table 5-9 is shown in Figure 5-29. F. Horizontally transferred genes are likely to be found at higher BSR cut-off values.

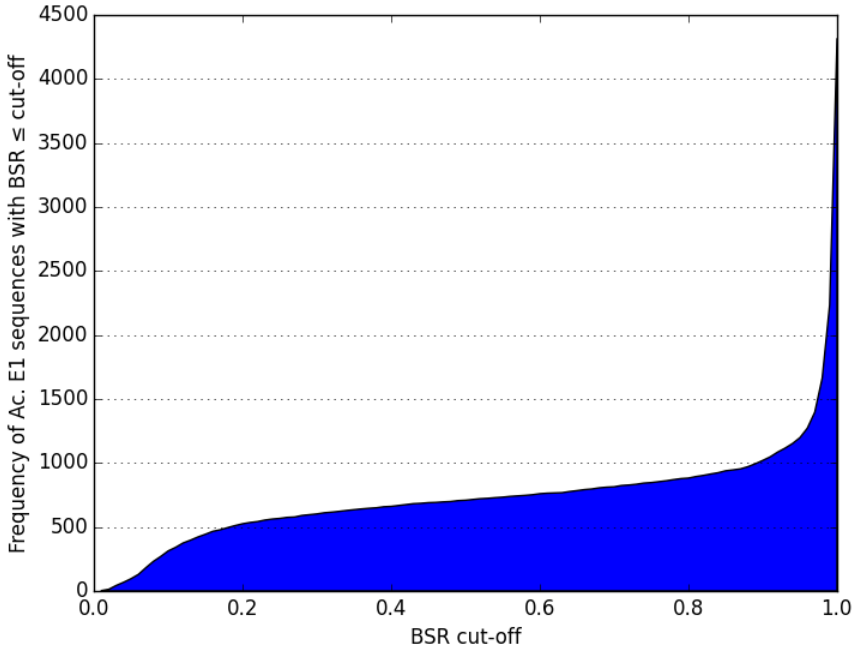


Figure 5-29. Frequency of predicted protein sequences in the A. calcoaceticus E1 genome that have BSR ≤ a range of cut-off values to all other Acinetobacter *tested.*

It is possible that the true MDP is undetectable by these methods. Should the reaction be catalysed by a domain that has not been investigated and does not have a signature in any protein database; if the protein does not contain domains recognised as catalysing a reaction similar to that which would be expected to degrade metaldehyde; or if MDP is very similar to proteins in strains that don't degrade metaldehyde (the ability in *A. calcoaceticus* E1 is the result in a change of enzyme regulation, for example) then MDP could never be identified by the methods used here.

As the cost of obtaining a bacterial genome continues to decrease, the library of prokaryotic genomes increases, and the number and diversity of characterised proteins increases the greater the potential this method has for detecting candidates for a metabolic activity of interest. Some enzymes would never be detected this way, but an automated pipeline could make the genomic comparison trivial to perform where the genomes are available.

Chapter 6: Final discussion

Prior to the start of this work it was known that metaldehyde is metabolised by soil microorganisms, producing acetaldehyde as an intermediate and carbon dioxide as the final product. Nothing was known about the identity or nature of the organisms performing the metabolism. In the work presented here, two domestic soils yielded bacterial isolates capable of utilising metaldehyde for growth. *Acinetobacter calcoaceticus* were isolated from both soils, while *Variovorax* was isolated from only the soil that had a history of metaldehyde treatment. Two isolated colonies (*A. calcoaceticus* E1 and *Variovorax* E3) were cultured from the soil with metaldehyde history.

To confirm that metaldehyde was being utilised for growth, and to quantify the degree to which the organisms could degrade metaldehyde required the application of liquid chromatography-mass spectrometry (LCMS). An improvement to the LCMS method used by the Food and Environment Research Agency to quantify metaldehyde was developed here. The alterations to solvation temperature and capillary voltage recommended in this work were adopted in their commercial and regulatory work. The lowered solvation temperature is likely to be an improvement over most previously published LCMS methods (Li et al., 2010; Zhang et al., 2011b).

Characterisation of the metabolic pathway by which metaldehyde is degraded was conducted in *A. calcoaceticus* E1. It showed the most promise for future work due to its faster growth rate, propensity to form distinct colonies (a convenient trait when performing genetic manipulations), and greater affinity for metaldehyde. Quantification by the improved LCMS method showed that the isolate is able to degrade metaldehyde to below the legal maximum concentration in drinking water of 0.1 µg/L. Evidence was also found that acetaldehyde is a metabolic intermediate in *A. calcoaceticus* E1; the metaldehyde dependent expression of an aldehyde dehydrogenase and consequent high maximum flux observed when oxidising acetaldehyde. This is in alignment with what was seen in microcosms using radioactively labelled metaldehyde. The acetaldehyde dependent reduction of β-NAD^+ by an acetaldehyde dehydrogenase is a possible method by which a metaldehyde detecting biosensor could operate, and so confirmation that this intermediate is produced is a confirmation of the potential future utility of the isolated *A. calcoaceticus* E1.

Exploitation of the metaldehyde degrading capacity of *A. calcoaceticus* E1 depends upon the identification of the primary metaldehyde degrading protein (MDP). Traditional genetics methods were attempted, but failed. Comparative genomics methods were developed based on hypotheses involving horizontal gene transfer. Initially these methods used absolute differences in the BLAST scores to identify genes more similar in the degrading strains. The BLAST score ratio (BSR) method was adopted after discovery of the paper describing it (Rasko et al., 2005). It was found that the two

metaldehyde degrading isolates do not appear to share any genetic material that could plausibly thought to degrade metaldehyde, indicating a separate origin for this ability in the isolates. One good candidate for MDP was found by characterising enzymes not found in *Acinetobacter* outside of *A. calcoaceticus* E1. Protein 1137 is predicted to have the conserved features of an ether hydrolysis, but a different substrate than similar proteins that have been previously studied. BSR has been used to identify genes associated with the development of pathogenicity in *A. baumannii*, but does not appear to have been used to identify genes with novel functions before. Should the isochorismatase-like Protein 1137 be found to be the primary metaldehyde degrading enzyme, then this would demonstrate this novel method has the potential to identify interesting candidate genes.

Investigations of bioremediation could be conducted using *A. calcoaceticus* E1 as is, perhaps with some artificial selection to improve its ability to remediate. Beyond this, the primary focus of any future work should be the identification of the primary enzyme that acts to depolymerise metaldehyde. To do this, it would be best to develop methods of genetic manipulation in *A. calcoaceticus* E1. The genetic plasticity of the *Acinetobacter* genus could be utilised to achieve this. The organism contains genes that are predicted to act as DNA transporters, and likely can be transformed by linear DNA. Constructs with homologous sequence to the gene of interest, interrupted by an antibiotic resistance gene, would be expected to cause the targeted disruption of genes by homologous recombination when transformed into the strain. This method could then be used to test the candidate genes identified by comparative genomics. Should this targeted method fail, the construction of a random knock-out library could be performed, as described by Metzgar *et al.* (2004). Once identified, the characterisation of the enzyme and development in a remediation technology, or in a biosensor could begin.

Appendix

7.1 General materials and methods

7.1.1 Media

7.1.1.1 Minimal media

Minimal salts media (MSM) were prepared using the concentration of salts given in Table 3-1 dissolved in ultrapure water. The salts solution was autoclaved. For solid plates 1.5 g agarose would be added to 200 ml MSM prior to autoclaving. Agarose was used in place of agar to eliminate carbohydrates from the plates that could be used for growth by organisms. 2 ml of a trace elements solution (Vishniac & Santer, 1957; Table 7-2) (Vishniac & Santer, 1957 Table 7-2) was added for each 1 litre of salts solution.

	Concentration (mM)
Na_2HPO_4	55
KH_2PO_4	11
NH_4Cl	6
$MgSO_4$	0.4

Table 7-1. Salts concentrations for minimal media. All manufactured by Fisher Scientific.

	Concentration (mM)
Na_2EDTA	140
$ZnSO_4$	7.6
$CaCl_2$	37
$MnCl_2$	25
$FeSO_4$	18
$(NH_4)Mo_7O_{24}$	0.9
$CuSO_4$	6.4
$CoCl_2$	6.7

Table 7-2. Concentration of compounds in trace elements solution.

Up to 150 mg/l (0.851 mM) metaldehyde was added to liquid media in powder form and dissolved overnight on a magnetic stirrer. When used in solid media molten MSM-agarose would be cooled in a water bath to 50°C and 100 mg metaldehyde would be added to 200 ml molten agarose. Agarose plates with metaldehyde added will be referred to as “metaldehyde plates” those with no added carbon substrate will be referred to as “no-carbon plates”.

Sodium acetate, when used, was dissolved in DI water at 1 M concentration, filter sterilised, and added to MSM to achieve the desired concentration, typically 10 mM.

7.1.1.2 Lysogeny broth (LB) and LB agar

This medium was made using 5 g yeast extract (Formedium, Hunstanton, UK), 10 g granulated triptone (Melford), and 10 g sodium chloride (Sigma-Aldrich) per 1 litre of deionised water. For solid media 3 g of granulated agar (bacteriological grade; Formedium) was added to 200 ml LB in 250 ml flasks and autoclaved. Molten agar may be cooled to 50°C before the addition of antibiotics, or used without antibiotics, and poured into 20 cm plates and allowed to set and dry in a laminar flow hood for 15 minutes.

7.1.2 Glycerol stocks

A single colony was streaked onto a media plate and incubated with antibiotics and media appropriate for the strain. Growth from the streak plate was collected using a sterile loop and suspended in 1 ml of filter sterilised 1:1 LB:glycerol (Sigma-Aldrich) mixture in a 1.5 ml tube. The mixture was flash frozen in liquid nitrogen and stored at -80°C. Stocks would be tested several days later by using a heated metal loop to melt some of the stock and streak it onto the appropriate solid medium. Plate stocks were made using the same heated loop method.

7.1.3 Polymerase chain reaction

The polymerase chain reaction (PCR) was used for amplifying specific sections of DNA from template DNA. A Px2 Thermal Cycler (Thermo Electron, MA, USA) was used. Two general PCR preparations were used, a GoTaq (Promega) preparation for low fidelity reactions and a Phusion High-Fidelity (Thermo Scientific) preparation was used when proofreading was necessary. GoTaq reactions were made up of reagents given in **Error! Reference source not found.** and Phusion reactions reagents are given in **Error! Reference source not found.**. Total reaction volumes were 20 or 50 µl.

Reagent	Final concentration
5X GoTaq Green Buffer	1X
dNTPs	0.2 mM each
Forward and reverse primers	0.5 µM each
GoTaq DNA polymerase	0.025 U/µl

Table 7-3. Final concentrations of reagents used in GoTaq PCR reactions.

Reagent	Final concentration
5X Phusion HF Buffer	1X
dNTPs	0.2 mM each

Forward and reverse primers	0.5 µM each
Phusion DNA polymerase	0.02 U/µl

Table 7-4. Final concentrations of reagents used in Phusion PCR reactions.

Specific thermocycle programs used in specific experiments are given in the text.

Where colony PCR was performed, template DNA was delivered to the reaction mixture by lightly inserting a small sterile pipette tip into a colony and immersing the tip into the reaction mixture.

7.1.3.1 Primers used for 16S amplicons

8F	AGA GTT TGA TCC TGG CTC AG
785R	GGA TTA GAT ACC CTG GTA GTC C
U1492R	GGT TAC CTT GTT ACG ACT T

Oligos purchased from Sigma-Aldrich, suspended to 100 mM in ultrapure water, 10 mM working stocks prepared by dilution with ultrapure water.

7.1.4 Gel electrophoresis of DNA

A stock of 10 × TBE buffer was prepared using 110 g tris(hydroxymethyl)aminomethane (Ultrapure Tris; Invitrogen), 55 g boric acid (Fisher), 12 g ethylenediaminetetraacetic acid (EDTA; Fisher) per litre of deionised water. 1 × working solution prepared by dilution with deionised water.

Agarose gels for electrophoresis of DNA were typically produced using 40 ml 1 × TBE buffer with 1.2% (w/v) agarose (Melford) was added. The mixture was brought to boiling in a microwave to dissolve the agarose and then cooled at room temperature for 15 minutes. 4 µl SYBRsafe (Invitrogen) DNA stain was added and mixed by manual rotation. The mixture was left to set in a mould, including well moulds, for 1 hour. Gels were submerged in 1 × TBE in the electrophoresis chamber. If required samples to be analysed were treated with 6 × loading dye (NEB). Typically, 5-20 µl of samples were added and the gel run at 100 V for 1 hour before being imaged using a GeneGenius Bio Imaging System (Syngene). Alterations to the overall brightness, contrast and gamma of the image were sometimes made using GeneSnap V6.00.19.

7.1.5 Restriction digests

Restriction digests were performed using New England Biolabs (NEB) buffers and enzyme preparations following the manufacturer's instructions unless otherwise stated in the results.

7.1.6 sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

Resolving gel mixtures were prepared with 2.4 ml dH_2O, 2.4 ml Tris pH 8.8, 0.1 ml 10% (w/v) SDS, 5 ml acrylamide (stock 30% acrylamide, 0.8% bis-acrylamide (w/v); Protogel, National Diagnostics – Atlanta, USA), 50 µl 10% (w/v) ammonium persulphate (APS; Sigma-Aldrich) and 8 µl tetramethylethylenediamine (TEMED). Stacking gel mixtures were prepared with 3.2 ml dH_2O, 1.25 ml 0.5 M Tris pH 6.8, 50 µl 10% SDS, 25 µl 10% APS and 8 µl TEMED.

2x Loading sample buffer was prepared using 100 mM Tris-Cl (pH 6.8), 4% (w/v) SDS, 0.2% (w/v) bromophenol blue (Sigma-Aldrich), 20% (v/v) glycerol and 200 mM β-mercaptoethanol (Sigma-Aldrich).

Cell lysate was prepared by dilution 1:1 with 2X sample loading buffer, followed by 10 minutes heating at 98°C on a heating block (Grant, Shepreth, UK).

7.1.7 Basic Local Search Alignment Tool (BLAST) methods

BLAST (Camacho et al., 2009) hosted by National Center for Biotechnology Information (NCBI) which is accessed using the URL http://blast.ncbi.nlm.nih.gov/Blast.cgi, was used to find homology with DNA and protein sequences in NCBI databases. Protein and DNA sequences were uploaded to query the non-redundant protein sequences database using blastp and the nucleotide collection database using megablast respectively, both using default settings unless otherwise stated.

Locally run BLAST searches of specific subject sequences (*e.g*. a FastA file containing genomic contigs) used BLAST+ V2.2.28 installed on a local personal computer.

BLAST results were typically saved or downloaded in XML format and parsed using Biopython v1.62 or higher.

7.2 Fits of oxygen utilisation data

Oxygen utilisation curves produced by *A. calcoaceticus* E1 cells grown with metaldehyde, washed and resuspended to OD_{600} 1.0 in 100 mM pH 7 phosphate buffer that have been treated with metaldehyde or acetaldehyde. Curves have been constructed by fitting Model B – a 3 reaction model, the first has Michaelis-Menten kinetics, the second linear kinetics and the 3rd is the oxidation factor; see Section 4.2.6.

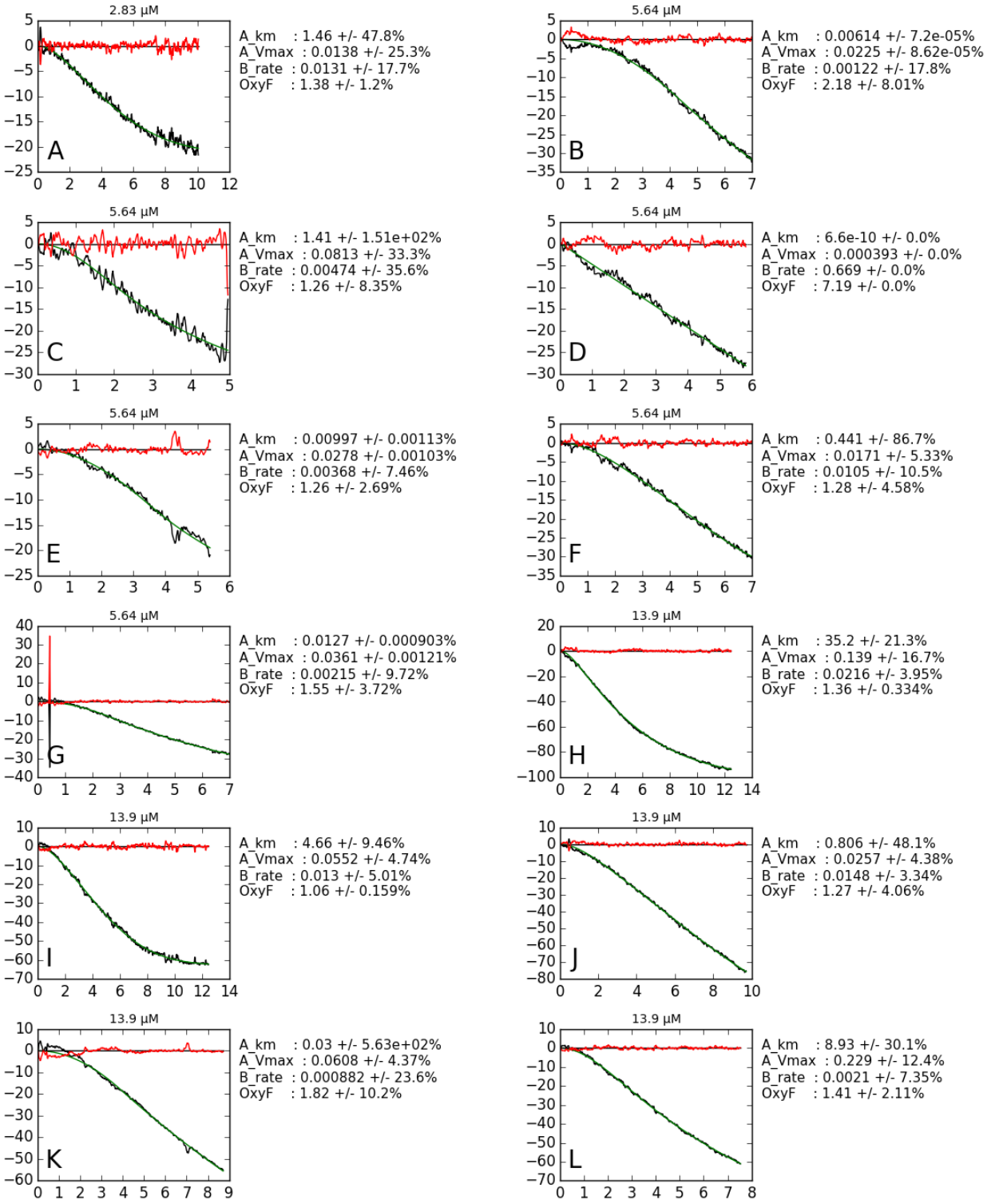


Figure 7-1. (Continued next page)

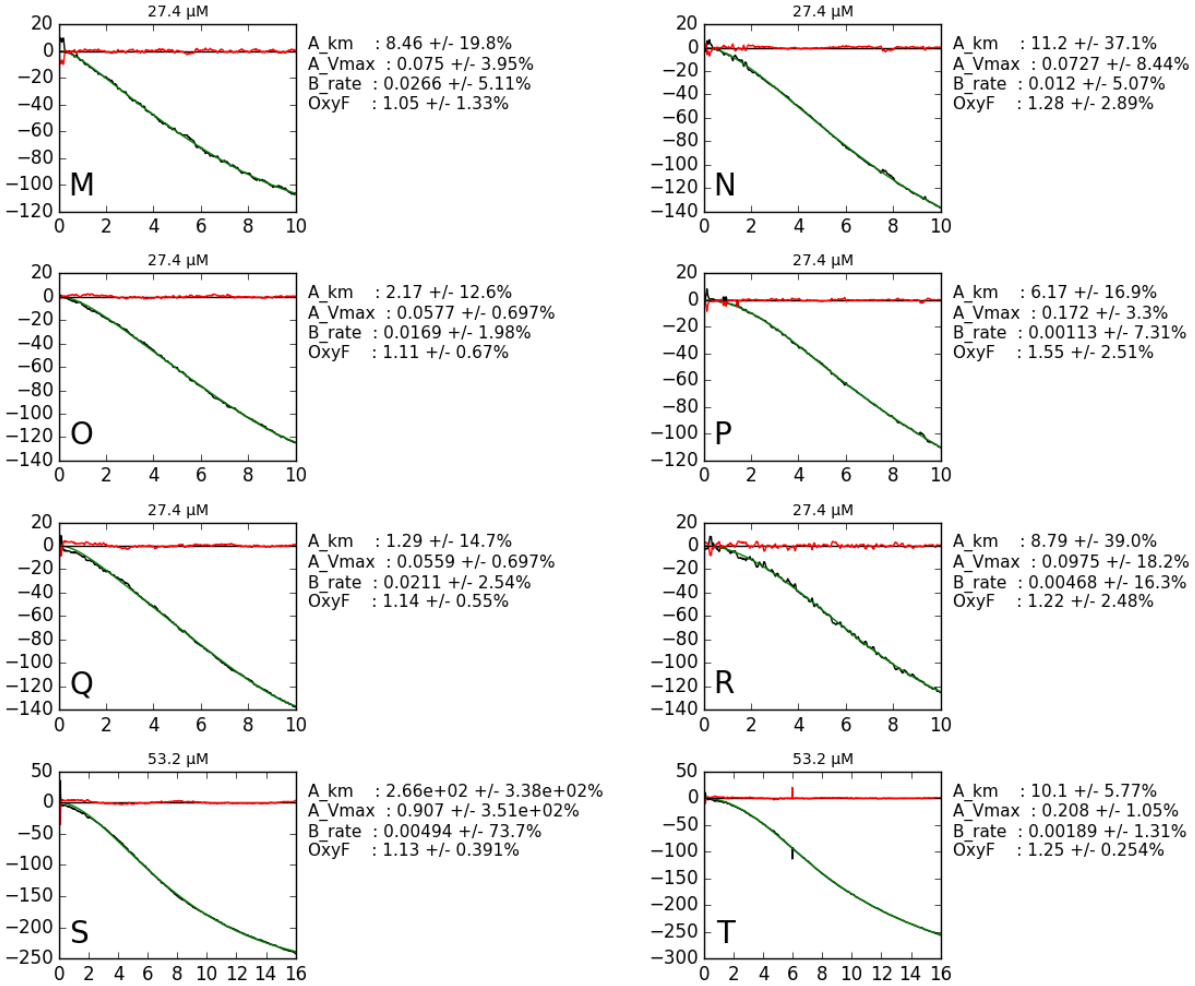


Figure 7-1. Oxygen utilisation data, first presented in Figure 4-10, of A. calcoaceticus E1 treated with metaldehyde (black curves); curves constructed by fitting Model B (green); residuals shown in red. Fitted parameter values are shown to the right of each chart with calculated standard deviations expressed as a percentage.

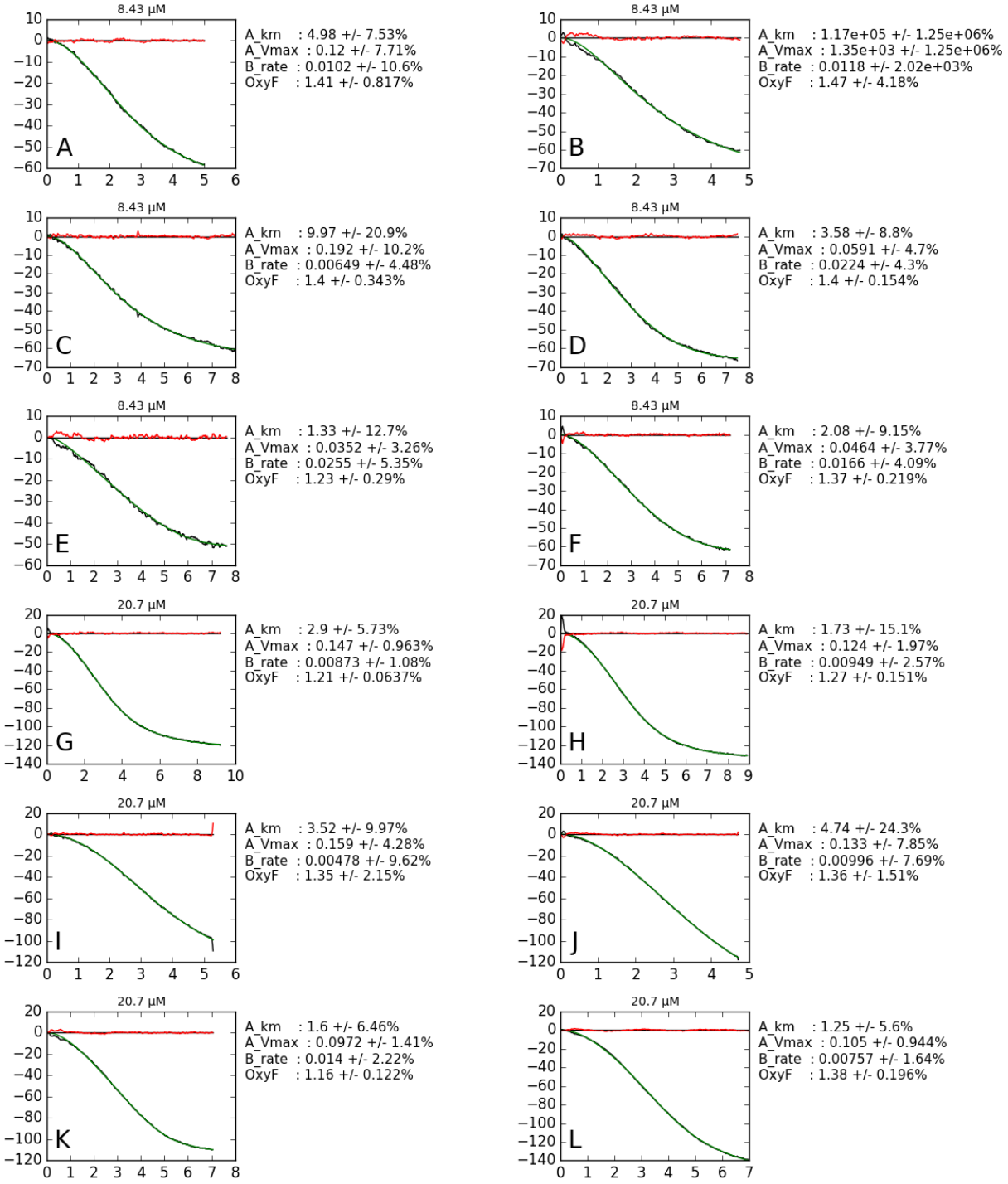


Figure 7-2. (Continued next page)

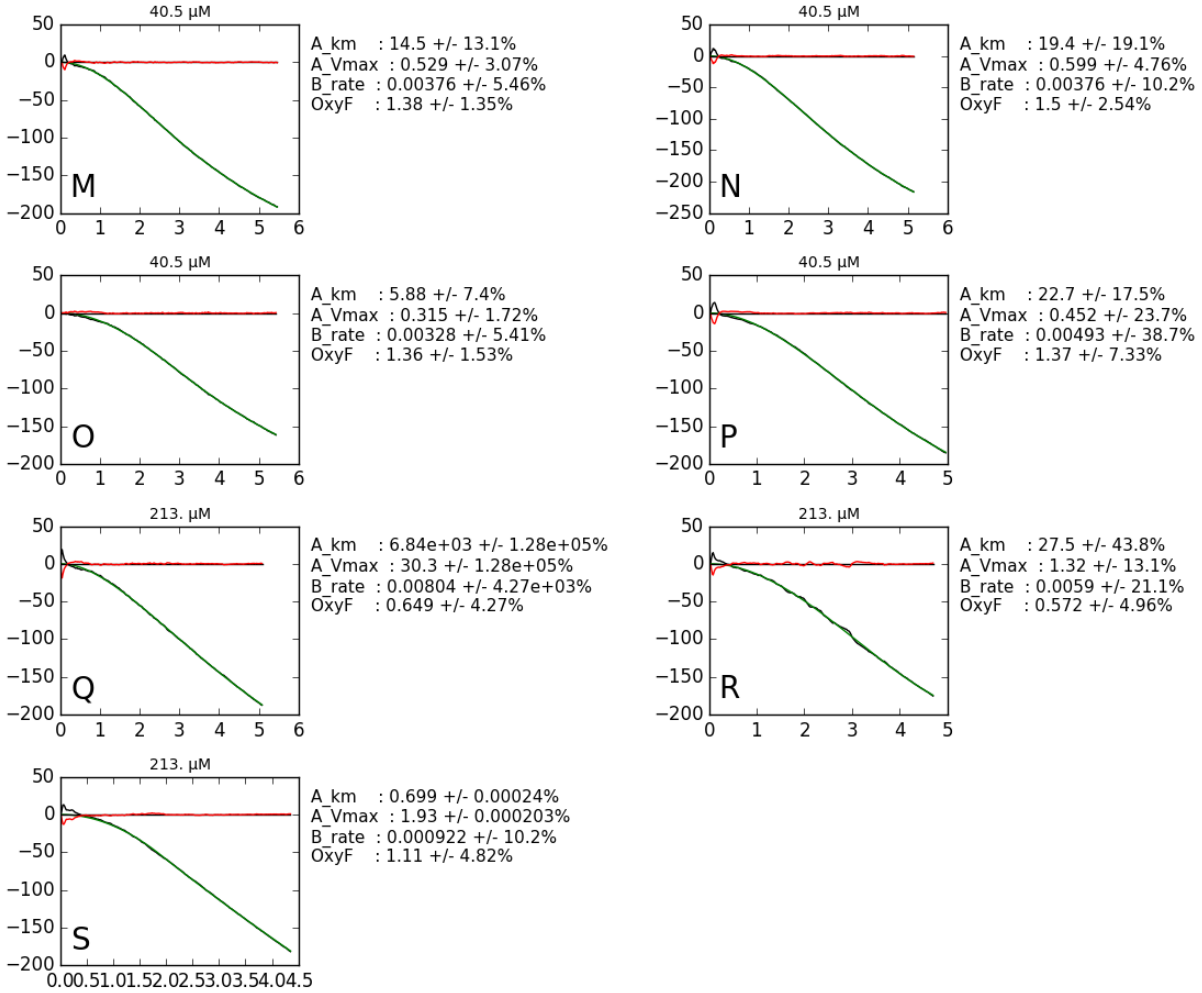


Figure 7-2. Oxygen utilisation data, first presented in Figure 4-11, of A. calcoaceticus E1 treated with metaldehyde (black curves); curves constructed by fitting Model B (green); residuals shown in red. Fitted parameter values are shown to the right of each chart with calculated standard deviations expressed as a percentage.

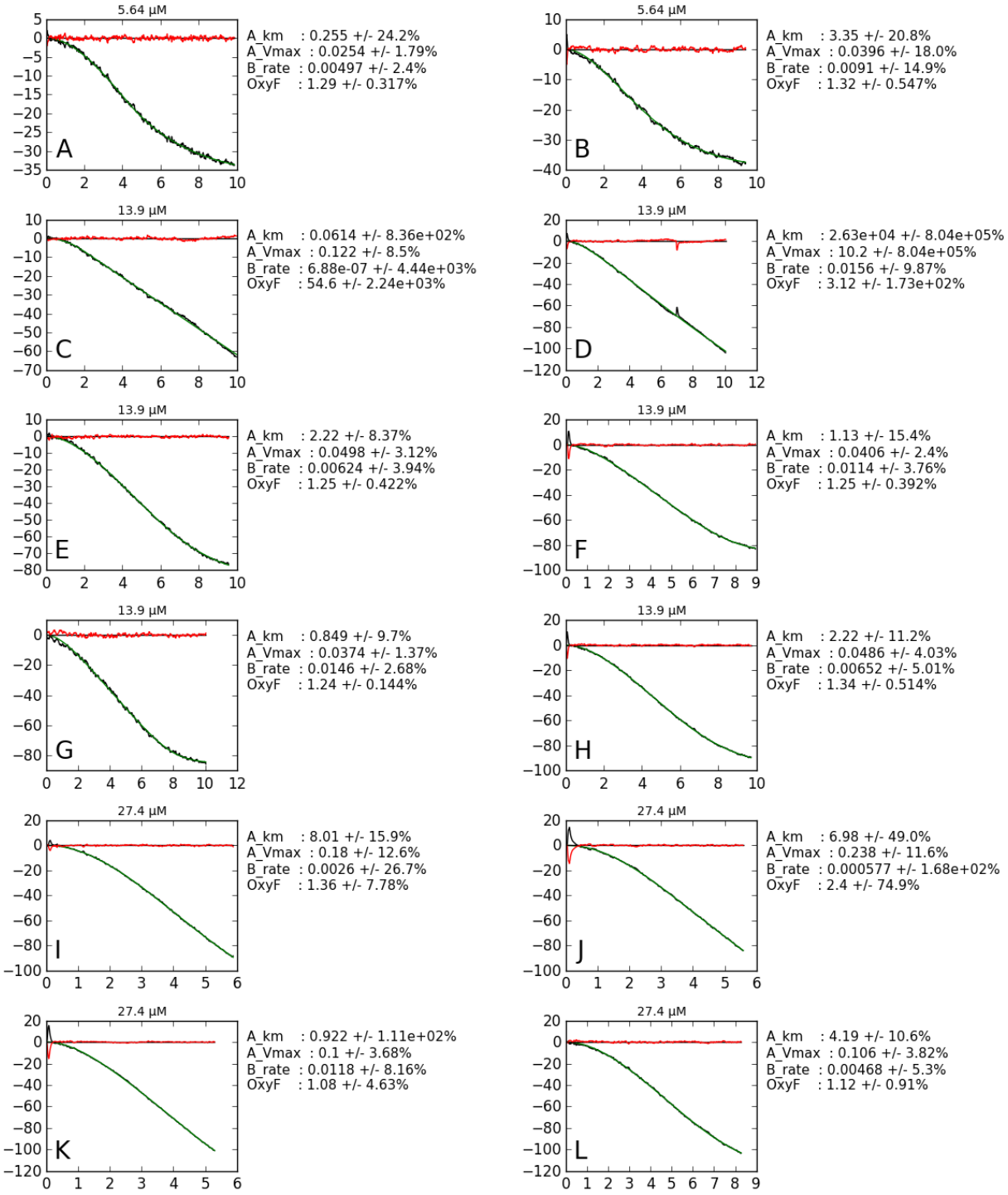


Figure 7-3. (Continued next page).

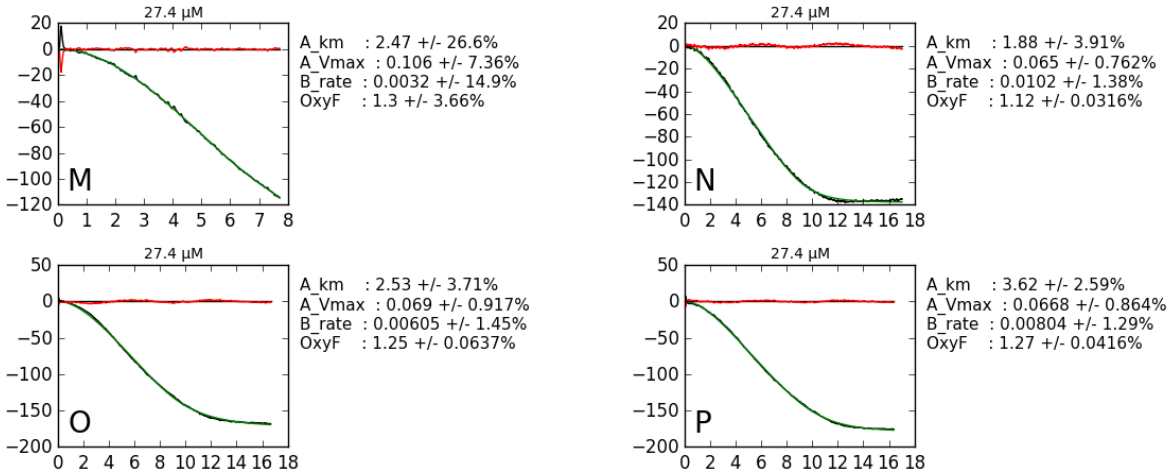


Figure 7-3. Oxygen utilisation data, first presented in Figure 4-12, of A. calcoaceticus E1 treated with metaldehyde (black curves); curves constructed by fitting Model B (green); residuals shown in red. Fitted parameter values are shown to the right of each chart with calculated standard deviations expressed as a percentage.

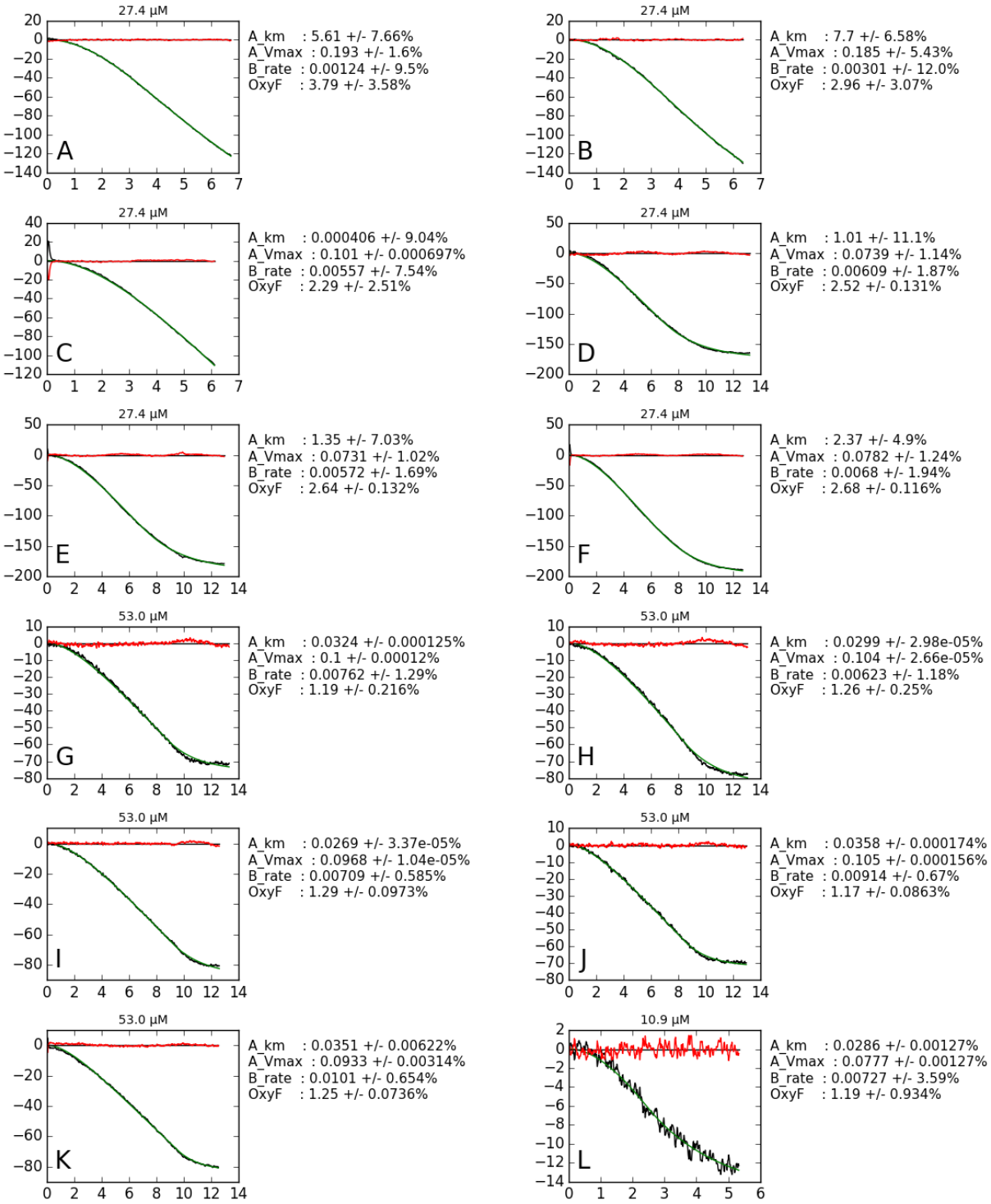


Figure 7-4. (Continued next page)

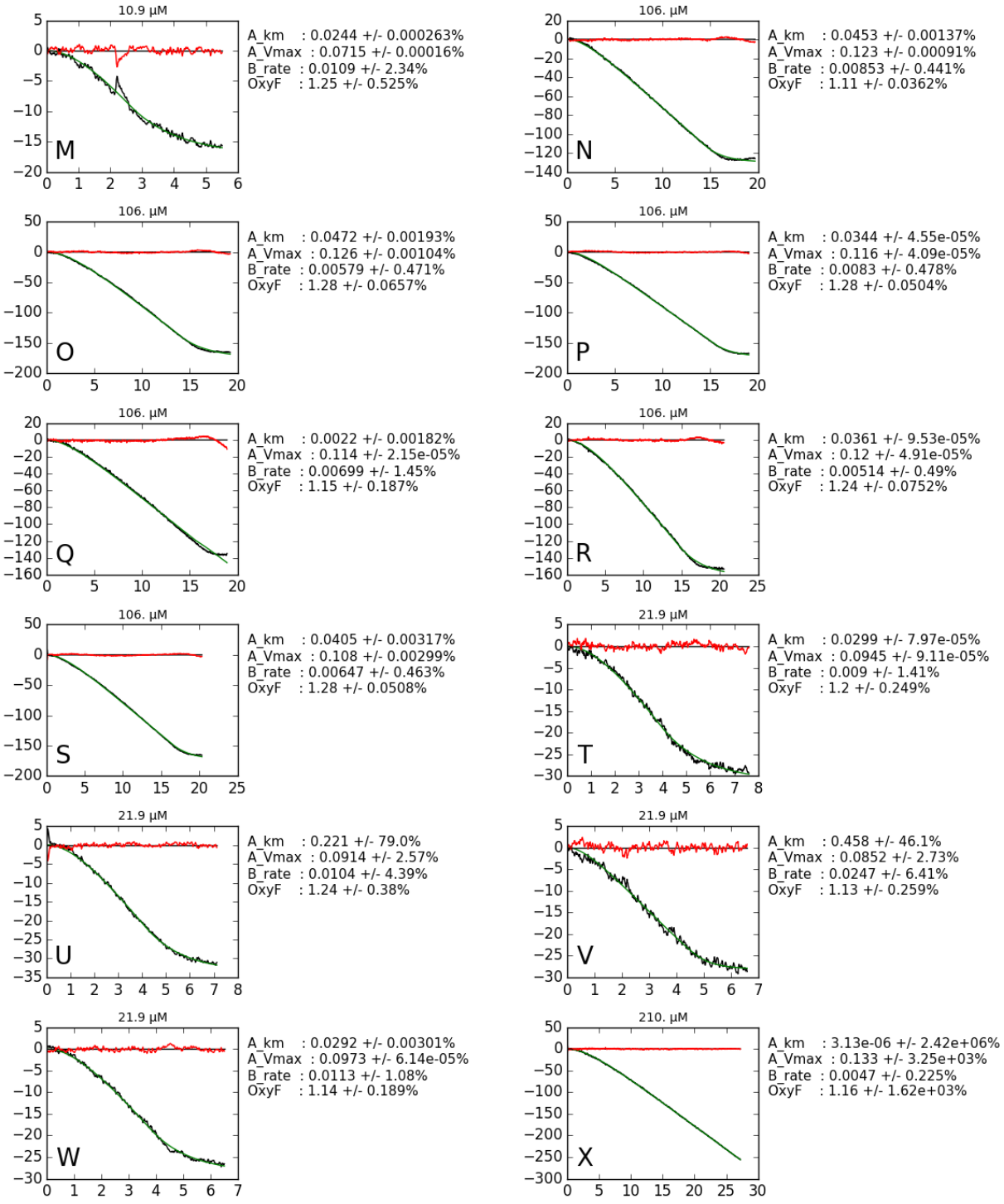


Figure 7-4. Oxygen utilisation data, first presented in Figure 4-14, of A. calcoaceticus E1 treated with metaldehyde (black curves); curves constructed by fitting Model B (green); residuals shown in red. Fitted parameter values are shown to the right of each chart with calculated standard deviations expressed as a percentage.

Abbreviations

AA	Amino acid
Abbr.	Abbreviation
Ace.	Acetate
ALDH	Aldehyde dehydrogenase
ANI	Average nucleotide identity
ANOVA	Analysis of variance
APS	Ammonium per sulfphate
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
BSR	BLAST score ratio
CC	Culture collection
CIP	Collection of the Institut Pasteur
CSD	Sulphino alanine
DNA	Deoxyribose nucleic acid
DDH	DNA-DNA hybridisation
DNT	2,4-dinitrotoluene
DT	Degradation time
DTT	Dithiothreitol
DWI	Drinking Water Inspectorate
EDTA	Ethylenediaminetetraacetic acid
ESI	Electro-spray ionisation
FERA	Food and Environment Research Agency
FWER	Family-wise error rate
GC	Guanine-cytosine
HGT	Horizontal gene transfer
HSD	Honestly significant difference
Inf	Infinity
IPR	InterPro
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IQD	Interquartile distance
KS	Kolmogorov-Smirnov
K_M	Michaelis-Menten constant
LB	Lysogeny broth
LC	Liquid chromatography
LCB	Locally colinear block
LCMS	Liquid chromatography-mass spectrometry
LD	Lethal dose
LOQ	Limit of quantification
m/z	Mass-charge ratio
MA	Metaldehyde
MBP	maltose binding protein
MDP	Metaldehyde degrading protein
Met.	Metaldehyde
Min.	Minute
MS	Mass spectrometry
MSM	Minimal salts media
NADH	Nicotinamide adenine dinucleotide
NCBI	National Center for Biotechnology Information
NEB	New England BioLabs

OD	Optical density
ORF	Open reading frame
OTU	Operational taxonomic unit
PB	Phosphate buffer
PCR	Polymerase chain reaction
PDB	Protein Database
PHO	Phenol hydroxylase
RNA	Ribose nucleic acid
RSD	Relative standard deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SE	Standard error
SPE	Solid phase extraction
SSF	Slow sand filter
TE	Tris-EDTA
TEMED	Tetramethylethylenediamine
Temp.	Temperature
UK	United Kingdom
UV	Ultraviolet
V_{max}	Maximum velocity of a reactio

Reference List

Abbott BJ, Laskin AI & McCoy CJ (1973). Growth of Acinetobacter calcoaceticus on Ethanol. *Appl. Microbiol.* 25, 787–792.

Agriculture & Environment Research Unit (2012). Pesticide Properties DataBase - Metaldehyde. Available at: http://sitem.herts.ac.uk/aeru/footprint/en/Reports/446.htm [Accessed January 8, 2013].

Allison R (2015). Cereals 2015: Monitoring shows rise in metaldehyde "spikes." *Farmers Wkly.* Available at: http://www.fwi.co.uk/arable/cereals-2015-monitoring-shows-rise-in-metaldehyde-spikes.htm [Accessed November 24, 2015].

Arnold K, Bordoli L, Kopp J & Schwede T (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinforma. Oxf. Engl.* 22, 195–201.

Ashida N, Ishii S, Hayano S, Tago K, Tsuji T, Yoshimura Y, Otsuka S & Senoo K (2010). Isolation of functional single cells from environments using a micromanipulator: application to study denitrifying bacteria. *Appl. Microbiol. Biotechnol.* 85, 1211–1217.

Autin O, Hart J, Jarvis P, MacAdam J, Parsons SA & Jefferson B (2013a). The impact of background organic matter and alkalinity on the degradation of the pesticide metaldehyde by two advanced oxidation processes: UV/H_2O_2 and UV/TiO_2. *Water Res.* 47, 2041–2049.

Autin O, Romelot C, Rust L, Hart J, Jarvis P, MacAdam J, Parsons SA & Jefferson B (2013b). Evaluation of a UV-light emitting diodes unit for the removal of micropollutants in water for low energy advanced oxidation processes. *Chemosphere* 92, 745–751.

Baker P, Ricer T, Moynihan PJ, Kitova EN, Walvoort MTC, Little DJ, Whitney JC, Dawson K, Weadge JT, Robinson H, Ohman DE, Codée JDC, Klassen JS, Clarke AJ & Howell PL (2014). P. aeruginosa SGNH Hydrolase-Like Proteins AlgJ and AlgX Have Similar Topology but Separate and Distinct Roles in Alginate Acetylation. *PLoS Pathog.* 10, e1004334.

Barnett SA, Hulme AT & Tocher DA (2005). A low-temperature redetermination of metaldehyde. *Acta Crystallogr. Sect. E Struct. Rep. Online* 61, o857–o859.

Bates NS, Sutton NM & Campbell A (2012). Suspected metaldehyde slug bait poisoning in dogs: a retrospective analysis of cases reported to the Veterinary Poisons Information Service. *Vet. Rec.* 171, 324–324.

Baumann P, Doudoroff M & Stanier RY (1968). A Study of the Moraxella Group II. Oxidative-negative Species (Genus Acinetobacter). *J. Bacteriol.* 95, 1520–1541.

Benkert P, Künzli M & Schwede T (2009). QMEAN server for protein model quality estimation. *Nucleic Acids Res.* 37, W510–W514.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN & Bourne PE (2000). The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242.

Besemer J, Lomsadze A & Borodovsky M (2001). GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 29, 2607–2618.

Bille L, Toson M, Mulatti P, Dalla Pozza M, Capolongo F, Casarotto C, Ferrè N, Angeletti R, Gallocchio F & Binato G (2016). Epidemiology of animal poisoning: An overview on the features and spatio-temporal distribution of the phenomenon in the north-eastern Italian regions. *Forensic Sci. Int.* 266, 440–448.

Bonnin RA, Poirel L & Nordmann P (2012). AbaR-type transposon structures in Acinetobacter baumannii. *J. Antimicrob. Chemother.* 67, 234–236.

Booze TF & Oehme FW (1986). An investigation of metaldehyde and acetaldehyde toxicities in dogs. *Fundam. Appl. Toxicol. Off. J. Soc. Toxicol.* 6, 440–446.

Bragg LM, Stone G, Butler MK, Hugenholtz P & Tyson GW (2013). Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data. *PLoS Comput. Biol.* 9, e1003031.

Breugelmans P, D'Huys P-J, De Mot R & Springael D (2007). Characterization of novel linuron-mineralizing bacterial consortia enriched from long-term linuron-treated agricultural soils. *FEMS Microbiol. Ecol.* 62, 374–385.

Brisou J & Prévot AR (1954). Etudes de systematique bacterienne. X. Revision des espéces réunies dans le genre Achromobacter. *Ann Inst Pasteur* 86, 722–728.

Brown P, Charlton A, Cuthbert M, Barnett L, Ross L, Green M, Gillies L, Shaw K & Fletcher M (1996). Identification of pesticide poisoning in wildlife. *J. Chromatogr. A* 754, 463–478.

Bruland N, Wübbeler JH & Steinbüchel A (2009). 3-Mercaptopropionate Dioxygenase, a Cysteine Dioxygenase Homologue, Catalyzes the Initial Step of 3-Mercaptopropionate Catabolism in the 3,3-Thiodipropionic Acid-degrading Bacterium Variovorax paradoxus. *J. Biol. Chem.* 284, 660–672.

Burke JE & Dennis EA (2009). Phospholipase A2 structure/function, mechanism, and signaling. *J. Lipid Res.* 50, S237–S242.

Burley SK, David PR, Taylor A & Lipscomb WN (1990). Molecular structure of leucine aminopeptidase at 2.7-A resolution., Molecular structure of leucine aminopeptidase at 2.7-A resolution. *Proc. Natl. Acad. Sci. U. S. Am. Proc. Natl. Acad. Sci. U. S. Am.* 87, 87, 6878, 6878–6882.

Busquets R, Kozynchenko OP, Whitby RLD, Tennison SR & Cundy AB (2014). Phenolic carbon tailored for the removal of polar organic contaminants from water: A solution to the metaldehyde problem? *Water Res.* 61, 46–56.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K & Madden TL (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421.

Canchaya C, Proux C, Fournous G, Bruttin A & Brüssow H (2003). Prophage Genomics. *Microbiol. Mol. Biol. Rev.* 67, 238–276.

Carbajal-Rodríguez I, Stöveken N, Satola B, Wübbeler JH & Steinbüchel A (2011). Aerobic degradation of mercaptosuccinate by the gram-negative bacterium Variovorax paradoxus strain B4. *J. Bacteriol.* 193, 527–539.

Carpenter M (1989a). Hydrolysis of metaldehyde as a function of pH at 25°C. In *Metaldehyde Draft Assessment Report, Vol. 3, B8*. European Food Safety Authority, pp.293–4.

Carpenter M (1989b). Photodegradation of metaldehyde in pH 7 buffered solution. In *Metaldehyde Draft Assessment Report, Vol. 3, B8*. European Food Safety Authority, pp.294–5.

Carr EL, Kämpfer P, Patel BKC, Gürtler V & Seviour RJ (2003). Seven novel species of Acinetobacter isolated from activated sludge. *Int. J. Syst. Evol. Microbiol.* 53, 953–963.

Cases I & Lorenzo V de (2001). The black cat/white cat principle of signal integration in bacterial promoters. *EMBO J.* 20, 1–11.

Chen I, Christie PJ & Dubnau D (2005). The Ins and Outs of DNA Transfer in Bacteria. *Science* 310, 1456–1460.

Cheng H, Shen N, Pei J & Grishin NV (2004). Double-stranded DNA bacteriophage prohead protease is homologous to herpesvirus protease. *Protein Sci. Publ. Protein Soc.* 13, 2260–2269.

Chief Inspector of Drinking Water (2009). *Drinking Water 2008 – Public water supplies (regional reports)*, Drinking Water Inspectorate. Available at: http://webarchive.nationalarchives.gov.uk/20120906081707/http://dwi.defra.gov.uk/about/annual-report/2008/index.htm [Accessed November 30, 2015].

Chief Inspector of Drinking Water (2016). *Drinking Water 2015 – Public water supplies (regional reports)*, Drinking Water Inspectorate. Available at: http://dwi.defra.gov.uk/about/annual-report/2015/index.html [Accessed November 30, 2015].

Colbourne J (2013). Letter to the Parliamentary Under-Secretary for Natural Environment and Fisheries. Available at: http://www.dwi.gov.uk/about/annual-report/2012/letter-english.pdf [Accessed August 20, 2016].

Cordero OX & Polz MF (2014). Explaining microbial genomic diversity in light of evolutionary ecology. *Nat. Rev. Microbiol.* 12, 263–273.

Cordova-Rosa SM, Dams RI, Cordova-Rosa EV, Radetski MR, Corrêa AXR & Radetski CM (2009). Remediation of phenol-contaminated soil by a bacterial consortium and Acinetobacter calcoaceticus isolated from an industrial wastewater treatment plant. *J. Hazard. Mater.* 164, 61–66.

Cranor W (1990a). Aerobic Soil Metabolism of 14C-Metaldehyde. In *Metaldehyde Draft Assessment Report, Vol. 3, B8*. European Food Safety Authority, pp.259–63.

Cranor W (1990b). Anerobic Soil Metabolism of 14C-Metaldehyde. In *Metaldehyde Draft Assessment Report, Vol. 3, B8*. European Food Safety Authority, pp.263–6.

Darmon E & Leach DRF (2014). Bacterial Genome Instability. *Microbiol. Mol. Biol. Rev.* 78, 1–39.

Davaji B, Jeong Bak H, Chang W-J & Hoon Lee C (2014). A novel on-chip three-dimensional micromachined calorimeter with fully enclosed and suspended thin-film chamber for thermal characterization of liquid samples. *Biomicrofluidics* 8, 34101.

Dealtry S, Nour EH, Holmsgaard PN, Ding G-C, Weichelt V, Dunon V, Heuer H, Hansen LH, Sørensen SJ, Springael D & Smalla K (2016). Exploring the complex response to linuron of bacterial

communities from biopurification systems by means of cultivation-independent methods. *FEMS Microbiol. Ecol.* 92.

De Palmenaer D, Vermeiren C & Mahillon J (2004). IS231–MIC231 elements from Bacillus cereus sensu lato are modular. *Mol. Microbiol.* 53, 457–467.

Dejonghe W, Berteloot E, Goris J, Boon N, Crul K, Maertens S, Höfte M, Vos PD, Verstraete W & Top EM (2003). Synergistic Degradation of Linuron by a Bacterial Consortium and Isolation of a Single Linuron-Degrading Variovorax Strain. *Appl. Environ. Microbiol.* 69, 1532–1541.

Department for Environment, Food & Rural Affairs (2016). *British survey of fertiliser practice 2015*, Available at: https://www.gov.uk/government/statistics/british-survey-of-fertiliser-practice-2015 [Accessed August 23, 2016].

Devers M, Azhari NE, Kolic N-U & Martin-Laurent F (2007). Detection and organization of atrazine-degrading genetic potential of seventeen bacterial isolates belonging to divergent taxa indicate a recent common origin of their catabolic functions. *FEMS Microbiol. Lett.* 273, 78–86.

Divari S, Valetti F, Caposio P, Pessione E, Cavaletto M, Griva E, Gribaudo G, Gilardi G & Giunta C (2003). The oxygenase component of phenol hydroxylase from Acinetobacter radioresistens S13. *Eur. J. Biochem.* 270, 2244–2253.

Doria FC, Borges AC, Kim JK, Nathan A, Joo JC & Campos LC (2013). Removal of Metaldehyde Through Photocatalytic Reactions Using Nano-Sized Zinc Oxide Composites. *Water. Air. Soil Pollut.* 224, 1–9.

Dunon V, Sniegowski K, Bers K, Lavigne R, Smalla K & Springael D (2013). High prevalence of IncP-1 plasmids and IS1071 insertion sequences in on-farm biopurification systems and other pesticide-polluted environments. *FEMS Microbiol. Ecol.* 86, 415–431.

Dunwell JM, Purvis A & Khuri S (2004). Cupins: the most functionally diverse protein superfamily? *Phytochemistry* 65, 7–17.

Eckert M, Fleischmann G, Jira R, Bolt HM & Golka K (2000). Acetaldehyde. In *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA.

Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797.

Ellis EC (2015). Ecology in an anthropogenic biosphere. *Ecol. Monogr.* 85, 287–331.

Elsas JD van, Trevors JT, Jansson JK & Nannipieri P (2006). *Modern Soil Microbiology, Second Edition*, CRC Press.

Environment Agency (2009). *The determination of metaldehyde in waters using chromatography with mass spectrometric detection*, Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/316782/Metaldehyde-226b.pdf [Accessed August 21, 2016].

European Food Safety Authority (2006). *Draft Assessment Report, Vol. 3, B.8: Metaldehyde*,

Eurostat (2016). Pesticide Sales. Available at: http://ec.europa.eu/eurostat/web/products-datasets/-/aei_fm_salpest09 [Accessed August 20, 2016].

Fanelli F, Chiara M, Liuzzi VC, Haidukowski M, Tristezza M, Caterina M, D'Erchia AM, Pesole G, Horner DS & Mule' G (2015). Draft genome sequence of Acinetobacter sp. neg1 capable of degrading ochratoxin A. *FEMS Microbiol. Lett.* 362, fnv004.

Ferrer M, Martínez-Abarca F & Golyshin PN (2005). Mining genomes and "metagenomes" for novel catalysts. *Curr. Opin. Biotechnol.* 16, 588–593.

Fewson CA (1967). The growth and metabolic versatility of the gram-negative Bacterium NCIB 8250 ("Vibrio 01"). *J. Gen. Microbiol.* 46, 255–266.

Finn JA, Parks AR & Peters JE (2007). Transposon Tn7 Directs Transposition into the Genome of Filamentous Bacteriophage M13 Using the Element-Encoded TnsE Protein. *J. Bacteriol.* 189, 9122–9125.

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J & Bateman A (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285.

Fondi M, Maida I, Perrin E, Orlandini V, La Torre L, Bosi E, Negroni A, Zanaroli G, Fava F, Decorosi F, Giovannetti L, Viti C, Vaneechoutte M, Dijkshoorn L & Fani R (2016). Genomic and phenotypic characterization of the species Acinetobacter venetianus. *Sci. Rep.* 6, 21985.

Fox GE, Pechman KR & Woese CR (1977). Comparative Cataloging of 16S Ribosomal Ribonucleic Acid: Molecular Approach to Procaryotic Systematics. *Int. J. Syst. Evol. Microbiol.* 27, 44–57.

Fromm H, Winter K, Filser J, Hantschel R & Beese F (1993). The influence of soil type and cultivation system on the spatial distributions of the soil fauna and microorganisms and their interactions. *Geoderma* 60, 109–118.

Fujii T, Takeo M & Maeda Y (1997). Plasmid-Encoded Genes Specifying Aniline Oxidation from Acinetobacter sp. Strain YAA. *Microbiology* 143, 93–99.

Garthwaite D, Barker I, Laybourn R, Huntly A, Parrish GP, Hudson S & Thygesen H (2015). *Pesticide Usage Survey Report 263 - Arable crops in the UK (Version 2)*, Department for Environment, Food & Rural Affairs. Available at: https://secure.fera.defra.gov.uk/pusstats/surveys/2014surveys.cfm [Accessed August 23, 2016].

Garthwaite DG, Thomas MR, Parrish G, Smith L & Barker I (2010). *Pesticide Usage Survey Report 224: Arable Crops In Great Britain 2008*, Department for Environment, Food & Rural Affairs.

Gentry T, Rensing C & Pepper I (2004). New Approaches for Bioaugmentation as a Remediation Technology. *Crit. Rev. Environ. Sci. Technol.* 34, 447–494.

Ghatak S, Blom J, Das S, Sanjukta R, Puro K, Mawlong M, Shakuntala I, Sen A, Goesmann A, Kumar A & Ngachan SV (2016). Pan-genome analysis of Aeromonas hydrophila. *Antonie Van Leeuwenhoek* 109, 945–956.

Gillespie D & Spiegelman S (1965). A quantitative assay for DNA-RNA hybrids with DNA immobilized on a membrane. *J. Mol. Biol.* 12, 829–842.

Gimingham CT (1940). Some Recent Contributions by English Workers to the Development of Methods of Insect Control. *Ann. Appl. Biol.* 27, 161–175.

Gohl O, Friedrich A, Hoppert M & Averhoff B (2006). The Thin Pili of Acinetobacter sp. Strain BD413 Mediate Adhesion to Biotic and Abiotic Surfaces. *Appl. Environ. Microbiol.* 72, 1394–1401.

Gonod LV, Chenu C & Soulas G (2003). Spatial variability of 2,4-dichlorophenoxyacetic acid (2,4-D) mineralisation potential at a millimetre scale in soil. *Soil Biol. Biochem.* 35, 373–382.

Goral AM, Tkaczuk KL, Chruszcz M, Kagan O, Savchenko A & Minor W (2012). Crystal structure of a putative isochorismatase hydrolase from Oleispira antarctica. *J. Struct. Funct. Genomics* 13, 27–36.

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P & Tiedje JM (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91.

Gough J, Karplus K, Hughey R & Chothia C (2001). Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J. Mol. Biol.* 313, 903–919.

Govantes F, Porrúa O, García-González V & Santero E (2009). Atrazine biodegradation in the lab and in the field: enzymatic activities and gene regulation. *Microb. Biotechnol.* 2, 178–185.

Gözdereliler E, Boon N, Aamand J, Roy KD, Granitsiotis MS, Albrechtsen H-J & Sørensen SR (2013). Comparing Metabolic Functionalities, Community Structures, and Dynamics of Herbicide-Degrading Communities Cultivated with Different Substrate Concentrations. *Appl. Environ. Microbiol.* 79, 367–375.

Greene EA, Beatty PH & Fedorak PM (2000). Sulfolane degradation by mixed cultures and a bacterial isolate identified as a Variovorax sp. *Arch. Microbiol.* 174, 111–119.

Hamaki T, Suzuki M, Fudou R, Jojima Y, Kajiura T, Tabuchi A, Sen K & Shibai H (2005). Isolation of novel bacteria and actinomycetes using soil-extract agar medium. *J. Biosci. Bioeng.* 99, 485–492.

Henrichsen J (1975). The Occurrence of Twitching Motility Among Gram-Negative Bacteria. *Acta Pathol. Microbiol. Scand. [B]* 83B, 171–178.

Holt JG (1994). *Bergey's Manual of Determinative Bacteriology*, Lippincott Williams & Wilkins.

Huerta-Cepas J, Serra F & Bork P (2016). ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol. Biol. Evol.*, msw046.

Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, de Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, McMenamin C, Mi H, Mutowo-Muellenet P, Mulder N, Natale D, Orengo C, Pesseat S, Punta M, Quinn AF, Rivoire C, Sangrador-Vegas A, Selengut JD, Sigrist CJA, Scheremetjew M, Tate J, Thimmajanarthanan M, Thomas PD, Wu CH, Yeats C & Yong S-Y (2012). InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res.* 40, D306-312.

Hwanga C-A & Draughon FA (1994). Degradation of Ochratoxin A by Acinetobacter calcoaceticus. *J. Food Prot.* 57, 410–414.

Janssen DB, Dinkla IJT, Poelarends GJ & Terpstra P (2005). Bacterial degradation of xenobiotic compounds: evolution and distribution of novel enzyme activities. *Environ. Microbiol.* 7, 1868–1882.

Janssen PH, Yates PS, Grinton BE, Taylor PM & Sait M (2002). Improved Culturability of Soil Bacteria and Isolation in Pure Culture of Novel Members of the Divisions Acidobacteria, Actinobacteria, Proteobacteria, and Verrucomicrobia. *Appl. Environ. Microbiol.* 68, 2391–2396.

Jennions MD & Møller AP (2002). Publication bias in ecology and evolution: an empirical assessment using the "trim and fill" method. *Biol. Rev.* 77, 211–222.

Joernvall H & Persson B (2006). PROSITE documentation PDOC00068 – Aldehyde dehydrogenases active sites. Available at: http://prosite.expasy.org/PDOC00068 [Accessed July 22, 2016].

Jolicoeur P & Morin A (1987). Isolation of Acinetobacter calcoaceticus strains degrading the volatile fatty acids of swine wastes. *Biol. Wastes* 19, 133–140.

Jones A & Charlton A (1999). Determination of Metaldehyde in Suspected Cases of Animal Poisoning Using Gas Chromatography−Ion Trap Mass Spectrometry. *J. Agric. Food Chem.* 47, 4675–4677.

Juni E & Janik A (1969). Transformation of Acinetobacter calco-aceticus (Bacterium anitratum). *J. Bacteriol.* 98, 281–288.

Kale A, Pijning T, Sonke T, Dijkstra BW & Thunnissen A-MWH (2010). Crystal Structure of the Leucine Aminopeptidase from Pseudomonas putida Reveals the Molecular Basis for its Enantioselectivity and Broad Substrate Specificity. *J. Mol. Biol.* 398, 703–714.

Kamagata Y & Tamaki H (2005). Cultivation of Uncultured Fastidious Microbes. *Microbes Environ.* 20, 85–91.

Kanehisa M, Goto S, Kawashima S, Okuno Y & Hattori M (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32, D277–D280.

Kang SG, Maurizi MR, Thompson M, Mueser T & Ahvazi B (2004). Crystallography and mutagenesis point to an essential role for the N-terminus of human mitochondrial ClpP. *J. Struct. Biol.* 148, 338–352.

Kay P & Grayson R (2014). Using water industry data to assess the metaldehyde pollution problem. *Water Environ. J.* 28, 410–417.

Kegley SE, Hill BR, Orme S & Choi AH (2016). PAN Pesticide Database, Pesticide Action Network, North America. Available at: http://www.pesticideinfo.org/ [Accessed August 15, 2016].

Kekulé A & Zincke T (1872). Ueber das sogenannte Chloraceten und die polymeren Modificationen des Aldehyds. *Justus Liebigs Ann. Chem.* 162, 125–150.

Kim K-S, Pelton JG, Inwood WB, Andersen U, Kustu S & Wemmer DE (2010). The Rut Pathway for Pyrimidine Degradation: Novel Chemistry and Toxicity Problems. *J. Bacteriol.* 192, 4089–4102.

Kim M, Oh H-S, Park S-C & Chun J (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 64, 346–351.

Kim SJ & Kim YS (Yonsei U (1985). Isolation of a malonate-utilizing Acinetobacter calcoaceticus from soil. *Korean J. Microbiol. Korea R* 23, 230–4.

Kirkman TW (1996). Statistics to use. Available at: http://www.physics.csbsju.edu/stats/ [Accessed June 7, 2016].

Klibanov AM, Tu T-M & Scott KP (1983). Peroxidase-Catalyzed Removal of Phenols from Coal-Conversion Waste Waters. *Science* 221, 259–261.

Koh J-S, Yamakawa T, Kodama T & Minoda Y (1985). Rapid and Dense Culture of Acinetobacter calcoaceticus on Palm Oil. *Agric. Biol. Chem.* 49, 1411–1416.

Konstantinidis KT & Tiedje JM (2005). Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 102, 2567–2572.

Kowalski M, Waylen C, Clist S, Wilcox S, Lynn S & Garrow D (2013). *Freshwater availability and use in the UK*, Waste and Resources Action Plan. Available at: http://www.wrap.org.uk/content/freshwater-availability-and-use-uk-0 [Accessed August 23, 2016].

Krahn T, Wibberg D, Maus I, Winkler A, Bontron S, Sczyrba A, Nordmann P, Pühler A, Poirel L & Schlüter A (2016). Intraspecies Transfer of the Chromosomal Acinetobacter baumannii blaNDM-1 Carbapenemase Gene. *Antimicrob. Agents Chemother.* 60, 3032–3040.

Kromidas S (2008). *More Practical Problem Solving in HPLC*, John Wiley & Sons.

Krutz JL, Shaner DL, Weaver MA, Webb RM, Zablotowicz RM, Reddy KN, Huang Y & Thomson SJ (2010). Agronomic and environmental implications of enhanced s-triazine degradation. *Pest Manag. Sci.* 66, 461–481.

Lang AS, Zhaxybayeva O & Beatty JT (2012). Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Microbiol.* 10, 472–482.

Larson HN, Zhou J, Chen Z, Stamler JS, Weiner H & Hurley TD (2007). Structural and functional consequences of coenzyme binding to the inactive asian variant of mitochondrial aldehyde dehydrogenase: roles of residues 475 and 487. *J. Biol. Chem.* 282, 12940–12950.

Lazartigues A, Banas D, Feidt C, Brun-Bellut J & Thomas M (2012). Pesticide pressure and fish farming in barrage pond in Northeastern France Part I: site characterization and water quality. *Environ. Sci. Pollut. Res.* 19, 2802–2812.

Li C, Wu Y-L, Yang T & Zhang Y (2010). Determination of Metaldehyde in Water by SPE and UPLC–MS–MS. *Chromatographia* 72, 987–991.

Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM, Buso N & Lopez R (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* 43, W580–W584.

Lindell D, Sullivan MB, Johnson ZI, Tolonen AC, Rohwer F & Chisholm SW (2004). Transfer of photosynthesis genes to and from Prochlorococcus viruses. *Proc. Natl. Acad. Sci. U. S. A.* 101, 11013–11018.

Locey KJ & Lennon JT (2016). Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci.* 113, 5970–5975.

Loenarz C & Schofield CJ (2011). Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends Biochem. Sci.* 36, 7–18.

Loh KD, Gyaneshwar P, Markenscoff Papadimitriou E, Fong R, Kim K-S, Parales R, Zhou Z, Inwood W & Kustu S (2006). A previously undescribed pathway for pyrimidine catabolism. *Proc. Natl. Acad. Sci.* 103, 5114–5119.

Luo Q, Olucha J & Lamb AL (2009). Structure−Function Analyses of Isochorismate-Pyruvate Lyase from Pseudomonas aeruginosa Suggest Differing Catalytic Mechanisms for the Two Pericyclic Reactions of This Bifunctional Enzyme. *Biochemistry (Mosc.)* 48, 5239–5245.

Luria SE, Adams JN & Ting RC (1960). Transduction of lactose-utilizing ability among strains of E. coli and S. dysenteriae and the properties of the transducing phage particles. *Virology* 12, 348–390.

Mandelbaum RT, Allan DL & Wackett LP (1995). Isolation and Characterization of a Pseudomonas sp. That Mineralizes the s-Triazine Herbicide Atrazine. *Appl. Environ. Microbiol.* 61, 1451–1457.

Martinez B, Tomkins J, Wackett LP, Wing R & Sadowsky MJ (2001). Complete Nucleotide Sequence and Organization of the Atrazine Catabolic Plasmid pADP-1 from Pseudomonas sp. Strain ADP. *J. Bacteriol.* 183, 5684–5697.

Melderen LV & Bast MSD (2009). Bacterial Toxin–Antitoxin Systems: More Than Selfish Entities? *PLOS Genet* 5, e1000437.

Mengoni A, Ricci S, Brilli M, Baldi F & Fani R (2007). Sequencing and analysis of plasmids pAV1 and pAV2 ofAcinetobacter venetianus VE-C3 involved in diesel fuel degradation. *Ann. Microbiol.* 57, 521–526.

Mergaert J, Webb A, Anderson C, Wouters A & Swings J (1993). Microbial degradation of poly(3-hydroxybutyrate) and poly(3-hydroxybutyrate-co-3-hydroxyvalerate) in soils. *Appl. Environ. Microbiol.* 59, 3233–3238.

Metzgar D, Bacher JM, Pezo V, Reader J, Döring V, Schimmel P, Marlière P & Crécy-Lagard V de (2004). Acinetobacter sp. ADP1: an ideal model organism for genetic analysis and genome engineering. *Nucleic Acids Res.* 32, 5780–5790.

Miller R (1928). Poisoning by “Meta Fuel” Tablets (Metacetaldehyde). *Arch. Dis. Child.* 3, 292–295.

Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, Shaw PD & Marshall D (2013). Using Tablet for visual exploration of second-generation sequencing data. *Brief. Bioinform.* 14, 193–202.

Mirgain I, Green G a. & Monteil H (1993). Degradation of atrazine in laboratory microcosms: Isolation and identification of the biodegrading bacteria. *Environ. Toxicol. Chem.* 12, 1627–1634.

Miwa H, Ahmed I, Yoon J, Yokota A & Fujiwara T (2008). Variovorax boronicumulans sp. nov., a boron-accumulating bacterium isolated from soil. *Int. J. Syst. Evol. Microbiol.* 58, 286–289.

Miyafusa T, Caaveiro JMM, Tanaka Y & Tsumoto K (2012). Crystal structure of the enzyme CapF of Staphylococcus aureus reveals a unique architecture composed of two functional domains. *Biochem. J.* 443, 671–680.

Mo X, Chen C, Pang H, Feng Y & Feng J (2010). Identification and characterization of a novel xylanase derived from a rice straw degrading enrichment culture. *Appl. Microbiol. Biotechnol.* 87, 2137–2146.

Möllerfeld J, Römbke J & Heller M (1993). Determination of the degradability and persistence of 14C-metaldehyde in the water/sediment-system. In *Metaldehyde Draft Assessment Report, Vol. 3, B8*. European Food Safety Authority, pp.296–300.

Moreau P, Burgeot T & Renault T (2014). In vivo effects of metaldehyde on Pacific oyster, Crassostrea gigas: comparing hemocyte parameters in two oyster families. *Environ. Sci. Pollut. Res.* 22, 8003–8009.

Murdoch RW & Hay AG (2015). The biotransformation of ibuprofen to trihydroxyibuprofen in activated sludge and by Variovorax. *Biodegradation* 26, 105–113.

Nabeerasool MA, Campen AK, Polya DA, Brown NW & van Dongen BE (2015). Removal of Metaldehyde from Water Using a Novel Coupled Adsorption and Electrochemical Destruction Technique. *Water* 7, 3057–3071.

National Center for Biotechnology Information PubChem Compound Database; CID = 61021. Available at: https://pubchem.ncbi.nlm.nih.gov/compound/metaldehyde [Accessed August 20, 2016].

Nemec A, Janda L, Melter O & Dijkshoorn L (1999). Genotypic and phenotypic similarity of multiresistant Acinetobacter baumannii isolates in the Czech Republic. *J. Med. Microbiol.* 48, 287–296.

Nemec A, Krizova L, Maixnerova M, van der Reijden TJK, Deschaght P, Passet V, Vaneechoutte M, Brisse S & Dijkshoorn L (2011). Genotypic and phenotypic characterization of the Acinetobacter calcoaceticus–Acinetobacter baumannii complex with the proposal of Acinetobacter pittii sp. nov. (formerly Acinetobacter genomic species 3) and Acinetobacter nosocomialis sp. nov. (formerly Acinetobacter genomic species 13TU). *Res. Microbiol.* 162, 393–404.

O'Brien PJ & Herschlag D (1999). Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol.* 6, R91–R105.

Ozer EA, Allen JP & Hauser AR (2014). Characterization of the core and accessory genomes of Pseudomonas aeruginosa using bioinformatic tools Spine and AGEnt. *BMC Genomics* 15, 737.

Parte AC (2014). LPSN—list of prokaryotic names with standing in nomenclature. *Nucleic Acids Res.* 42, D613–D616.

Pauling L & Carpenter DC (1936). The Crystal Structure of Metaldehyde. *J. Am. Chem. Soc.* 58, 1274–1278.

Platt A, Shingler V, Taylor SC & Williams PA (1995). The 4-hydroxy-2-oxovalerate aldolase and acetaldehyde dehydrogenase (acylating) encoded by the nahM and nahO genes of the naphthalene catabolic plasmid pWW60-22 provide further evidence of conservation of meta-cleavage pathway gene sequences. *Microbiol. Read. Engl.* 141, 2223–2233.

Purcell M (2014). Letter re: Metaldehyde and other pesticides. Available at: http://dwi.defra.gov.uk/stakeholders/information-letters/2014/01-2014.pdf [Accessed August 20, 2016].

Radajewski S, Ineson P, Parekh NR & Murrell JC (2000). Stable-isotope probing as a tool in microbial ecology. *Nature* 403, 646–649.

Rahman MM, Shiddiky MJA, Rahman MA & Shim Y-B (2009). A lactate biosensor based on lactate dehydrogenase/nictotinamide adenine dinucleotide (oxidized form) immobilized on a conducting polymer/multiwall carbon nanotube composite film. *Anal. Biochem.* 384, 159–165.

Rajoo S, Ahn JO, Lee HW & Jung JK (2013). Isolation and characterization of a novel ε-caprolactam-degrading microbe, Acinetobacter calcoaceticus, from industrial wastewater by chemostat-enrichment. *Biotechnol. Lett.* 35, 2069–2072.

Rappé MS & Giovannoni SJ (2003). The uncultured microbial majority. *Annu. Rev. Microbiol.* 57, 369–394.

Rasko DA, Myers GS & Ravel J (2005). Visualization of comparative genomic analyses by BLAST score ratio. *BMC Bioinformatics* 6, 2.

Rather JA, Pilehvar S & Wael KD (2012). A biosensor fabricated by incorporation of a redox mediator into a carbon nanotube/nafion composite for tyrosinase immobilization: detection of matairesinol, an endocrine disruptor. *Analyst* 138, 204–210.

Ravin AW (1963). Experimental Approaches to the Study of Bacterial Phylogeny. *Am. Nat.* 97, 307–318.

Raynaud X & Nunan N (2014). Spatial Ecology of Bacteria at the Microscale in Soil. *PLOS ONE* 9, e87217.

Riveros-Rosas H, González-Segura L, Julián-Sánchez A, Díaz-Sánchez ÁG & Muñoz-Clares RA (2013). Structural determinants of substrate specificity in aldehyde dehydrogenases. *Chem. Biol. Interact.* 202, 51–61.

Rodriguez-R LM & Konstantinidis KT (2014). Bypassing cultivation to identify bacterial species. *Microbe* 9, 111–118.

Rolph CA, Jefferson B & Villa R (2014). Switching on Pesticide Degraders in Biological Filters Used in Drinking Water Production. In N. Nakamoto, N. Graham, M. R. Collins, & R. Gimbel, eds.

Progress in Slow Sand and Alternative Biofiltration Processes: Further Developments and Applications. IWA Publishing.

Rousk J, Bååth E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG, Knight R & Fierer N (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME J.* 4, 1340–1351.

Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream M-A & Barrell B (2000). Artemis: sequence visualization and annotation. *Bioinformatics* 16, 944–945.

Saito T, Morita S, Motojyuku M, Akieda K, Otsuka H, Yamamoto I & Inokuchi S (2008). Determination of metaldehyde in human serum by headspace solid-phase microextraction and gas chromatography–mass spectrometry. *J. Chromatogr. B* 875, 573–576.

Satsuma K (2010). Mineralisation of the herbicide linuron by Variovorax sp. strain RA8 isolated from Japanese river sediment using an ecosystem model (microcosm). *Pest Manag. Sci.* 66, 847–852.

Sayler GS, Hooper SW, Layton AC & King JMH (1990). Catabolic plasmids of environmental and ecological significance. *Microb. Ecol.* 19, 1–20.

Schumacher M, Castle G, Gravell A, Mills GA & Fones GR (2016). An improved method for measuring metaldehyde in surface water using liquid chromatography tandem mass spectrometry. *MethodsX* 3, 188–194.

Schwede T, Kopp J, Guex N & Peitsch MC (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* 31, 3381–3385.

Scott C, Jackson CJ, Coppin CW, Mourant RG, Hilton ME, Sutherland TD, Russell RJ & Oakeshott JG (2009). Catalytic Improvement and Evolution of Atrazine Chlorohydrolase. *Appl. Environ. Microbiol.* 75, 2184–2191.

Seffernick JL, Souza ML de, Sadowsky MJ & Wackett LP (2001). Melamine Deaminase and Atrazine Chlorohydrolase: 98 Percent Identical but Functionally Different. *J. Bacteriol.* 183, 2405–2410.

Seffernick JL & Wackett LP (2016). Ancient Evolution and Recent Evolution Converge for the Biodegradation of Cyanuric Acid and Related Triazines. *Appl. Environ. Microbiol.* 82, 1638–1645.

Seifert H, Schulze A, Baginski R & Pulverer G (1994). Plasmid DNA fingerprinting of Acinetobacter species other than Acinetobacter baumannii. *J. Clin. Microbiol.* 32, 82–86.

Šekuljica NŽ, Prlainović NŽ, Stefanović AB, Žuža MG, Čičkarić DZ, Mijin DŽ & Knežević-Jugović ZD (2015). Decolorization of Anthraquinonic Dyes from Textile Effluent Using Horseradish Peroxidase: Optimization and Kinetic Study. *Sci. World J.* 2015, ID: 371625.

Shapir N, Mongodin EF, Sadowsky MJ, Daugherty SC, Nelson KE & Wackett LP (2007). Evolution of Catabolic Pathways: Genomic Insights into Microbial s-Triazine Metabolism. *J. Bacteriol.* 189, 674–682.

Sigrist CJA, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L & Xenarios I (2013). New and continuing developments at PROSITE. *Nucleic Acids Res.* 41, D344-347.

Sillitoe I, Lewis TE, Cuff A, Das S, Ashford P, Dawson NL, Furnham N, Laskowski RA, Lee D, Lees JG, Lehtinen S, Studer RA, Thornton J & Orengo CA (2015). CATH: comprehensive structural and functional annotations for genome sequences. *Nucleic Acids Res.* 43, D376–D381.

Silver S & Misra TK (1988). Plasmid-Mediated Heavy Metal Resistances. *Annu. Rev. Microbiol.* 42, 717–743.

Singh C & Lin J (2008). Isolation and characterization of diesel oil degrading indigenous microrganisms in Kwazulu-Natal, South Africa. *Afr. J. Biotechnol.* 7, 1927–1932.

Skiebe E, de Berardinis V, Morczinek P, Kerrinnes T, Faber F, Lepka D, Hammer B, Zimmermann O, Ziesing S, Wichelhaus TA, Hunfeld K-P, Borgmann S, Gröbner S, Higgins PG, Seifert H, Busse H-J, Witte W, Pfeifer Y & Wilharm G (2012). Surface-associated motility, a common trait of clinical isolates of Acinetobacter baumannii, depends on 1,3-diaminopropane. *Int. J. Med. Microbiol.* 302, 117–128.

Smith D, Alvey S & Crowley DE (2005). Cooperative catabolic pathways within an atrazine-degrading enrichment culture isolated from soil. *FEMS Microbiol. Ecol.* 53, 265–275.

Smyth HF, Carpenter CP & Weil CS (1951). Range-finding toxicity data: List IV. *AMA Arch. Ind. Hyg. Occup. Med.* 4, 119–122.

Snellinx Z, Taghavi S, Vangronsveld J & Lelie D van der (2003). Microbial consortia that degrade 2,4-DNT by interspecies metabolism: isolation and characterisation. *Biodegradation* 14, 19–29.

Sørensen SR, Rasmussen J, Jacobsen CS, Jacobsen OS, Juhler RK & Aamand J (2005). Elucidating the Key Member of a Linuron-Mineralizing Bacterial Community by PCR and Reverse Transcription-PCR Denaturing Gradient Gel Electrophoresis 16S rRNA Gene Fingerprinting and Cultivation. *Appl. Environ. Microbiol.* 71, 4144–4148.

de Souza ML, Newcombe D, Alvey S, Crowley DE, Hay A, Sadowsky MJ & Wackett LP (1998). Molecular Basis of a Bacterial Consortium: Interspecies Catabolism of Atrazine. *Appl. Environ. Microbiol.* 64, 178–184.

Sprince H, Parker CM, Smith GG & Gonzales LJ (1974). Protection against Acetaldehyde Toxicity in the rat byl-cysteine, thiamin andl-2-Methylthiazolidine-4-carboxylic acid. *Agents Actions* 4, 125–130.

Stackebrandt E & Goebel BM (1994). Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. *Int. J. Syst. Evol. Microbiol.* 44, 846–849.

Stamatakis A (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.

Stamatakis A, Hoover P & Rougemont J (2008). A rapid bootstrap algorithm for the RAxML Web servers. *Syst. Biol.* 57, 758–771.

Stevenson BS, Eichorst SA, Wertz JT, Schmidt TM & Breznak JA (2004). New Strategies for Cultivation and Detection of Previously Uncultured Microbes. *Appl. Environ. Microbiol.* 70, 4748–4755.

Stipanuk MH, Simmons CR, Karplus PA & Dominy JE (2011). Thiol dioxygenases: unique families of cupin proteins. *Amino Acids* 41, 91–102.

Stolarski R, Kierdaszuk B, Hagberg CE & Shugar D (1987). Mechanism of hydroxylamine mutagenesis: tautomeric shifts and proton exchange between the promutagen N6-methoxyadenosine and cytidine. *Biochemistry (Mosc.)* 26, 4332–4337.

Stoyan H, De-Polli H, Böhm S, Robertson GP & Paul EA (2000). Spatial heterogeneity of soil respiration and related properties at the plant scale. *Plant Soil* 222, 203–214.

Sutherland T d., Horne I, Harcourt R l., Russell R j. & Oakeshott J g. (2002). Isolation and characterization of a Mycobacterium strain that metabolizes the insecticide endosulfan. *J. Appl. Microbiol.* 93, 380–389.

Takenaka S, Nomura R, Minegishi A & Yoshida K (2013). Enrichment and characterization of a bacterial culture that can degrade 4-aminopyridine. *BMC Microbiol.* 13, 62.

Takeuchi N, Kaneko K & Koonin EV (2014). Horizontal Gene Transfer Can Rescue Prokaryotes from Muller's Ratchet: Benefit of DNA from Dead Cells and Population Subdivision. *G3 GenesGenomesGenetics* 4, 325–339.

Tang LL, DeNardo MA, Gayathri C, Gil RR, Kanda R & Collins TJ (2016). TAML/H2O2 Oxidative Degradation of Metaldehyde: Pursuing Better Water Treatment for the Most Persistent Pollutants. *Environ. Sci. Technol.* 50, 5261–8.

Tao B & Fletcher AJ (2014). Catalytic degradation and adsorption of metaldehyde from drinking water by functionalized mesoporous silicas and ion-exchange resin. *Sep. Purif. Technol.* 124, 195–200.

Tao B & Fletcher AJ (2013). Metaldehyde removal from aqueous solution by adsorption and ion exchange mechanisms onto activated carbon and polymeric sorbents. *J. Hazard. Mater.* 244–245, 240–250.

Tettelin H, Masignani V, Cieslewicz MJ, Eisen JA, Peterson S, Wessels MR, Paulsen IT, Nelson KE, Margarit I, Read TD, Madoff LC, Wolf AM, Beanan MJ, Brinkac LM, Daugherty SC, DeBoy RT, Durkin AS, Kolonay JF, Madupu R, Lewis MR, Radune D, Fedorova NB, Scanlan D, Khouri H, Mulligan S, Carty HA, Cline RT, Van Aken SE, Gill J, Scarselli M, Mora M, Iacobini ET, Brettoni C, Galli G, Mariani M, Vegni F, Maione D, Rinaudo D, Rappuoli R, Telford JL, Kasper DL, Grandi G & Fraser CM (2002). Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V Streptococcus agalactiae. *Proc. Natl. Acad. Sci. U. S. A.* 99, 12391–12396.

Tettelin H, Riley D, Cattuto C & Medini D (2008). Comparative genomics: the bacterial pan-genome. *Curr. Opin. Microbiol.* 11, 472–477.

Thiergart T, Landan G & Martin WF (2014). Concatenated alignments and the case of the disappearing tree. *BMC Evol. Biol.* 14, 266.

Tinberg CE & Lippard SJ (2010). Oxidation Reactions Performed by Soluble Methane Monooxygenase Hydroxylase Intermediates Hperoxo and Q Proceed by Distinct Mechanisms. *Biochemistry (Mosc.)* 49, 7902–7912.

Tindall BJ, Rosselló-Móra R, Busse H-J, Ludwig W & Kämpfer P (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int. J. Syst. Evol. Microbiol.* 60, 249–266.

Todd JD, Curson ARJ, Kirkwood M, Sullivan MJ, Green RT & Johnston AWB (2011). DddQ, a novel, cupin-containing, dimethylsulfoniopropionate lyase in marine roseobacters and in uncultured marine bacteria. *Environ. Microbiol.* 13, 427–438.

Touchon M, Cury J, Yoon E-J, Krizova L, Cerqueira GC, Murphy C, Feldgarden M, Wortman J, Clermont D, Lambert T, Grillot-Courvalin C, Nemec A, Courvalin P & Rocha EPC (2014). The Genomic Diversification of the Whole Acinetobacter Genus: Origins, Mechanisms, and Consequences. *Genome Biol. Evol.* 6, 2866–2882.

Triebskorn R, Christensen K & Heim G (1998). Effects of Orally and Dermally Applied Metaldehyde on Mucus Cells of Slugs (deroceras Reticulatum) Depending on Temperature and Duration of Exposure. *J. Molluscan Stud.* 64, 467–487.

Tukey JW (1977). *Exploratory data analysis*, Addison-Wesley.

U. S. Geological Survey (2011). DOTABLES. Available at: http://water.usgs.gov/software/DOTABLES/ [Accessed August 8, 2016].

Uchiyama T & Watanabe K (2008). Substrate-induced gene expression (SIGEX) screening of metagenome libraries. *Nat. Protoc.* 3, 1202–1212.

United States Environmental Protection Agency (2006). *Reregistration Eligibility Decision for Metaldehyde*, Available at: https://swap.stanford.edu/20141016105705/http://www.epa.gov/oppsrrd1/REDs/metaldehyde_red.pdf.

Vallaeys T, Albino L, Soulas G, Wright AD & Weightman AJ (1998). Isolation and characterization of a stable 2,4-dichlorophenoxyacetic acid degrading bacterium, Variovorax paradoxus, using chemostat culture. *Biotechnol. Lett.* 20, 1073–1076.

Vaneechoutte M, De Baere T, Nemec A, Musílek M, van der Reijden TJK & Dijkshoorn L (2008). Reclassification of Acinetobacter grimontii Carr et al. 2003 as a later synonym of Acinetobacter junii Bouvet and Grimont 1986. *Int. J. Syst. Evol. Microbiol.* 58, 937–940.

Vaneechoutte M, Young DM, Ornston LN, Baere TD, Nemec A, Reijden TVD, Carr E, Tjernberg I & Dijkshoorn L (2006). Naturally Transformable Acinetobacter sp. Strain ADP1 Belongs to the Newly Described Species Acinetobacter baylyi. *Appl. Environ. Microbiol.* 72, 932–936.

Vishniac W & Santer M (1957). THE THIOBACILLI, 12. *Bacteriol. Rev.* 21, 195–213.

Wallace GG, Smyth M & Zhao H (1999). Conducting electroactive polymer-based biosensors. *TrAC Trends Anal. Chem.* 18, 245–251.

Wang F, Zhou J, Li Z, Dong W, Hou Y, Huang Y & Cui Z (2015). Involvement of the Cytochrome P450 System EthBAD in the N-Deethoxymethylation of Acetochlor by Rhodococcus sp. Strain T3-1. *Appl. Environ. Microbiol.* 81, 2182–2188.

Wang G, Meng K, Luo H, Wang Y, Huang H, Shi P, Pan X, Yang P & Yao B (2011). Molecular cloning and characterization of a novel SGNH arylesterase from the goat rumen contents. *Appl. Microbiol. Biotechnol.* 91, 1561–1570.

Wang G, Wang Q, Lin X, Bun Ng T, Yan R, Lin J & Ye X (2016). A novel cold-adapted and highly salt-tolerant esterase from Alkalibacterium sp. SL3 from the sediment of a soda lake. *Sci. Rep.* 6, 19494.

Wang YP & Gu J-D (2006). Degradability of dimethyl terephthalate by Variovorax paradoxus. *Ecotoxicology* 15, 549–557.

Waste Water Treatment Online (2014). Project Focus: Hall claims UK first in Water Treatment. *Waste Water Treat. Online*. Available at: http://wwtonline.co.uk/features/project-focus-hall-claims-uk-first-in-water-treatment [Accessed August 12, 2016].

Water UK (2015). *Industry facts and figures 2015 – Company volumetrics*, Available at: http://www.water.org.uk/publications/reports/industry-facts-and-figures-2015 [Accessed August 23, 2016].

Wedgwood MA & Bailey SER (1988). The inhibitory effects of the molluscicide metaldehyde on feeding, locomotion and faecal elimination of three pest species of terrestrial slug. *Ann. Appl. Biol.* 112, 439–457.

Whitman WB, Coleman DC & Wiebe WJ (1998). Prokaryotes: The unseen majority. *Proc. Natl. Acad. Sci.* 95, 6578–6583.

Wilder RS (1947). Process for preparing metaldehyde. *US Pat. 2426961*.

Willems A, De Ley J, Gillis M & Kersters K (1991). NOTES: Comamonadaceae, a New Family Encompassing the Acidovorans rRNA Complex, Including Variovorax paradoxus gen. nov., comb. nov., for Alcaligenes paradoxus (Davis 1969). *Int. J. Syst. Evol. Microbiol.* 41, 445–450.

Wilson MJ, Wilson DJ, Burch G & Cotching B (2014). Effects of simulated rainfall on two contrasting metaldehyde slug baits. *Agron. N. Z.* 44, 95–101.

Woese CR & Fox GE (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc. Natl. Acad. Sci.* 74, 5088–5090.

Yan Y & Moult J (2006). Detection of operons. *Proteins Struct. Funct. Bioinforma.* 64, 615–628.

Yeom S, Mutlu BR, Aksan A & Wackett LP (2015). Bacterial Cyanuric Acid Hydrolase for Water Treatment. *Appl. Environ. Microbiol.* 81, 6660–6668.

Young NS, Ioannidis JPA & Al-Ubaydli O (2008). Why Current Publication Practices May Distort Science. *PLOS Med* 5, e201.

Yuquan X, Wei Z, Ming C, Min L, Junming L & Xuanjun F (1999). Isolation and identification of a phenol-degrading bacterial strain. *Acta Sci. Circumstantiae* 20, 450–455.

Zhang H, Wang C, Lu H, Guan W & Ma Y (2011a). Residues and dissipation dynamics of molluscicide metaldehyde in cabbage and soil. *Ecotoxicol. Environ. Saf.* 74, 1653–1658.

Zhang H-J, Zhou Q-W, Zhou G-C, Cao Y-M, Dai Y-J, Ji W-W, Shang G-D & Yuan S (2012). Biotransformation of the Neonicotinoid Insecticide Thiacloprid by the Bacterium Variovorax boronicumulans Strain J1 and Mediation of the Major Metabolic Pathway by Nitrile Hydratase. *J. Agric. Food Chem.* 60, 153–159.

Zhang HY, Wang C, Xu PJ & Ma YQ (2011b). Analysis of molluscicide metaldehyde in vegetables by dispersive solid-phase extraction and liquid chromatography-tandem mass spectrometry. *Food Addit. Contam. Part A* 28, 1034–1040.

Zhang XY & Dai XF (2006). Degradation and Determination of the Residue of Metaldehyde in Tobacco and Soil. *Chin. J. Pestic. Sci.* 4, 344–8.