PHD

Characterisation of the enzymes involved in methanol dissimilation in Bacillus methanolicus: Assessment of enzyme-based biohydrogen production for a cell-free biofuel cell

Mozzanega, Philippe

Award date:
2010

Awarding institution:
University of Bath

Characterisation of the enzymes involved in methanol dissimilation in *Bacillus methanolicus*

Assessment of enzyme-based biohydrogen production for a cell-free biofuel cell

Submitted by

Philippe MOZZANEGA

for the degree of Doctor of Philosophy
University of Bath,
Department of Biology and Biochemistry

November 2010

Characterisation of the enzymes involved in methanol dissimilation in *Bacillus methanolicus*

Assessment of enzyme-based biohydrogen production for a cell-free biofuel cell

Submitted by

Philippe MOZZANEGA

for the degree of Doctor of Philosophy

University of Bath,

Department of Biology and Biochemistry

2010

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my (most excellent!) supervisors, Professor Michael J. Danson and Dr David W. Hough, for the opportunity to do my PhD in their lab, the support and guidance throughout the last four years, dragging me into their offices at times. I would also like to thank Professor Anthony Atkinson, Dr Kirstin Eley and Dr Roger Cripps of TMO-Renewables Ltd, for the helpful discussions throughout. I would not have been able to do this research without contributions towards the funding of this PhD by the BBSRC and TMO-Renewables Ltd. Kind personal regards to Dr Simon Warner (Syngenta) who has certainly ignited my desire to embark upon doing a PhD.

A great big Thank You goes to past and present members of the Centre for Extremophile Research (CER). Special mention to (and in no particular order) (Drs) Karl 'Shrek' Payne, Sylvain Royer and Tracey 'knee-rubbing' Goult for the helpful discussions and general help, and Dr Albert Bolhuis for sharing his simple protocol for *Bacillus* transformation. Huge thanks to Jonathan Extance ("Dr"-in waiting): you've been a brilliant project student and of great help for the enzymology (you'll have your *Daisy* one day, even if just for the front badge). Gratitude is also extended to the now Drs Sarah Lamble and Winnie Wu, and soon to be Drs Natasha Voina and Nia Hughes. Chris Vennard and Carolyn Williamson: it has been a pleasure to "share" the bench next to you and Natasha under the watchful eye of the Meerkat Galaxy (Carolyn, thanks for the curtains). Also, thanks to Doc, Marty and the DeLorean (*Back To The Future II*) for providing me with a good introduction for seminars.

Finally, thanks to my post-grad companions in hardship and others met over these last four years: in no particular order (but still) Phil, Kathy, Freddie, Jennie, Fred, Iwan, Philippa, John, Zoe, Kiri, Vicky, Klaus, Steph, Maria (too many to cite). To my friends back home, KptN, MéchantBarbu, MoutMout, Goyou, Fox, JP and JM: to more feasts. And last but certainly not least, Hayley: Thank You for the support, the fun and Love, for being there for me, always, far but still so close.

- À mes parents -

"*I prefer to have a bottle in front of me than a frontal lobotomy*" –
M^{me} Chabouret, 19th Nov. 1953.

CONTENTS

"*Ever tried. Ever failed. No matter. Try again. Fail again. Fail better.*" –
Samuel Beckett, Worstward Ho (1983)

ABBREVIATIONS

°C	degree Celsius
A_{340nm}	absorbance at 340nm (e.g.)
Amp / Ap^r	ampicillin / ampicillin resistance gene (*bla*)
APS	ammonium persulphate
ATCC	American Type Culture Collection
BGSC	Bacillus Genetic Stock Centre
BLAST	Basic Local Alignment Search Tool
Bm	*Bacillus methanolicus*
BmMDH	methanol dehydrogenase from *B. methanolicus* (chromosome-bourne)
pBMMDH	methanol dehydrogenase from *B. methanolicus* (plasmid-bourne)
bp	base pair
Bs	*Bacillus subtilis*
BSA	bovine serum albumin
ca.	*circa*
carb	carbenicillin
CDS	coding sequence
CO_2	carbon dioxide
dgPCR	degenerate PCR
DMSO	dimethyl sulphoxide
DNA	deoxyribo-nucleic acid
dATP	deoxyadenosine triphosphate
dNTPs	deoxynucleoside triphosphates
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen (GmbH)
DTT	dithiothreitol
Ec	*Escherichia coli*
E.C.	Enzyme Commission number
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
et al.	and others (*et alia*)
g	g-force
gDNA	genomic DNA
GCHE	glucose casein hydrolysate
genta	gentamycin
gge	gallon gasoline equivalent ($1.3\ 10^8$ Joules)
GSH, GSSG	glutathione (reduced, oxidised)
Gth	*Geobacillus thermoglucosidasius*
Gtoe	Gigaton oil equivalent
H^+ / H_2	proton / dihydrogen

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
i.e.	*id est*
IPCC	Intergovernmental Panel on Climate Change
IPTG	isopropyl β-D-1-thiogalactopyranoside
kan	kanamycin
kb	kilobase
kDa	kilodalton
LB	lysogeny broth medium
MFC	microbial fuel cell
MilliQ (water)	ultrapurified water
M_r	relative molecular mass
NAD^+ / NADH	nicotinamide adenine dinucleotide (oxidized / reduced)
NB	nutrient broth
NCIMB	National Collection of Industrial, Food and Marine Bacteria
OD_{600}	optical density at 600nm (e.g.)
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PNS	purple non-sulphur (bacteria)
RBS	ribosome binding site
RE	restriction endonuclease (or restriction enzyme)
rpm	revolutions per minute
sdm	site-directed mutagenesis
SDS	sodium dodecyl sulphate
strep	streptomycin
TAE	Tris acetate EDTA
Taq	*Thermophilus aquaticus* (DNA polymerase)
TB	Terrific broth
TEMED	*N,N,N',N'* tetramethylethylenediamine
Tris	Tris(hydroxymethyl)methylamine
TS	tryptone soya broth (casein soybean digest)
TY	tryptone yeast extract broth
U	unit (of enzyme activity)
US$	US dollar
vs.	*versus*
v/v	volume per volume
w/v	weight per volume
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
YE	yeast extract

ABSTRACT

Hydrogen is a clean, high energy content, low emission energy carrier. The sustainable production of hydrogen from renewable sources is of particular interest in a world with greater emphasis on green credentials for its growing energy use.

The generation of molecular hydrogen by both whole cells and isolated enzymes has been proposed as a means of fuelling conventional fuel cells. The use of isolated enzymes to produce H_2 has been the subject of few studies, with the construction od synthetic enzymatic pathways. Here, starting from methanol feedstock, a cheap and high energy density fuel, we propose a three enzyme synthetic pathway built around enzymes from methanol-utilising thermophiles. Methanol requires only three successive dehydrogenase enzymes to remove all of its H atoms. From the thermophilic methanol-utiliser *Bacillus methanolicus*, a methanol dehydrogenase and a formaldehyde dehydrogenase were identified, cloned and expressed, while a formate dehydrogenase was sourced from *Geobacillus thermoglucosidasius*. Enzymes were expressed heterologously in *E. coli* as well as homologously in a *Bacillus subtilis*-based system.

Soluble expression was achieved for the methanol dehydrogenase and its activator protein. The enzyme was only active in the backwards direction (formaldehyde reductase) in vitro. The kinetics parameters were estimated at $K_M = 1.2$ mM and $V_{max} = 2.5$ U.mg^{-1} and were similar for the heterologously and homologously-expressed enzyme. The formaldehyde dehydrogenase and formate dehydrogenase could not be purified in a soluble form and no kinetics data could be obtained.

Chapter 1- Introduction

1.1 *"The Stone Age didn't end for the lack of stones"*

1.1.1 General considerations

The current energy mix that fuels the global economy is mostly composed of fossil sources. According to the International Energy Agency (IEA), energy demand is expected to double from 12 Gtoe (Giga ton oil equivalent) in 2008 to 25 Gtoe by 2050, CO_2 emissions production equivalent rising from 7 to 14 Gton of carbon, with one-third of this energy being sourced from renewables according to forecasts (http://www.worldenergy.org). At the same time, oil production is expected to peak at around 2040 if not earlier, while environmental policies are aiming at reducing drastically greenhouse gases and CO_2 emissions by 2050 (Kyoto protocol, 1997; Stern review, 2006, Copenhagen Accord, 2009) (Figure 1.1).

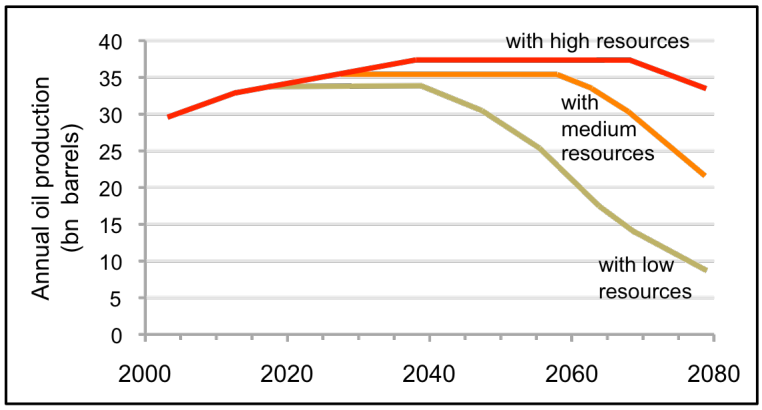


Figure 1.1. The three oil peak scenarios. (Adapted from Kerr, 2007).

The debate over the energy supply has been revived by the recent steep increases in oil prices. However, oil remains for now the fuel of choice. But as consumption still increases, so does the gap between production and consumption (Figure 1.2). Production capacities of 10 million barrels a day are possible up until 2040, if

reserves are proven and materialise, with an expected 2.25 to 3.9 trillion barrels potentially recoverable.

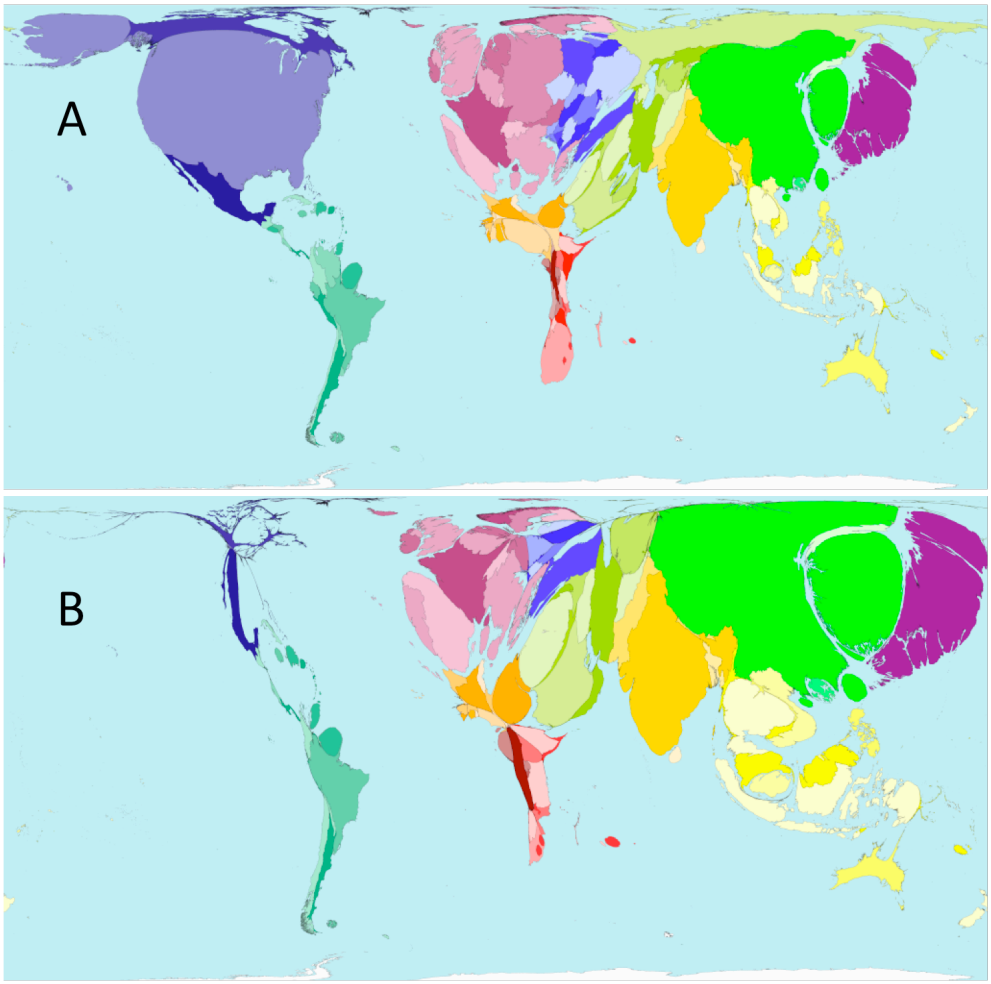


Figure 1.2. Representation of the world territories with (A) land area proportional to fuel use (2001), and (B) fuel use increase (1980-2001). (Adapted from www.worldmapper.org).

Following eras of solid fuel (wood, coal) and liquid fuel (oil) consumption, mankind is transitioning into an era of mixed fuels consumption, and especially gaseous fuel, namely natural gas and hydrogen. Yet, as Sheikh Zaki Yamani, the larger-than-life Saudi oil minister of the 1970s, put it during the first oil crisis in 1973, *"the Stone Age didn't end for the lack of stones"*.

Hydrogen is the ultimate fuel to replace today's fossil fuels. It has the highest energy per mass than any other fuel, with a low heat value (LHV) of 241 $kJ.mol^{-1}$ (2.4 x that of methane, 2.8 x that of gasoline, 4 x that of coal), but also the lowest per volume (Table 1.1.).

Table 1.1. Energy content and potential CO_2 emissions of different types of fuels. ([#] wastes: organic fraction of municipal and industrial wastes and agricultural residues; [†] potentially nil). (UK National Physical Laboratory (NPL)).

	Carbon content (kg_C/kg_{fuel})	**Energy content** (kWh/kg_{fuel})	**CO_2 emissions** (kg_{CO2}/kg)	(kg_{CO2}/kWh)
Coal	0.75	2.8 – 8.3	4.5	(0.37)
Oil	0.85	12.5	3.2	
Natural gas	0.75	11.8	2.7	(0.23)
LPG	0.82	12.8	-	
Gasoline	0.9	13.0	2.9	(0.27)
Diesel	0.86	11.9	3.0	(0.24)
Biodiesel	0.86	10.9	-	
Ethanol	0.52	8.3	-	
Methanol	0.375	6.4	-	
Wood	24	2.5 – 4.4	1.5	(0.18)
Wastes[#]	-	2.5 – 5.0	-	
^{235}Uranium	0	$22.8\ 10^6$	-	
Hydrogen	0	39.4	[†]	

The move towards cleaner energy systems is supported by four factors. As outlined in the Kyoto protocol, there is now a political obligation derived from an environmental need to reduce CO_2 emissions. Oil nowadays represents over a third of the world primary energy use, creating worldwide energy dependence. Furthermore, oil is a scarce commodity with the end of constant supply modelled for *ca*. 2050. Finally, the incorporation into models of the burgeoning energy demands of emerging economies, many of which have an enormous and growing population, it is clear that these countries are putting further strains on these finite fossil resources (for example, China has now 100 times as many cars as 20 years ago) (also see Figure 1.2-B).

The transition from a fossil fuel-powered economy to one sustained by new and more environmentally-friendly energy sources is being driven by the prospect of declining oil reserves. The speed at which biomass-derived fuels (e.g. methyl/ethyl ethers, ethanol, hydrogen) are implemented will be determined by the adaptability of the existing energy chain and the level of resistance to the emergence of the new value chains, the minimisation of costs and the ability to retrofit existing energy infrastructures to new fuels.

1.1.2 Climate change

Earth's 4.5 bn year history is one long story of climate change. Yet, none has been so intense and rapid as the one we may be witnessing today. The 2007 Intergovernmental Panel on Climate Change (IPCC) climate change model suggests that by 2100, an increase of between 2.5 and 6 °C in the world's global temperature is possible (www.ipcc.ch). Even if these figures have since been reviewed, a question one may ask about climate change is "how can a relatively small change in temperature be the trigger for dramatic changes in the patterns of our climates?" Answering that question can be approached from a thermodynamic point of view. An increase of 1 °C in surface water temperature equates to an increase of 1.5 $g.m^{-3}$ in the concentration of water vapour in the atmosphere (or 4.5 Gtons), releasing an energy of over 10^{11} MJ. The global climate system is displaced from its equilibrium state following a non-linear phase displacement of non-equilibrium states. A change of 5 °C in temperature results in a 300% increase in the atmospheric water vapour content. This, in combination with an increased atmospheric concentration of greenhouse gases, may trigger drastic changes in the behaviour of our climates, such as more sudden and intense precipitations, and the warming of areas of lands and seas while other are cooled. This in turn plays a part in affecting long-term weather patterns such as the cloud cover, humidity and how heat energy from the oceans interacts with the atmosphere.

1.1.3 (Geo-)Political aspects

On top of the climate change issue, concerns over energy supply security and increasing prices of energy services are having an impact on policy making. Oil is the largest primary fuel accounting for over one-third of the global primary energy mix. The transport sector alone accounts for 18% of the primary energy use, for 17% of the global CO_2 emissions, and is responsible for 20% of the projected increase in both energy demand and greenhouse gases over the next 20 years; moreover, it is also still reliant on oil for 95% its needs (Ball and Wietschel, 2009). In his January 2003 State of the Union address, President G. W. Bush declared, "America is addicted to oil". A US$ 1.2bn 5 years initiative was launched to reverse

America's dependence on foreign energy imports. The Hydrogen Fuel Initiative (HFI), part of a US Department of Energy (DoE) Hydrogen program, aims at developing and accelerating research and technology to overcome obstacles towards taking the hydrogen fuel cell from the laboratory to commercially-viable hydrogen-powered fuel cells to power cars, trucks, homes and businesses with no (or little) pollution or greenhouse gases (www.hydrogen.energy.gov). It integrates hydrogen activities between a number of DoE program offices and implement activities critical to the HFI. Through partnerships with the private and academic sectors, Advanced Energy Initiative (AEI) launched in 2006 seeks to develop hydrogen fuel cells and infrastructure technologies needed to make it practical and cost-effective. A key element is to develop advanced hydrogen production and delivery technologies that can supply tomorrow's energy transportation systems with affordable hydrogen with significantly reduced or near-zero emissions.

Actual technologies can produce and deliver enough hydrogen to meet the demands of the US's refineries and chemical industries. But to meet the expectations laid in the HFI or the European Commission Fuel Cells and Hydrogen Joint Technology Initiative (FCH JU), and supply the much larger volumes needed for a hydrogen economy, the costs of producing, delivering and storing hydrogen must be reduced. Targets in term of hydrogen cost (US$ 3 gge), storage capacity, fuel-cell cost (US$ 30/kWh) and durability have been set (Figure 1.3).

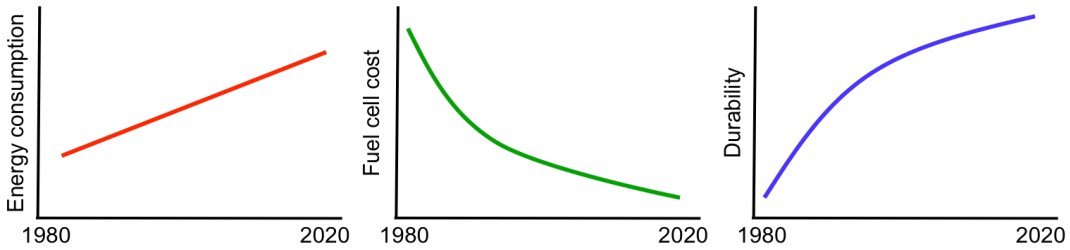


Figure 1.3. Predicted evolutionary trends of hydrogen-related items. (Adapted from www.eere.energy.gov).

Hydrogen production costs are getting closer to US$ 1.50/kg. But to achieve the targets fully and to allow hydrogen to compete on a par with fossil fuels, a leap forward in production technologies is needed, as well as resolving storage and delivery issues.

1.2 Hydrogen as a fuel - Biohydrogen

1.2.1 Conventional hydrogen production

It is important to note that, in the same way as electricity, hydrogen is not an energy source but a secondary form of energy, and has to be manufactured. Hydrogen can store and deliver usable energy. The potential of hydrogen as an energy carrier that can provide pollution-free, carbon-free power makes it a promising player in the energy future.

Global hydrogen production stands at 0.1 Gton, representing an energy of 2,500 TWh with a market value of over US$ 140 bn/year (Marbán and Valdés-Solís, 2007). However, 98% of this hydrogen is produced using conventional CO_2-emitting technologies (steam reforming and/or gasification of fossil fuels, water hydrolysis). Furthermore, to fulfil the needs of the transport industry alone, current hydrogen production levels would have to be multiplied by a factor of 10-20. Given the current production schemes, this would utilise all the natural gas available, and at the same time increase global carbon emissions by over one-third.

There are numerous research projects, and even commercial ventures, aimed at producing hydrogen using renewable, non-polluting technologies, or using untapped energy sources. Examples might be the production of hydrogen by water splitting using nuclear, wind, solar or geothermal energy (O'Brien *et al*., 2010; Stoots *et al.*, 2010; Ullberg *et al.,* 2010; Steinfeld, 2002; Liu *et al*., 2010; Ignason *et al*., 2008; Tolga Balta *et al*., 2009), or by biological means; the latter technology, generating biohydrogen, is explored in the next section.

1.2.2 Biohydrogen

Interest in the industrial application of microorganisms to generate energy via hydrogen emerged in the 1970s as a result of the then oil crisis. Hydrogen produced from biological processes was then termed Biohydrogen. Biological approaches to hydrogen could contribute to large-scale production of hydrogen as various

microorganisms can produce H_2 under normal conditions of temperature and pressure from readily available substrates. Biohydrogen is in effect 'carbon neutral' as its substrates are carbohydrate compounds produced by photosynthetic fixation of atmospheric CO_2. Feedstocks for biohydrogen production by microorganisms include (but are not limited to) agricultural residues (Argun *et al*., 2009; Akutsu *et al*., 2009), effluents from industrial processes of the food industry such as sugar refining, confectionery, olive processing, palm oil, cheese production or kitchen waste (Yetis *et al*., 2000; Penfold *et al.*, 2003; Eroğlu *et al*., 2004; Chong *et al*., 2009; Davila-Vazquez *et al*., 2009; Rosales-Colunga *et al*., 2010; Jayalakshmi *et al*., 2009). Lloyd *et al.* (2002) even reported hydrogen production in *Giardia intestinalis*, a protozoal intestinal parasite. Microbial processes are therefore particularly adapted to biohydrogen generation. Furthermore, the ability to evolve H_2 is relatively widespread in the microbial world and biological systems provide a wide range of approaches for the production of hydrogen: direct and indirect biophotolysis, photo-fermentation and dark-fermentation.

1.2.2.1 Photobiological hydrogen

Photobiolysis

Photobiological production of hydrogen from water (photolysis) is the conversion of solar energy into chemical energy by splitting water into H_2 and CO_2. Many microorganisms can harness solar energy for biomass generation, and several photosynthetic ones can evolve H_2 and be of use in a context of biohydrogen production.

Photoautotrophic microorganisms produce H_2 by direct or indirect photolysis. Photolysis uses the same processes happening in plant photosynthesis, with adaptation for H_2 generation instead of O_2 and carbon biomass. In direct photolysis, photosynthesis in microalgae (green algae) and cyanobacteria (blue-green algae) involves the absorption of light by two distinct photosystems (PSI and PSII) operating in series, with PSII involved in water splitting and O_2 and H_2 evolution while PSI generates reducing equivalents for CO_2 reduction and ATP generation. The electrons removed from water are used in CO_2 reduction (Miyake *et al*., 1999).

Green algae under anaerobic conditions can either use H_2 as an electron donor in the

fixation of CO_2 or evolve H_2. Light energy absorbed by PSII generates electrons. These are passed onto ferredoxin via PSI. Hydrogenase accepts the electrons directly from the reduced ferredoxin and generates H_2. *Chlamydomonas reinhardtii* has been the organism of choice in biophotolysis, with rates of up to 8 mmol $H_2.L_{culture}^{-1}$ (Kosourov *et al*., 2002).

Cyanobacteria (blue-green algae) are a diverse group of Gram-positive photoautotrophic microorganisms containing photosynthetic pigments (chlorophyll *a* (chl *a*), carotenoids). They use indirect biophotolysis to evolve H_2, as they possess a similar type of photosynthetic machinery as higher plants, with CO_2 being fixed via the Calvin cycle to synthesise sugars generating biomass, and the carbohydrates are fermented in a second process to produce H_2.

Biophotolysis can also involve nitrogenase and hydrogenase enzymes, which can enhance H_2 evolution. However, the photosynthetically-generated O_2 is detrimental to the activity of nitrogenase and favorises H_2-uptake. Nitrogenases in cyanobacteria are localized in heterocysts, which provide O_2 protection, and therefore indirect biophotolysis compartmentalizes, spatially and/or temporally, the O_2 production from the H_2 production.

Metabolic engineering of microalgae has proved more difficult than that of cyanobacteria due to the greater complexity of eukaryotic genetics. The main area of research for improvement is factors in light conversion efficiency, as well as oxygen-tolerant hydrogenases. The cyanobacterial *Anbaena* spp. has been shown to achieve *ca*. 100 μmol $H_2.mg_{chl\ a}^{-1}.h^{-1}$, for a light conversion efficiency (to H_2) of less than 4% (Dutta *et al*., 2005; Yoon *et al*., 2006). Rates were increased up to seven-fold in mutants lacking the uptake-hydrogenase activity (Masukawa *et al*., 2002), yet under non-artificial light conditions the light energy conversion dropped to 0.1% (Borodin *et al*., 2000).

Purple non-sulfur bacteria

Purple non-sulfur (PNS) bacteria (*Rhodobacter*, *Rhodopseudomonas*, *Rhodospirillum*) produce H_2 under photoheterotrophic conditions (light, anaerobic, organic electron donor). H_2 is produced anaerobically by the nitrogenase enzyme (Vignais, 1985). Nitrogenase-mediated H_2 formation is irreversible, as opposed to

the hydrogenase-mediated H_2 production that is inhibited by its own hydrogen production. Uptake-hydrogenase deficient mutant strains have allowed a more than doubling of H_2 production yields (Jahn *et al*., 1994; Worin *et al*., 1996; Öztürk *et al*., 2006). Nitrogenase reduces $2H^+$ to H_2 with the concomitant oxidation of electron carriers, although any electron sink will compete and affect the yield of H_2 production. Identification and mutagenesis of these competing reducing processes, such as poly-β-hydroxybutyrate (PHB), further improved H_2 production. Double mutants lacking the uptake hydrogenase and PHB synthase saw a two-fold increase in H_2 production yield (Kim *et al*., 2006).

PNS bacteria constitute an attractive candidate for biohydrogen production as they can be used on a variety of organic acids and alcohols.

1.2.2.2 Fermentative hydrogen production – dark fermentation

The quantity and quality of substrates (sugars, simple and complex carbohydrates) available for fermentation are varied, being sourced in agricultural and food processing wastes and residues, some of higher value than other (e.g. confectionery, sugar refining or distillation). Fermentative bacteria can be employed to reclaim this often unused resource and generate energy in the form of hydrogen. A wide range of microorganisms, strict and facultative anaerobes (*Clostridia, Enterobacter* etc), aerobes (*Bacillus* etc), can be used as candidates.

Dark-fermentation refers to the production of biohydrogen by bacteria in the dark under anaerobic conditions from a carbon-rich substrate. The anaerobic fermentation of organic matter by microorganisms can generate high rates of H_2, although the accumulation of acetate or other reduced by-products of the fermentation process inhibit H_2 production (van Ginkel and Logan, 2005). Also, if photolysis can produce a pure exhaust of H_2, dark fermentation processes generate a mixture of H_2 and CO_2 primarily. Thermodynamically, 4 mol H_2/mol hexose is the theoretical maximum yield, as some of the substrate will be used for ATP generation (at a minimum of 1 mol/mol for survival) (Thauer, 1977).

Pure cultures are being used in the majority of studies. This has generated a large amount of literature on model organisms and has helped to understand their particular fermentative metabolism.

The formate hydrogen lyase complex (FHL)

The initial steps in dark fermentation lead to the generation of ATP, NADH and pyruvate, the end-products of fermentation of glucose, and it is the fate of pyruvate which determines the yield of H_2. Pyruvate can be the substrate for three different enzyme systems: the pyruvate:ferredoxin oxidoreductase (PFOR), the pyruvate:formate lyase (PFL) and the lactate dehydrogenase (LDH). Formate arises from mixed-sugar fermentation. In mixed-acid fermentation, performed by facultative anaerobes (e.g. *Escherichia coli*), the main enzymes are PFL and the formate:hydrogen lyase complex (FHL), the FHL multi-enzyme complex consisting of a formate dehydrogenase, an hydrogenase and electron transport proteins. In *E. coli* under fermentative conditions, these are Hyd3 (*hyc* operon) and FDH-H, while Hyd1 and Hyd2 are involved in H uptake and FDH-N is the respiratory FDH. Pyruvate is transformed to acetyl-coA and formate by PFL, and formate is cleaved into equimolar amounts of H_2 and CO_2 by FHL, with acetyl-coA contributing to ATP and NAD regeneration (via acetate and ethanol respectively). The theoretical yield is 2 mol H_2/mol glucose (Clark, 1989), but in practice yields of *ca*. 1 mol/mol are observed (Bisaillon *et al*., 2006). Production of lactate as a by-product of fermentation generally hinders the observed yield achieved, though this can be addressed by metabolic engineering to mutate or knock-out the lactate dehydrogenase (Taylor *et al.*, 2009). Furthermore, while the formate dehydrogenase reaction is practically irreversible in cleaving formate into H_2 and CO_2, the fate of H_2 produced and its recycling can further reduce yields. Hyd1 (*hya* operon) in *E. coli* is an uptake hydrogenase and is expressed under fermentative conditions to recycle H_2 produced from formate (Sawers *et al*., 1985). Increasing expression of FHL can also increase H_2 evolution. *E. coli* HD701 lacks the FHL complex repressor *hycA* (Δ*hycA*), and therefore possesses an upregulated FHL (Penfold *et al*., 2003; Yoshida *et al*., 2007). This is because HycA competes with formate for binding sites on the FHL. Penfold *et al*. (2003) observed a more than doubling of the H_2 evolution rate in *E. coli* HD701 compared to its parental strain MC4100, from 0.88 to 2.2 mL H_2.mg $_{\text{dry weight}}^{-1}$.L $_{\text{culture}}^{-1}$. Further increases in the yield of H_2 production were also sought. The *E. coli* hydrogenase and FDH_H being located on the periplasmic side of the cytoplasmic membrane, they are transported by the twin-arginine translocation (Tat) protein. The Tat translocase comprises the TatA-E

membrane proteins (Berks *et al*., 2003). Inactivation of the Tat export system in *E. coli* MC4100 yielded strains B1LK0 (Δ*tatC*) and DADE (Δ*tatABCD,* Δ*tatE*) with H_2 evolution rates of 1.70 and 1.75 mL H_2.mg $_{\text{dry weight}}^{-1}$.L $_{\text{culture}}^{-1}$ respectively (Penfold *et al*., 2006). However, the FTD701 strain (Δ*hycA*, Δ*tatC*) did not have a significant cumulative effect. *E. coli* is incapable of utilizing sucrose due to the absence of an invertase. The 70 kb pUR400 plasmid, which carries the genes for sucrose transport and metabolism (*scrK*, ATP-dependent fructokinase; *scrY*, sucrose-specific porine; *scrA*, enzyme IIScr of the phosphotrasnferase system of sucrose uptake; *scrB*, β-D-fructofuranoside fructohydrolase, *scrR*, repressor) was conjugated into *E. coli* HD701 (Δ*hycA*). With sucrose for substrate, the rate of H_2 evolution was 1.27 mL H_2.mg $_{\text{dry weight}}^{-1}$.L $_{\text{culture}}^{-1}$ for HD701 pUR400 while the parental strain HD701 evolves H_2 at 0.03 mL H_2.mg $_{\text{dry weight}}^{-1}$.L $_{\text{culture}}^{-1}$ for HD701 (Penfold and Macaskie, 2004).

Facultative anaerobes of the genus *Enterobacter* have been seen as candidates for industrial scale biohydrogen production. They too produce H_2 from formate but this is paired with the stimulation of an NAD regeneration pathway (Zhao *et al*., 2009; Zhang *et al*., 2009). Knockout of the FHL in *E. aerogenes* (Δ*hycA*, Δ*hybO*) was targeted as a mean of enhancing hydrogen evolution. Maximum specific hydrogen productivity was 50% higher in the double mutant (Δ*hycA*, Δ*hybO*) than in the wild-type strain (Zhao *et al*., 2009). In this organism, regeneration of NAD is coupled with the reduction of ferredoxin by an NADH:ferredoxin oxidoreductase (NFOR), and the reduced ferredoxin can subsequently transfer electrons to H^+ generating further H_2. The maximum theoretical yield is then of 10 mol H_2/mol glucose, vastly superior to the 4 mol/mol from fermentation processes alone (as described above).

In the anaerobic Clostridia, the NADH pathway is also used. Pyruvate is cleaved into acetyl-CoA and CO_2 by PFOR, transferring the electrons to a ferredoxin:hydrogenase producing H_2. *Clostridium butyricum* has a theoretical H_2 yield of 3.3 mol H_2/mol hexose (Chen *et al*., 2006), but practical yields are in the region of 2 mol/mol. *C. butyricum* isolated from palm oil mill effluent yielded 2.2 mol/mol glucose (Chong *et al*., 2009) and hydrogen evolution is positively proportional to *hydA* expression (encoding for hydrogenase) (Wang *et al*., 2008a; 2008b). *C. acetobutylicum* ATCC824 and its mutant M5, which lacks a

megaplasmid responsible for butanol and acetone production, evolved H_2 at respective rates of 1.79 and 2.64 mol/mol glucose, with the increased yield in M5 due to a decrease in biomass production.

Use of thermophilic organisms

Recently, interest has stemmed from thermophiles and hyperthermophiles that thrive and produce H_2 at temperatures above 60 °C as they have been reported to be capable of higher yields than mesophiles (Hallenbeck, 2005). Yields of 3.2-3.7 mol/mol glucose have been reported for *Caldicellulosiruptor saccharolyticus* (Kadar *et al*., 2007).

The marine genus *Thermotoga* (anaerobe) contains some of the most thermophilic bacteria, most growing at temperatures between 60 and 70 °C, with some growing at above 90 °C (Huber *et al*., 1986). A source for a variety of thermophilic enzymes, *Thermotoga* has delivered highly thermostable hydrolytic enzymes. Furthermore, *Thermotoga* accounts for some hydrogen producers. Carbohydrates are fermented to lactic acid, and protons acting as electron acceptors are further reduced to H_2; the overall hydrogen production yield is close to the theoretical maximum of 4 mol/mol glucose. In *Thermotoga*, glucose is oxidized twice, leading to the formation of NADH and ferredoxin^{red}. Reduced ferredoxin is a strong enough reducing agent to reduce H^+ to H, while NADH alone is not ($\Delta G_0{}' = 19.3$ kJ.mol^{-1}). *Thermotoga maritima* has been shown to use its [Fe-Fe]-hydrogenase in which the oxidation of ferredoxin is used to drive the unfavourable oxidation of NADH to produce H_2 (Schut and Adams, 2009). *Thermotoga neapolitana* has been shown to achieve hydrogen yields of 3.9 mol H_2/mol glucose (d'Ippolito *et al*., 2010). The difference from the theoretical maximum of 4 mol/mol has been attributed mainly to the production of lactic acid, affecting the availability of pyruvate/NADH for the hydrogenase. In cultures where the ratio of acetic to lactic acid was 2:1, the yield of H_2 was only 2.6 mol/mol glucose. Also, accumulation of hydrogen in culture vessels inhibits its own production, as ferredoxin and NADH become increasingly reduced with increasing partial pressure of hydrogen, leading to the favoured production of lactic acid.

The hydrogenase systems of other hyperthermophiles have also been investigated

with a view to bettering H_2 production. The FHL complex of the hyperthermophilic archaeon *Thermococcus litoralis* consisting of the formate dehydrogenase *fdhAB* and a [Ni-Fe]-hydrogenase *mhyCDEFG* has been investigated with potential use for H_2 evolution (Takàcs *et al*., 2008). Similarly, *Pyrococcus furiosus* has been shown to contain three hydrogenases (Adams *et al*., 2001). Cao *et al*. (2009) reported hydrogen production by *Thermoanaerobacter thermosaccharolyticum* from sugar from acid-hydrolysed corn stover.

Complex substrates and mixed inoculums

Several mesophilic, thermophilic and hyperthermophilic Clostridia have the ability to use complex carbohydrates (starch, cellulose etc). *Clostridium thermocellum*, an thermophilic cellulolytic and ethanogenic bacterium capable of directly converting cellulosic substrate into ethanol, was used with delignified wood fibers as substrate and evolved H_2 at 1.6 mol/mol glucose equivalent (Levin *et al*., 2006), and can use dextrins amongst others. Also, *Caldicellulosiruptor saccharolyticus* is able to use xylose and paper sludge (Kadar *et al*., 2007).

Mixed cultures offer the advantages of compiling the advantages of specific strains. Hsiao *et al*. (2008) used mixed *Clostridium* cultures to evolved H_2 from molasses at rates 200 times higher than individual strains. Sludges from municipal or brewery waste (Abreu *et al*., 2009), mixed cultures of cellulose hydrolytic bacteria (*Cellulomonas* sp and *Cellulomicrobium cellulans*) (Lo *et al*., 2009), or dairy manure compost (Guo *et al*., 2010) have all been investigated for biohydrogen production purposes. Mixed cultures of thermophiles have also been used, offering flexibility in treating difficult and more complex substrates. Cheong *et al*. (2006) use filtered raw sludge from a cattle manure treatment plant, at 55 °C, with synthetic wastewater. *Thermanaerobacterium* sp. (*T. thermosaccharolyticum*, *T. polysaccharolyticum*, *T. aoteatoense*) contributed to the main species in the inocula. Zhao *et al*. (2009) used an extremely thermophilic (70 °C) mixed culture, including *Thermoanemobacter*, *Thermoanemobacterium* and *Caldanaerobaceter* species, resulting in high yields of simultaneous hydrogen and ethanol production (1.58 mol H_2/mol glucose and 0.9 mol EtOH/mol glucose respectively). *Caldicellulosiruptor saccharolyticus* was used individually for thermophilic hydrogen production from samples of sweet sorghum, surgarcane bagasse, wheat

straw and maize leaves that had not been pre-treated, yielding up to 3.8 mol H_2.mol glucose (Ivanova *et al*., 2009).

1.2.2.3 Two-stage processes - combining dark and light fermentation

Unlike with bio-ethanol production (Taylor *et al*., 2009), no single microorganism system can produce competitive yields of H_2. Multi-organism systems and combining strategies in two-stages components offer increased return on the inputs that would make this a realistic option of future energy generation.

There are numerous possibilities of combinations for two-stage biohydrogen processes, with factors of compatibilities and spatial and/or temporal separation to consider. However, the vast majority of systems tested combine a PNS second stage with a fermentative first stage.

Redwood and Macaskie (2006) set up a two-stage process combining dark fermentation and PNS bacteria. *E. coli* HD701 fermented glucose in the first stage to produce H_2, organic acids (acetic acid) and CO_2. In the second stage, the PNS *Rhodobacter sphaeroides* furthered the hydrogen yield by photo-converting the acetic acid into more H_2. Such dark fermentation-photofermentation (DF-PF) two-stage coupled systems have been trialed in different configurations. The first stage can be operated by mesophilic or thermophilic strains, with dark fermentation of carbohydrates into reduced organic acids, but it may or may not produce H_2, while the second stage of PNS bacteria ensures the conversion of the fermentation products disposing of the reducing power generated as further H_2.

Different first stage associations with a PNS second stage have been trialled. Microalgae (*C. rheinhardtii*) and PNS *Rhodospirullum* have been combined (Melis and Melnicki, 2006) as have been cyanobacteria and PNS. The obligate anaerobic (*Clostridium butyricum*) and PNS Rhodobacter have also been used in systems, with *C. butyricum* and *R. sphaeroides* using tofu or rice wastewaters or glucose for yields of 1.1-1.64 mol H_2/mol glucose (Kim *et al*., 2001), *C. butyricum*, *Enterobacter* and *Rhodobacter* spp. on sweet potato starch residues for a yield of 7.2 mol H_2/mol glucose (Yokoi *et al*., 2002), and *Caldicellulosiruptor* and *R. capsulatus* associated on potato peel hydrolysate with a yield of 5.6 mol H_2/mol glucose (Claassen *et al*., 2004). Dark fermentation and PNS have been assessed on glucose

feed with *Lactobacillus delbruckii* and *R. sphaeroides* (7.1 mol H_2/mol glucose) (Asada *et al*., 2006), while Redwood and Macaskie (2007) have been improving their *E. coli* HD701/*R. sphaeroides* coupled system, taking the H_2 yield from 0.4 to 2.4 mol H_2/mol glucose, with further system improvements looking at 10.1 mol H_2/mol glucose. Mixed cultures have also been used in first stages and associated in a PNS second stage (*Rhodopseudomonas*, *Rhodobacter*). These mixed cultures mainly consisted of sludges from agricultural wastes (cow or poultry manure, palm oil mill effluents, kitchen waste etc) (Ali Hassan *et al*., 1997; Eroğlu *et al*., 2006). Recently, Waks and Silver (2009) developed a dual-organism system where *Saccharomyces cerevisiae* was engineered to possess a formate-overproducing pathway taken from the *E. coli* PFL. Formate over-production in the yeast was even enhanced by the introduction of the AdhE of *E. coli*, resulting in an 18-fold increase in formate production by *S. cerevisiae*. Hydrogen was produced from the formate-containing spent medium by *E. coli.*

1.2.2.4 Enzymatic Biohydrogen

The stoichiometric maximum yield of hydrogen from glucose is 6 mol H_2.mol^{-1}, or 12 mol H_2.mol^{-1} for glucose in aqueous solution (1).

$$C_6H_{12}O_6 + 6\,H_2O \longrightarrow 12\,H_2 + 6\,CO_2 \qquad (1)$$

However, as presented above, in biological systems only 25% (or 3 mol/mol) of that yield can be achieved. Therefore, the prospect of designing cell-free synthetic enzyme systems is attractive as it frees the system from non-productive by-products that hinder the overall H_2 yield. Synthetic biology is a recent development and is driven by the engineering of specific artificial pathways for specific applications.

The glucose dehydrogenase from *Thermoplasma acidophilum* (TaGDH) has been associated with the hydrogenase from *Pyrococcus furiosus* (PfHyd) and *in vitro* hydrogen evolution from a glucose substrate demonstrated (Woodward *et al*., 1996; 2000a). This synthetic pathway was enlarged with the inclusion of an amyloglucosidase from *Rhizopus* mold, along with commercial cellulase and xylanase enzymes and an invertase and β-glucosidase from *Thermoanaerobacter ethanolicus*. All were combined with TaGDH and PfHydrogenase, and hydrogen evolution measured *in vitro* from a variety of

carbohydrate substrates. Hydrogen evolution was best from glucose, although galactose, xylose, mannose, fructose, lactose, starch and steam-exploded aspen wood were also positively assayed in the system (Woodward *et al*., 2000a). Woodward *et al.* (2000b) also associated a glucose-6-phosphate and 6-phosphogluconate dehydrogenases with PfHyd, with glucose-6-phosphate (G6P) as substrate. The *in vitro* H_2 molar yield was 2 mol H_2/mol G6P. Yet in the same study, combining PfHyd with the enzymes involved in the oxidative pentose phosphate cycle, the molar yield of hydrogen was 11.6 mol H_2/mol G6P, or 98% of the stoichiometric maximum.

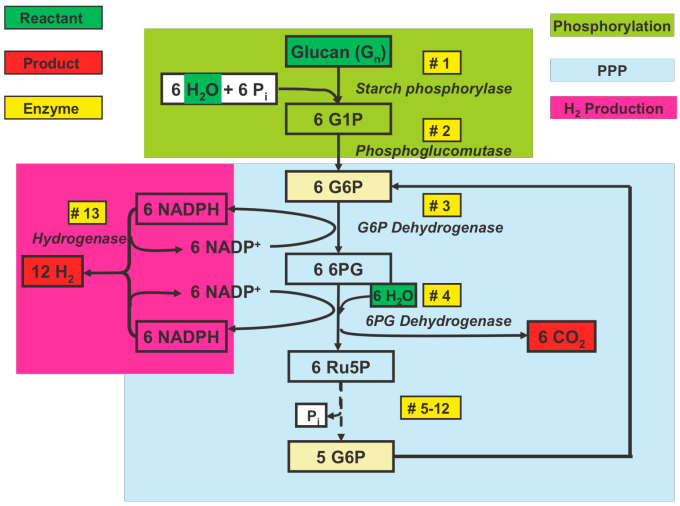
A synthetic enzymatic pathway consisting of 13 commercially-available enzymes for producing hydrogen from starch and water was recently reported (2) (Zhang *et al*., 2007) (Figure 1.4).

$$C_6H_{10}O_5 + 7\,H_2O \longrightarrow 12\,H_2 + 6\,CO_2 \qquad (2)$$

with starch being a polymer of $C_6H_{10}O_5$ (n), being converted to G1P, then G6P and through to hydrogen and CO_2 via the RuMP and a hydrogenase.

The pathway was divided into five task modules: #1 conversion of the substrate into glucose-1-phosphate, #2 generation of glucose-6-phosphate from glucose-1-phosphate (phosphoglucomutase), #3 production of NADPH via the two dehydrogenases of the oxidative part of the pentose phosphate pathway (PPP), #4 regeneration of glucose-6-phosphate from ribulose-5-phosphate via the eight enzymes of the non-oxidative part of the PPP, and #5 generation of H_2 from NADPH via a hydrogenase.

This synthetic enzyme pathway, which does not occur in nature, was assembled from 12 commercially-available mesophilic enzymes from animal, plant, yeast and bacterial sources and consisting in a starch phosphorylase, phosphoglucomutase, and the enzyme suite involved in the pentose phosphate pathway. The hydrogenase was from the hyperthermophilic archaeon *P. furiosus*. The measured yield of hydrogen generated was 8.4 mol H_2.mol^{-1} G6P (5.2 mol H_2.mol^{-1} glucose equivalent from starch), or 70% of the theoretical maximum.



E.C.	Enzyme Name	Reaction	Vender	Origin	Unit
2.4.1.1	glycogen phosphorylase	$(C_6H_{10}O_5)_n+P_i+H_2O\rightarrow(C_6H_{10}O_5)_{n-1}$+glucose-1-P	Sigma	rabbit muscle	10
5.4.2.2	phosphoglucomutase	G-1-P→G-6-P	Sigma	rabbit muscle	10
1.1.1.49	glucose-6-phosphate dehydrogenase	G-6-P+$NADP^+$→6-phosphogluconate+NADPH	Sigma	*S. cerevisiae*	1
1.1.1.44	6-phosphogluconic dehydrogenase	6-phosphogluconate+H_2O+$NADP^+$→ribulose-5-phosphate+NADPH+CO_2	Sigma	*S. cerevisiae*	1
5.3.1.6	ribose 5-phosphate isomerase	ribulose-5-phosphate→ribose-5-phosphate	Sigma	spinach	1
5.1.3.1	ribulose-5-phosphate 3-epimerase	ribulose-5-phosphate→xylulose-5-phosphate	Sigma	*S. cerevisiae*	1
2.2.1.1	transketolase	xylulose-5-phosphate+ribose-5-phosphate→sedoheptulose-7-phosphate+glyceraldehyde-3-phosphate	Sigma	E. coli	1
		xylulose-5-phosphate+erythrose-4-phosphate→fructose-6-phosphate+glyceraldehyde-3-phosphate			
2.2.1.2	transaldolase	sedoheptulose-7-phosphate+glyceraldehyde-3-phosphate→fructose-6-phosphate+erythrose-4-phosphate	Sigma	*S. cerevisiae*	1
5.3.1.1	triose-phosphate isomerase	glyceraldehyde 3-phosphate→dihydroxacetone phosphate	Sigma	rabbit muscle	1
4.1.2.13	aldolase	glyceraldehyde 3-phosphate+dihydroxacetone phosphate→fructose-1,6-bisphosphate	Sigma	rabbit muscle	1
3.1.3.11	fructose-1,6-bisphosphate	fructose-1,6-bisphosphate+H_2O→fructose-6-phosphate+Pi	[41]	*E. coli*	1
5.3.1.9	phosphoglucose Isomerase	fructose 6-phosphate→glucose-6-P	Sigma	*S. cerevisiae*	1
1.12.1.3	*P. furiosus* hydrogenase I	NADPH+H^+→$NADP^+$+H_2	[22.42]	*P. furiosus*	~70

Figure 1.4. The synthetic metabolic pathway for conversion of polysaccharides and water to H and CO_2. Abbreviations: PPP, pentose phosphate pathway; G1P, glucose-1-phosphate; G6P, glucose-6-phosphate; 6PG, 6-phosphogluconate; Ru5P, ribulose-5-phosphate. The enzymes are: #1 glucan phosphorylase; #2 phosphoglucomutase; #3 G6P dehydrogenase; #4 6PG dehydrogenase; #5 phosphoribose isomerase; #6 Ru5P epimerase; #7 transaldolase; #8 transketolase; #9 triose phosphate isomerase; #10 aldolase; #11 phosphoglucose isomerase; #12 fructose-1-6-bisphosphatase; #13 hydrogenase. (Zhang *et al.*, 2007).

1.3 Methylotrophs and methanol metabolism

1.3.1 Methylotrophic microorganisms

Methylotrophic bacteria are a diverse group of microorganisms capable of growth on reduced C_1 compounds such as methanol or methane as their carbon and energy source. They possess a number of specialised enzymes and specific metabolic pathways allowing them to grow on these reduced carbon substrates that have no carbon-carbon bonds, and to use these as their energy and carbon source for the synthesis of compounds and biomass from the C_1 substrate assimilation. The most common C_1 substrate for Methylotrophs is methanol, whereas some members have the ability to use methane, methylated amines, halogenated methanes and methylated sulphur species, which are ubiquitous in nature, either in addition to methanol or exclusively. They have been found in a variety of environments including in extreme pH and temperature (Hanson and Hanson, 1996; Bodrossy *et al*., 1997; Bowman *et al*., 1997; Dedysh *et al*., 2000).

Methylotrophic organisms have frequently been isolated from soil samples. However, the general techniques employed for isolation of methane- or methanol-growing organisms select for the fast-growing populations, being in most cases Gram-negative methylotrophic bacteria. The large diversity of slow-growing Gram-positive methylotrophic bacteria has been largely ignored, requiring specific growth conditions and screening techniques (such as for thermotolerant bacilli). General studies on the alcohol metabolism of actinomycetes and bacilli have highlighted the involvement of different alcohol dehydrogenase (ADH) enzyme systems.

Phylogenetically, methylotrophs are found amongst the Protobacteria and the high G+C or low G+C Gram-positive bacteria. Many are obligate methylotrophs (incapable of growth on multicarbon compounds), some are facultative methylotrophs, while some are restricted obligate methylotrophs (capable of poor growth on a restricted range of multicarbon compounds).

The first step in assimilating carbon is the oxidation of the C_1 species to the level of formaldehyde. In methylotrophs, formaldehyde is a key central intermediate

metabolite. Part of it is further oxidised to CO_2 for energy production, and part of it is assimilated in the cell cycle for the net synthesis of the C_2 and C_3 skeleton via the serine cycle, the ribulose monophosphate cycle (RuMP) or the xylulose monophosphate cycle (XuMP) (Figure 1.5). These unique pathways are mirrored in the phylogenetic divisions: serine cycle containing methylotrophs cluster into the α-Protobacteria and are referred to as type I methylotrophs, the restricted obligate methylotrophs fall into the β-Protobacteria, the methanotrophs using the RuMP cycle are clustered into the γ-Protobacteria (type II strains), and all Gram-positive methylotrophs contain the RuMP cycle (Hanson and Hanson, 1996). The Xylulose monophosphate cycle is used by methylotrophic yeasts.

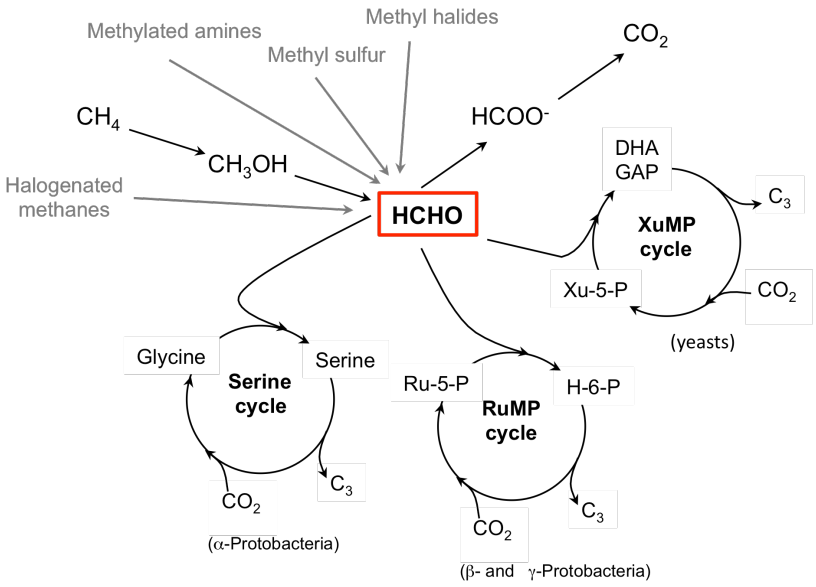


Figure 1.5. Schematic representation of methylotrophic metabolism of C_1 compounds. Abbrev.: Ru-5-P, ribulose-5-phosphate; H-6-P, hexulose-6-phosphate; Xu-5-P, xylulose-5-phosphate; DHA, dihydroxyacetone; GAP, glyceraldehyde-3-phosphate.

1.3.2 Methanol

There is a diverse range of C_1 compounds, methane, methanol, formaldehyde, methylated amines, and methyl halogenates that are ubiquitous in nature and represent the principal compounds of the C_1 carbon cycle (methane cycle) in the biosphere. Vast quantities of methanol are formed in mineralisation processes, mostly from the degradation of methyl-esters and -ethers (demethylation reactions)

that can be found in plant material such as pectin and lignin (Anthony, 1986; Hagel and Facchini, 2010).

Methanol is a valuable raw material and already an important carbon feedstock of the chemical and fuel industry. Therefore, a methanol-based industry has long been advocated, with amongst its prominent voices the 1994 Nobel Prize in Chemistry laureate George Olah (Olah, 2005). One way to reduce CO_2 emissions is by switching from petroleum to renewable biomass. Methanol, in the same way as petroleum, is a transportable liquid fuel but can be produced from materials of biological origin, making it a promising feedstock for green chemistry. Biotechnological interest in the microbial utilisation of methanol has increased because it is an ideal carbon source and can be produced from renewable biomass.

1.3.3 Oxidation of the C_1 compound (methanol) to formaldehyde

Enzymes catalysing the oxidation of methanol to formaldehyde fall into three classes: a quinoprotein methanol dehydrogenase found in Gram-negative methylotrophs (Goodwin and Anthony, 1998), an NAD-linked methanol dehydrogenase found in *Bacillus* strains (Arfman *et. al*., 1997), and a methanol:N,N'-dimethyl-4-nitrosoaniline oxidoreductase (MNO) found in other Gram-positive strains (Bystrykh *et. al.*, 1993a; 1997).

All known Gram-negative methane- and methanol-utilising bacteria contain a periplasmic methanol dehydrogenase that catalyses the oxidation of primary alcohols to their corresponding aldehydes. The enzyme possesses an $\alpha_2\beta_2$ structure and contains a pyrroloquinolinone quinone (PQQ) cofactor and a Ca^{2+} near the active site. The electron transfer chain goes from the oxidised PQQ to a cytochrome *c* (cyt_c) to a terminal oxidase. PQQ-MDH sequences are highly conserved within methylotrophs (Goodwin and Anthony, 1998).

Gram-positive bacteria lack the periplasmic space of Gram-negative bacteria, and therefore their methanol dehydrogenase is cytoplasmic. An NAD-linked MDH has been characterised from thermotolerant methylotrophic *Bacillus* strains (Arfman *et al*., 1997). The enzyme oxidises C_1-C_4 primary alcohols and is a

homodecamer of 43 kDa subunits, with 1 Zn^{2+} and 1-2 Mg^{2+} per subunit. Each subunit also contains a tightly but non-covalently bound NAD(H). A 50 kDa activator protein interacts with the enzyme enhancing the re-oxidation of the NADH cofactor in MDH (Arfman *et al*., 1991; 1997). Other Gram-positive methylotrophs oxidise methanol to formaldehyde via an MNO (methanol:N,N'-dimethyl-4-nitrosoaniline oxidoreductase), a decameric enzyme of 50 kDa subunits, but part of a three component complex. In a similar manner to the NAD-dependent MDH in *Bacillus*, each subunit has a tightly but non-covalently bound NADP(H) (Bystrykh *et. al.*, 1993a; 1997).

Finally, methylotrophic yeasts use an alcohol oxidase (AOD) to convert methanol to formaldehyde.

1.3.4 Formaldehyde oxidation

Formaldehyde is formed as a central intermediate in methylotrophic metabolism, and stands at a metabolic branch point. Theoretically, it is possible for methylotrophs to grow on formaldehyde, but it is generally too toxic and, where reported, growth is poor (Hirt, 1978; Whittenbury and Dalton, 1981). Cooper *et al.* (2001) reported evidence of formaldehyde chemistry (e.g. formose reaction) on the much-studied Murchison and Murray meteorites and, with formaldehyde a relatively abundant and ubiquitous molecule in comets and interstellar space, it was also probably abundant on the early Earth (Irvine, 1999). The toxicity of formaldehyde arises from its non-specific reactivity with proteins and nucleic acids (Merk and Speit, 1998; Quievryn and Zhitkhovich, 2000).

A number of formaldehyde oxidation systems are known in methylotrophs.

1.3.4.1 Linear pathways

A linear pathway converts formaldehyde to formate via a formaldehyde dehydrogenase. Usually, this enzyme is NAD- and/or factor-linked. Some Gram-positive methylotrophs (*Streptomyces* sp., *Rhodococcus erythropolis*, *Amycolatopsis methanolica*) utilise an NAD/factor-dependent formaldehyde dehydrogenase (van Ophem *et al*., 1992; Duine, 1999). This homotrimeric zinc

containing enzyme uses a sugar-thiol, mycothiol (MySH; 1-*O*-(2'-[*N*-acetyl-L-cysteinyl]amido-2'-deoxy-α-D-glucopyranosyl)-D-*myo*-inositol) as a coenzyme. In Gram-negative methylotrophs (autotrophic), including *Paracocccus denitrificans* and the non-sulfur purple bacteria *Rhodopseudomonas acidophilum* and *Rhodobacter spheroides*, an analogous NAD-linked enzyme that is also linked to glutathione (GSH) performs the oxidation of formaldehyde to formate (van Ophem *et al*., 1992; Ras *et al*., 1995; Barber *et al*., 1996; Barber and Donohue, 1998). These thiol coenzymes spontaneously form adducts with formaldehyde (S-hydroxymethyl glutathione in the case of GSH), and these thiohemiketal adducts are the real substrates of the enzyme, being oxidised to the formate ester and hydrolysed to formate and the thiol coenzyme regenerated.

Two other linear formaldehyde oxidation pathways are known to methylotrophs, both linked to folate species and found in α-Protobacteria and serine and RuMP cycle methylotrophs. One involves tetrahydrofolate (H_4F) and results in the oxidation of the methylene-H_4F into formate and H_4F. The other, which appears to be the major dissimilatory folate-linked pathway, involves a tetrahydromethopterin (H_4MPT)-dependent enzyme. The H_4MPT cofactor was thought to be specific to archaea, but the genes encoding the H_4MPT-linked enzyme in *Methylobacterium extorquens* show significant identity to the corresponding archaeal genes. It was suggested that this pathway was acquired by an early methylotroph by horizontal gene transfer from an archaeon (Chistoserdova *et al*., 1998).

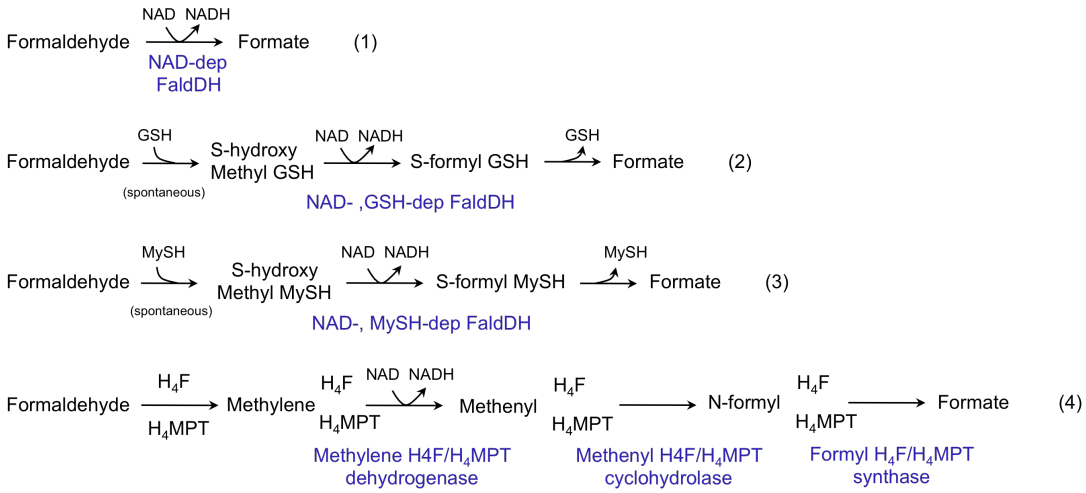


Figure 1.6. Linear pathways for formaldehyde oxidation in (aerobic) methylotrophic bacteria. (1) NAD-dependent formaldehyde dehydrogenase; (2) NAD-dependent gluthatione-linked formaldehyde dehydrogenase; (3) NAD-dependent mycothiol-linked formaldehyde dehydrogenase; (4) tetrahydrofolate- (H_4F) and tertahydromethanopterin- (H_4MPT) linked pathways. (Adapted from Anthony, 1982; Harms, 1996; Misset-Smith *et al*., 1997; Chistoserdova *et al*., 1998).

Burkholderia fungorum, a free-living bacterium, contains three pathways for formaldehyde oxidation: an NAD-linked, glutathione (GSH)-independent formaldehyde dehydrogenase, an NAD-linked, GSH-dependent formaldehyde dehydrogenase and an H_4MPT-dependent formaldehyde oxidation system (Marx *et al*., 2004). *B. mallei*, *B. pseudomallei* and *B. cepacia* are predicted to contain only the first two pathways, suggesting that the three pathways are involved in formaldehyde detoxification and are functionally redundant. *Methylococcus capsulatus* (Bath), an obligate methane-oxidising methylotroph, contains an NAD(P)-dependent PQQ-linked formaldehyde dehydrogenase, a homotetramer of subunit 49.5 kDa that requires a heat-stable activator 8.6 kDa protein called modifin for activity (Zahn *et al*., 2001; Tate and Dalton, 1999). *M. capsulatus* (Bath) also possesses a methylene-H_4MPT dehydrogenase source of formaldehyde dehydrogenase (Adeosun *et al*., 2004).

A GSH-independent NAD-linked formaldehyde dehydrogenase has been cloned from *Pseudomonas putida*, and shown to be a member of the zinc-containing medium-chain alcohol dehydrogenase family with the presence of the conserved

GXGXXG motif. The 170 kDa homo-tetrameric protein possesses a tightly but non-covalently bound NAD(H) cofactor distinct from the NAD coenzyme.

1.3.4.2 The serine cycle

The serine cycle (Figure 1.7) starts with the condensation of methylene tetrahydrofolate (H_4F-methylene) and glycine (C_2) into serine (C_3). This serine undergoes in turn a series of transformations to phosphoenolpyruvate, which is carboxylated to malate. The malate is cleaved into two glyoxylate (C_2) molecules, which condense back into glycine. Starting from two glycine and two formaldehyde (or H_4F-methylene) molecules, the cycle yields a net production of a 3-phosphoglycerate (C_3) (Anthony, 1982).

Most of the enzymes involved in the serine cycle have been characterised as recombinant proteins from *Methylobacterium extorquens* AM1 in *E. coli* (Chistersodova and Lidstrom, 1994a; 1994b; 1996; 1997), with the exception of the enzyme converting the acetyl CoA to glyoxylate.

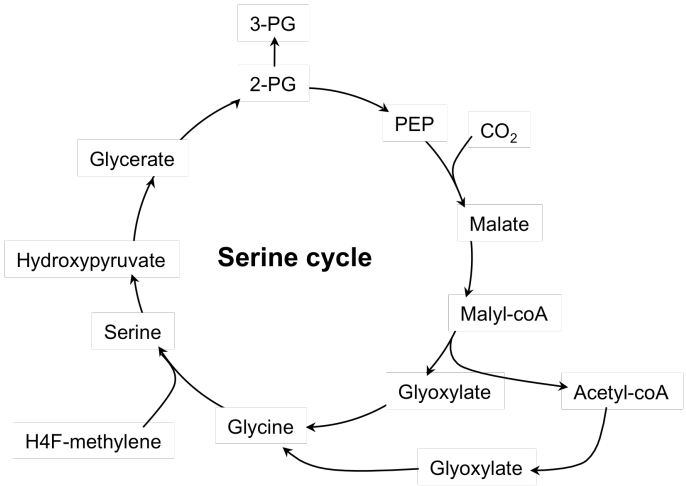


Figure 1.7. Schematic representation of the serine cycle for formaldehyde assimilation. (Adapted from Anthony, 1982).

1.3.4.3 The ribulose monophosphate (RuMP) cycle

The ribulose monophosphate (RuMP) cycle (Figure 1.8) initiates with the fixation of formaldehyde (C_1) with ribulose-5-monophosphate (C_5) into hexulose-6-phosphate (C_6). The condensation reaction is catalysed by the hexulose phosphate synthase

(HPS). This is then isomerised to fructose-6-phosphate by action of a phosphohexulose isomerase (PHI). A series of reactions results in the regeneration of the C_5 acceptor for formaldehyde via glyceraldehyde-3-phosphate. The condensation of three formaldehydes entering the cycle results in the net production of a C_3 product, dihydroxyacetone-phosphate (Anthony, 1982).

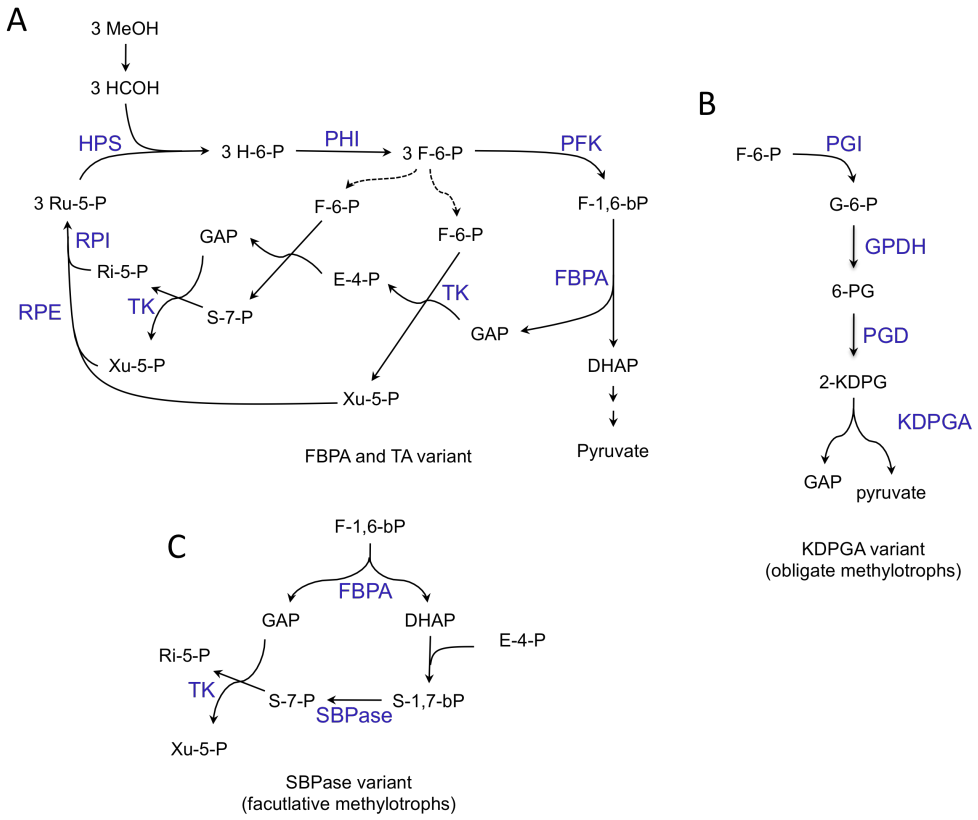


Figure 1.8. Schematic representation of the RuMP cycle for formaldehyde assimilation with (A) the FBPA (fructose bisphosphate aldolase) and TA (transaldolase) variants for cleavage and regeneration; (B) the KPDGA variant for cleavage; (C) the SBPase variant for regeneration. Enzymes (gene): HPS (*hps*) hexulose phosphate synthase; PHI (*phi*) phosphohexulose isomerase; PKF (*pfk*) phosphofructokinase; FBPA (*fba*) fructose bisphosphe aldolase; TK (*tkt*) transketolase; RPE (*rpe*) ribulose-5-phosphate 3-epimerase; RPI (*rpi*) ribose-5-phosphate isomerase; PGI (*pgi*) phosphoglucose isomerase; GPDH (*gpdh*) glucose-6-phosphate dehydrogenase; PGD (*pgd*) 6-phsphogluconate dehydratase; KDPGA 2-keto-3-deoxy-6-phosphogluconate aldolase; SBPase (*glpX*) fructose/sedoheptulose bisphosphatase. Substrates: Ru-5-P ribulose-5-phosphate; H-3-P hexulose-6-phosphate; F-6-P fructose-6-phosphate; F-1,6-bP fructose-1,6-bisphosphate; DHAP dihydroxyacteone phosphate; GAP glyceraldehyde-3-phosphate; Xu-5-P xylulose-5-phosphate; E-4-P eryhtrose-4-phosphate; S-7-P sedoheptulose-7-phosphate; S-1,7-bP sedoheptulose-1,7-bisphosphate; Ri-5-P ribose-5-phosphate. (Adapted from Anthony, 1982; Brautaset *et al*., 2004).

Different variants of the RuMP pathway exist. In obligate methylotrophs, the fructose-6-phosphate is transformed to glucose-6-phosphate and then 6-phosphogluconate. A 2-keto-3-deoxy-6-phosphogluconate aldolase and transladolase take the substrate further to 2-keto-3-deoxy-phosphogluconate and glyceraldehyde-3-phosphate and pyruvate, with the glyceraldehyde carrying on in the cycle to ribulose-5-phosphate, and pyruvate being the net C_3 product. In facultative methylotrophs (e.g. *Bacillus*), the generic RuMP cycle is complemented by a variant that involves sedoheptulose bisphosphate. Key enzymes are a fructose bisphosphatase and a sedoheptulose bisphosphatase.

The genes encoding for the HPS and PHI enzymes have been cloned and sequenced from both Gram-negative and Gram-positive methylotrophs and homologs of these genes have been found in non-methylotrophic bacteria as well as in archaea. In *Bacillus subtilis*, orthologs to both genes may encode functional enzymes with roles in protection against formaldehyde. Furthermore, the *hps* and *phi* genes, encoding the two enzymes, cluster in all methylotrophic bacteria, while their organisation and regulation varies between obligate and facultative methylotrophs (Figure 1.9). In *Methylomonas aminofaciens* (an obligate methylotroph), a transposase (IS*10*-R) is located between and regulates the *hps* (down-regulated) and *phi* (up-regulated) genes (Sakai *et al*., 1999). In the archaeon *Pyrococcus horikoshii*, a single open-reading frame encodes both *hps* and *phi*, resulting in a bifunctional enzyme, and this peculiar gene organisation has also been found in *P. abyssi* and *P. furiosus* (Orita *et al*., 2005; Yurimoto *et al*., 2009).

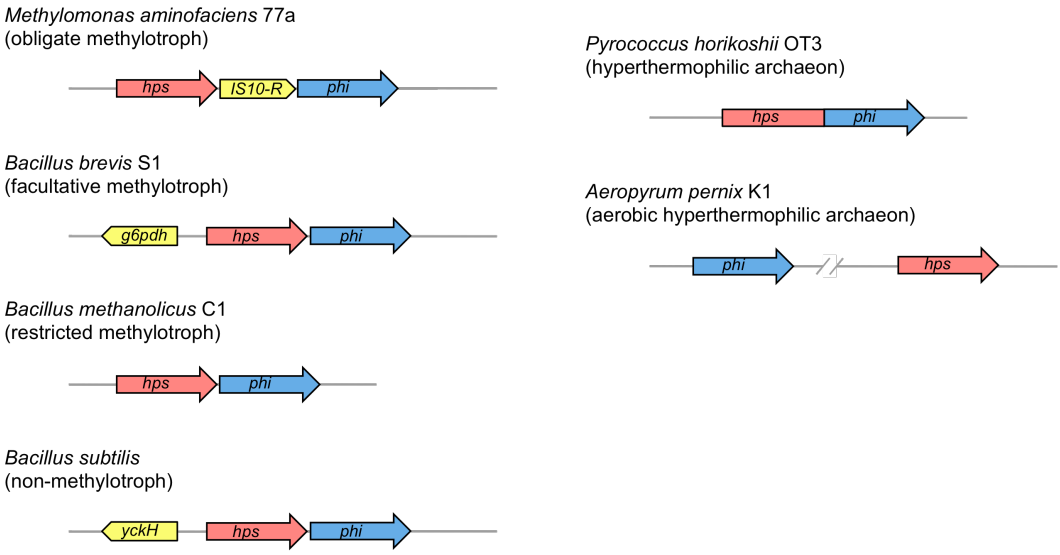


Figure 1.9. Gene clusters for the RuMP pathway. Genes: *IS10-R* transposase; *hps*, hexulose phosphate synthase; *phi*, phoshohexulose isomerase; *g6pdh*, glucose-6-phosphate dehydrogenase; *yckH*, regulator protein. (Adapted from Sakai *et al*., 1999; Arfman and Dijkhuizen, 1990; Jakobsen *et al*., 2006; Orita *et al*., 2005; Yurimoto *et al*., 2009)

1.3.5 Formate dehydrogenase

Formate dehydrogenases (FDH) cover a heterogenous group of enzymes found in Prokaryotes as well as Eukaryotes, which catalyse the oxidation of formate to CO_2 and H_2O. These enzymes vary in their quarternary structure, the presence and the type of prosthetic group, and in their substrate specificity. One group, the NAD-dependent FDH, is a homodimer, does not contain metal ion or prosthetic groups, and is highly specific for formate and NAD^+.

Formate dehydrogenase was first characterised in pea seeds (Mathews and Vennesland, 1950). Intense studies since the 1970s have partially focused on FDH from methylotrophic bacteria and yeasts, with the application for NADH regeneration as well as fundamental research.

All formate dehydrogenases belong to the super family of D-specific 2-hydroxy acid dehydrogenases with predicted conserved regions important to the catalytic and coenzyme binding domains (Vinals *et al*., 1993). Over 80 NAD-dependent FDH sequences are available, and these highlight the enzymes' extremely conserved

character as 60 amino acids are identical, especially catalytically essential residues, and an overall similarity of >75% (Popov and Lamzin, 1994).

In plants, FDH activity has been linked to a stress response, while in the pathenogenic *Staphylococcus aureus* one of three FDHs is upregulated in biofilm growth conditions (Colas des Francs-Small *et al*., 1993; Resch *et al*., 2005). In aerobic organisms, the enzyme is NAD-dependent, while anaerobic species possess a variety of proteins containing a complex range of redox centres and a high sensitivity to O_2. As presented above, methylotrophic bacteria oxidise methanol to CO_2 either through a cyclic mechanism operating at the formaldehyde level, or through linear chains of dehydrogenases. NAD-dependent FDHs (EC 1.2.1.2) have the key role of catalysing the last step in the catabolism of C_1 compounds, *i.e.* formate to CO_2, supplying the organism with energy and reducing equivalents. Some methylotrophs possess more than one FDH.

The catalytic mechanism of the enzyme is specified by the direct transfer of a hybrid ion from the substrate in the active centre of the dehydrogenase, resulting from the cleavage of a single C-H bond, onto the C4-atom (*pro-R* region) of the nicotinamide moiety of NAD^+, without stages of acid/base catalysis (proton transfer steps) as in other dehydrogenases and the product results in the formation of a single atom-atom bond. It also benefits from a very simple substrate. Thus, the enzyme is highly suitable for model mechanistic studies. FDH follows a Michaelis-Menten bi-bi two substrates kinetics scheme, with NAD^+ as the first substrate. All FDHs have similar K_M (formate) values of 3-10 mM and K_M (NAD^+) of 35-90 µM. The k_{cat} of bacterial FDH is approximately twice that of methylotrophic yeasts.

Interest in formate dehydrogenases stems from its value as a tool for NADH regeneration. The reaction it catalyses is almost irreversible, it is able to operate over a wide range of pH values, uses a cheap substrate, and methylotrophs provide a high scale enzyme production matrix for a relatively low cost, making the enzyme a versatile catalyst. From that point of view, the FDH from the methylotrophic yeast *Candida boidinii* has been studied further (Schirwitz *et al*., 2007; Andreadeli *et al*., 2008). The coenzyme specificity of CboFDH has been changed by saturation site-directed mutagenesis, and mutant D195S/Y196H/K356T had altered kinetics and coenzyme specificity. The specific activity towards NAD^+ halved in the

mutants compared to the wild-type, while the specific activity towards $NADP^+$ increased by over two orders of magnitude in the double and triple mutants.

FDH is present in all methylotrophs (Asano *et al*., 1988): yeasts (*Hansenula polymorpha*, *Candida boidinii*, *Komagataella pastoris*), fungi (*Aspergillus niger*, *Neurospora crassa*) and bacteria (*Pseudomonas* sp., *Moraxella*, *Paracoccus*, *Thiobacillus*) (Hollenberg *et al*., 1989; Sakai *et al*., 1997; Goldberg *et al*., 2004; Saleeba *et al*., 1992; Chow and RajBhandary, 1993; Tishkov *et al*., 1991; Shinoda *et al*., 2002; Nanba *et al*., 2003). It plays an essential role in energy regeneration during growth on C_1 compounds. Bacterial FDHs are generally homodimers of subunit M_r 70-100 kDa. They have a longer N-terminal region than that of yeasts and fungi, resulting in an elongated loop that covers a significant region of the subunit. The loop does not seem to contain elements with secondary structure but is relatively rigid due to the presence of 7 proline residues. The coverage offered by this loop might in part explain the higher thermostability of bacterial FDHs.

In *Moraxella*, a facultative methylotroph, the FDH is a homodimer of subunit M_r 98 kDa (Asano *et al*., 1988), while in *Sulfurospirillum multivorans* the *fdhA* gene encodes a 100 kDa enzyme (Schmitz and Diekert, 2003). A whole-genome analysis approach identified three different FDH systems in the facultative methylotroph *Methylobacterium extorquens* (Chistoserdova *et al*, 2004).

1.4 *Bacillus methanolicus* and methanol metabolism

1.4.1 *Bacillus methanolicus*

As has been presented before, a wide variety of bacteria and yeasts is able to grow on synthetic media with methanol as the sole or major source of carbon and energy. This is due to the presence of a few unique enzymes that enable these organisms to generate their energy and their first metabolites from this C_1 substrate. In the

chemical industry there is constant interest in the production of fuels and chemicals from methanol. Methanol is an attractive feedstock because of its low cost, ease of handling and availability. Methanol-utilising microbes are being studied for their potential utility in biotechnological processes (Dijkhuizen, 1985), and in this context, bacteria able to grow on methanol at elevated temperatures are of particular technological interest.

Snedecor and Cooney (1974) reported growth of a mixed population of spore-forming bacteria capable of utilising methanol as its sole source of carbon-energy at temperatures up to 65 °C. Inocula from varied sources and mixed cultures of methanol-utilising *Bacillus* strains were enriched by culture at 55 °C. Isolation of pure cultures was difficult but six strains were eventually obtained in pure cultures: all possessed the RuMP pathway, grew rapidly on methanol at 60 °C and were tolerant to high methanol concentrations, with the apparent involvement of an alcohol dehydrogenase in the initial oxidation step of methanol (Dijkhuizen *et al*., 1988). Further studies mainly focused on the novel NAD-dependent methanol dehydrogenase of these thermotolerant methylotrophic *Bacillus* strains (Arfman *et al*., 1989; 1990; 1991, 1992a; Dijkhuizen and Arfman, 1990). Some of these strains were first classified as *B. brevis* but the new species of thermotolerant methylotroph *Bacillus*, with fourteen strains identified, was subsequently named *Bacillus methanolicus* sp. nov. (Arfman *et al*., 1992b).

A Gram-positive, endospore-forming bacterium, *B. methanolicus* is an obligatory aerobic, thermotolerant organism capable of growth between 30 and 60 °C, with a temperature optimum at 50-53 °C and a doubling time of 40-80 min. It is a methylotroph, able to use methanol as its sole source of carbon and energy and displays a strong resistance to high concentrations of methanol. Thus, the generic position of 14 strains was determined. DNA-DNA hybridisation studies, 5S rRNA sequence analysis and physiological characteristics confirmed that these 14 strains clustered in a well-defined group forming a new distinct genospecies. Analysis of the 16S rRNA sequence indicated that it is closest to *B. azotoformans* and *B. firmus* (96.3 and 96.1% identity), with 95.4% identity with *B. subtilis* and *B. megaterium*, and 93.7% identity with *B. stearothermophilus*, *B. kaustophilus* and

B. thermoglucosidasius (Figure 1.10). All strains tested are sensitive to the different classes of antibiotics.

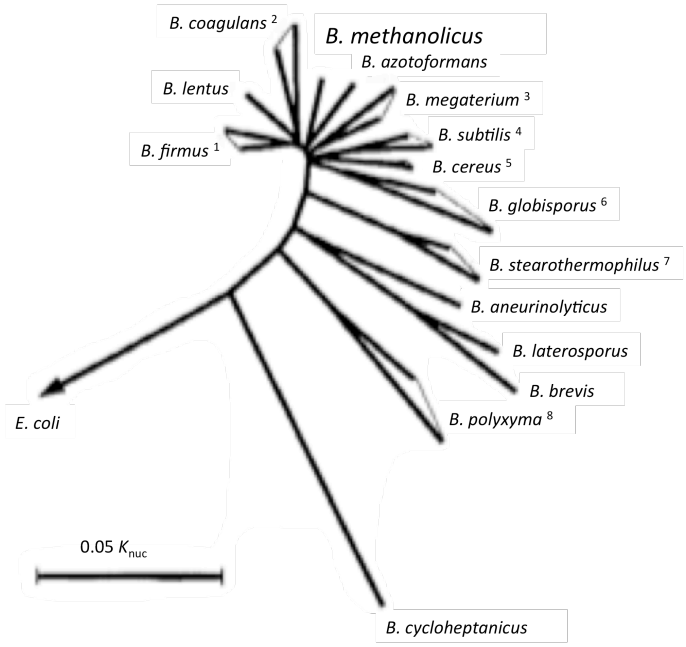


Figure 1.10. 16S rRNA distance matrix tree showing the relationships between methanol-utilising *B. methanolicus* C1 and other bacilli. Groups of more closely related bacilli are indicated by triangles. 1- *B. firmus*, *B. benzoevorans* and *B. circulans*; 2- *B. coagulans*, *B. acidoterrestris*, *B. badius* and *B. smithii*; 3- *B. megaterium*, *B. fastidiosus*, *B. maroccanus*, *B. psychosaccharolyticus* and *B. simplex*; 4- *B. subtilis*, *B. amyloliquefaciens*, *B. atrophaeus*, *B. lautus*, *B. lentimorbus*, *B. licheniformis*, *B. popillae* and *B. pumilus*; 5- *B. cereus*, *B. anthracis*, *B. medusa*, *B. mycoides* and *B. thuringiensis*; 6- *B. globisporus*, *B. fusiformis*, *B. insolitus*, *B. pasteurii*, *B. psychrophilus* and *B. sphaericus*; 7- *B. stearothermophilus*, *B. kaustophilus* and *B. thermoglucosidasius*; 8- *B. polymyxa*, *B. amylolyticus*, *B. alvei*, *B. azotofixans*, *B. gordonae*, *B. larvae*, *B. marquariensis*, *B. macerans*, *B. pabuli* and *B. pulvifaciens*. (Adapted from Arfman *et al*., 1992b).

The description of the species is as follows. The cells are rod shaped, stain Gram-positive and are non-motile. Sporulating cells are swollen and possess oval spores at the sub-terminal-to-central position, although *B. methanolicus* sporulates poorly in liquid growth medium at 50 °C (Schendel *et al*., 1990). Growth is obligatory aerobic and occurs at temperatures between 35 and 60 °C, with an optimum at around 55 °C. Most strains can grow in 2% NaCl and wild-type strains require biotin and vitamin B_{12} for growth (Arfman *et al*., 1992b; Schendel *et al*., 1990). Strains are restricted methylotrophs indicating that methylotrophy is an important trait while adaptation to

growth in seawater-based media (Komives *et al*., 2005) points to low cost requirement for growth components. The G+C content of the DNA is 48-50 mol.%.

The type strain *B. methanolicus* PB1 was deposited with NCIMB, accession 13113.

1.4.2 Methanol metabolism in *Bacillus methanolicus*

Thermotolerant methylotrophic *Bacillus* oxidizes methanol by way of an NAD-dependent MDH and utilizes the RuMP pathway for formaldehyde fixation (Dijkhuizen *et al*., 1988). NAD-dependent ADHs are widely distributed, but with generally low, if any, affinity for methanol. In *Bacillus methanolicus*, methanol oxidation into formaldehyde is catalysed by a cytoplasmic NAD-dependent MDH (BmMDH) and the *mdh* gene has been cloned in *B. methanolicus* C1 (Arfman *et al*., 1989; de Vries *et al*., 1992). BmMDH oxidizes C_1-C_4 short-chain primary alcohols to their respective aldehydes but with decreasing relative rates as the chain gets longer (100% for methanol). An MDH was also found to be encoded by an *mdh* gene harboured on a 19 kb plasmid (pBM19) in *B. methanolicus* MGA3, together with five genes with deduced functions in the RuMP pathway (Brautaset *et al*., 2004; Jakobsen *et al*., 2006). Interestingly, only the first two enzymes of the RuMP cycle, HPS and PHI, are encoded on the chromosome; consequently, *B. methanolicus* strains cured of pBM19 are incapable of growth on methanol (Brautaset *et al*., 2004). BmMDH showed similarities with, and displayed unique motifs of, family III of NAD-dependent alcohol dehydrogenases, which contain an iron or zinc ion in the active site and are best exemplified by the ADH2 from *Zymomonas mobilis* and the ADH4 from *Saccharomyces cerevisiae* (Arfman *et al*., 1997). Sequence analysis of BmMDH also identified the NAD-binding motif GXGXXG (de Vries *et al*., 1992). Electron microscopic analysis and characterization of the enzyme revealed that it is a homodecameric nicotinoprotein, with the subunits (M_r = 43 kDa) arranged in a sandwich of two pentagonal rings, each subunit containing one zinc and one or two magnesium ions (Vonck *et al*., 1991) (Figure 1.11). Zinc is often present in type III ADH, but magnesium had not been reported before. A tightly but non-covalently-bound NADH cofactor is also present in each subunit (Arfman *et al*., 1997).

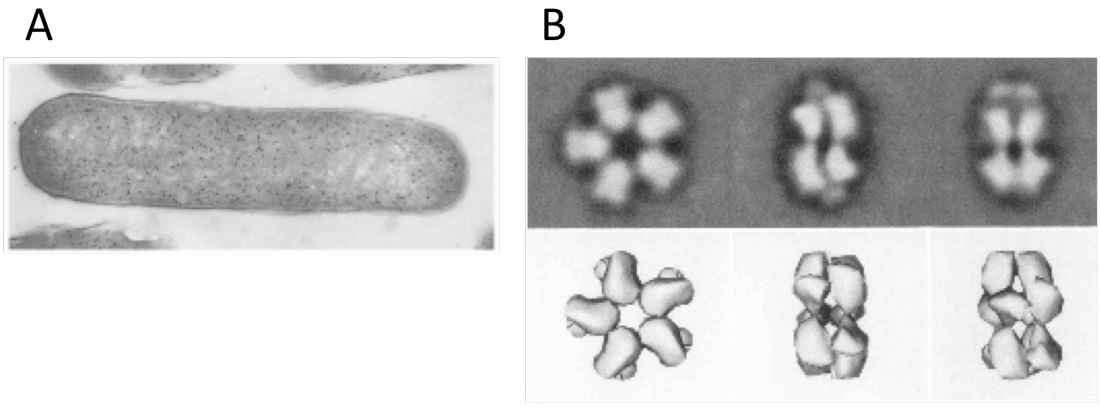


Figure 1.11. (A) Immunogold labeling of methanol dehydrogenase in whole cells of methanol-grown B. methanolicus C1, showing the cytoplasmic localization of the enzyme. (B) Three views of the MDH by electron microscopy and models. (Adapted from Arfman *et al*., 1989; Vonck *et al*., 1991).

MDH activity is highly stimulated by a soluble 50 kDa activator protein, ACT, resulting in a 40-fold increase in methanol turnover (Arfman *et al*., 1991). ACT increases the V_{max} of MDH, while the K_M for methanol only slightly decreases. The ACT protein, a homodimer of subunit M_r 27 kDa, has been shown to belong to the NUDIX family of proteins, and showed clear NAD hydrolyzing activity (Kloosterman *et al*., 2002). ACT activation involves the removal of the nicotinamide mononucleoside (NMN) moiety of NAD, with one NAD(H) per subunit. MDH and ACT are coordinately expressed in *B. methanolicus*. Increased intra-cellular NADH/NAD ratios reduce ACT-mediated activation of MDH, possibly preventing accumulation of toxic formaldehyde. Mutations in the Mg^{2+}-dependent NAD binding domain ^{94}GGGSXXDXXK^{103} of MDH lead to insensitivity to ACT (S97G) or impaired cofactor binding (G95A) (Hektor *et al*., 2002). Mg^{2+} has a profound effect on NAD(H) cofactor binding by MDH and consequently on ACT activation. A model for ACT:MDH interaction was proposed (Figure 1.12) (Kloosterman *et al*., 2002). In the non-activated state, MDH displays a ping-pong mechanism with the cofactor being the electron acceptor. On cleavage of the NAD(H) cofactor by ACT, MDH enters the activated state and displays a cofactor independent mechanism due to the absence of the NMN moiety of NAD(H). The activated MDH has a higher affinity and activity for methanol (K_M = 140 mM and V_{max} = 12 U.mg^{-1}) than the nonactivated MDH (230 mM and 1.3 U.mg^{-1}, respectively) (Arfman *et al*., 1991). Under scarce methanol conditions, MDH will be

in the activated state allowing efficient catalytic oxidation of the substrate. The activated MDH can re-enter the nonactivated cycle by binding NADH in the cofactor binding site.

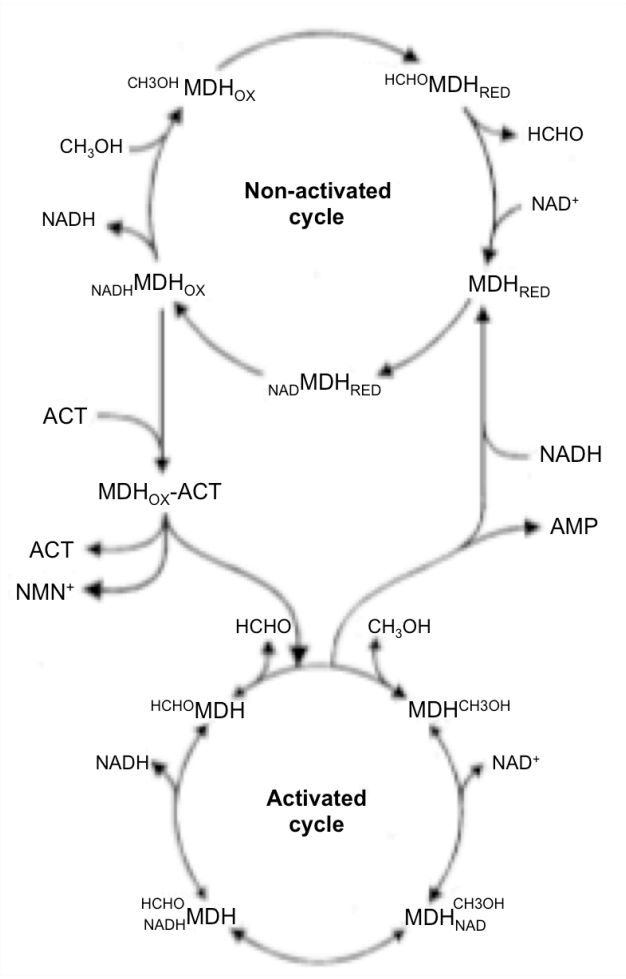


Figure 1.12. Model of the effects of ACT on MDH reaction cycles in *B. methanolicus.* In the non-activated cycle (top), NAD(H) (cofactor) acts as a temporary electron deposit, and the reaction obeys a ping-pong mechanism. Cleavage of the NMN(H) moiety of NAD(H) cofactor by ACT results in the MDH entering the activated reaction characterised by a ternary complex mechanism. An activated MDH molecule can re-enter the non-activated cycle by binding NADH. (Kloosterman *et al*., 2002).

Amongst the known different routes for formaldehyde assimilation, *B. methanolicus* possesses the enzymatic machinery to utilise the RuMP cycle (de Vries *et al*., 1990; Kato *et al*., 2006) (Figure 1.13). Common to these pathways is the C_1 formaldehyde fixed to multi-carbon entities that are regenerated after one cycle, with the net production of one C_3 compound from three C_1 moieties entering the cycle. In the

fixation part, the condensation of three formaldehydes with three ribulose-5-phosphates (Ru-5-P) is catalysed by 3-hexulose-6-phosphate synthase (HPS) generating three hexulose-6-phosphate (H-6-P). Each H-6-P is isomerised to fructose-6-phosphate (F-6-P) by a 6-phospho-3-hexuloisomerase (PHI). Of the different RuMP pathway variants, *B. methanolicus* converts F6P to fructose-1,6-bisphosphate (F-1,6-bP) by phosphofructokinase (PFK) with the consumption of one ATP. F-1,6-bP is cleaved by fructose bisphosphate aldolase (FBPA) into glyceraldehyde-3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP). DHAP enters glycolysis, producing pyruvate with the concomitant generation of one NAD(H) and two ATP. GAP enters the final part of the cycle where three Ru-5-P molecules are regenerated through the actions of transketolase (TK), fructose/sedoheptulose bisphosphatase (GLPX) and ribulose-5-phosphate epimerase (RPE).

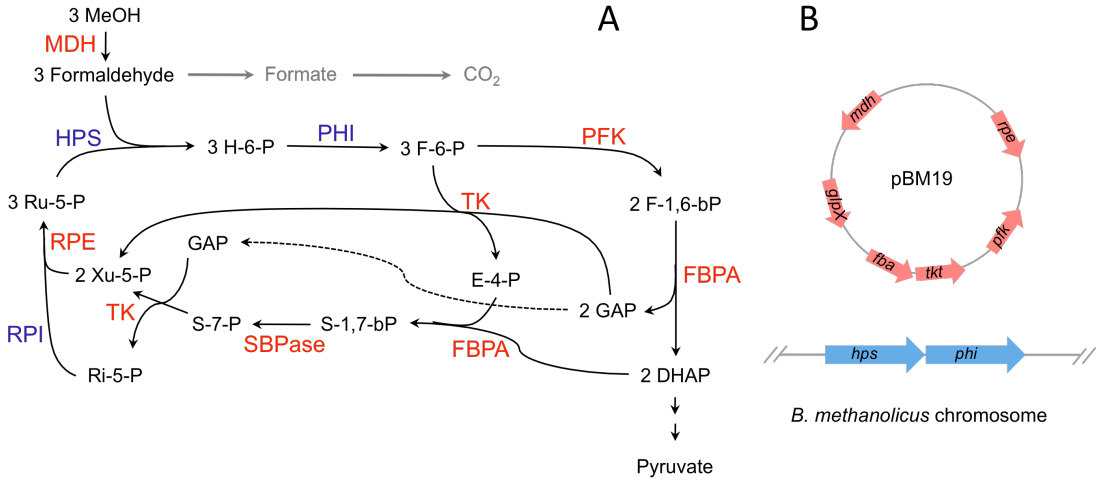


Figure 1.13. Schematic representation of methanol metabolism in *B. methanolicus* (A) with the physical location of the genes involved (B). Chromosomal genes: *hps*, hexulose phosphate synthase HPS; *phi*, phosphohexulose isomerase; mdh, methanol dehydrogenase MDH. pBM19 genes: *mdh*, methanol dehydrogenase MDH; *pfk*, phosphofructokinase PFK; *fba*, fructose bisphosphate aldolase FBPA; *tkt*, transketolase TK; *glpX*, fructose/sedoheptulosebisphosphatase; *rpe*, ribulose phosphate epimerase RPE. (Adapted from Brautaset *et al.*, 2004).

As mentioned above, pBM19 harbours *mdh* and five of the genes with deduced roles in the RuMP cycle (*pfk*, *fba*, *tkt*, *glpX* and *rpe*). This, coupled with the fact that curing of pBM19 resulted in the loss of the ability to grow on methanol, explains why methylotrophy in *B. methanolicus* is plasmid-dependent.

Genes or enzymes involved in the dissimilatory pathway of formaldehyde oxidation to CO_2 have not yet been identified in *B. methanolicus*. However, evidence of an alternative pathway for formaldehyde dissimilation has been reported when accumulation of labeled formaldehyde, formate and CO_2 was observed in cultures supplemented with ^{13}C methanol or ^{13}C formate (Pluschkell and Flickinger, 2002), indicating that formaldehyde could potentially be oxidized to formate and CO_2 in a linear pathway.

Much of the interest in the studies of *B. methanolicus* and its metabolism to methanol relate to its potential for industrial production of amino acids at 50 °C. Being a thermotolerant methylotroph, *B. methanolicus* is an interesting candidate for overproduction of amino acids from methanol. The wild-type strain can secrete 58 $g.L^{-1}$ of L-glutamate and mutants have been selected that secrete 37 $g.L^{-1}$ of L-lysine, with the high growth temperature being an advantage in reactor cooling requirements and also lower contamination risks in cultures (Brautaset *et al*., 2007).

1.5 Aims of the project

The generation of molecular H_2 by both whole cells and isolated enzymes has been proposed as a means of fuelling conventional fuel cells. It has also been shown to be possible to feed electrons generated by bio-oxidations to the electrodes of a fuel cell without the intermediary of molecular hydrogen. These systems are simpler in concept and raise the possibility of compact biofuel cells that could be fuelled by cheap carbon substrates.

However, the goal of producing molecular H_2 for fuel purposes remains desirable. The production of hydrogen by growing cultures of bacteria is well known and has been studied increasingly in recent times. The economics of such a system remain a problem and a major concern is that not all the H atoms can be released from most substrate molecules, giving unwanted by-products.

The use of isolated enzymes to produce H_2 has also been studied. Woodward *et al.* (1996) have linked glucose dehydrogenase and hydrogenase and shown *in vitro* H_2 production from glucose. A greater yield of H_2 production *in-vitro* has been achieved by associating isolated enzymes of the pentose phosphate pathway in synthetic enzymatic pathways (Woodward *et al.*, 2000a; 2000b; Zhang *et al.*, 2007). These approaches also suffer, in many cases, from incomplete recovery of the H atoms as H_2 and, in addition, to the relative lack of stability of available enzymes.

The use of methanol as a feedstock has potential advantages in that its metabolism to CO_2 requires three dehydrogenase reactions that successively remove all H atoms of the molecule as reduced cofactors (Figure 1.14). Methanol possesses additional advantages as an energy source in a biofuel cell, being cheap, easily available, and a high energy density liquid. Furthermore, some methanol-utilising thermophilic organisms are known and can be expected to contain robust enzymes suitable for fuel cell use and biohydrogen production.

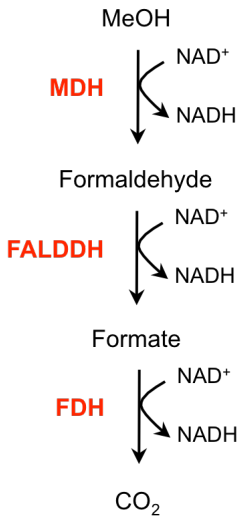


Figure 1.14. Putative synthetic enzymatic pathway for hydrogen production from methanol (MeOH). H_2 can be recovered from the reduced NADH cofactor by way of a hydrogenase.

The overall aims of the project were to:

a- clone and express the 3 dehydrogenases involved in methanol dissimilation,

b- characterise the dehydrogenases and assess their suitability for biofuel cell use and biohydrogen production,

c- study thermophilic methanol-utilising bacteria derived from culture collections and from environmental sources,

d- study available thermostable hydrogenases, cloning them into suitable hosts as necessary, including into the wild-type thermophilic methanol utiliser.

Chapter 2- General Materials and Methods

2.1 Materials

2.1.1 General laboratory reagents

Tris(hydroxymethyl)methylamine (Tris), sodium hydroxide (NaOH), sodium chloride (NaCl), and sodium dodecyl sulphate (SDS) were from Fisher Scientific (Loughborough, UK). EDTA was from Arcos Organics (Geel, Belgium). Potassium hydroxide (KOH), potassium chloride (KCl), hydrochloric acid (HCl), MES, MOPS and Triton X-100 were supplied by Sigma-Aldrich (Gilligham, UK). Ethanol and propan-2-ol were also from Fisher Scientific (Loughborough, UK). Double-distilled water ($_{dd}H_2O$) and ultra-purified water (MilliQ water) were produced in the lab from an Elix system and a Milli-Q® Integral system (Millipore (UK) Ltd, Watford, UK), respectively.

2.1.1.1 Bacterial cell culture reagents

Yeast extract and tryptone were obtained from Melford Laboratories Ltd (Ipswich, UK). Tryptone soya broth was from Oxoid (Oxford, UK). Sucrose, D-glucose, di-potassium hydrogen orthophosphate (K_2HPO_4), potassium dihydrogen orthophosphate (KH_2PO_4), sodium dihydrogen orthophosphate (NaH_2PO_4), di-sodium hydrogen orthophosphate (Na_2HPO_4), tri-sodium citrate ($Na_3C_6H_5O_7$), magnesium sulphate ($MgSO_4$), DL-tryptophan, iron sulphate ($FeSO_4.7H_2O$), copper sulphate ($CuSO_4.5H_2O$), manganese sulphate ($MnSO_4.H_2O$), zinc sulphate ($ZnSO_4.7H_2O$), sodium molybdate ($Na_2MoO_4.2H_2O$), calcium chloride ($CaCl_2.2H_2O$), cobalt chloride ($CoCl_2.6H_2O$), thiamine hydrochloride (vitamin B1), calcium D-pantothenate (vitamin B5), (-)-riboflavin (vitamin B2), nicotinamide (vitamin B3, niacin), vitamin B12 (cyanocobalamin), and biotin (vitamin H) were from Sigma-Aldrich (Gilligham, UK). Carbenicillin, kanamycin, gentamycin, streptomycin and chloramphenicol were

also from Sigma-Aldrich (Gillingham, UK). L-Histidine and L-arginine were from Duchefa Biochemie B.V. (The Netherlands). Casein hydrolysate, boric acid (H_3BO_3) and L-glutamate were from Fluka (Gilligham, UK). Zinc acetate ($(CH_3COO)_2Zn.2H_2O$) was from Arcos Organics (Geel, Belgium). Methanol was from Fisher Scientific (Loughborough, UK). Casamino acids, Difco™ nutrient broth and bacto-agar were obtained from BD-Biosciences (Oxford, UK). L-Rhamnose monohydrate was supplied by Promega (Southampton, UK).

2.1.1.2 Molecular biology reagents

Custom oligonucleotide primers were synthesised by Bioneer (Alameda, CA, USA), Integrated DNA Technologies (Leuven, Belgium) or Invitrogen (Paisley, UK). *Taq* DNA polymerase, *Taq* 10x buffer and 2x Red MaxaTaq mastermix were obtained courtesy of GeneSys Ltd (Camberley, UK). Finnzymes' Phusion® high-fidelity DNA polymerase and 2-Log DNA ladder were purchased from New England Biolabs Ltd (NEB, Hitchin, UK). The dNTPs, DNA ladder HyperLadder I and HyperLadder IV were from Bioline (London, UK). The pGEM®-T Easy Vector system, T4 DNA ligase and JM109 cells were supplied by Promega (Southampton, UK). Restriction enzymes and BSA were supplied by Promega and NEB. Sodium acetate ($CH_3COONa.3H_2O$), isopropyl-β-D-thiogalactopyranoside (IPTG), bromo-chloro-indolyl-galactopyranoside (X-Gal), lysosyme and blue dextran were from Sigma-Aldrich. The DNeasy tissue kit was from Qiagen (Crawley, UK). The Wizard® Plus SV miniprep DNA purification kit, Wizard® SV gel and the PCR clean-up system were from Promega. Agarose was from Melford Laboratories Ltd (Ipswich, UK) and low-melting point agarose was Seakem LE agarose from Lonza UK (Slough, UK). The pET vectors were supplied by Novagen (Nottingham, UK).

2.1.1.3 Protein related materials

The 30% acrylamide/bis solution, ammonium persulphate (APS), protein SDS-PAGE standard broad-range markers, protein assay dye reagent concentrate and Poly-Prep® chromatography column were from BioRad Laboratories Ltd (Hemel Hampstead, UK). β-Mercaptoethanol was from BDH Chemicals Ltd (Poole, UK). Serum albumin protein standard was supplied by Pierce Chemical Co. (USA). Sodium dodecyl sulphate (SDS), nickel sulphate ($NiSO_4$), glycine, dithiothreitol (DTT),

tetramethylethylenediamine (TEMED) and Coomassie stain Brilliant Blue R were from Sigma-Aldrich (Gillingham, UK). The His•Bind® resin and Chitin beads were from Novagen (Nottingham, UK) and NEB UK (Hitchin, UK), respectively. Metal chelating cellulose was from Bioline (London, UK). Imidazole was from Acros Organics (Geel, Belgium). Glacial acetic acid was from Fisher Scientific (Loughborough, UK).

2.1.1.4 Equipment

PCRs were carried out in an Eppendorf Mastercycler® (Eppendorf UK, Histon, UK). Electroporation cuvettes were purchased from Bio-Rad Laboratories Ltd (Hemel Hempstead, UK). BD 14mL Falcon polypropylene round-bottom tubes were from BD Biosciences (Oxford, UK). Centrifugations were carried out, accordingly to needs, in a benchtop Picofuge, centrifuges 5402R or 5810R (all Eppendorf UK), or centrifuges Allegra 25R or Avanti J-25 (Beckman-Coulter UK Ltd, High Wycombe, UK) using rotors TA 10-250 and JLA-16.250 (Beckman-Coulter). Cell disruption was carried out using a One Shot cell disrupter (Constant Systems Ltd, Warwick, UK) and/or sonication in a 150W Ultrasonic Disintegrator (MSE Scientific Instruments, London, UK). SDS-PAGE was carried out with a mini-Protean® 3 system (Bio-Rad). The ÄKTAexplorer™ system was from GE Healthcare (Amersham, UK). Spectroscopy was carried out with a Varian Cary 50 (Varian Ltd, Oxford, UK).

2.1.2 Specific materials

2.1.2.1 Bacterial cultures

Bacillus methanolicus PB1 (ATCC^{51375}, NCIMB 13113) was purchased from the National Collection of Industrial Bacteria (NCIMB Ltd, Aberdeen, UK).

Escherichia coli strain JM109 (*e14*$^{-}$(*Mcr*A^{-}) *rec*A1 *end*A1 *gyrA96 thi-1 hsd*R17 (*rK*$^{-}$ *mK*$^{+}$) *sup*E44 *rel*A1 Δ(*lac-pro*AB) [F' *tra*D36 *pro*AB *laq*I^{q}ZΔM15]) was originally purchased from Promega (Southampton, UK).

Escherichia coli BL21(DE3)pLysS (F^{-}, *omp*T, *hsd*S$_B$ (r$_B^-$, m$_B^-$), *dcm*(DE3), *gal*(*λc1857*, *indl*, *Sam7*, *nin5*, *lacUV5-T7genel*), λ(DE3), pLysS(Cm^r)) and *E. coli* KRX (K12 [F' *tra*D36 Δ*omp*P

*pro*A+B+ *laq*I^{q} Δ(*lac*Z)M15] Δ*omp*T *end*A1 *rec*A1 *gyr*A96 (Nal^{r}) *thi*-1 *hsd*R17 (r_K^- m_K^+)e14$^-$ (McrA$^-$) *rel*A1 *sup*E44 Δ(*lac-pro*AB) Δ(*rha*BAD)::T7 RNA polymerase) were from Promega. *E. coli* ArcticExpress™ (DE3) (*E. coli* B F$^-$ *omp*T *hsd*S(r_B^- m_B^+) *dcm*$^+$ Tet^r *gal* λ(DE3) *end*A *Hte* [*cpn10 cpn60* Gent^r]) was purchased from Strategene (La Jolla, USA).

Bacillus subtilis subsp. *subtilis* strain 168 (NCIMB 10106, DSM 402) (trp$^-$ or ind$^-$) was purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen (German Collection of Microorganisms and Cell Cultures; DSMZ GmbH, Braunschweig, Germany).

Geobacillus thermoglucosidasius (NCIMB 11955) was obtained from the National Collection of Industrial, Marine and Food Bacteria (NCIMB).

2.1.2.2 Gene cloning

The pGEM®-T Easy vector system was purchased from Promega (Southampton, UK). The pET28a(+) vector was purchased from Novagen (Merck-Biosciences, Nottingham, UK). The pTYB11 vector was purchased from NEB UK (Hitchin, UK).

Insert-containing pGEM®-T Easy vectors were cloned routinely in *E. coli* JM109.

Insert-containing pET vectors were first cloned into *E. coli* JM109 for screening, and then into the *E. coli* expression strain of choice.

2.1.2.3 Assay materials

Methanol was from Fisher Scientific (Loughborough, UK). Formaldehyde (36.5 - 38% aqueous solution) and formate (sodium salt), the enzymes alcohol dehydrogenase (ADH, horse liver), formaldehyde dehydrogenase (*Pseudomonas putida*) and formate dehydrogenase (*Candida boidinii*) were also from Sigma-Aldrich (Gillingham, UK). Benzonase® nuclease was from Novagen. The Complete™ mini EDTA-free Protease inhibitor cocktail was from Roche Diagnostics (Burgess Hill, UK). Glutathione was from Acros Organics (Geel, Belgium).

2.2 Methods

2.2.1 Bacterial strains, media and growth conditions

Unless otherwise specified, culture media were sterilised by autoclaving (121 °C, 20 min) before use. For solid media, 2% (w/v) agar was added to the broth before autoclaving. Filter-purified components were filtered on a 0.22 µm syringe filter (Millipore).

2.2.1.1 *Bacillus methanolicus*

Bacillus methanolicus PB1 ATCC^{51375} was grown at 50 °C in the following media. Resurrection of the freeze-dried culture was in tryptone soya (TS) medium (30 g.L^{-1}). The freeze-dried strip was rehydrated in 1 mL TS broth. An aliquote was used to seed a 10 mL TS culture while the wet strip was placed on a TS agar plate. Both were incubated overnight at 50 °C, with shaking for the liquid culture.

B. methanolicus was maintained in TS broth supplemented with 1% (v/v) methanol (filter-sterilised) with shaking (200 rpm). *B. methanolicus* was also maintained on TS agar plates supplemented with 1% (v/v) methanol.

Growth of *B. methanolicus* on methanol was performed in $MeOH_{200}$ medium (salt buffer, 1 mM $MgSO_4$, vitamins, trace metals, 0.025% yeast extract and methanol 200 mM, pH 7.2) (Jakobsen *et al*., 2006). Salt buffer, vitamins and trace metals were essentially as in MV medium (Schendel *et al*., 1990). MV medium is minimal salt (MS) medium supplemented with thiamine hydrochloride (vitamin B1), D-calcium pantothenate (vitamin B5), (-)-riboflavin (vitamin B2), and nicotinamide (vitamin B3), each at 20 µg.L^{-1}, and cobalamin (vitamin B_{12}) at 1 µg.L^{-1}; MS medium contained (per litre): K_2HPO_4, 3.8 g; $NaH_2PO_4.H_2O$, 2.8 g; $(NH_4)_2SO_4$, 3.6 g; $MgSO_4.7H_2O$, 0.5 g; $FeSO_4.7H_2O$, 2 mg; $CuSO_4.5H_2O$, 40 µg; H_3BO_3, 30 µg; $MnSO_4.H_2O$, 200 µg; $ZnSO_4.7H_2O$, 200 µg; $Na_2MoO_4.2H_2O$, 47 µg; $CaCl_2.2H_2O$, 5.3 µg; and $CoCl_2.6H_2O$, 40 µg. The pH was adjusted prior to autoclaving. Biotin 30 µL/100 mL (0.1 g/10 mL 0.1 M KOH), filter-purified trace elements (500 µL/100 mL) and 1% (v/v) methanol were added last.

2.2.1.2 *Escherichia coli*

Escherichia coli was routinely cultured at 37 °C in lysogeny broth (LB) medium (yeast extract, 5 $g.L^{-1}$; tryptone, 10 $g.L^{-1}$; NaCl, 10 $g.L^{-1}$), adjusted to pH 7.0 with NaOH.

Escherichia coli JM109 was cultured overnight in LB broth or agar plates at 37 °C (shaking at 200 rpm for the liquid cultures). The medium was supplemented with antibiotic for selection when appropriate (carbenicillin, 50 $\mu g.mL^{-1}$ or ampicillin, 100 $\mu g.mL^{-1}$; kanamycin, 30 $\mu g.mL^{-1}$).

E. coli BL21(DE3) were grown on LB at 37 °C and *E. coli* BL21(DE3) pLysS were grown on LB supplemented with chloramphenicol (34 $\mu g.mL^{-1}$) at 37 °C. ArcticExpress™ were grown on LB supplemented with gentamycin and streptomycin (each 50 $\mu g.mL^{-1}$) at 30 °C, and *E. coli* KRX on LB supplemented with chloramphenicol (34 $\mu g.mL^{-1}$) at 30 °C. All cultures were grown with shaking at 200 rpm. Cultures were supplemented with the appropriate antibiotic for selection of transformants (plasmid construct).

2.2.1.3 *Bacillus subtilis*

Bacillus subtilis DSM 402 was grown at 28 °C in TS medium. The lyophilised material was revived by culture in nutrient broth (NB). The freeze-dried strip was rehydrated with 1 mL NB; an aliquot used to seed a 10 mL NB culture and the strip placed onto a NB agar plate. Incubation was at 28 °C overnight, with shaking (200 rpm). Cells were streaked onto a TS agar plate.

Further cultures of *B. subtilis* were on TS at 28 °C, with shaking.

2.2.1.4 Glycerol stocks and culture storage

For short-term storage, cultures were kept on plates at 4 °C. For longer-term storage, glycerol stocks were prepared. Glycerol stocks were prepared in cryogenic vials (Nalgene, USA). A fresh 5 mL overnight culture was spun down and the pelleted cells resuspendend in 1mL of medium. The culture and 40% (v/v) glycerol (sterile) were combined to a 1:1 ratio (*i.e.* final concentration of glycerol was 20%), mixed, snap-frozen in liquid nitrogen and stored at -80 °C.

2.2.2 Cloning of a gene of interest into *E. coli*

Most of the methods used for molecular cloning were based on those described in Sambrook and Russell (2001).

2.2.2.1 Preparation of electro-competent *E. coli*

Electro-competent *E. coli* were prepared as follows. A single colony of a freshly streaked *E. coli* culture was inoculated aseptically into 10 mL YENB (8% (w/v) yeast extract, 7.5% (w/v) nutrient broth) and incubated overnight at 37 °C with shaking (200 rpm). The following day, the entire culture was used to inoculate a 1 L flask containing 250 mL YENB and incubated at 37 °C with shaking to an $OD_{600nm} \approx 0.6$. The culture was chilled on ice for 5 min before centrifugation in two pre-chilled 250 mL centrifugation bottles at 5,000*g* for 10 min. The supernantants were discarded and the pellets gently resuspended in one-tenth of the volume with ice-cold sterile MilliQ water. The cells were pelleted by centrifugation as before and washed 3 times in ever decreasing volumes of sterile MilliQ water. Following the third wash, the cells were resuspended in 5 mL 10 % (v/v) glycerol and centrifuged at 5,000*g* for 10 min. The pelleted cells were finally resuspended in 0.75 mL 10% (v/v) glycerol (sterile) (*i.e.* 0.75 mL/250 mL starting culture), and dispensed in 40 μL aliquots in 0.5 mL tubes pre-chilled in a bath of ethanol/dry-ice. Electro-competent cells were stored at -80 °C.

2.2.2.2 DNA extraction

Genomic DNA extraction

Genomic DNA was extracted from Gram-positive bacteria, *B. methanolicus* and *Geobacillus thermoglucosidasius*, following a method adapted from the DNeasy Tissue kit (Qiagen). Cells from a 10 mL overnight culture were harvested by centrifugation at 5,000*g* for 10 min. The pellet was resuspended in 180 μL lysis buffer (20 mM Tris-HCl, pH 8, 2 mM EDTA, 1.2% Triton X100, 20 $mg.mL^{-1}$ lysosyme) and incubated for 30 min at 37 °C. Twenty-five microlitres of proteinase K and 200 μL buffer AL (Qiagen) were added and the mixture incubated for 30 min at 70 °C. The next steps of the genomic DNA extraction were performed following Qiagen's

DNeasy tissue kit protocol for animal tissues. Genomic DNA was eluted in 100 µL $_{dd}H_2O$ and the concentration was measured on an biophotometer (Eppendorf).

Plasmid DNA preparation

All plasmid preparations were performed using Promega's Wizard™ *Plus* SV Miniprep DNA purification system.

E. coli A single colony was inoculated into 5 mL LB supplemented with the appropriate antibiotic and incubated overnight at 37 °C with shaking. The following day, the culture was centrifuged (5,000*g*, 10 min) and the supernatant removed. Plasmid DNA was purified following Promega's protocol. The final elution was in 35 µL $_{dd}H_2O$.

Bacillus A single colony was inoculated into 5 mL TS supplemented with the appropriate antibiotic when needed and incubated overnight with shaking at 50 °C (*B. methanolicus*) or 28 °C (*B. subtilis*). The following day, the culture was centrifuged (5,000*g*, 10 min) and the supernatant removed. Cell lysis was performed with the lysis buffer used for Gram-positive genomic DNA extraction as detailed above. The remainder of the plasmid DNA preparation was as described in Promega's protocol. The final elution was in 35 µL $_{dd}H_2O$.

2.2.2.3 Gel electrophoresis and purification

DNA, restriction digests and PCR products were visualised on a 1% (w/v) agarose gel in 1x TAE buffer (40 mM Tris acetate, pH 8.2, 1 mM EDTA), and supplemented with ethidium bromide (1 µL of 10mg.mL^{-1}). Loading dye (blue/orange loading dye 6x, Promega) was added to the samples prior to loading. Samples were run alongside 5 µL HyperLadder I. Gels were run at 100 V for 45 min or until adequate separation/segregation of the bands had been achieved, and visualised on a UV transilluminator. Bands of the expected size product were cut out using a clean razor blade and placed in a 1.5 mL microfuge tube.

The extraction of DNA was performed using the Wizard® SV gel and PCR clean-up system (Promega) following the manufacturer's instruction. The agarose slices were not weighed but considered to be *ca.* 150 mg for calculation of buffer volumes to add. The final elution volume was 35 µL $_{dd}H_2O$.

2.2.2.4 Restriction digests

Digestion of DNA was carried out using restriction endonucleases. Single or double restriction digests of DNA were performed with the enzymes of choice, as available and/or determined by the cloning strategy, using the recommended reaction buffer (supplied as a 10x concentrate) as per the manufacturer's instructions. In double digests, the buffer used was the one giving maximum activity for both enzymes as highlighted in Promega's compatible buffer resource and NEB's double-digest finder. Digests were set up in 1.5 mL microfuge tubes, to a final volume of 30 µL (with $_{dd}H_2O$), with up to 2 µg DNA and 20 U enzyme(s) and incubated for 1 h at 37 °C unless otherwise stated. The glycerol concentration was kept below 10% (v/v) to avoid star activity. Table 2.1 presents the set-up of a typical reaction. When needed and possible, the enzyme(s) were heat-inactivated by incubation at 65 °C for 20 min.

Table 2.1. Typical set up for a restriction digest. (RE: restriction enzyme.)

Reagents	Volume (µL)	
	Single digest	Double digest
10x RE buffer	3	3
10x BSA	3	3
DNA (plasmid)	5	5
RE #1	1	1
RE #2	0	1
MilliQ water	to V_f 30µL	

Products were separated by agarose gel electrophoresis and purified as described above (in 2.2.2.3).

2.2.2.5 Polymerase Chain Reaction

Oligonucleotide design

Oligonucleotides primers of *ca*. 20-*mers* were designed based on the sequence of the gene of interest. When performing PCR over full-length coding sequence (CDS), the primer pair was designed to share 100% identity with the 5'-end starting from the start codon ATG (forward primer) and the 3'-end from the stop codon (reverse primer).

Both the forward and reverse primers may incorporate at their respective 5'-ends a recognition sequence for a restriction enzyme determined by the cloning strategy for introduction into the expression vector. If that was the case, an extra 4-6 bp were added upstream of the restriction sequence introduced according to the Cleavage Close to the End of DNA Fragments table (NEB). The choice of restriction enzyme was based on two conditions: a- that they are present in the multiple cloning site (MCS) of the target expression vector, and b- that they are not present within the coding sequence of the genes of interest. When occurring, an extra 3-6 bp were added to flank the restriction site as highlighted in NEB's table Cleavage Close to End of DNA fragments.

PCR

Taq DNA polymerase or Phusion® High-fidelity DNA polymerase were used in PCRs. The reactions' final volume was always set at 25 µL with PCR-grade (or sterile MilliQ) water, unless otherwise stated (Table 2.2). Template DNA (*ca*. 20 ng), 1 µL of each forward and reverse primers (5 pmol.µL^{-1}), 2.5 µL 10x *Taq* polymerase buffer or 5 µL 5x Phusion® HF buffer, 1 µL 25 mM dNTPs and 1 U polymerase were combined in a 0.2 mL thin-walled tube sitting on ice in a fumehood.

Table 2.2. PCR set up and running conditions

Reagents	Volume (µL)
10x *Taq* pol buffer	2.5
DNA template	eq. 20ng
Primer_Fwd (5 pM)	1
Primer_Rev (5 pM)	1
dNTPs (25 mM)	1
Taq polymerase	1U
PCR-grade water	to V_f 25 µL

PCR was performed in an Eppendorf Mastercycler® (Eppendorf AG, Germany). The reactions specific running conditions were adapted to the template, primers and expected product-length requirements. However, a default running program was set up: for *Taq* polymerase, Hot start 3 min 95 °C, 35 cycles [94 °C for 30 s (denaturation), 55 °C for 45 s (annealing), 72 °C for 1 min (extension)], and a final

extension for 10 min at 72 °C; for Phusion®, denaturation was at 98 °C for 15 s, and annealing and extension steps were reduced to 15 s and 30 s, respectively. An extension time of 1 kb per min was used with *Taq* polymerase, and 20 s per kb with Phusion® polymerase.

2.2.2.6 Cloning into pGEM®-T Easy vector system

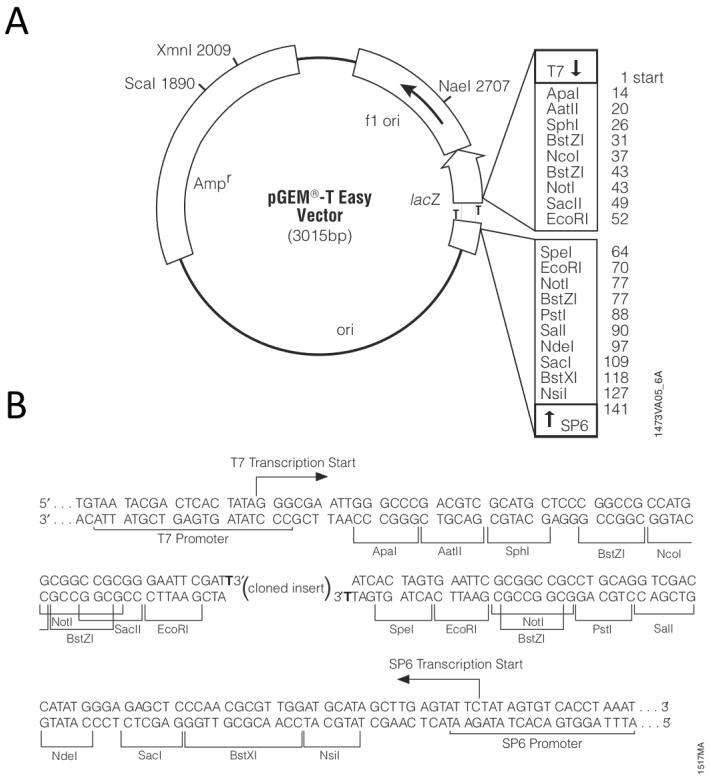


Figure 2.1. The pGEM®-T Easy vector. (A) Circular map and reference points; (B) promoters and multiple cloning site sequences. (Adapted from Promega).

A-tailing

Phusion® DNA polymerase generates blunt-ended products during PCR amplification. In order to clone blunt-ended dsDNA fragment into the pGEM®-T Easy vector, these PCR products needed to be modified with the addition of a single adenosine phosphate (dATP) at each 5'-end of the double stranded DNA. The A-tailing reaction was performed as follows: in a 0.2 mL thin-walled tube, 1 µL 10x *Taq* buffer, 1 µL 25mM dATP and 0.5 U *Taq* polymerase were combined with 5 µL blunt-end PCR product, to a final volume of 10 µL with PCR-grade water. The reactions were incubated at 70 °C for 30 min in a thermalcycler.

Ligation

The pGEM®-T Easy vector system (Figure 2.1) was used for the cloning of A-tailed PCR products, following the protocol outlined in Promega's technical manual for pGEM®-T and pGEM®-T Easy vector systems.

Ligation reactions (final volume 10 µl with MilliQ water) were set up in 0.2 µL thin-walled tubes by combining 1 µl (50 ng.µL^{-1}) vector, 3 µL eluted insert DNA, 1 µL 10x T4 ligase buffer and 1 U T4 DNA ligase. Ligations were incubated for at least 2 hours in a thermalcycler, cycling at 14 °C for 30 s, and 30 °C for 30 s.

2.2.2.7 Transformation of *E. coli* JM109 cells

A 40 µL aliquot of electrocompetent *E. coli* cells was thawed on ice. One microlitre of insert-ligated pGEM®-T Easy plasmid DNA was added to the cells and the mixture transferred to a pre-chilled gene-pulser cuvette with a 0.1 cm gap. The pulse was set at 1.8 kV, 200 Ω. The cuvette was transferred to the gene-pulser chamber and the pulse given. The pulse time was recorded (4-6 ms). Following the pulse, 1 mL SOC (0.2 g.L^{-1} Bacto-tryptone, 0.05 g.L^{-1} Bacto-yeast extract, 10 mM NaCl, 2.5 mM KCl and 10 mM $MgCl_2$, with 10 mM $MgSO_4$ and 20 mM glucose added after autoclaving) was added to the cuvette, the cells transferred to a 14 mL BD Falcon polypropylene round-bottom tube and incubated for 1 h at 37 °C with shaking (200 rpm). Typically, 100 µL cells were spread onto an LB agar plate supplemented with 0.5 mM (or 80 µL 100 mM) IPTG, 80 µg.mL^{-1} (or 20 µL 50 mg.mL^{-1}) X-Gal and carbenicillin (50 µg.mL^{-1}) and incubated overnight at 37 °C. X-Gal and IPTG allow for blue/white screening selection of transformants. White colonies were picked.

2.2.2.8 Screening of transformants

A white colony was picked and grown overnight in 5 mL LB at 37 °C. Plasmid DNA was then extracted as previously described.

Screening of transformants was done via different methods: colony PCR, PCR and restriction digest. The products were run on an agarose gel for analysis.

Colony PCR

Colony PCR was performed directly on the colony to be screened. The reaction was set up with *Taq* polymerase as previously described. For the template, "a visible amount of colony" was picked with a pipette tip and transferred into a thin-walled tube containing 20 µL MilliQ water, thoroughly mixed, and incubated at 95 °C for 5 min; 2 µL were then added to the PCR mix. The PCR running conditions were also slightly modified with a longer hot start (10 min, 95 °C) and an increase to 40 cycles.

Plasmid PCR

PCR was performed as described previously on the purified plasmids.

In both plasmid PCR and colony PCR, the primer pair used was that of the vector region flanking the putative insert, *i.e.* pSP6 (5'-ATTTGGTGACACTATAG) and pT7 (5'-TAATACGACTCACTATAGGG) in the case of pGEM®-T Easy, and pT7 and tT7 (5'-GCTAGTTATTGCTCAGCGG) in the case of the pET vector.

Restriction digest

Single and/or double restriction digests were performed on plasmids to determine the presence of the insert by either releasing it (i.e. the restriction sites used to introduce the insert as per the cloning strategy) or generating a tel-tale sized fragment. In pGEM®-T Easy, *Eco*RI was routinely used as it cuts on either side of the multiple cloning site, immediately next to the edges of the insert.

Restriction digests were set up as described previously.

2.2.2.9 Sequencing

Sequencing of insert-containing plasmids was performed by Geneservices, London. A minimum of 5 µL purified plasmid DNA (100 ng.µL^{-1}) and 5 µL primer (3.2 pmol.µL^{-1}) per reaction were supplied. The sequencing reads were analysed using Seqman (DNAStar, USA).

2.2.3 Sub-Cloning (of a gene of interest) into the expression vector pET

2.2.3.1 The pET vector family

The pET vectors pET19b and pET28a(+) are presented in Figure 2.2.

2.2.3.2 Cloning strategy

Cloning strategies were devised using the bioinformatics suite of the VectorNTI® software (Invitrogen).

The inserts were PCR-amplified from the source DNA using gene-specific forward and reverse primers containing at their 5'-ends the recognition sequence of the restriction enzymes to be used in the cloning. Restriction sites for cloning procedures were chosen so that they were unique to the vector and the insert, *i.e*. did not have any other restriction sites in the insert and the rest of the vector. Extra nucleotides (usually 4) were added to the primer sequence at the 5'-end so that the restriction enzyme recognition site does not sit at the 5'-end of the primer, a feature that enhances the yield of subsequent restriction digest of the PCR product with these enzymes if needed.

Usually, the inserts were first cloned into pGEM®-T Easy, screened, excised and ligated into the pET vector of choice.

2.2.3.3 Restriction digest

Single or double restriction digests were performed in parallel on both the insert-containing purified plasmid pGEM®-T Easy and the empty pET vector. The setting of restriction digests is detailed above (2.2.2.4). The excised insert and the linearised pET vector were gel purified as previously described.

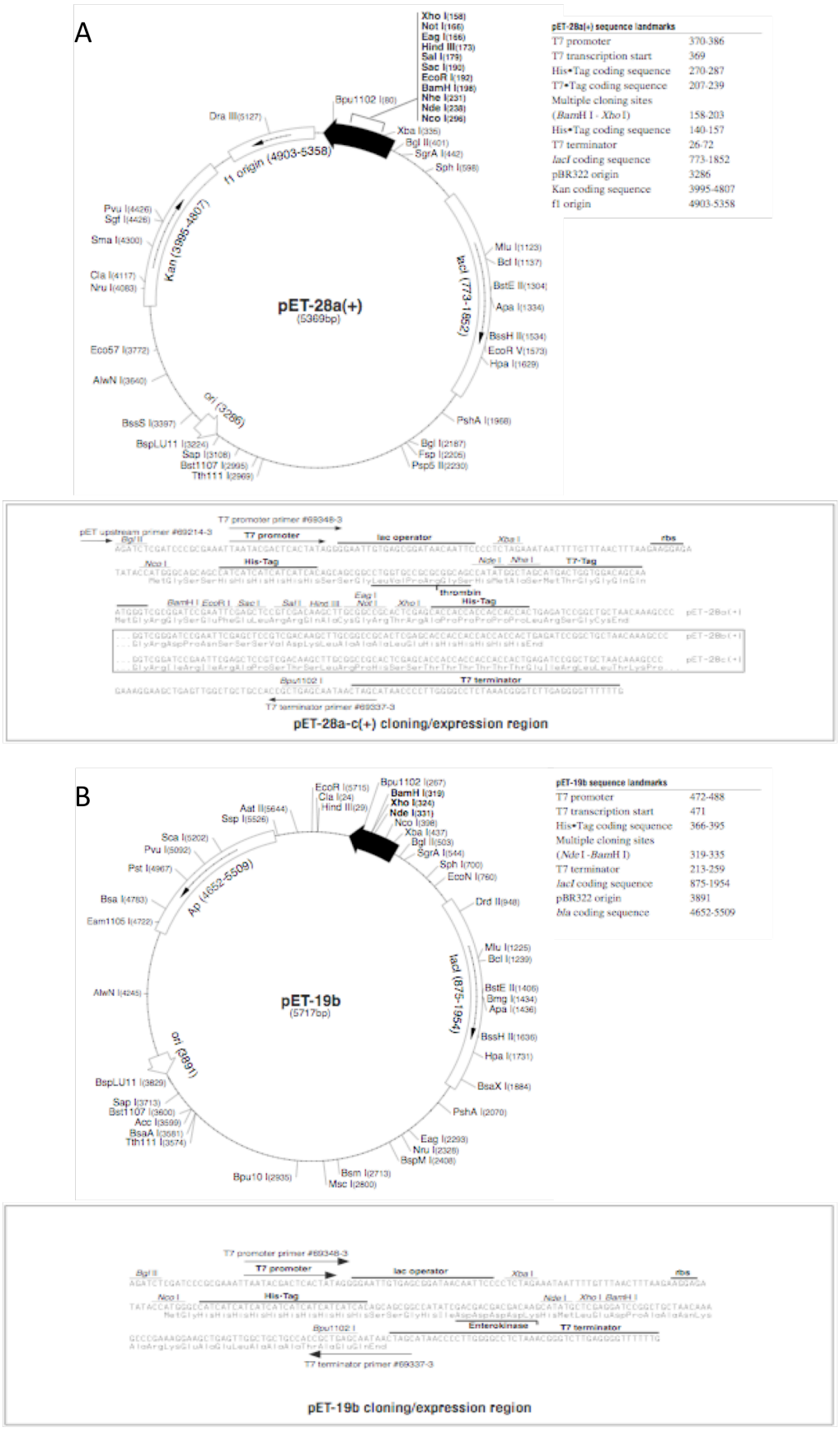


Figure 2.2. Vector maps for (A) pET-19b and (B) pET-28a(+). (Novagen).

2.2.3.4 Ligation

Ligations in expression vectors of the pET series were performed as follows. The linearised pET vector was treated with Antarctic phosphatase (NEB) at 37 °C for 1 h, afterwhich the enzyme was heat-inactivated by incubation at 65 °C for 10 min. Alkaline phosphatases catalyse the removal of 5'-phosphate groups from DNA. Since phosphatase-treated fragments lack the 5'-phosphoryl termini required by ligases, dephosphorylation prevents religation of linearised plasmid DNA. The amount of gel-purified restriction digested insert (GoI) and plasmid (pET) was estimated on agarose gel against a known amount of DNA loaded and run alongside within HyperLadder I. The ligation reaction was set up as described above, but amounts of insert and vector were adjusted according to equation (1):

$$\frac{\text{ng vector x kb insert}}{\text{kb vector}} \text{ x insert:vector ratio} = \text{ng insert} \qquad (1)$$

Typically, 50 ng vector was used as default, and a 3:1 insert to vector molar ratio. A control reaction containing no insert DNA was set up.

2.2.3.5 Transformation of pET:insert into JM109

Ethanol precipitation

Ethanol precipitation of DNA was preformed as follows. The volume of aqueous phase was measured and 1/10V of 3 M Na acetate (pH 5.2) and 2.5 volumes of cold ethanol were added. The sample was vigorously mixed and left to precipitate at -20 °C for 30 min. The sample was centrifuged at 13,000*g*, the ethanol removed and the pellet washed in 70% (v/v) ethanol. The resulting pellet was left to air-dry before being resuspended in $_{dd}H_2O$.

Transformation

Transformation of the ligated pET vectors into JM109 *E. coli* was as described previously. Typically, 100 µL cells were spread onto an LB plate supplemented with kanamycin (50 $\mu g.mL^{-1}$) and incubated overnight at 37 °C.

Screening of transformants was by colony PCR, PCR or restriction digest as described previously. In PCR screening of pET vectors, primers were pT7

(5'-TAATACGACTCACTATAGGG) and tT7 (5'-GCTAGTTATTGCTCAGCGG). Confirmed pET vectors containing the insert were purified by miniprep.

2.2.3.6 Transformation into the *E. coli* expression strain

Purified insert-containing pET vectors were transformed into an *E. coli* expression strain, BL21(DE3)pLysS (chloramphenicol, 34 $\mu g.mL^{-1}$), ArcticExpress™ (gentamycin 20 $\mu g.mL^{-1}$, streptomycin 50 $\mu g.mL^{-1}$) or KRX (chloramphenicol 34 $\mu g.mL^{-1}$). Transformation was as previously described. The cultures were supplemented with the appropriate antibiotic for selection of the pET vector (pET19b: ampicillin 100 $\mu g.mL^{-1}$ or carbemicillin 50 $\mu g.mL^{-1}$; pET28a: kanamycin 50 $\mu g.mL^{-1}$).

Transformants were either PCR screened, *i.e.* colony screen and/or PCR on plasmid DNA as previously described, or screened by restriction digest.

2.2.4 Preparation of competent *Bacillus subtilis* and transformation

Fresh cultures of *B. subtilis* were made for each transformation, based on the protocol of Bolhius *et al*. (1999). A single colony was used to inoculate 5 mL GCHE (1% glucose, 0.2% potassium L-glutamate, 100 mM trisodium citrate, 3 mM $MgSO_4$, 22 $mg.L^{-1}$ ferric ammonium citrate, 50 $mg.L^{-1}$ L-tryptophan, and 0.1% casein hydrolysate) (Kunst and Rapoport, 1995) and incubated at 28 °C with shaking.

B. subtilis was transformed by growth on GCHE medium to an $OD_{600} \approx 1.0$; after addition of DNA to the culture, growth was continued for 4 h. Antibiotic selection was carried out.

2.2.5 Protein expression

2.2.5.1 Induction of expression of pET vector

Fresh, single colonies of *E. coli* BL21(DE3) and BL21(DE3)pLysS from a transformation plate were used to inoculate 10 mL LB with kanamycin (50 $\mu g.mL^{-1}$) and grown overnight at 37 °C with shaking (200 rpm). When other *E. coli* expression strains were used, the antibiotic selection was customised according to the strain's and

harboured plasmids genotype. The following day, 1 or 5 mL of culture were used to seed 50 or 250 mL fresh LB medium (with kanamycin, 50 $\mu g.mL^{-1}$), which was then grown at 37 °C with shaking to an $OD_{600} \approx 0.6$, induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and further grown for up to 12 h at 37 °C with shaking.

Similarly, a colony of *E. coli* ArcticExpress™ taken from a transformation plate was grown overnight at 37 °C in 10 mL LB with gentamycin (20 $\mu g.mL^{-1}$) and streptomycin (50 $\mu g.mL^{-1}$) (for selection of the host strain), and kanamycin (50 $\mu g.mL^{-1}$) (for selection of the pET28a plasmid). The following day 1mL of the culture was used to seed a 50 mL LB broth without antibiotics, and incubated at 30 °C with shaking for 3 h. The temperature was lowered to 10-13 °C and the cultures incubated for 10 min to equilibrate. IPTG was then added to 1 mM, and the culture incubated further with shaking for 24 h.

2.2.5.2 Preparation of cell extracts and fractionation

Cultures for expression trials were centrifuged (3,800*g*, 20 min) and the supernatant discarded. The pellets were re-suspended in one-tenth of the original culture volume of appropriate buffer (His-bind buffer, assay buffer or else). Cells were then either passed through a cell disrupter at 25 kpsi or sonicated. Sonication (amplitude 14 micron) was performed on ice, in three successive 15 s bursts with 30 s rests in-between. For large preparations, the number and/or the duration of the bursts were increased. Benzonase® nuclease and C*O*mplete™ protease inhibitor were added. An aliquote was taken to serve as the total cell extract. The cell debris was pelleted by centrifugation (13,000*g*, 15 min), and the soluble protein fraction (supernatant) was removed and transferred to a fresh tube. The pellet (insoluble protein fraction) was re-suspended in the same volume (*i.e.* one-tenth the original culture volume) of 8 M urea.

2.2.5.3 Protein estimation

Protein estimation was based on the method of Bradford (1976). Protein assay dye reagent was made by adding 2 mL concentrate (Bio-Rad) to 7 mL $_{dd}H_2O$. The Coomassie blue dye in the assay binds to primarily basic and aromatic amino acids residues, especially arginine.

Protein estimation was measured by spectrophotometry at 595 nm: in a 1 mL cuvette, 100 µL protein samples were combined to 900 µL dye reagent and incubated for 15 min before a reading was taken. Absorbance readings were compared against BSA standards of 0-20 µg.

2.2.5.4 SDS-PAGE

SDS-PAGE was set up using a mini-Protean®3 system (Bio-Rad) and run in Tris/Glycine buffer, pH 8.3 (10x Tris/glycine buffer: 100 mL 10% (w/v) SDS, 144 g glycine, 30 g Tris) according to the method of Laemmli (1970). Typically, the resolving gel (10% or 12% respectively) was made as 4 mL or 3.3 mL $_{dd}H_2O$, 3.3 mL or 4 mL 30% acrylamide/bis solution, 2.5 mL resolving buffer (1.5 M Tris-HCl pH 8.8), and 100 µL 10% SDS, followed by polymerisation with 100 µL 10% APS and 10 µL TEMED. The stacking gel was made as 1.1 mL $_{dd}H_2O$, 0.33 mL 30% (w/v) acrylamide/bis solution, 0.5 mL stacking gel buffer (0.5 M Tris-HCl, pH 6.8), and 20 µL 10% SDS, and was polymerised with 20 µL 10% APS and 5 µL TEMED.

Protein samples were prepared with the addition of 4x loading buffer (4 mL 0.5 M Tris-HCl, pH 6.8, 2 mL 10% (w/v) SDS, 4 mL glycerol, 100 µL β-mercaptoethanol, and 8 mg bromophenol blue) and denatured by heating at 95 °C for 3 min prior to loading. Electrophoresis was carried out at 80 V through the stacking gel and 200 V thereafter until the dye had reached the bottom of the gel. Gels were stained with Coomassie staining solution (water: methanol: Coomassie brilliant blue R 250: glacial acetic acid, 45.4:45.4:0.2:9) for 30 min and destained with water: methanol: glacial acetic acid (6:3:1) overnight. The relative molecular weight of proteins was estimated by co-electrophoresis with 5 µL broad-range molecular weight markers (Bio-Rad).

2.2.5.5 Purification of recombinant enzyme from cell extract

Metal-affinity chromatography (His-Bind purification)

His-tagged recombinant proteins were purified by metal affinity chromatography.

The Poly-Prep® chromatography column (Bio-Rad) was prepared under renaturing conditions. The His•Bind® resin (Novagen) was gently mixed by inversion and 2 mL

of the slurry was transferred to the column and left to pack by gravity flow to prepare a 1 mL final bed volume. The resin was then washed 3 times with 5 column volumes of $_{dd}H_2O$ and charged with 2 column volumes of 400 mM $NiSO_4$. The column was then further washed with 3 column volumes of binding buffer (50 mM Tris pH 8.0, 300 mM NaCl, 20 mM imidazole,).

The buffer was allowed to drain to the top of the resin and 1 to 5 mL of the cell lysate (soluble fraction) was loaded on top of the column and allowed to drain. The flow-through was reloaded. The 2nd flow-through was collected. The column was then washed with binding buffer, and the flow-through collected. The column was washed in sequence with 2 column volumes of buffers of increasing imidazole concentration (70 mM, 140 mM, 400 mM) (7%, 14%, 40% eluting buffer in binding buffer). The column was then washed with 5 column volumes of eluting buffer (50 mM Tris pH 8.0, 2 M KCl, 1 M imidazole, 8 M urea). Two millilitre fractions were collected at each stage and stored at 4 °C until further analysis.

The column was washed with 5 column volumes of $_{dd}H_2O$ and stored in 20% (v/v) ethanol. Prior to further use, the column is washed with $_{dd}H_2O$ and binding buffer.

Gel filtration

Gel filtration was used to elucidate the form of assembly of the enzymes. A Hiload 16/60 Superdex 200 gel filtration column and a Superdex 200 10/300GL gel filtration column were linked to an ÄKTAexplorer™ system (GE Healthcare). Columns were calibrated prior to use with M_r standards (Ribonuclease, M_r 13,700; Ovalbumin, M_r 43,000; Conalbumin, M_r 75,000; Aldolase, M_r 158,000 and Ferritin, M_r 440,000) (GE Healthcare), and a standard curve was produced allowing estimation of the M_r and thus the state of assembly of the protein of interest. The markers selected were dependent on the M_r of protein sample run. The plot used to estimate the M_r of the protein of interest is log M_r as a function of K_{av} with $K_{av} = (V_e-V_o)/(V_c-V_o)$, where V_e is the elution volume, V_o is the column void volume and V_c is the geometric column volume. Where required, protein samples were concentrated using an Amicon Ultra-10k NMWL centrifugal filter device (Millipore) following the manufacturer's instructions (pre-rinsing with $_{dd}H_2O$). Approximately 0.6 mL of sample was loaded onto the gel filtration columns. The elution profile of the protein was recorded by the

ÄKTA system by measuring the A_{280} and those fractions over which the protein eluted were collected and stored on ice. K_{av} was calculated using the equation above. Depending on the calculated K_{av} of the protein of interest, the equation of the low M_r or the high M_r calibration line was used to calculate the log M_r of the native protein and therefore estimate the state of assembly.

2.2.6 Enzyme assays

Enzyme assays on the putative protein of interest were run in parallel with controls consisting of substrate alone, and of a soluble fraction of a culture transformed with an empty expression vector, to ensure the measured rates of activity were solely due to the activity of the recombinant enzyme.

All assays were performed in 1mL (V_f) at 50 °C unless otherwise stated, using pre-warmed buffer solutions. All assays were designed to follow the conversion of NAD^+ to NADH, or *vice versa*, by spectrophotometry at 340 nm. Phosphate buffer and pyrophosphate buffer (1 mM) were at pH 8.5. Glycine-KOH (100mM) buffer was pH 9.5 and HEPES buffer (50 mM) was pH 8.0. Buffers were always supplemented with 5 mM $MgSO_4$ and 5 mM $ZnSO_4$. The amount of substrates and enzyme were varied extensively. The specific activities were determined from the change in $A_{340\ nm}$ using the molar absorption coefficient of NAD(H) of 6,220 $M^{-1}.cm^{-1}$. Kinetic parameters were estimated by the direct linear method of Eisenthal and Cornish-Bowden (1974) and Hanes transformation (Hanes, 1932).

2.2.6.1 NADH oxidase assay

NADH oxidase activity was measured spectrophotometrically at 340nm. The assay mixture combined 940 µL HEPES buffer, 50µL 4mM NADH. The reaction was started by the addition of 10 µL cell extract.

2.2.6.2 Methanol dehydrogenase assay

Methanol dehydrogenase or formaldehyde reductase activity was measured spectrophotometrically at 340 nm. A typical assay comprised 940 µL HEPES buffer, 50 µL 10 mM NAD^+, and 10 µL purified recombinant enzyme or cell extract. After

equilibration at the assay temperature, the reaction was started with the addition of 1 µL 6 M methanol (C_f = 6mM).

A formaldehyde reductase assay was also performed, corresponding to the reverse reaction. The assay was set up as described with 4 mM NADH and 6 mM formaldehyde as substrates.

2.2.6.3 Formaldehyde dehydrogenase assay

The formaldehyde dehydrogenase assay was similar to the methanol dehydrogenase assay with the exception that the reaction was started with the addition of 6 mM formaldehyde.

Similarly, a formate reductase assay, following the reverse direction, was also performed with NADH and formate as substrates.

2.2.6.4 Formate dehydrogenase assay

Formate dehydrogenase assay was similar to the formaldehyde dehydrogenase assay, but for the substrates being NAD^+ and 4 mM formate (sodium salt).

2.2.6.5 Temperature optimum and thermal inactivation

Temperature optimum

The temperature optimum for the enzyme activity was determined as follows. Enzyme assays were performed as previously described except for the temperature, which was adjusted to cover a range from 30 °C to 90 °C, in 5 °C increments. Each activity was recorded over a 3 min period.

Thermal inactivation studies

Thermostability studies were performed by heating enzyme aliquots in assay buffer over a range of temperatures, usually in 5 °C increments. At known time points, 20 µL aliquots were removed and immediately cooled on ice. The control, or time-zero point, consisted of enzyme in non-preheated assay buffer, and removing a sample immediately. Assays were performed at 50 °C and recording activity over a 3 min period.

Chapter 3- Cloning the Genes Encoding the Three Dehydrogenases Involved in Methanol Dissimilation in *Bacillus methanolicus* PB1

3.1 Introduction

The dissimilation of methanol to CO_2 can be achieved in three steps via three successive dehydrogenase reactions. The first of these consists of the oxidation of methanol to formaldehyde, while the next steps are further oxidations of formaldehyde to formate and finally to CO_2.

All thermotolerant methanol-utilising *Bacillus spp.* studied to date oxidise methanol by means of a novel NAD-dependent methanol dehydrogenase (MDH). This oxidises C_1-C_4 primary alcohols and also catalyses the NADH-dependent reduction of the corresponding aldehydes. The methanol dehydrogenase from thermotolerant methylotrophic *Bacillus* sp. was first isolated by Arfman *et al*. (1989) and the methanol dehydrogenase gene from *B. methanolicus* C1 was cloned and sequenced (de Vries *et al*., 1992). The deduced amino acid sequence exhibited similarity and the unique sequence motifs to those of other alcohol dehydrogenase (ADH) type III enzymes, which are distinct from the long-chain zinc-containing (type I) or short-chain zinc-lacking (type II) enzymes. In *B. methanolicus* C1, the *mdh* gene is 1,146 bp long, translating into a 381 amino acid protein. Analysis of the sequence revealed a new magnesium-dependent NAD(P)(H)-binding domain (Hektor *et al*., 2002).

The activity of MDH is strongly stimulated by a protein present in the soluble fraction of *B. methanolicus* C1 cells (Arfman *et al*., 1991). This activator was purified, and the molecular and biochemical characterisation of the 558 bp *act* gene revealed the presence of the highly conserved motif typical of Nudix hydrolase proteins in the deduced ACT 185 amino acid sequence (Kloosterman *et al*., 2002).

Recently, a natural 19 kb plasmid was isolated from *B. methanolicus* MGA3 (Brautaset *et al*., 2004). Sequence analysis of this pBM19 plasmid showed that it harboured a methanol dehydrogenase gene *mdh*, crucial for methanol consumption. In addition, five genes (*pfk*, phosphofructokinase; *rpe*, ribulose-5-phosphate 3-epimerase; *tkt*, transketolase; *glpX*, fructose-1,6-bisphosphatase; *fba*, fructose-1,6-bisphosphate aldolase) with deduced roles in methanol assimilation via the ribulose monophosphate (RuMP) pathway are also encoded by pBM19. An MGA3 strain cured of pBM19 could not maintain growth on methanol as sole source of carbon, while a pTB1.9 plasmid harbouring a complete *mdh* gene could not restore growth on methanol when introduced into the pBM19-cured strain, suggesting that the additional genes found on pBM19 are required for the organism to sustain growth on this carbon source. A survey of 13 different strains of thermotolerant methanol-utilising *Bacillus* spp found that they all carried plasmids similar to pBM19, suggesting plasmid-linked methylotrophy (Lidstrom and Wopat, 1984; Jakobsen *et al*., 2006).

Carbon assimilation following oxidation of methanol to formaldehyde was thought to be by way of the fructose-1,6-bisphosphate aldolase cleavage and a variant of the RuMP cycle. Failure of whole cells to oxidise formate and the absence of formaldehyde- and formate dehydrogenases indicated the operation of a non-linear oxidation sequence for formaldehyde (Arfman *et al*., 1989). However, a linear pathway for the dissimilation of formaldehyde to CO_2 was proposed based on ^{13}C-NMR data (Pluschkell *et al*., 2002). The accumulation of ^{13}C-formate in cultures fed on ^{13}C-labelled methanol indicated that *B. methanolicus* possesses formaldehyde dehydrogenase and formate dehydrogenase activities resulting in a methanol dissimilation pathway via CO_2. Accumulation of $^{13}CO_2$ in the vessel's exhaust gas confirmed that *B. methanolicus* dissimilates methanol to CO_2 by both a cyclic mechanism as part of the RuMP pathway and a proposed linear pathway to formate (Figure 3.1).

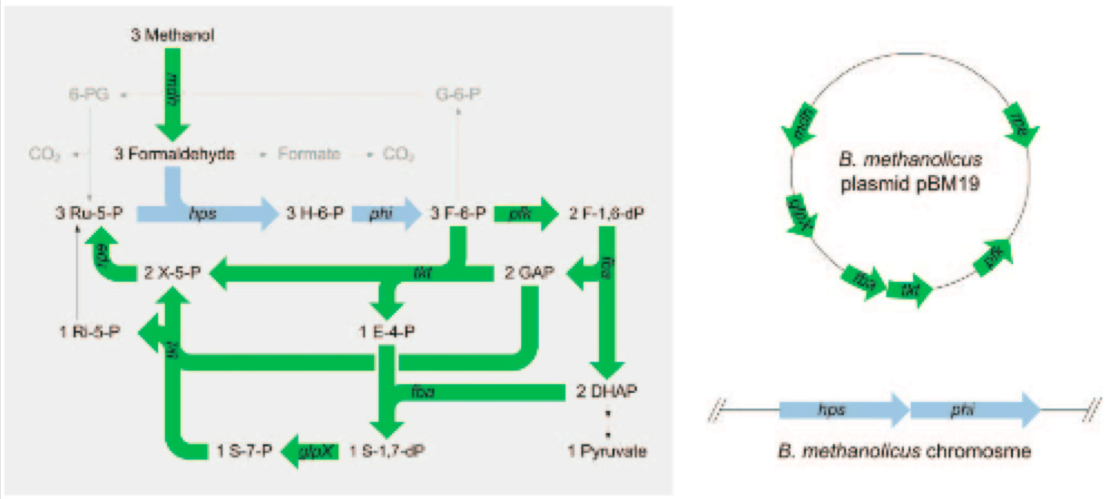


Figure 3.1. Graphical map of biochemical reactions and genes involved in methanol oxidation and assimilation in *B. methanolicus*. Pathways for linear and cyclic formaldehyde dissimilation are indicated (from Jakobsen *et al*., 2006).

pBM19 genes: *mdh*, methanol dehydrogenase; *pfk*, phosphofructokinase; *fba*, fructose bisphosphate aldolase; *tkt*, transketolase; *glpX*, fructose/sedoheptulose bisphosphatase; *rpe*, ribulose phosphate epimerase. Chromosomal genes: *hps*, hexulose phosphate synthase; *phi*, phosphohexuloisomerase. Metabolites: H-6-P, hexulose-6-phosphate; F-6-P, fructose-6-phosphate; F-1,6-dP, fructose-1,6-bisphosphate; DHAP, dihydroxyactetone phosphate; GAP, glyceraldehyde-3-phosphate; E-4-P, erythrose-4-phosphate; S-7-P, sedoheptulose-7-phosphate; S-1,7-dP, sedoheptulose-1,7-bisphosphate; X-5-P, xylulose-5-phosphate; Ri-5-P, ribose-5-phosphate; Ru-5-P, ribulose-5-phosphate; G-6-P, glucose-6-phosphate; 6-PG, 6-phosphogluconate.

To date, no dissimilatory pathway genes have been cloned from any *B. methanolicus* strain.

The purpose of this chapter is therefore to identify and clone, from *B. methanolicus*, the three dehydrogenase genes involved in the linear pathway of methanol dissimilation to CO_2.

3.2 Materials and Methods

3.2.1 Microbial cultures

Escherichia coli was cultured on LB medium, supplemented with antibiotics when appropriate (carbenicillin, 50µg.mL^{-1}; kanamycin, 50µg.mL^{-1}) at 37 °C with shaking.

Bacillus methanolicus PB1 NCIMB 13113 (ATCC 51375) was grown on TS medium supplemented with 1% (v/v) methanol, or on $MeOH_{200}$ medium, at 50 °C for 30 h with shaking.

Geobacillus thermoglucosidasius NCIMB 11955 was cultured on TS broth medium at 55 °C with shaking.

3.2.2 Genomic and plasmidic DNA extraction

B. methanolicus PB1 genomic and plasmid DNA and *Geobacillus thermoglucosidasius* genomic DNA were extracted as described in Chapter 2 – Materials and Methods.

3.2.3 Restriction digest of pBM19

pBM19 was extracted from *B. methanolicus* PB1, digested with *Eco*RI, *Bsa*I and *Pml*I. Restriction digests were set up and analysed as described in Chapter 2 – Materials and Methods (2.2.2.6).

3.2.4 pGEM®-T Easy cloning

All gene products were cloned into the pGEM®-T Easy vector (Figure 2.1 in 2.2.2.6)). The procedure is described in Chapter 2 – Materials and Methods.

3.2.4.1 *mdh* and *act*

The genes encoding the methanol dehydrogenase (MDH) and the activator ACT, *mdh* and *act* respectively, were PCR-amplified from *B. methanolicus* PB1 genomic DNA using the primer pairs Bmmdh_F (5'-ATGACAAACTTTTTCATTCC) and Bmmdh_R (5'-TTACAGAGCGTTTTTGATG), and Bmact_F (5'-ATGGGAAAATTATTGAGG) and Bmact_R (5'-TCATTTATGTTTGAGAGC), respectively. Plasmid-borne *mdh* was PCR-amplified from *B. methanolicus* PB1 pBM plasmid DNA using the primer pair pBMmdh_F (5'-ATGACAACAAACTTTTTCATTCC) and pBMmdh_R (5'-TTACATAGCGTTTTTGATG).

3.2.4.2 *falddh*

The gene encoding for the formaldehyde dehydrogenase FALDDH, *falddh*, was PCR-amplified from *B. methanolicus* PB1 genomic DNA using the primer pair Bmfalddh_F (5'-ATGAAGGCGGTTACG) and Bmfalddh_R (5'-TTACGGTTTAAAAACGAC).

3.2.4.3 *fdh*

The gene encoding for the formate dehydrogenase FDHa, *fdh*A, was PCR-amplified from *G. thermoglucosidasius* genomic DNA, using the primer pair Gthfdha_F (5'-ATGCTAAAAAACTACTCCACC) and Gthfdha_R (5'-TTAGCCACGCGGCTCTGC).

Table 3.1. Deoxyoligonucleotide primers used for PCR (Stop and Start codons in bold).

Primer name	Sequence 5'→3'	Tm (°C)
B. methanolicus		
pBM primers		
pBMmdh_F	**ATG**ACAACAAACTTTTTCATTCC	58.8
pBMmdh_R	**TTA**CATAGCGTTTTTGATG	53.7
genomic primers		
Bmmdh_F	**ATG**ACAAACTTTTTCATTCC	55.5
Bmmdh_R	**TTA**CAGAGCGTTTTTGATG	56.3
Bmact_F	**ATG**GGAAAATTATTGAGG	53.7
Bmact_R	**TCA**TTTATGTTTGAGAGC	50.7
Bmfalddh_F	**ATG**AAAGCTGTAACATATCAGG	57.3
Bmfalddh_R	**TCA**TGGTTTAAAAACGACTTTGATGC	67.2
G. thermoglucosidasius		
Gthfdha_F	**ATG**CTAAAAAACTACTCCACC	56.5
Gthfdha_R	**TTA**GCCACGCGGCTCTGC	66.5

PCR and cloning

The PCRs were performed using Phusion polymerase (NEB), followed by electrophoretic separation on a 0.8-1% agarose gel, and subsequent gel extraction using Promega's Wizard® kit. The purified amplified DNA was A-tailed using *Taq* polymerase and ligated into the vector pGEM®-T Easy using T4 ligase; the vector was then transformed into *E. coli* JM109, transformants were screened and the

vectors subsequently purified using Promega's Wizard® kit and their insert sequence determined.

3.2.5 Degenerate PCR

Degenerate PCR was performed when no sequence data were available for the gene of interest in the organism of interest.

Degenerate primers were designed based on conserved regions identified from protein sequence alignments of the gene of interest from different genera and species. The motifs were back-translated into nucleotides, and the primers further codon optimised using a codon usage table to reduce the degree of degeneracy.

Degenerate PCR was performed as described for PCR but using the degenerate primers and the FailSafe™ PreMix PCR kit (Epicentre®).

3.2.5.1 Formaldehyde dehydrogenase, *falddh*

Formaldehyde dehydrogenase genes in *B. cereus*, *B. clausii* and *B. thuringiensis* identified from the ENTREZ nucleotide database were aligned (http://www.ncbi.nlm.nih.gov/entrez). The sequence of the formaldehyde dehydrogenase from *Pseudomonas putida* was identified from the PDB (1KOL) and was added to an alignment of the *Bacilli* FALDDH accessions using ESPript (http://espript.ibcp.fr/). Degenerate primers (> 20 mers) were designed at the 5' and 3'-ends to allow the amplification of the full-length coding sequence. Internal degenerate primers were also designed at regions identified as offering a high degree of conserved residues. These internal degenerate primers sat at 1/3 and 2/3 of the length of the gene.

3.2.5.2 Formate dehydrogenase, *fdh*

Formate dehydrogenase genes were identified in *B. halodurans*, *B. cereus*, *B. clausii*, *B. subtilis* and *Geobacillus kaustophilus* from the ENTREZ database. The protein sequences were aligned in BOXSHADE. Degenerate primers were designed in the same way as for the formaldehyde dehydrogenase, at the 5' and 3'-end as well as internally along the sequence. The Boxshade server can be found at http://www.ch.embnet.org/software/BOX_form.html.

3.2.6 Genome walking

A genomic DNA library was created using *B. methanolicus* genomic DNA as a template. Up to 100 ng of genomic DNA was digested with 10 U of restriction enzyme in a final volume of 25 µL ($_{dd}H_2O$) and incubated for 3 h at 37 °C. Restriction enzymes were heat killed by incubation at 65 °C for 20 min. *Ca*. 50 ng digested genomic DNA was self-ligated using 10 U T4 DNA ligase (NEB) to a final volume of 100 µL ($_{dd}H_2O$), and incubated at 16 °C overnight. The ligase was then denatured. Two-rounds of nested PCR were performed on the circular DNA, using gene specific primer (GSP) pairs positioned on the edges of the known sequence of the gene of interest and facing outwards (*i.e.* 5'_Reverse and 3'_Forward) (Figure 3.2). The first round of genome walking PCR (gwPCR) was set up as previously described for PCR, using Phusion™ DNA polymerase, with the exception of the use of 2 µL digested gDNA template and 25 pM outer GSP, made up to a final volume of 50 µL (MilliQ water). Thermocycling conditions were adjusted for the 1st round PCR to the following: hot start 98 °C for 5 min, 35 cycles [98 °C for 15 s, 55 °C for 15s, 72 °C for 5 min] and final extension at 72 °C for 10min. In the 2nd round (nested) PCR, the conditions were the same but for using 25pM inner nested GSP, 2 µL 1st round PCR product as template, 30 amplification cycles, and allowing for a 15min final extension. An aliquot (5 µL) was run on a 0.8% agarose gel. If a product was detected, the remainder of the reaction was A-tailed and cloned into the cloning vector pGEM®-T Easy, transformed into *E. coli* JM109, and the insert sequenced with primers pT7 and pSP6.

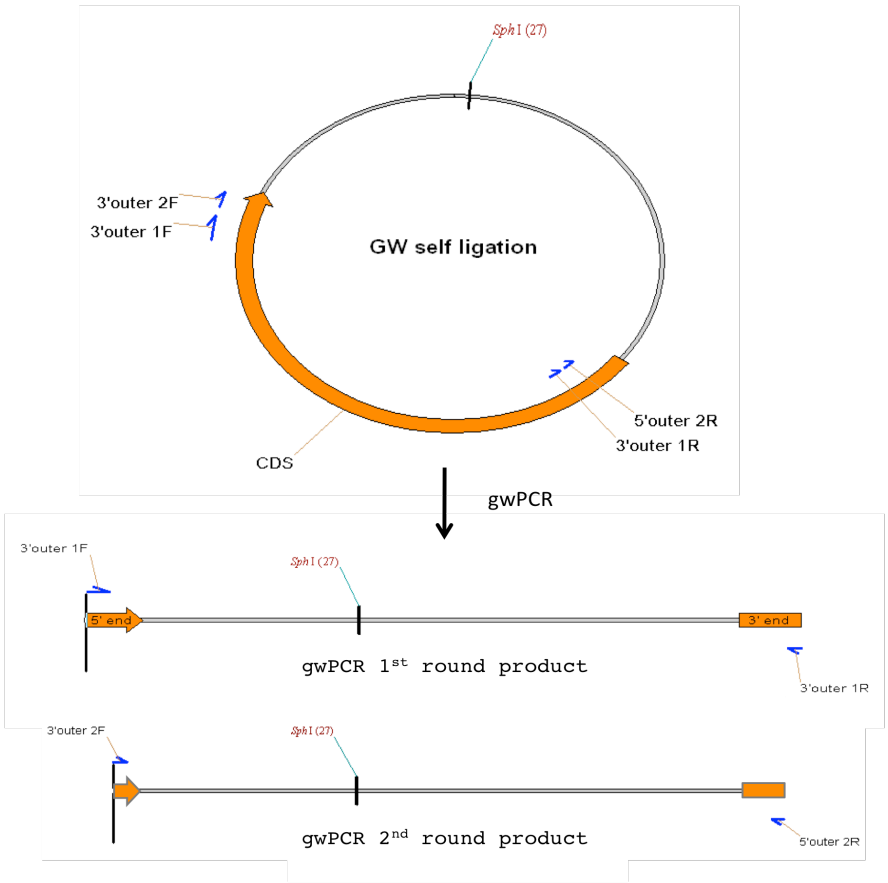


Figure 3.2. Genome-walking procedure following self-ligation of the genomic DNA library. Outer gene-specific primers are used for nested PCR: 3'outer_1F vs. 5'outer_1R, and 3'outer_2F *vs.* 5'outer_2R respectively (primers are annotated with the blue arrows, the gene of interest is shaded orange). The resulting PCR product starts with the 3'end of the gene reading into the 3' flanking sequence. At the other end, the 5'flanking sequence reads into the 5'end of the gene. The border between the 3' and 5' flanking sequences is marked by the restriction site used for generating the library (here, *Sph*I).

3.3 Results

3.3.1 *Bacillus methanolicus* PB1 (ATCC 51375) does contain a pBM19-like plasmid

To elucidate whether *B. methanolicus* PB1 carries a pBM19-like plasmid, the published sequence of the full plasmid pBM19 from *B. methanolicus* MGA3 was

studied (Brautaset *et al*., 2004; GenBank: AY386314/NC_005328). The plasmid pBM19 features three *Bsa*I and a single *Eco*RI restriction sites, generating fragments of 8 kb, 6 kb and 5 kb (*Bsa*I), and 19 kb (*Eco*RI). The restriction enzyme *Pml*I also cuts at a unique site within pBM19, therefore linearising it as does *Eco*RI, but its recognition site lies within the *mdh* gene on the plasmid. If strain PB1 contains a pBM19-like plasmid, it is likely that it will follow the restriction pattern predicted for strain MGA3.

The plasmid DNA purified from *B. methanolicus* PB1 was digested with *Eco*RI, *Bsa*I and *Pml*I respectively, and the pattern produced was observed on a 0.8% low-melting point agarose gel (Figure 3.3). Restriction digests with *Eco*RI or *Pml*I produced a single band that only migrated high above the 10 kb marker, and that may be at *ca.* 19kb. The size was estimated only through running it alongside the uncut plasmid DNA as the size is above the top marker of the 1 kb DNA ladder at 10 kb. *Eco*RI and *Pml*I appeared to linearise the plasmid DNA extracted from *B. methanolicus* PB1, supporting the fact that *Eco*RI and *Pml*I are single cutters within this plasmid. However, restriction with *Bsa*I produced three bands of estimated sizes 8 kb, 6 kb and 5 kb, indicating the presence of (at least) three *Bsa*I restriction sites within the plasmid sequence. The total size of the plasmid is therefore confirmed as being of *ca*. 19 kb, by adding the sizes of these individual contiguous fragments.

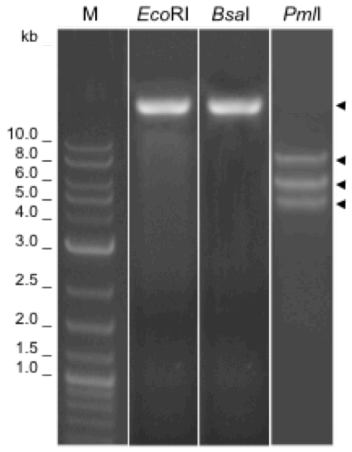


Figure 3.3. 0.8% low melting point agarose gel electrophoresis of the restriction digest with *Eco*RI, *Bsa*I and *Pml*I of the plasmid extracted from *B. methanolicus* PB1. M: 1 kb DNA ladder.

3.3.2 Cloning of the *mdh* and its activator *act* (pGEM®-T Easy)

3.3.2.1 *Bacillus methanolicus* harbours two copies of the *mdh* gene

The protein sequences for both copies of *B. methanolicus* MDH were aligned in Clustal W (www.ebi.ac.uk) (Figure 3.4). Whereas the chromosome-encoded MDH of strain C1 (GenBank: P31005) is 381 amino acids long, the plasmid-borne MDH of strain MGA3 (GenBank: NP_957659) is 382 amino acids long, due to the insertion of an extra threonine at position 3. Over the full-length of the protein sequences, the two MDHs differ by 10 amino acids overall (9 substitutions and one insertion) resulting in a 97.4% identity and 99.0% similarity. Conversely, alignment of the corresponding nucleotide sequences (GenBank accessions M65004 and NC_005328) revealed a number of mutations between the two *mdh* sequences. Sharing 97.2% identity, most of these mutations are silent, while the few non-silent mutations result in the amino acids substitutions and one insertion observed.

```
P31005|BmMDH       M-TNFFIPPASVIGRGAVKEVGTRLKQIGAKKALIVTDAFLHSTGLSEEVAKNIREAGLD 59
NP_957659|pBMMDH   MTTNFFIPPASVIGRGAVKEVGTRLKQIGAKKALIVTDAFLHSTGLSEEVAKNIREAGVD 60
                   * ************************************************************:*

P31005|BmMDH       VAIFPKAQPDPADTQVHEGVDVFKQENCDALVSIGGGSSHDTAKAIGLVAANGGRINDYQ 119
NP_957659|pBMMDH   VAIFPKAQPDPADTQVHEGVDVFKQENCDSLVSIGGGSSHDTAKAIGLVAANGGRINDYQ 120
                   ****************************:*****************************

P31005|BmMDH       GVNSVEKPVVPVVAITTTAGTGSETTSLAVITDSARKVKMPVIDEKITPTVAIVDPELMV 179
NP_957659|pBMMDH   GVNSVEKPVVPVVAITTTAGTGSETTSLAVITDSARKVKMPVIDEKITPTVAIVDPELMV 180
                   ************************************************************

P31005|BmMDH       KKPAGLTIATGMDALSHAIEAYVAKGATPVTDAFAIQAMKLINEYLPKAVANGEDIEARE 239
NP_957659|pBMMDH   KKPAGLTIATGMDALSHAIEAYVAKGATPVTDAFAIQAMKLINEYLPKAVANGEDIEARE 240
                   ************************************************************

P31005|BmMDH       AMAYAQYMAGVAFNNGGLGLVHSISHQVGGVYKLQHGICNSVNMPHVCAFNLIAKTERFA 299
NP_957659|pBMMDH   KMAYAQYMAGVAFNNGGLGLVHSISHQVGGVYKLQHGICNSVNMPHVCAFNLIAKTERFA 300
                    ************************************************************

P31005|BmMDH       HIAELLGENVSGLSTAAAAERAIVALERYNKNFGIPSGYAEMGVKEEDIELLAKNAFEDV 359
NP_957659|pBMMDH   HIAELLGENVAGLSTAAAAERAIVALERINKSFGIPSGYAEMGVKEEDIELLAKNAYEDV 360
                   **********:***************** **.************************:***

P31005|BmMDH       CTQSNPRVATVQDIAQIIKNAL 381
NP_957659|pBMMDH   CTQSNPRVETVQDIAQIIKNAM 382
                   ********.************:
```

Figure 3.4. Protein sequence alignment for *B. methanolicus* C1 MDH (P31005) and *B. methanolicus* MGA3 pBM MDH (NP_957695). Amino acids are annotated according to their single letter code (SLC). Conserved residues are indicated with a (*). Conflicting but similar residues are indicated with (:) (BLOSUM62 score no less than -1) (grey shading) or (.) (BLOSUM62 score of -2) (grey shading, white lettering). Conflicting residues with a BLOSUM62 score of -3 or less are marked with ( ) (black shading, white lettering). The regions used for primer design are underlined with an arrow.

Primer pairs were designed at the 5' (Bmmdh_F and pBMmdh_F) (5'-**ATG**ACAAACTTTTTCATTCC and 5'-**ATG**ACAACAAACTTTTTCATTCC) and 3'end (Bmmdh_R and pBMmdh_R) (5'-**TTA**CAGAGCGTTTTTGATG and 5'-**TTA**CATAGCGTTTTTGATG) for the chromosomal and plasmid-bourne *mdh* genes, based on the gene sequences in strain C1 (*mdh*: M65004) and strain MGA3 (*pBMmdh*: NC_00532).

PCR amplification yielded fragments of the expected 1.15 kb for both *Bmmdh* and *pBMmdh* from their respective DNA sources, *B. methanolicus* PB1 genomic and plasmid DNA. Figure 3.5 shows the discrete bands obtained for the products when separated on a 1% agarose gel.

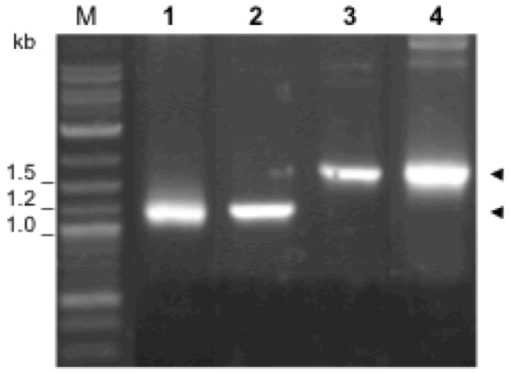


Figure 3.5. 1% agarose gel electrophoresis of PCR-amplified *mdh* gene products from *B. methanolicus* PB1 genomic and plasmid DNA. Lanes: 1- *Bmmdh*, 2- *pBMmdh*, 3- long *Bmmdh*, 4- long *pBMmdh*; M: 1 kb DNA ladder. Phusion® was used as the DNA polymerase.

The gel-purified bands for *Bmmdh* and *pBMmdh* were A-tailed and cloned into pGEM®-T Easy and transformed into JM109 *E. coli* cells. Sequencing data revealed that *Bmmdh* and *pBMmdh* exhibit 99.5% identity over the full-length coding sequence, with one mutation (C1141/A1144) and an in-frame insertion of three nucleotides at bp 7 (^{7}ACA) in *pBMmdh*. Over the aligned translated products, MDH and pBMMDH, these correspond to two amino acid differences: the threonine insertion at position 3 in pBMMDH and the L381/M382 substitution. Over the remainder of the protein sequence, MDH and pBMMDH are identical.

The conflicts observed in the sequences of MDH and pBMMDH could have been introduced by the primers used for PCR since these primers covered the T3 and L381/M382 regions and their design was based on the available sequence for MDH

(strain C1) and pBMMDH (strain MGA3). Therefore, these mutations could have been artefacts and imported from the strains C1 and MGA3, on which the design of the primers was based. To elucidate this point, two PCRs were set up with genomic and plasmid DNA of strain PB1 as templates. Since the full sequence of the *mdh*-harbouring pBM19 plasmid is known in strain MGA3 (NC_005328) and the region flanking the chromosomal *mdh* in strain C1 has also been sequenced (M65004) (Brautaset *et al*., 2004; de Vries *et al*., 1992), primer pairs were designed for each DNA source so that they would cover the full-length coding sequence of either variant plus an additional *ca*. 200 bp at the 5'- and 3' ends generating the "long" version of the genes. The following primer pairs were then synthesised: Bmmdh(--230)_F 5'-CCCTTCCACTTTAATCCTCC and Bmmdh(+136)_R 5'-CGTACCGCCTTTGTTTTTCG, pBMmdh(-239)_F 5'-CCCTTCCACCTTAACC and pBMmdh(+173)_R 5'-CCTATGGCGGGATTCG. Across their respective overlap, Bmmdh(-230)_F and pBMmdh(-239)_F possess only two substitutions (-221T/-230C, -216T/-225C), reinforcing the observation that the two *mdh* genes are extremely similar.

PCR yielded the expected 1.5 kb products: 1.1 kb of coding sequence flanked by 200 bp either side (Figure 3.5, lanes 3 and 4). Sequence analysis also confirmed the observations ascertained above that, in strain PB1, both copies of the *mdh* are extremely similar, and the uncertainty over the differences observed at the amino acid level are indeed real.

Pair-wise alignment of these *mdh* genes from strain PB1 with those of strains C1 and MGA3 shared similar percentage identities: the *mdh* genes of strains PB1 and C1 share 94% identity while the *pBMmdh* genes shared 93% identity in strain PB1 and MGA3 (95% and 94% identity, 97% and 96% similarity, in the deduced amino acid sequences, respectively).

3.3.2.2 An activator *act*

PCR performed on the genomic DNA of *B. methanolicus* PB1 using primers Bmact_F and Bmact_R yielded a *ca*. 500 bp product (Figure 3.6). Sequence analysis revealed that the product was 558 bp long (inclusive of the stop codon), the same as the length of the activator gene described in strain C1, with a deduced 185 amino acid sequence. The activator sequences from strain PB1 and strain C1 were aligned

and 30 nucleotides substitutions could be observed, resulting in a 95% identity over the full-length coding sequence. At the protein level, these translated into a 97% identity (100% similarity), with the 5 amino acid substitutions observed being semi-conservative: I52V, V88I, E101Q, D153E and K162Q.

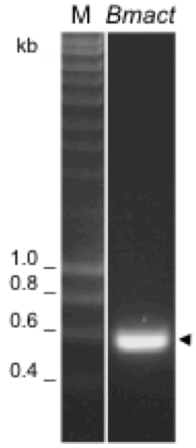


Figure 3.6. 1% agarose gel electrophoresis of PCR-amplified *act* gene product from *B. methanolicus* PB1. M: 1 kb DNA ladder. Phusion® was used as the DNA polymerase.

To elucidate whether the 5'- and 3' ends of the activator gene *act* product sequence were carried over from strain C1 via the primers themselves, a new set of primers was produced that would generate a PCR product containing the full-length *act* gene flanked by *ca*. 100 bp either side. Sequence analysis of the 800 bp product generated confirmed the 5'- and 3' end sequences obtained as real PB1.

3.3.3 Identification and cloning of a formaldehyde dehydrogenase from *Bacillus methanolicus*

3.3.3.1 Degenerate PCR

Formaldehyde dehydrogenase (FALDDH) protein sequences from different *Bacillus* species were identified on the ENTREZ database. A formaldehyde dehydrogenase from *Pseudomonas putida* was identified within the RSCB Protein Data Bank (PDB: 1KOL); it is the only formaldehyde dehydrogenase whose crystal structure has been solved.

Formaldehyde dehydrogenases from *Bacillus cereus* (ATCC[14579] and EL33) (GenBank: AAP10036, YP_084425), *B. clausii* (BAD65961) and *B. thuringiensis* (YP_037205) were aligned on ESPript, along with the formaldehyde dehydrogenase from *Pseudomonas putida* (1KOL). Conserved motifs were highlighted in the alignment and the regions selected for designing degenerate primers (Figure 3.7).

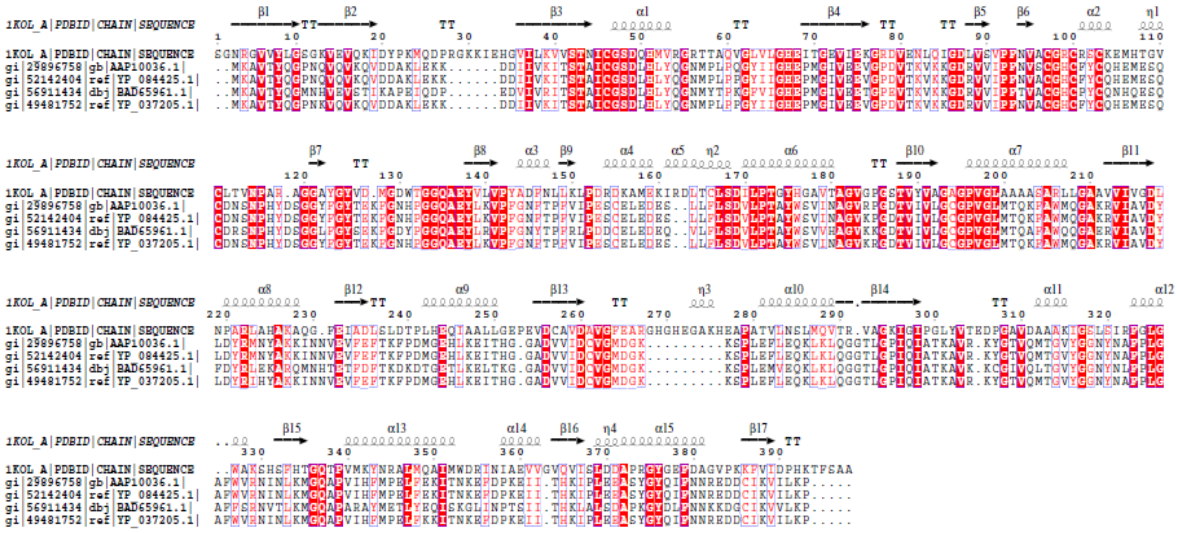


Figure 3.7. Alignment of formaldehyde dehydrogenase sequences from *B. cereus* ATCC14579 (AAP10036) and EL33 (YP_084425), *B. clausii* (BAD65691) and *B. thuringiensis* (YP_037205). The top line of the alignment corresponds to the formaldehyde dehydrogenase form *P. putida* (1KOL) annotated for its tertiary structure. Identical residues are shaded red, and similar residues are in red. The *Bmfalddh* degenerate forward and reverse primers were positioned over [1]MKAVTYQGP (Bmfalddh_1Fdg), [131]EYLK/RVDF (Bmfalddh_4F/Rdg), [251]IDCVGMDGK (Bmfalddh_6F/Rdg), and [369]CIKVI/VLKP• (Bmfalddh_•Rdg),. The alignment was performed on ESPript (http://espript.ibcp.fr/ESPript).

A first set of degenerate primers was designed at the 5' and 3'-end of the alignment, *i.e.* Bmfalddh_1Fdg and Bmfaddh_•Rdg, acting as the extremities of a putative full-length *falddh* from *B. methanolicus* PB1. These were respectively positioned over the following regions: [1]MKAVTYQGP (N-terminus) and [369]CIKVI/VLKP• (C-terminus). Sets of internal degenerate primer pairs were also designed at two sites of highly conserved regions, [131]EYLK/RVDF and [251]IDCVGMDGK, respectively, yielding Bmfalddh_4Fdg and _4Rdg, and Bmfalddh_6Fdg and _6Rdg. The amino acid sequence from *Pseudomonas putida* was largely ignored in the design of these degenerate primers as the sequences of the four FALDDHs from the

Bacilli selected showed a higher degree of conservation between themselves than with that from *P. putida*.

Table 3.2. *Bacillus* formaldehyde dehydrogenase degenerate primers (IUPAC degeneracy code: R= A+G, Y= C+T, K= G+T, M= A+C, S= G+C, W= A+T, B= G+C+T, D= A+G+C, H= A+C+T, V= A+C+G, N= A+C+G+T).

Primer name	Sequence (5' → 3')	dg	Tm (°C)
Bmfalddh_1Fdg	**ATG**AARGCNGTNACNTAYCARGG	512	64.0
Bmfalddh_4Fdg	GARTAYCTNARRGTNGAYTTT	512	55.5
Bmfaldh_4Rdg	AAARTCNACYYTNAGRTAYTC	512	52.5
Bmfalddhd_6Fdg	GAYTGYGTNGGNATGGAYGGNAA	512	68.6
Bmfalddh_6Rdg	TTNCCRTCCATNCCNACRCARTC	512	68.6
Bmfalddh_•Rdg	**TTA**NGGYTTNAGDAYNACYTTDATRCA	9216	67.8

A first attempt at PCR with primers Bmfaldh_1Fdg and Bmfalddh_•Rdg over *B. methanolicus* PB1 genomic DNA proved unsuccessful. No full-length product was generated. However, having a set of degenerate primers evenly distributed along the sequence of the FALDDH, PCRs were set up using pair-wise combinations of forward and reverse degenerate primers with all five combinations covered (Table 3.2). Expected products covered the entire putative coding sequence for a *B. methanolicus falddh* gene (Figure 3.8-A). Consequently, these overlapping PCR products could be fused and used to generate a putative full-length product. Using the FailSafe™ PCR system (Epicentre®), all five products were obtained but again no full-length product generated when pairing primer Bmfalddh_1Fdg with Bmfalddh_•Rdg (Figure 3.8-B-E).

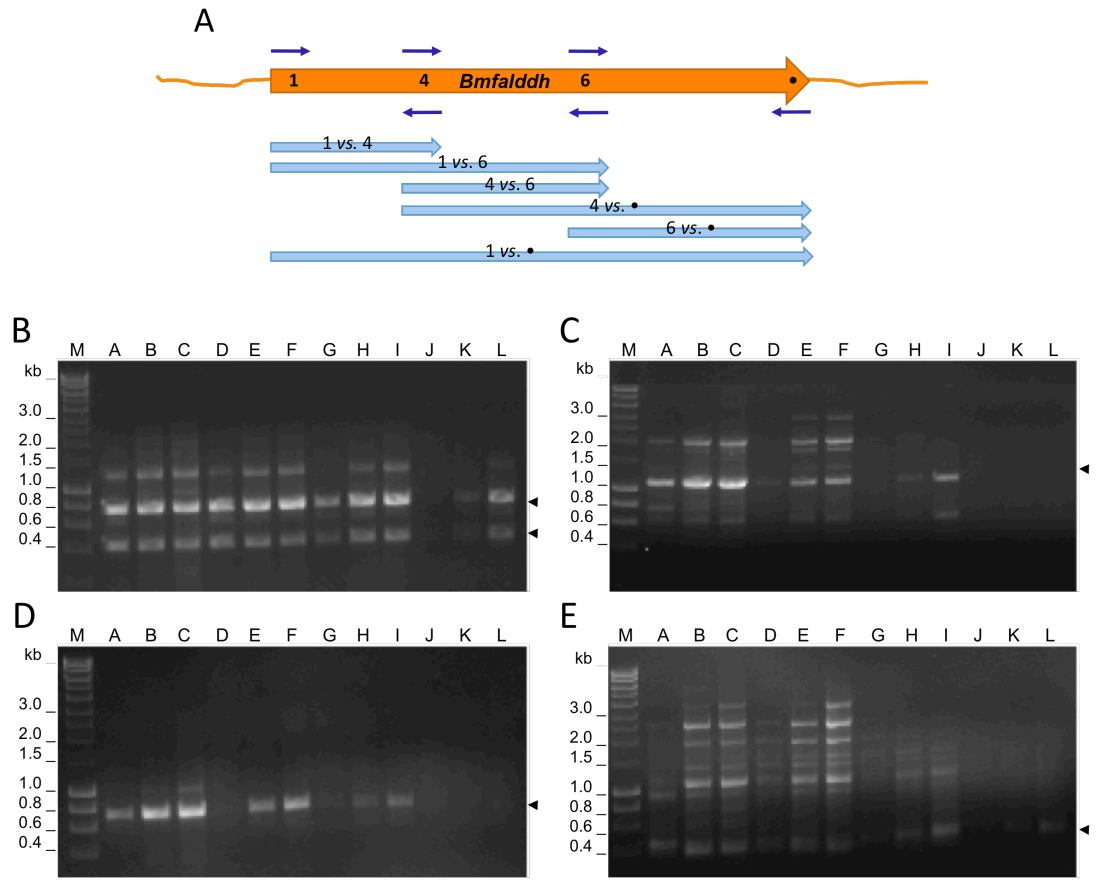


Figure 3.8. (A) Schematics of the pair-wise and overlapping PCR. Primer positions are given and the individual PCR products shown. (B-E) 1% agarose gel electrophoresis of dgPCR-amplified *falddh* gene product from *B. methanolicus* PB1. (B) Bmfalddh_1Fdg *vs.* _4Rdg (*ca.* 400 bp) and _1Fdg *vs.* _6Rdg (*ca.* 800 bp); (C) Bmfalddh_1Fdg *vs.* _8Rdg (*ca.* 1.2 kb); (D) Bmfalddh_4Rdg *vs.* _8Rdg (*ca.* 800 bp); (E) Bmfalddh_6Fdg *vs.* _8Rdg (*ca.* 400 bp). Lanes A-L :FailSafe buffers A-L; M: 1 kb DNA ladder (kb). FailSafe™ was used as the DNA polymerase.

Following electrophoretic purification and gel extraction, the inserts were cloned into pGEM®-T Easy, sequenced and analysed, revealing overlapping sections of the sequence for a putative *falddh* in *B. methanolicus* PB1. By analysing the sequence of the overlapping products obtained with primer pairs Bmfaldh_1Fdg and Bmfaldh_6Rdg, and Bmfaldh4Fdg and Bmfalddh_•Rdg, respectively, the regions corresponding to the degenerate internal primers _4Fdg and _6Rdg were also resolved. In doing so, a full-length putative *falddh* gene was constructed for *B. methanolicus* with only the 5'- and 3' ends still degenerate.

Since PCR performed using pair-wise combinations of degenerate primers did not generate the putative full-length gene product, the degenerate primers were refined. Looking at the codon usage data collated from the sequences obtained for the *mdh*, *act* and *falddh* genes in *B. methanolicus* PB1, and correlating these with a codon usage table for *Bacillus*, the degree of degeneracy of the primers was decreased even further by eliminating rare and under-represented codons. The refined degenerate primers became Bmfalddh_1Fdg (5'-ATGAAGGCGGTNACGTAYCAGGG, dg = 8) and Bmfalddh_•Rdg (5'-TTAGGGTTTRAAAACYAATTTGATGC, dg = 4).

When PCR was performed using these refined primers Bmfalddh_1Fdg and Bmfalddh_•Rdg, and Phusion™ as the polymerase, a 1.2 kb product was amplified from *B. methanolicus* PB1 genomic DNA (Figure 3.9). Sequence of the cloned product was confirmed as the full-length 1134 bp coding sequence for a *falddh*.

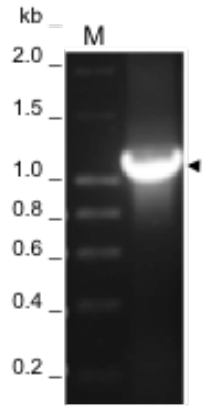


Figure 3.9. 1% agarose gel electrophoresis of dgPCR-amplified *falddh* gene product from *B. methanolicus* PB1 genomic DNA, using primers Bmfalddh_1Fdg *vs.* Bmfalddh_•Rdg. M: 1 kb DNA ladder (kb). Phusion® was used as the DNA polymerase.

This product length is also compatible with the length observed in other *Bacilli* formaldehyde dehydrogenase genes, translating into a 377 amino acid protein as noted previously. Furthermore, when translated and aligned against the other FALDDHs from *Bacilli*, the protein from *B. methanolicus* PB1 showed high similarity with the alcohol dehydrogenase glutathione-dependent formaldehyde dehydrogenases (adhB) from *B. cereus* (80% identity, 92% similarity) and *B. thuringiensis* (81% identity, 91% similarity) and the glutathione-dependant formaldehyde dehydrogenase from *B. clausii* (67% identity, 81% similarity). The annotated NAD(H) binding site, the Zn binding and catalytic sites are also conserved

throughout (NAD(H) binding site at residues 38-40, 43, 85, 161, 165, 184-185, 187-189, 208-210, 214, 229, 253-255, 279, 296-298, 321-323, 363; Zn binding site: residues 38, 40, 59, 161 and 89-90, 92, 95, 103).

However, this still left the regions underlying the primers themselves as unresolved regarding their degeneracy, and whether they were truly part of the *falddh* gene of *B. methanolicus* PB1, its true 5'- and 3' end, or were imported into the PCR product through the design of the primers. In such cases, the *falddh* gene sequenced so far would be a chimera with a large core sequence of *B. methanolicus* origin, but the 5'- and 3' ends from a "*Bacilli*" background as observed on the sequence alignment of *Bacilli* FALDDH (Figure 3.7). In order to resolve these degenerate regions and reveal the true start and end of the putative *falddh* in *B. methanolicus*, and potentially extend the sequence knowledge into the adjacent upstream (5'-) and downstream (3'-) regions of the CDS, genome walking was performed.

3.3.3.2 Resolving the ends

Starting from the DNA sequence for a formaldehyde dehydrogenase in *B. methanolicus* obtained above, sets of internal primers were designed along its sequence, especially close to the extremities and facing outwards. Figure 3.10 and Table 3.3 detail the position and sequence of these primers.

Table 3.3. Internal deoxyoligonucleotide primers used for genome walking and nested PCR for *falddh* in *B. methanolicus* genomic DNA.

Primer name	Sequence 5'→3'	Tm (°C)
Bmfalddh_2eR	CCTTTACTACTAAGGTTTTAGGTCC	58.4
Bmfalddh _2bR	GACCCGCATCTGGCTGTAGATG	69.8
Bmfalddh _2R	GTTTCCTTGATAGATGTGAAGGTC	61.6
Bmfalddh _2F	GACCTTCACATCTATCAAGGAAAC	61.6
Bmfalddh _3cR	CCTTCAGGCAACGGAAAGTTTCC	70.3
Bmfalddh _3bR	CGTGTCCAATAATATAACCTTCAGG	63.1
Bmfalddh _3R	CTTCTTCCACGATCCCCATCG	69.7
Bmfalddh _3F	CGATGGGGATCGTGGAAGAAG	69.7
Bmfalddh _4dR	CATTAAACGGGACAACAACAC	61.6
Bmfalddh _4R	GGGGTGAAATTTCCAAAAGGCAC	69.8
Bmfalddh _4F	GTGCCTTTTGGAAATTTCACCCC	69.8
Bmfalddh _6R	CCATCCATACCGACGGCATCAAT	69.2
Bmfalddh _6F	ATTGATGCCGTCGGTATGGATGG	69.2
Bmfalddh _6dR	GTAGTTGCCTCCATATACACC	57.5
Bmfalddh _6dF	GGTGTATATGGAGGCAACTAC	57.5
Bmfalddh _7R	GGTGCCTGTCCCATTTTAATTGTAAC	67.5
Bmfalddh _7F	GTTACAATTAAAATGGGACAGGCACC	67.5
Bmfalddh _7bF	CCAATTATGCCGAAGCTATTTG	64.0
Bmfalddh _7cF	GATCCAAAAGAAATTATAACACAC	57.3
Bmfalddh _7dF	CCTCTTAATGATGCAAGCCATGGC	69.0

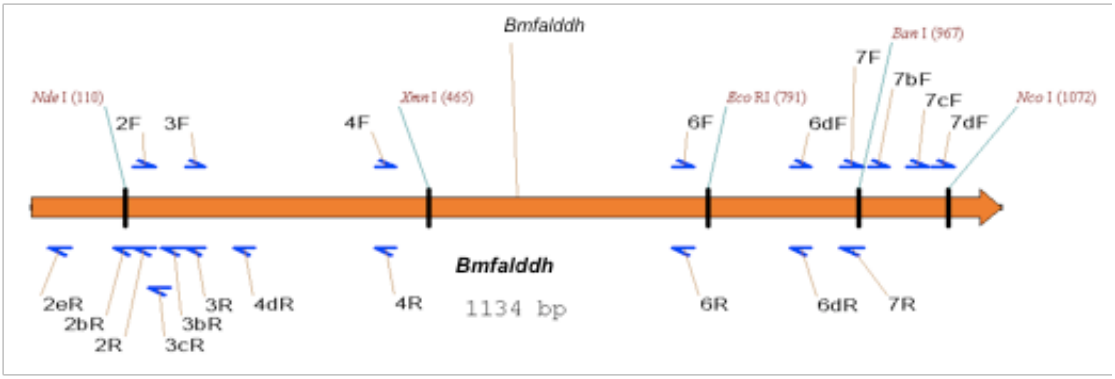


Figure 3.10. Schematic representaion of the formaldehyde dehydrogenase ORF (orange arrow) in *B. methanolicus* and position of the internal Bmfalddh primers designed.

A *B. methanolicus* genomic DNA library was constructed by restriction of the genomic DNA with the following enzymes: *Ban*I, *Eco*RI, *Hind*III, *Kpn*I, *Mlu*I, *Not*I, *Nsi*I, *Pst*I, *Sac*II, *Sph*I, *Xma*I and *Xmn*I. *Xmn*I cuts within *Bmfalddh* towards its 5' end, while *Eco*RI is cutting in the middle of the sequence and *Ban*I near the

3' end. Nested PCR was performed on the individual self-ligated libraries. Primer pairs used for the libraries were Bmfalddh_3R and Bmfalddh_7F in the 1st round, and Bmfalddh_2R and Bmfalddh_7bF in the 2nd round. For the *Xmn*I, *Eco*RI and *Ban*I libraries, two sets of nested PCR could be set up, each one directed towards the 5' or the 3' end. For the *Xmn*I library, primer pairs were Bmfalddh_7F and Bmfalddh_6dR (3' end), and Bmfalddh_3F and Bamfalddh_3bR (5' end) for the 1st round, and Bmfalddh_7bF and Bmfalddh_6R (3' end), and Bmfalddh_4F and Bmfalddh_2R (5' end) in the 2nd round. For the *Eco*RI libraries, primer pairs were, for walking towards the 3' end, Bmfalddh_7bF and Bmfalddh_7R for the 1st round, and Bmfalddh_7cF and Bmfalddh_6dR for the 2nd round; towards the 5' end, Bmfalddh_3F and Bmfalddh_3bR (1st round) and Bmfalddh_4F and Bmfalddh_2R (2nd round). Similarly, for the *Ban*I library, primer pairs were Bmfalddh_6F and Bmfalddh_3R, and Bmfalddh_6dF and Bmfalddh_2R (1st and 2nd round, 5' end).

Following the two rounds of nested PCR, the following discrete bands were observed. A 2 kb band was produced with the *Hind*III library, a 1.8 kb band in the *Nsi*I library and a 2.5 kb band in the *Sph*I library (Figure 3.11). These were gel-extracted, purified, cloned in the pGEM-T easy vector and sequenced.

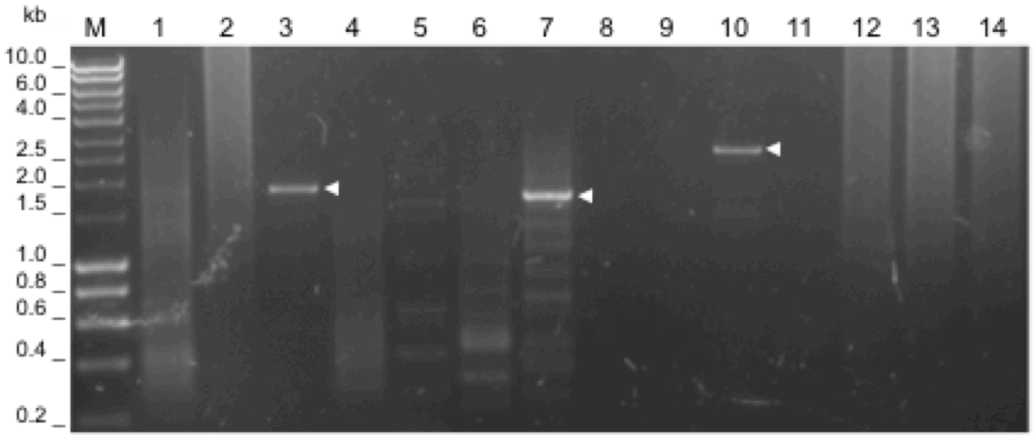


Figure 3.11. Genome walking on *B. methanolicus* genomic DNA for *Bmfalddh*. 2nd round PCR. Lanes are M-1 kb ladder, 1-*Ban*I (5'), 2-*Eco*RI (5'), 3-*Hind*III, 4-*Kpn*I, 5-*Mlu*I, 6-*Not*I, 7-*Nsi*I, 8-*Pst*I, 9-*Sac*II, 10-*Sph*I, 11-*Xma*I, 12-*Xmn*I (5'), 13-*Eco*RI (3') and 14-*Xmn*I (3'). Primer pair used were: lane 1: Bmfalddh_3F *vs*. 3bR; lane 2 and 12: Bmfalddh_4F *vs*. 2R; lanes 3-11: Bmfalddh_7R *vs*. 2bR; lane 13: Bmfalddh_7cF *vs*. 6dR; lane 14: Bmfalddh_7bF *vs*. 6R. Phusion® was used as the polymerase.

When assembled, the contigs generated coverage of over 2.2 kb of the *B. methanolicus* PB1 *falddh* gene region, revealing the start ATG and end of the gene.

At the 3' end, over 750 bp of sequence was unravelled. The stop codon TAA was identified by extending the open reading frame, while the degeneracy of the few nucleotides that were part of the Bmfalddh_•Rdg primer was resolved. The C-terminus of the protein was thus confirmed as DNCIKLVFKP•. This is very similar to the sequence consensus observed for the FALDDH from *B. cereus, B. thuringiensis* and *B. clausii* where the C-termini of their respective FALDDHs read ^{367}DDCIKVILKP• (*Bce* and *Bth*) and ^{367}DGCIKVVLKP• (*Bca*) (Figure 3.7). Similarly at the 5' end, the sequence data were extended by over 300 bp, the degeneracy of the Bmfalddh_1Fdg primer was resolved, and the N-terminus of the FALDDH from *B. methanolicus* PB1 confirmed as ^{1}MKAVTYQG, identical to that in *B. cereus, B. thuringiensis* and *B. clausii*. A putative ribosome binding site (AAAGGAGG) was identified 9 bp upstream of the start codon ATG along with a possible -35 promoter motif (TAATATT) 30 bp upstream of the start codon. Yet, another start codon was observed, 48 bp upstream and in-frame of the first ATG translating into the extended open reading frame MAAENIVGDLLERCSR^{1}MKAV. This longer ORF translates into a 393 amino acid protein. A putative ribosome binding site motif AGAAGA and a possible -35 promoter motif TAATATA were identified but respectively at -15 bp and -53 bp of this second ATG. However, the accepted distance for a RBS motif is 8 - 12bp upstream of the start codon, and *ca*. 35 bp for the promoter motif (Shine and Dalgarno, 1975; Rocha *et al*., 1999). A third in-frame start codon was identified 237 bp upstream of the primary ATG. The corresponding ORF would yield 79 extra amino acids at the N-terminus of the FALDDH, making it 455 residues long. However, neither a RBS nor a -35 motif could be identified in the 87 bp of upstream sequence data available and therefore this 3rd putative start was discarded.

The formaldehyde dehydrogenase gene cloned from *B. methanolicus* PB1 is 1134 bp long, with a 376 amino acids deduced translated protein. When comparing this FALDDH from *B. methanolicus* to formaldehyde dehydrogenases from other *Bacilli* species, it shares 80% identity (91% similarity) with that of *B. cereus* and *B. thuringiensis*, and 67% identity (81% similarity) with that of *B. clausii*.

3.3.4 Identification and cloning of a Formate dehydrogenase

3.3.4.1 From *Bacillus methanolicus*

Formate dehydrogenase (FDH) sequences from different *Bacillus* and *Geobacillus* species were identified on the ENTREZ database. FDH from *Bacillus halodurans*, *B. cereus*, *B. clausii*, *B. subtilis* and *Geobacillus kaustophilus* were aligned on Boxshade (http://www.ch.embnet.org/software/BOX_form.html) (Figure 3.12). Most formate dehydrogenases contain 900 to 1000 residues, corresponding to ORFs of *ca*. 3 kb. Based on the multi-species alignment, degenerate primers were designed over the regions of most conserved residues and spaced accordingly in order to cover the widest range over a potential ORF in *B. methanolicus* genome. Conserved regions identified consisted of the following motifs (with for reference the position in *B. cereus*): ^{1}MA/NE/NQ/NM/TVT/R/H/SV/IT, ^{97}LLYCTVCD, ^{204}CPCNAMLEK, ^{307}KFGWDFVNS, ^{433}EGHPVLATR, ^{543}ADGTCI/VLWG, ^{659}MALVDSNANHV, ^{774}LEGWNS, ^{854}VFVEI/VSPE/DLA, and ^{925}TPAYKQT. Degenerate primers were designed and PCRs were performed using pair-wise combinations of these primers. Though many reactions were set up with *B. methanolicus* genomic DNA as template, the only product obtained was a partial *ca*. 600 bp fragment for a putative *fdh*, corresponding to the motif ranging from ^{97}LLYCTVCD to ^{307}KFGWDFVNS. Sequenced and translated, this fragment showed a high percentage similarity with the FDH from the other *Bacilli* within the given region. Based on this stretch of sequence, new primers were designed and paired against the other degenerate primers in PCR. However, no further products were obtained. FailSafe™ PCR was also performed but was unsuccessful. Although the generation of this 600 bp fragment that bears a high similarity with FDH from *Bacilli* species suggests that *B. methanolicus* does harbour a formate dehydrogenase gene, it does not prove it.

Attempts at genome walking on *B. methanolicus* providing no further sequence data, another source for a FDH was sought. TMO-Renwables benefiting from the genome sequence of *Geobacillus thermoglucosidasius*, made of it an obvious source choice.

Figure 3.12. Boxshade alignment of formate dehydrogenase sequences from Bacilli species. Exact matches (conserved residues) across all entries are shaded black, similar residues are shaded grey. (http://ch.embnet.org/software/BOX_form.html).

3.3.4.2 FDH from *Geobacillus*

Analyses of the *Geobacillus kaustophilus* and *Geobacillus thermoglucosidasius* genomes were performed on ERGO™ (http://ergo.integratedgenomics.com/) and identified formate dehydrogenase genes. In both genomes, the ORF orientation was on the (-) strand, encoding a 987 amino acid protein, and corresponding in each case to a formate dehydrogenase alpha chain (FDHa, EC1.2.1.2) (Figure 3.13). In close proximity to this formate dehydrogenase A, an *fdhD* gene encoding a formate dehydrogenase accessory protein was found. In *G. thermoglucosidasius*, two other genes that could be linked to the activity of the formate dehydrogenase were identified close to *fdhA*: *yrhD* and a gene encoding a formate/nitrite transporter; however, they do not appear to constitute an operon.

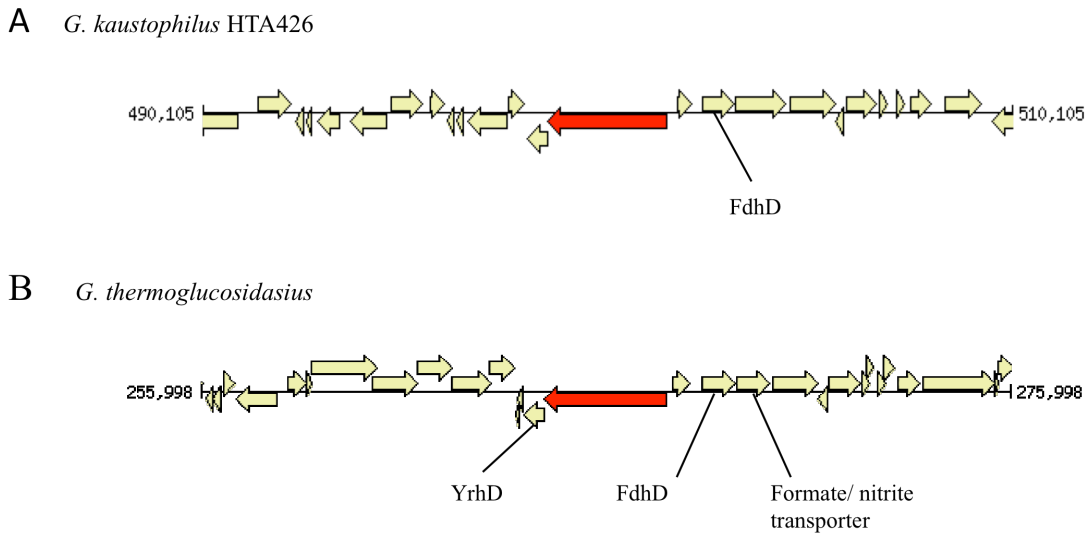
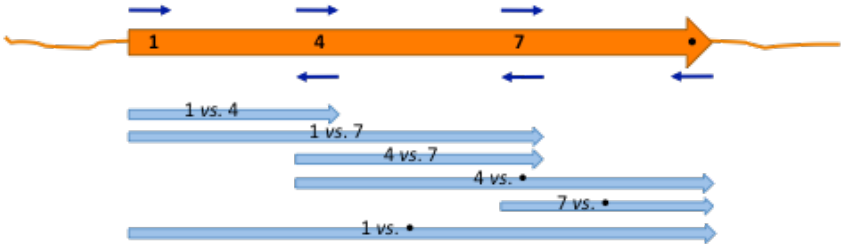


Figure 3.13. Contig region encompassing a formate dehydrogenase in (A) *Geobacillus kaustophillus* HTA426 and (B) *Geobacillus glucosidasius*. The formate dehydrogenase *fdhA* ORF is coloured in red, while significant other genes are annotated amongst all ORFs identified. The *fdhD* gene encodes a formate dehydrogenase accessory protein, FdhD.

The sequence was obtained from the ERGO database: the protein is 987 amino acids long, encoded by a 2964 bp ORF. Primers were designed at the 5' and 3' ends of the gene (GthfdhA_1F and GthfdhA_•R) as well as internally spaced roughly at each 1000 bp interval (GthfdhA_4F and _4R at the 1 kb mark, and GthfdhA_7F and _7R near the 2 kb mark).



Primer	Sequence	Tm (°C)
GthfdhA_1F	ATGCTAAAAAACTACTCCACCA	59.2
GthfdhA_4F	CGGCTGGGATTTCGTCAA	66.5
GthfdhA_4R	GACGAAATCCCAGCCGAA	66.1
GthfdhA_7F	CCAGGGCTTGACAACATTCAA	67.1
GthfdhA_7R	GAATGTTGTCAAGCCCTGGTT	64.9
GthfdhA_•R	TTAGCCACGCGGCTCT	63.5

Figure 3.14. Primers (sequences in Table), positions over the *fdhA* gene in *G. thermoglucosidasius*, and expected PCR products (*ca*. 1, 2 and 3 kb).

Geobacillus thermoglucosidasius genomic DNA was extracted and PCR performed. Pair-wise combinations of the primers were included in the set up of these PCRs, including the pairing of the 5' GthfdhA_1F and 3' GthfdhA_•R primers. Products of the expected sizes were obtained for all combination of primers as observed on a 0.8% agarose gel (Figure 3.15), but most importantly, a *ca.* 3 kb band was observed corresponding to the full-length coding sequence.

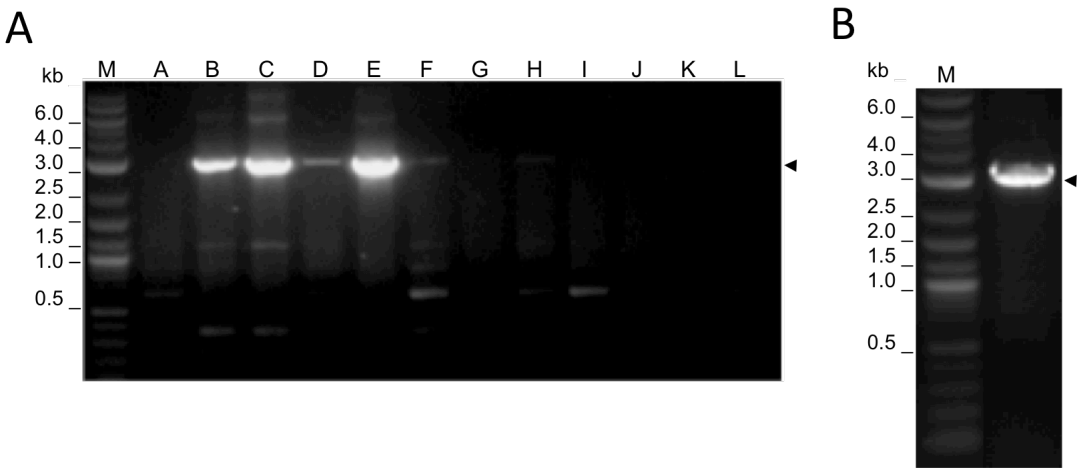


Figure 3.15. 0.8% agarose gel electrophoresis of PCR-amplified *fdh*A gene product from *G. thermoglucosidasius* gDNA, with primers GthfhdA_1F *vs*. GthfdhA_•R. (A) FailSafe™ PCR, buffers A-L; (B) PCR with Phusion® polymerase. M: 1 kb DNA ladder (kb).

The full-length product was gel-purified, cloned and the sequence confirmed as that of the *fdh*A from *G. thermoglucosidasius*.

The *fdh*D (formate dehydrogenase accessory protein) and *Yrh*D genes were also successfully PCR-amplified, cloned and sequenced in the same manner, since their functional annotation in the *G. thermoglucosidasius* genome possibly links them to the activity of the formate dehydrogenase A.

3.4 Discussion

Bacillus methanolicus, a thermotolerant methanol-utilising bacterium, harbours two copies of its methanol dehydrogenase gene, one a chromosomal copy and the other positioned on its pBM19 plasmid (de Vries *et al*., 1992; Brautaset *et al*., 2004). While these two genes have previously been individually described from different strains, this work involved the identification and cloning of both the *mdh* genes within the same *B. methanolicus* strain PB1.

pBM19 plasmids have been reported in all eleven different thermotolerant methylotrophic *B. methanolicus* wild-type strains screened (Brautaset *et al*., 2004).

While these strains (DFS2, HEN9, TSL32, CFS, RCP, SC6, NIWA, BVD, DGS, JCP, and N2) exhibit strong physiological variations, restriction analysis of their plasmid DNA indicated that they all possess pBM19-like plasmids, but the restriction profiles were not identical for all. Indeed, restriction analysis of plasmid DNA extracted from *B. methanolicus* strain PB1, used in this study, produced a profile similar to that expected of pBM19 from strain MGA3, corroborating that this strain harbours a pBM19-like plasmid. Furthermore, the choice of restriction enzymes used also gave evidence towards the presence of a plasmid-encoded *mdh* in strain PB1.

The two *mdh* genes cloned and sequenced from strain PB1, one chromosomal and the other form its pBM19 plasmid, showed remarkable identity with each other, sharing 99.5% identity and differing by only two amino acids with no further nucleotides mutations observed. This is the first report of both *mdh* genes being sequenced from the same strain of *B. methanolicus* and it is therefore difficult to assess whether similar identity would be observed within other strains. In both sets of sequence data, a strong ribosome binding site could be identified (AGGAGG). The unique motifs characteristic of the family III NAD-dependent ADH were also confirmed (^{94}GGGSXXDXXK^{103}) (Hektor *et al*., 2002).

The activator gene *act* that was cloned and sequenced from strain PB1 displayed 95% identity with that of strain C1, and 97% identity in the deduced amino acid sequence. The gene product ACT also displayed a strong similarity to the ADP-ribose pyrophosphatase protein of *B. subtilis* (P54570; 63% identity and 83% similarity), an ADP-ribose pyrophosphatase (Dunn *et al*., 1999). The Nudix hydrolase family emcompasses enzymes hydrolytically active against substrates containing a nucleotide diphosphate group linked to a moiety X. The highly conserved Nudix hydrolase sequence motif was identified in the deduced amino acids sequence of BmACT (Table 3.4) (Koonin *et al*., 1993; Bessman *et al*., 1996; Kloosterman *et al*., 2002).

Organism	Partial sequence	Gene
E.Coli	EPFG**G**KIEMG**E**TPEQAVV**RE**LQ**EE**V**G**ITPQHFSLFEKLEYEFPDRH	*mutT*
H. influenza	EFPG**G**KVDAG**E**TPEQALK**RE**LE**EE**I**G**IVALMAELYERFQFEYPTKI	*mutT*
P. vulgaris	EFPG**G**KLEDN**E**TPEQALL**RE**LQ**EE**I**G**IDVTQCTLLDTVAHDFPDRH	*mutT*
B. subtilis	EIPA**G**KLEKG**E**EPEYTAL**RE**LE**EE**T**G**YTAKKLTKITA-FYTS**P**GFA	*yqkg*
B.methanolicus C1	EIPA**G**KLEKG**E**DPRVTAL**RE**LE**EE**T**G**YECEQMEWLIS-FATS**P**GFA	*act*
B. methanolicus PB1	EIPA**G**KLEKG**E**DPRITAL**RE**LE**EE**T**G**YQCEQMEWLIS-FATS**P**GFA	*act*
	G E RE EE G (P)	

Figure 3.16. Partial sequence alignment of members of the Nudix hydrolase family. The characteristic Nudix box sequence motif is shown in bold. (Adapted from Kloosterman *et al*., 2002).

ACT has been shown to stimulate the activity of MDH in *B. methanolicus*, an NAD-dependant enzyme (Arfman *et al*., 1991; Arfman *et al*., 1992; de Vries *et al*., 1992; Hektor *et al*., 2000; Brautaset *et al*., 2006), by removal of the NMN(H) moiety of the NAD(H) co-factor in the protein.

Sequence analysis of the flanking regions of both *mdh* and *act* genes did not revealing any clustering of the genes as the partial sequence of the *act* gene could not be identified in the flanking region of *mdh* (and vice versa), reflecting the observations made in strain C1 (Kloosterman *et al*., 2002).

Until recently, it was also assumed that the lack of formaldehyde- and formate dehydrogenase activities in *B. methanolicus* indicated the operation of a non-linear oxidation sequence for formaldehyde (Arfman *et al*., 1989). However, based upon ^{13}C-NMR data, a linear pathway for dissimilation of formaldehyde into CO_2 was proposed (Pluschkell *et al*., 2002). However, neither a formaldehyde dehydrogenase nor a formate dehydrogenase had yet been identified in (and cloned from) *B. methanolicus*.

A formaldehyde dehydrogenase gene with high percentage identity to the glutathione-dependent formaldehyde dehydrogenase genes from *Bacillus clausii*, *B. thuringiensis* and *B. cereus* was isolated and cloned from *B. methanolicus* PB1 genomic DNA. The degenerate PCR approach in isolating this gene proved successful and the full-length ORF was characterised. The characterisation of the sequence ends and flanking regions used a simple method of genome walking that

avoided the need for adaptors (Zhang *et al*., 2000; Kilstrup and Kristiansen, 2000; Rishi *et al*., 2004) or use of probes (Zhao *et al*., 2007). A strong ribosome binding site AGGAGG was identified in *Bmfalddh* 9 bp upstream of the translation start codon. Ribosome binding sites (RBS) are very strong in *B. subtilis* and alternative downstream boxes do not seem to exist (Vellanoweth *et al*., 1992; Rocha *et al*., 1999). About 90% of the RBSs are distributed between the positions -5 and -11 bp, with an optimal position at 7-9 bp. Most RBSs are very strong, *i.e.* close to the consensus sequence AAAGGAGG, and their strength is not correlated with the codon usage class of the gene. The strong AGGAGG motif found in *Bmfalddh* 9 bp upstream of the translation start codon is echoed in the research by Hyatt *et al*. (2010), who developed a gene finding software – Prodigial - for microbial genomes, whereby predicted initiation sites were given scores and used in the matrix to identify genes on a genome. In the methanol dehydrogenase gene, *Bmmdh*, a similar RBS AGGAGG was identified 7 bp upstream of the translation start (this work and de Vries *et al*., 1992).

Despite repeated attempts, a full-length formate dehydrogenase gene could not be isolated in *B. methanolicus* PB1. The choosing of *Geobacillus thermoglucosidasius* (NCIMB11955) as an alternative source for an *fdh* gene was made as its genome had been sequenced and partially annotated, and it also is the candidate organism at the base of TMO-Renewables processes for bioethanol production. The 3 kb formate dehydrogenase A gene, *fdhA*, cloned from *G. thermoglucosidasius* also possessed a putative RBS, AAGGGGGAG, 6bp upstream of the start.

When showing accumulation of formaldehyde, formate and CO_2 in growing cultures of *B. methanolicus* growing on methanol as sole carbon source, Pluschkell *et al*. (2002) proposed a linear pathway for the detoxification of these metabolites, with the organism expected to harbour a formaldehyde and formate dehydrogenase. A formaldehyde dehydrogenase gene has now been cloned and sequenced from *B. methanolicus* PB1, but only part of a putative formate dehydrogenase gene. If a full-length formate dehydrogenase gene was to be characterised from *B. methanolicus*, the partial sequence obtained and the sequence from the formate dehydrogenase *fdhA* gene from *G. thermoglucosidasius* could be used to design *fdh*-specific probes to search the *B. methanolicus* genome.

Chapter 4- Heterologous Expression in *E. coli*

4.1 Introduction

Among the many available hosts for heterologous protein expression, the Gram-negative *Escherichia coli* remains one of the most attractive. *E. coli* is highly efficient at producing large quantities of recombinant proteins as it benefits from a short generation time enabling rapid growth to high densities. Its well-characterised genetics, easy transformation and the availability of an ever-increasing number of expression vectors and mutant expression strains make it an attractive choice. Promoters used to drive recombinant protein expression have long been based on *lac* operon-derived elements. The pET vector system (Novagen) used here relies on the bacteriophage T7 promoter (pT7) and its associated T7 RNA polymerase, along with *lac*-operon elements making it IPTG-inducible to drive gene expression. Numerous *E. coli* strains have been developed for tailored protein expression requirements. The *E. coli* strain BL21(DE3)pLysS allows high-efficiency protein expression of a gene that is under the control of a T7 promoter, with the co-expression of a T7 lysosyme (in plasmid pLysS) that lowers the background expression of the transgene under the control of pT7 without interfering with the level of expression achieved following induction with IPTG (Studier, 1990). In KRX, another engineered *E. coli* expression strain, the T7 RNA polymease is under the control of the more stringent rhamnose promoter (*rha*P_{BAD}) to provide tight control of the recombinant protein expression.

However, with high levels of over-expression, target proteins can often fail to adopt their native conformation, leading to misfolding, segregation into insoluble inclusion bodies, and a loss of enzymatic activity during attempts at *in vitro* refolding. Therefore, strains have been engineered for *in vivo* refolding by using chaperones (Richardson *et al*., 1998; Bukau and Horwich, 1998), in addition to the use of temperature-inducible promoters (protein folding is often favoured under low temperature culture conditions) (Phadtare *et al*., 1999). Finally, an alternative to

expressing a toxic or insoluble protein is to express it as a fusion protein, where the fusion partner improves the solubility and also promotes folding (Zhang *et al*., 1998; Ferrer *et al*., 2004).

Dehydrogenase enzymes of all types and sources have been expressed and characterised following expression in an *E. coli* host. Particularly, the alcohol and aldehyde dehydrogenases from the thermophilic alkane-degrading *Geobacillus thermoleovorans* have been expressed recombinantly in *E. coli* (Kato *et al*., 2001; Kato *et al*., 2010). The ADH1 from the thermophilic ethanologen *Geobacillus thermoglucosidasius* and the ADH from the Archaeon *Thermoplasma acidophilum* have also been heterologously expressed in *E. coli* and characterised (Jeon *et al*., 2008; Marino-Marmolejo *et al*., 2009).

In *Bacillus methanolicus*, characterisation of its MDH and ACT enzymes have involved expression in an *E. coli* host albeit using expression vectors and *E. coli* strains allowing the genes to be under the control of their own promoters (Hektor *et al*., 2002; Kloosterman *et al*., 2002).

This chapter describes the expression and assays of the three dehydrogenase enzymes of interest in this study in the heterogeneous expression host *Escherichia coli*, for the production of high quantities of soluble and active recombinant proteins.

4.2 Materials and Methods

The pET28a and pET19b vectors, and the *E. coli* expression strains BL21(DE3)pLysS and Rosetta™, were from Novagen. The *E. coli* expression strain ArcticExpress™ was supplied by Stratagene. The *E. coli* expression strain KRX was supplied by Promega. The pTYB11 vector was from NEB.

Details of manipulations are given in Chapter 2 - General Materials and Methods.

PCR amplification of the four genes of interest, *Bmmdh*, *Bmact*, *Bmfalddh* and *GthfdhA* were performed as previously described (see 2.2.2 and 3.3). Restriction sites were included at the 5'ends of the forward and reverse primers to allow complementary cloning as required. The stop codon (•) was also omitted in the reverse primer sequence in the case of C-terminal tagging in the expression vector.

4.2.1 cloning into pET

The *E. coli* expression vectors pET19b (Ap) and pET28a (Kan) were used (Figure 2.2 in Chapter 2 – 2.2.3.1).

4.2.1.1 General cloning strategy

The genes were first PCR-amplified using the appropriate primer pairs, each containing the restriction enzyme recognition sequence as defined in the cloning strategy, and Phusion® as the DNA polymerase. PCR products of the right size were separated on a 1% agarose gel, purified, A-tailed, cloned into pGEM®-T Easy and sequenced. The correct inserts were then excised from pGEM®-T Easy by double restriction digestion of 5 µL of purifed vector using the restriction enzymes whose sites have been introduced at either end of the insert, further separated from the vector backbone on a 1% agarose gel and gel purified. Simultaneously, the pET vector was digested using these same restriction enzymes, the enzymes heat-killed (65 °C, 20 min) and the vector backbone treated with alkaline phosphatase (37 °C, 30 min). The vector was then separated on a 1% agarose gel and gel purified. The linearised pET vector and insert were ligated using T4 DNA ligase and transformed into *E. coli* JM109 for screening and maintenance. A plasmid showing the correct pattern in screening (PCR or restriction digest screening) was sequenced (pT7_F and tT7_R primers) and transformed into the *E. coli* expression strain of choice.

A list of cloning primers and restriction sites used is presented in Table 4.1.

Table 4.1. Cloning primers list.

Primer	Sequence (5' → 3')	Tm (°C)
NcoI_Bmmdh_F	ATGACAAACTTTTTCATTCC	55.6
XhoI_Bmmdh_R	CCGCTCGAGCAGAGCGTTTTTGATG	66.4
NcoI_pBMmdh_F	ATGACAACAAACTTTTTCATTCC	60.6
XhoI_pBMmdh_R	CCGCTCGAGCATAGCGTTTTTGATGATTTGTGC	80.3
NcoI_Bmact_F	ATGGGAAAATTATTGAGC	53.8
XhoI_Bmact_R	CCGCTCGAGTTTATGTTTGAGAGC	68.5
Nhe_Bmfalddh_F	ATGAAAGCTGTAACATATCAAGGACC	64.0
SacII_Bmmdh_F	TCCCCGCGGATGACAAACTTTTTCATTCC	78.4
XhoI_Bmmdh•_R	CCGCTCGAGTTACAGAGCGTTTTTGATG	80.3
SacII_pBMmdh_F	TCCCCGCGGATGACAACAAACTTTTTCATTCC	80.1
XhoI_pBMmdh•_R	CCGCTCGAGTTACATAGCGTTTTTGATGATTTGTGC	78.6
SacII_Bmact_F	TCCCCGCGGATGGGAAAATTATTGAGC	77.8
XhoI_Bmact•_R	CCGCTCGAGTCATTTATGTTTGAGAGC	71.7
SacII_Bmfalddh_F	TCCCCGCGGATGAAAGCTGTAACATATCAAGGACC	81.0
XhoI_Bmfalddh•_R	CCGCTCGAGTTACGGTTTAAAAACGACTTTGATGC	77.9
SacII_GthfdhA_F	TCCCCGCGGATGCTAAAAAACTACTCCACC	77.9
XhoI_GthfdhA•_R	CCGCTCGAGTTAGCCACGCGGCTCTGCCTCC	85.5

4.2.1.2 Specifics regarding the genes of interest

The methanol dehydrogenase genes *Bmmdh* and *pBMmdh* (short of their stop codons) were PCR-amplified using the respective primer pairs *Nco*I_Bmmdh_F *vs.* *Xho*I_Bmmdh_R and *Nco*I_pBMmdh_F *vs.* *Xho*I_pBMmdh_R, and cloned in pET28a as *Nco*I/*Xho*I fragments, resulting in the vectors pET28a *Bmmdh*:His• and pET28a *pBMmdh*:His•; these are non-cleavable C-terminal histidine tags.

Similarly, the activator *Bmact* (without stop codon) was cloned in as a *Nco*I/*Xho*I insert, resulting in the vector pET28a *Bmact*:His• (primers used for PCR were *Nco*I_Bmact_F *vs.* *Xho*I_Bmact_R). The PCR-amplified activator, including its stop codon (with primer pair *Nco*I_Bmact_F *vs.* *Xho*I_Bmact•_R), was cloned into the vector pET19b as a *Nco*I/*Xho*I insert resulting in the vector pET19b *Bmact*•; no tag was incorporated.

The formaldehyde dehydrogenase gene was cloned in the vector pET28a as a *Nhe*I/*Xho*I insert (primer pair in PCR was *Nhe*I_Bmfalddh_F *vs.*

*Xho*I_Bmfalddh•_R), resulting in the vector pET28a His:*Bmfalddh*• specifying an N-terminal histidine tag on the protein.

The vector pET28a was modified with the insertion of a *Sac*II restriction enzyme recognition sequence immediately upstream of the *Nde*I site. Cloning in the genes of interest therefore generates constructs with a cleavable N-terminal histidine tag. A description and the making of pET28a*Sac*II are given below (4.2.1.3).

All four genes, *Bmmdh*, *Bmact*, *Bmfalddh* and *GthfdhA*, stop codons included, were cloned as *Sac*II/*Xho*I inserts into the modified vector pET28a*Sac*II, resulting in the vectors pET28a Histhr:*Bmmdh*•, pET28a Histhr:*Bmact*•, pET28a Histhr:*Bmfalddh*• and pET28a Histhr:*GthfdhA*•. The respective primer pairs used for PCR-amplifying the genes of interest were *Sac*II_Bmmdh_F *vs.* *Xho*I_Bmmdh•_R (*Bmmdh*), *Sac*II_Bmact_F *vs.* *Xho*I_Bmact•_R (*Bmact*), *Sac*II_Bmfalddh_F *vs.* *Xho*I_Bmfalddh•_R (*Bmfalddh*) and *Sac*II_GthfdhA_F *vs.* *Xho*I_GthfdhA•_R (*GthfdhA*). Primer sequences are detailed in Table 4.1 above.

4.2.1.3 Construction of pET28a*Sac*II

pET28a*Sac*II is a pET28a vector with a *Sac*II restriction site inserted immediately upstream of the *Nde*I site. It enables the cloning in of the four genes of interest here in the same manner, using the enzymes *Sac*II at the 5' end and *Xho*I at the 3' end, resulting in pET28a vectors conferring an N-terminal His tag while retaining the thrombin cleavage site, *i.e.* pET28a Histhr:GoI•.

The *Sac*II restriction site was inserted into pET28a via site-directed mutagenesis. Site-directed mutagenesis forward and reverse primers were designed over the designated region in pET28a, containing the *Sac*II restriction recognition sequence CCGCGG in their middle. *Ca*. 20bp of upstream and downstream pET28a sequence flanked the *Sac*II site on either side in the primers (Figure 4.1).

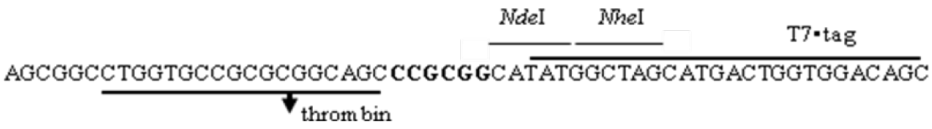


Figure 4.1. pET28a*Sac*II site-directed mutagenesis primers. Region of pET28a upon which site directed mutagenesis forward and reverse complementary primers were designed and position of the *Sac*II restriction site (bold). The *Nde*I and *Nhe*I sites are underlined. The thrombin recognition sequence is underlined and its cleavage site marked (arrow). The T7 tag is also annotated.

Primers thr*Sac*IIT7_F and thr*Sac*IIT7_R were synthesised (48 mers, T_m = 87.9 °C):

thr*Sac*IIT7_F: 5' GCCTGGTGCCGCGCGGCAGC**CCGCGG**CATATGGCTAGCATGACTGGTG,
thr*Sac*IIT7_R: 5' CACCAGTCATGCTAGCCATATG**CCGCGG**GCTGCCGCGCGGCACCAGGC.

Site-directed mutagenesis PCR was performed as described below (adapted from Stratagene). In a thin-walled tube were assembled: 5 µL 5x buffer, 10 ng dsDNA template (pET28a), 1 µL (5 pmol.µL^{-1}) each forward and reverse sdm primer, 1 µL (5 mM) dNTP mix, 1 µL (5 U.µL^{-1}) Phusion™, to a final volume of 25 µL with MilliQ water. Thermalcycling conditions were set as: 95 °C for 15 s, 18 cycles of 95 °C for 10 s, 55 °C for 10 s, and 72 °C for 20 s, per kb of plasmid length (*i.e.* 2 min). The reaction was then placed on ice, 1 µL (10 U.µL^{-1}) *Dpn*I restriction enzyme added, and the reaction incubated at 37 °C for an hour. Then, 1 µL of the *Dpn*I-treated DNA was transformed into *E. coli* JM109 competent cells (as previously described; Chapter 2 - 2.2.2.7). In a 14 mL Falcon tube, 0.8 mL SOC was added to the transformed cells and incubated at 37 °C for an hour. Up to 100 µL of the cells were then plated onto an LB agar plate supplemented with kanamycin (50 µg.mL^{-1}) and incubated overnight at 37 °C. Colonies were picked, plasmid extracted and checked for the presence of the *Sac*II site by restriction digestion with *Sac*II and *Bgl*I. Restriction digestion with *Sac*II linearises the *Sac*II-containing pET28a*Sac*II (5.4 kb) while the unmodified pET28a remains circular, and the *Sac*II/*Bgl*I double digest produces two fragments of sizes 2 kb and 3.4 kb with pET28a*Sac*II, while linearising pET28a.

4.2.2 cloning into pTYB11

The vector pTYB11 belongs to the IMPACT family of expression vectors developed by NEB (Figure 4.2). Like the pET vectors, it is based on the strong IPTG-inducible T7 promoter (pT7).

The pTYB11 vector utilises an intein from *Saccharomyces cerevisiae* VMA1 gene (SceVMA1). The target protein is fused at its N-terminus to a self-cleavable 56 kDa VMA1 intein-CDB tag (chitin-binding domain); the tag allows the affinity purification of the fusion head on a chitin column. The vector is also designed to allow purification of the protein of interest without any extra amino acids by cloning its 5' end into the *Sap*I site. The gene of interest must therefore be amplified to incorporate a *Sap*I site at its 5' end, and another restriction site at its 3' end chosen amongst those available in the multiple cloning site of the vector (namely *Spe*I, *Nru*I, *Sal*I, *Not*I, *Eco*RI, *Xho*I and *Xma*I). The *Sap*I recognition sequence reads as 5' GCTCTTC but the enzyme cuts 1 bp further along, generating a 3 bp 5' overhang (Figure 4.2-D). Since *Sap*I generates this peculiar overhang, extra care must be paid in the design of the *Sap*I-containing forward primers.

In pTYB11, the 3'end of the intein protein into the MCS reads as 5' GTACAGAAC▼**AGAAGAGC**, with the *Sap*I site in bold and the cleavage site marked. Therefore, in order to generate a fusion protein with the gene of interest in frame, an extra nucleotide is inserted in-between the SapI recognition sequence and the ATG start. The gene specific *Sap*I_forward primers designed then read as 5' $(N)_6$GAAGAGCT**ATG**$(N)_n$; the *Sap*I site is underlined and the start codon of the gene of interest is in bold. As the overhang sequence of the *Sap*I recognition site GCN encodes an alanine in the fusion protein, the extra nucleotide inserted was set as a thymine, creating the codon GCT, as is in the pTYB11 sequence.

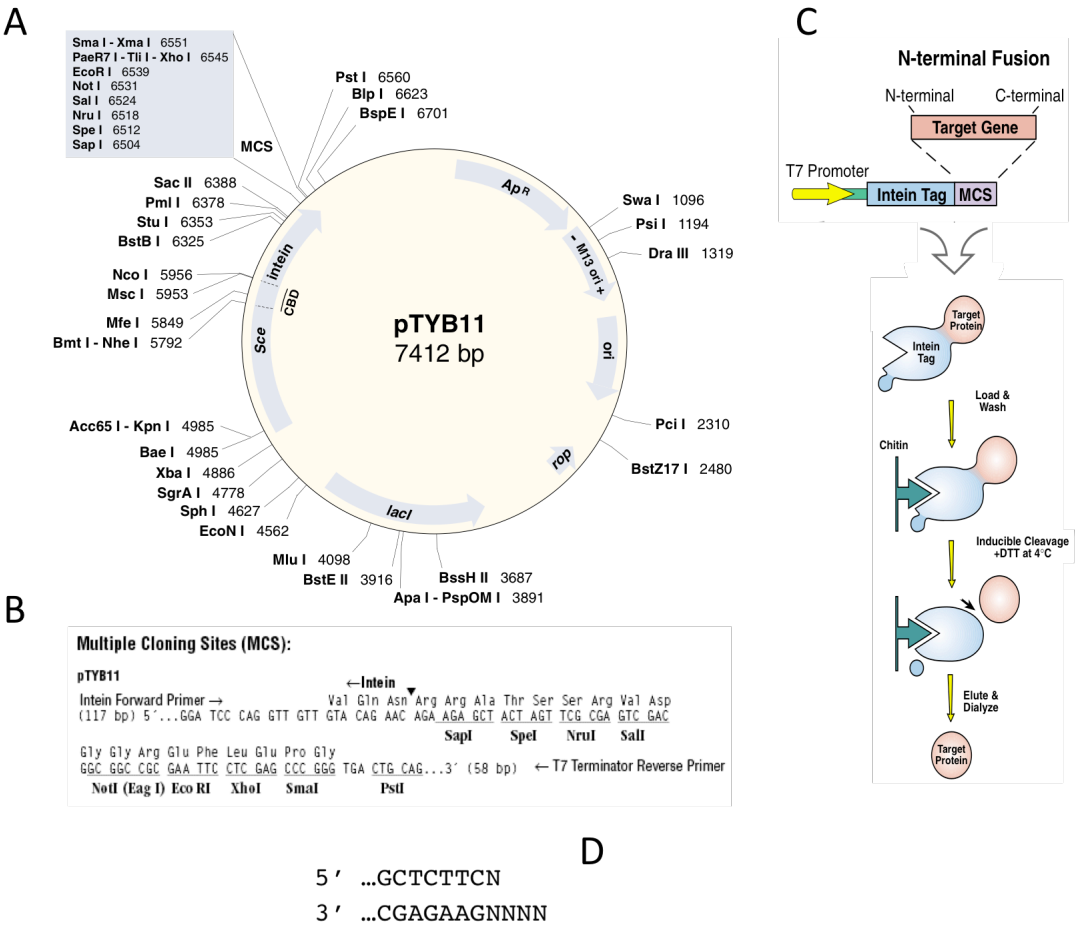


Figure 4.2. (A) pTYB11 vector map, (B) multiple cloning site and (C) diagram showing the basis of the IMPACT system. (D) *Sap*I recognition sequence and 5' overhang. (Adapted from NEB).

All four genes, *Bmmdh*, *Bmact*, *Bmfalddh* and *GthfdhA*, were cloned into the vector pTYB11 as *Sap*I/*Xho*I inserts. The respective primer pairs used for PCR-amplifying the genes of interest were *Sap*I_*Bmmdh*_F *vs.* *Xho*I_*Bmmdh*•_R (*Bmmdh*), *Sap*I_*Bmact*_F *vs.* *Xho*I_*Bmact*•_R (*Bmact*), *Sap*I_*Bmfalddh*_F *vs.* *Xho*I_*Bmfalddh*•_R (*Bmfalddh*) and *Sap*I_*GthfdhA*_F *vs.* *Xho*I_*GthfdhA*•_R (*GthfdhA*). Primers used are detailed in Table 4.2.

Table 4.2. List of *Sap*I_F and •*Xho*I_R primers. Start and Stop codons in bold, *Sap*I and *Xho*I sites underlined.

Primer	Sequence (5' → 3')	Tm (°C)
SapI_Bmmdh_F	GCGCGAAGAGCT**ATG**ACAAACTTTTTCATTCC	71.5
XhoI_Bmmdh•_R	GAGCCTCGAG**TTA**CAGAGCGTTTTTGATG	70.1
SapI_Bmact_F	GAGCGAAGAGCT**ATG**GGAAAATTATTGAGG	71.4
XhoI_Bmact•_R	GAGCCTCGAG**TCA**TTTATGTTTGAGAGC	70.1
SapI_Bmfalddh_F	GAGCGAAGAGCT**ATG**AAAGCTGTAACATATCAGG	70.6
XhoI_Bmfalddh•_R	GAGCCTCGAG**TCA**TGGTTTAAAAACGACTTTGATGC	71.4
SapI_GthfdhA_F	GAGCGAAGAGCT**ATG**CTAAAAAACTACTCCACC	71.9
XhoI_GthfdhA•_R	GAGCCTCGAG**TTA**GCCACGCGGCTCTGC	71.7

The genes were first PCR-amplified as *Sap*I/*Xho*I fragments, using the primer pair *Sap*I_GoI_F *vs.* *Xho*I_GoI•_R and Phusion™ as the DNA polymerase. Appropriately-sized PCR products were gel-purified, A-tailed, cloned into pGEM®-T Easy and sequenced. The correct inserts were then excised from pGEM®-T Easy by restriction digest of 5 µL purified vector with *Sap*I and *Xho*I (37 °C, 1 h), further segregated from the vector backbone on a 1% agarose gel and gel purified. The vector pTYB11 was simultaneously restricted with *Sap*I and *Xho*I (37 °C, 2 h), the enzymes heat-inactivated (65 °C, 20 min) and the linearised vector treated with alkaline phosphatase (37 °C, 30 min). The vector backbone was then gel purified, ligated with the purified insert using T4 DNA ligase, and transformed into JM109 cells before screening. A plasmid showing the correct pattern in screening was then sequenced (3' Intein_F (5' CCCGCCGCTGCTTTTGCACGTGAG), tT7_R (5' GCTAGTTATTGCTCAGCGG)) and finally transformed into the *E. coli* expression strain of choice.

The resulting constructs were vectors pTYB11 *Bmmdh*•, pTYB11 *Bmact*•, pTYB11 *Bmfalddh*• and pTYB11 *GthfdhA*•. The respective primer pairs used for PCR-amplifying the genes of interest were *Sap*I_*Bmmdh*_F *vs.* *Xho*I_*Bmmdh*•_R (*Bmmdh*), *Sap*I_*Bmact*_F *vs.* *Xho*I_*Bmact*•_R (*Bmact*), *Sap*I_*Bmfalddh*_F *vs.* *Xho*I_*Bmfalddh*•_R (*Bmfalddh*) and *Sap*I_*GthfdhA*_F *vs.* *Xho*I_*GthfdhA*•_R (*GthfdhA*).

4.2.3 Expression in *E. coli*

Transformants were confirmed in *E. coli* JM109 by growth on LB supplemented with the appropriate antibiotic for vector selection (Ampicillin 100 $\mu g.mL^{-1}$ or Carbenicillin 50 $\mu g.mL^{-1}$ for pET19b and pTYB11, Kanamycin 50 $\mu g.mL^{-1}$ for pET28a) and the presence of the correct insert size confirmed by restriction digestion or PCR analysis. Purified vectors were then transformed into *E. coli* expression host strains BL21(DE3)pLysS, Rosetta™, KRX and/or ArcticExpress™.

Growth of transformants and induction of expression were performed as described previously (in Chapter 2 – Materials and Methods (2.2.5.1)). *E. coli* was routinely grown at 37 °C overnight. ArcticExpress™ was grown at 30 °C until an OD_{600} of ~ 0.6, the temperature lowered to 13 °C and IPTG added (final concentration 0.1 mM (pET) or 0.5 mM (pTYB)), and the cultures further incubated overnight to allow the over-expression of a soluble product. KRX was grown at 37 °C in TB until an OD_{600} of 0.8-1.0. The cultures were then shifted to 15-25 °C. Once an OD_{600} of 1.0-1.5 was reached, protein expression was induced by adding rhamnose and IPTG to concentrations of 0.1% and 1 mM respectively. The cultures were further incubated at 20 °C overnight.

The cultures were harvested by centrifugation at 3,800*g*, the supernatant discarded and the pellet re-suspended in 1/10 volume of His-binding buffer (in the case of His-tagged pET product), chitin buffer (intein-tagged pTYB product) or phosphate buffer (50 mM HEPES, pH 8.0). Cells were lysed as previously described (in Chapter 2 – Materials and Methods (2.2.5.2)), and the cell debris and insoluble fractions segregated by centrifugation (13,000*g*, 20 min). The supernatant (soluble fraction) was retained and the pellet (insoluble fraction) was re-suspended in 8 M urea.

Samples of each fraction were run on SDS-PAGE to assess the levels of over-expression of the target protein.

4.2.4 Metal affinity chromatography

Nickel-based purification of His-tagged proteins is described previously (Chapter 2 – Materials and Methods (2.2.5.5)).

4.2.5 Chitin column chromatography

Affinity purification of intein-tagged proteins was performed as follows. A Poly-Prep® chromatography column (Bio-Rad) was prepared by transferring 2 mL of the chitin beads slurry (NEB) to give a 1 mL packed final bed volume. The column was washed with 3 column volumes of $_{dd}H_2O$ and equilibrated with 5 column volumes of Chitin column buffer (20 mM HEPES, pH 8.5, 500 mM NaCl). The clarified lysate was then loaded onto the column, allowed to clear, and the column washed with at least 10 volumes of buffer. On-column cleavage of the intein-CBD tag with the protein of interest was induced by the addition of a thiol agent. The column was quick-flushed with 3 volumes of chitin-cleavage buffer (20 mM HEPES, pH 8.5, 500 mM NaCl, 50 mM DTT), the flow stopped, and the column stored at 4 °C for at least 24 h. The protein of interest was eluted by restoring the flow and further washing of the column with chitin column buffer, the intein-CBD tag remaining bound to the column. The column was stripped with 1% (w/v) SDS, 0.3 M NaOH.

4.2.6 Protein estimation and SDS-PAGE

Protein estimation and SDS-PAGE gels were performed as previously described (Chapter 2 – Materials and Methods (2.2.5.3 and .2.2.5.4)).

4.2.7 Enzyme assays

Enzyme assays were as described previously (Chapter 2 – Materials and Methods (2.2.6)).

Methanol dehydrogenase, formaldehyde reducatse, NADH oxidase, formaldehyde dehydrogenase, formate reductase and formate dehydrogenase activities were assayed. Kinetic parameters were calculated when appropriate. Temperature optimum and thermal inactivation studies were performed as previously described (see 2.2.6.5).

4.2.8 NADH spectrophotometric scan

An aliquot of purified-enzyme was added to a spectrophotometric cuvette, in MDH assay buffer, with or without 5 mM DHA (dihydroxyacetone), and incubated at room temperature for 10 min. When DHA was added, the sample was incubated overnight. The sample was scanned in a spectrophotometer across the 250 - 400 nm range for a peak at 340 nm.

4.3 Results

4.3.1 Amplification and cloning of the dehydrogenase genes

A pre-requisite to pET cloning was the PCR amplification of the three dehydrogenase genes of interest, *Bmmdh*, *Bmfalddh*, *GthfdhA* and *Bmact* from their respective sources.

4.3.1.1. *Bmmdh* and *Bmact* pET constructs

Bmmdh, with and without its stop codon, was first amplified from *B. methanolicus* PB1 genomic DNA as *Nco*I/*Xho*I fragments (Figure 4.3-A). The products (1,166 and 1,163 bp respectively) were gel-purified, cloned into pGEM®-T Easy and transformed into *E. coli* JM109 cells. White colonies were screened for the presence of insert-containing plasmid by restriction digestion (Figure 4.3-B). Sequencing confirmed the sequence as the methanol dehydrogenase from *B. methanolicus* PB1. The insert was then excised from pGEM®-T Easy, ligated into pET28a and transformed into JM109 cells. Colonies were screened for the presence of an insert-containing pET by PCR (Figure 4.3-D). Constructs pET28a *Bmmdh*• and pET28a *Bmmdh*:His• were generated.

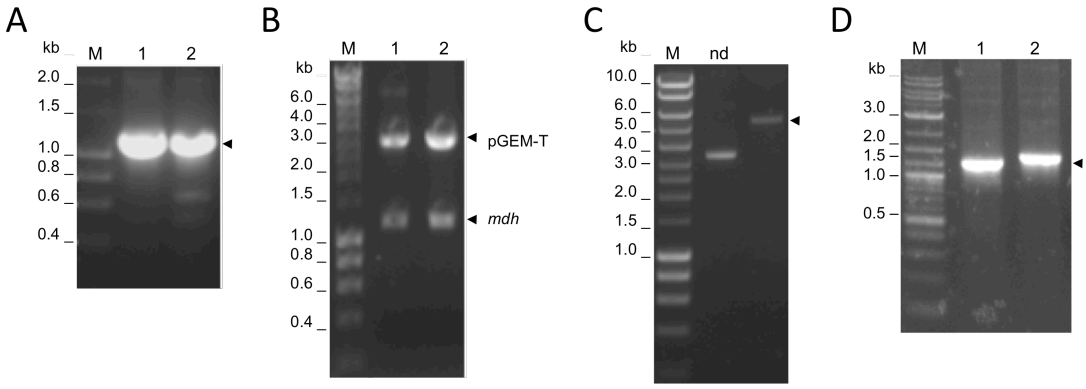


Figure 4.3. pET28a BmMDH:His and pET28a pBMMDH:His, 1% agarose gel electrophoresis analysis. (A) PCR of the *Bmmdh* (1166 bp) and *pBMmdh* (1169 bp) genes as NcoI/XhoI inserts; primers were NcoI_Bmmdh_F *vs*. XhoI_Bmmdh_R (*Bmmdh*) and NcoI_pBMmdh_F *vs.* XhoI_pBMmdh_R (*pBMmdh*); (B) pGEM-T restriction analysis and excision of the inserts (*Nco*I/*Xho*I): 1152 bp (*Bmmdh*) and 1155 bp (*pBMmdh*); (C) *Nco*I/*Xho*I restriction of pET28a (5.2 kb) (nd: non-digested); (D) PCR analysis of pET28a-MDH:His constructs; primers were pT7_F *vs* Bmmdh_R or pBMMDH_R. Lane 1: *Bmmdh*, lane 2: *pBMmdh*. M: 1 kb DNA ladder (kb). Phusion™ was used as DNA polymerase in PCR.

Similarly, the plasmid-borne copy of the *mdh* gene, with or without its stop codon, was amplified from *B. methanolicus* PB1 pBM19 plasmid as NcoI/XhoI inserts (Figure 4.3-A). The fragments (1,169 and 1,166 bp) were treated in the same way as *Bmmdh* (Figure 4.3-B), finally resulting in constructs pET28a *pBMmdh*• and pET28a *pBMmdh*:His• (Figure 4.3-D).

Similarly, the activator *Bmact* was first amplified with and without its stop codon as *Nco*I/*Xho*I fragments from *B. methanolicus* PB1 genomic DNA (Figure 4.4-A,E). The fragments of 578 and 575 bp respectively were gel-purified, cloned into pGEM®-T Easy and transformed into *E. coli* JM109 cells. White colonies were screened for the presence of activator insert-containing plasmid by restriction digestion (Figure 4.4-B,F). Inserts, confirmed correct by sequencing, were excised from pGEM™-T Easy. The activator without its stop codon was ligated into pET28a, while the insert with its stop codon was ligated into pET19b. Both were transformed into JM109 cells and subsequent colonies were screened for the presence of the insert by PCR (Figure 4.4-D,H), thereby generating constructs pET28a Bmact:His• and pET19b *Bmact*•.

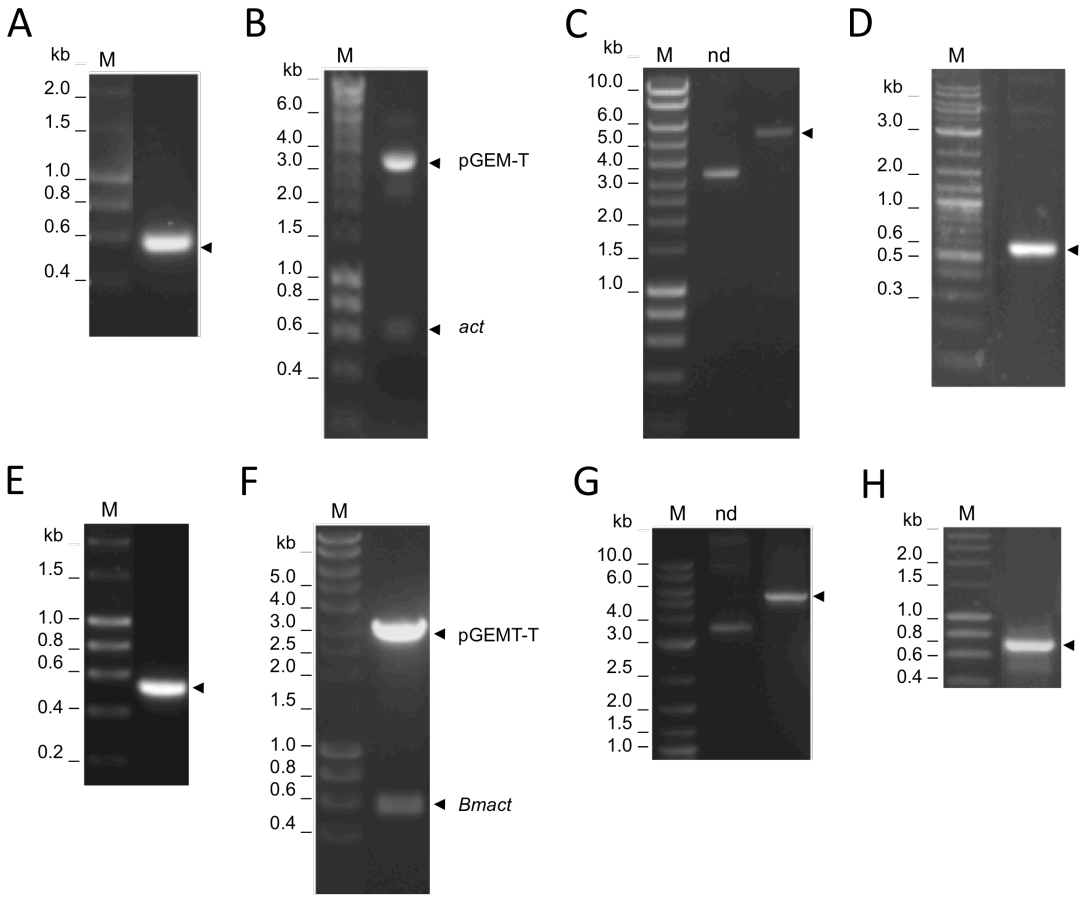


Figure 4.4. (A-D) pET28a BmACT:His and (E-H) pET19b BmACT 1% agarose gel electrophoresis analysis. PCR of the *Bmact* gene as NcoI/XhoI inserts (A) without (575 bp) and (E) with its stop codon (578 bp); primers were NcoI_Bmact_F *vs*. XhoI_Bmact_R or XhoI_Bmact_•R. pGEM-T restriction analysis and excision of the inserts (*Nco*I/*Xho*I): (B) 561 bp and (F) 564bp. *Nco*I/*Xho*I restriction digest of (C) pET28a (5.2 kb) and (G) pET19b (5.7 kb) (nd: non digested). PCR analysis of (D) pET28a BmACT:His and (H) pET19b BmACT constructs; primers were pT7_F *vs* Bmact_R. M: 1 kb DNA ladder (kb). Phusion™ was used as DNA polymerase in PCR.

Following the same processes, *Bmmdh* and *Bmact* were also successfully cloned into pET28a*Sac*II as *Sac*II/*Xho*I inserts, generating the C-terminal, His-tagged constructs pET28a Histhr:Bmmdh• and pET28a Histhr:Bmact• (Figure 4.5).

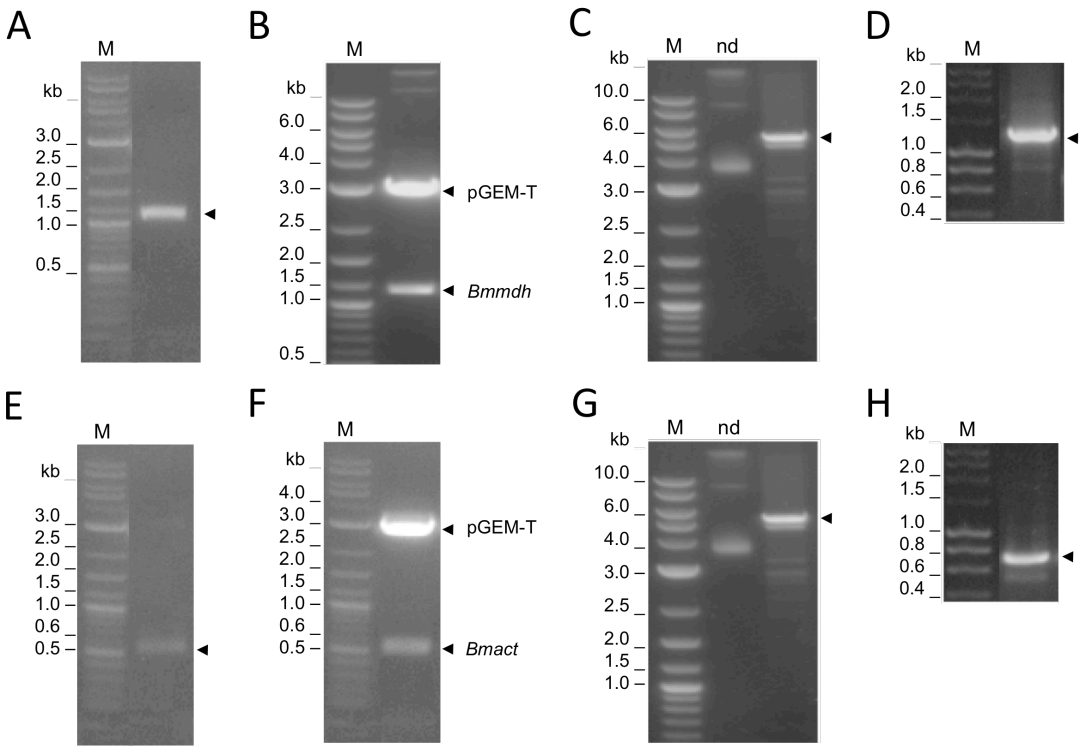


Figure 4.5. (A-D) pET28a Histhr:BmMDH and (E-H) pET28a Histhr:BmACT 1% gel electrophoresis analysis. PCR of (A) Bmmdh (1169 bp) as *Sac*II/*Xho*I (SacII_Bmmdh_F vs. XhoI_Bmmdh_•R) and (E) Bmact (578 bp) as *Sac*II/*Xho*I (SacII_Bmact_F vs. XhoI_Bmact_•R); pGEM-T restriction analysis and excision of the *Sac*II/*Xho*I inserts: (B) 1155 bp Bmmdh and (F) 564 bp Bmact. *Sac*II/*Xho*I restriction digest of (C) pET28aSacII (5.2 kb) (nd: non-digested). PCR analysis of (D) pET28a Histhr:BmMDH and (H) pET28a Histhr:BmACT; primers pT7_F vs. Bmmdh_F or Bmact_R. M: 1 kb DNA ladder (kb). Phusion was used as the DNA polymerase.

4.3.1.2 *Bmfalddh* pET constructs

The design of the primers used for amplifying the formaldehyde dehydrogenase gene out of *B. methanolicus* PB1 genomic DNA benefitted from the sequence established earlier (this work; Chapter 3 – Cloning (3.3.3)).

The formaldehyde dehydrogenase gene, complete with its stop codon, was amplified by PCR from *B. methanolicus* PB1 genomic DNA as an *Nhe*I/*Xho*I fragment (Figure 4.6-A). The 1,154 bp DNA product was gel-purified, cloned into pGEM®- Easy and transformed into *E. coli* JM109 cells. White colonies were screened for the presence of an insert-containing cloning vector by restriction

digestion, and confirmed by sequencing (Figure 4.6-B). The *Bmfalddh* insert was excised, gel-purified, ligated into pET28a and transformed into *E. coli* JM109 cells. Colonies were screened for the presence of an insert-containing pET by restriction digestion (Figure 4.6-C), generating constructs pET28a His:Bmfalddh•.

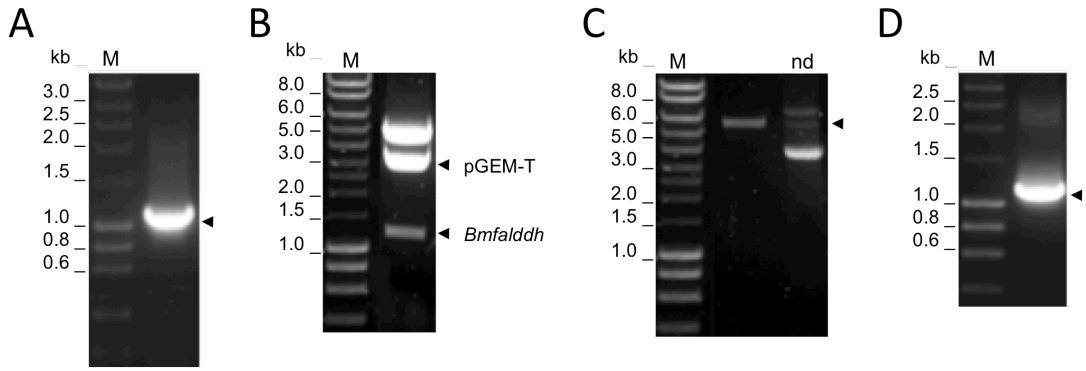


Figure 4.6. 1% agarose gel electrophoresis analysis of pET28a His:BmFALDDH. (A) PCR of the *Bmfalddh* (1154 bp) gene as *Nhe*I/*Xho*I inserts; primers were NheI_Bmfalddh_F *vs*. XhoI_Bmfalddh_R; (B) pGEM-T restriction analysis and excision of the insert (*Nhe*I/*Xho*I): 1140 bp; (C) *Nhe*I/*Xho*I restriction of pET28a (5.2 kb) (nd: non digested); (D) PCR analysis of pET28a-His:BmFALDDH construct; primers were pT7_F *vs* Bmfalddh_R (exp. 1.2 kb). M: 1 kb DNA ladder (kb). Phusion™ was used as DNA polymerase in PCR.

Following the same processes, *Bmfalddh* was also successfully cloned into pET28a*Sac*II as a *Sac*II/*Xho*I insert, generating the C-terminal, His-tagged construct pET28a Histhr:Bmfalddh• (Figure 4.7).

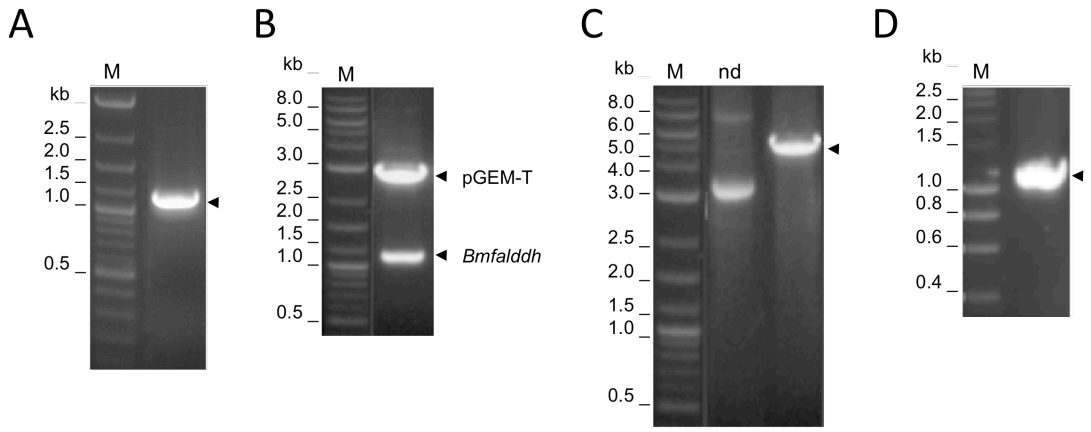


Figure 4.7. 1% agarose gel electrophoresis analysis of pET28a-Histhr:BmFALDDH. (A) PCR of the *Bmfalddh* (1154 bp) gene as *Sac*II/*Xho*I insert; primers were SacII_Bmfalddh_F *vs*. XhoI_Bmfalddh_•R; (B) pGEM-T restriction analysis and excision of the insert (*Sac*II/*Xho*I): 1140 bp; (C) *Sac*II/*Xho*I restriction of pET28a (5.2 kb) (nd: non digested); (D) PCR analysis of pET28a-Histhr:FALDDH construct; primers were pT7_F *vs* Bmfalddh_R (exp. 1.2 kb). M: 1 kb DNA ladder (kb). Phusion™ was used as DNA polymerase in PCR.

4.3.1.3 *GthfdhA* pET28a Histhr construct

GthfdhA• was PCR-amplified from *G. thermoglucosidasius* genomic DNA (Figure 4.8) as a *Sac*II/*Xho*I fragment. The PCR product was gel-purified, cloned into pGEM®-T Easy and transformed into *E. coli* JM109 cells. White colonies were screened for the presence of insert-containing plasmid by triple restriction digestion with *Sac*II, *Xho*I and *Bsa*I. Since the gene *GthfdhA* and the pGEM®-T backbone are both *ca*. 3 kb in length, a SacII/XhoI restriction digest would not have allowed adequate separation of the two bands. Instead, digestion of the pGEM®-T backbone with *Bsa*I generated two fragments of sizes 1.4 and a 1.6 kb. The insert was then *Sac*II/*Xho*I excised from a *Bsa*I pre-treated pGEM®-T Easy, ligated into pET28a*Sac*II and transformed into JM109 cells.

The construct pET28a Histhr:*GthfdhA*• was successfully generated.

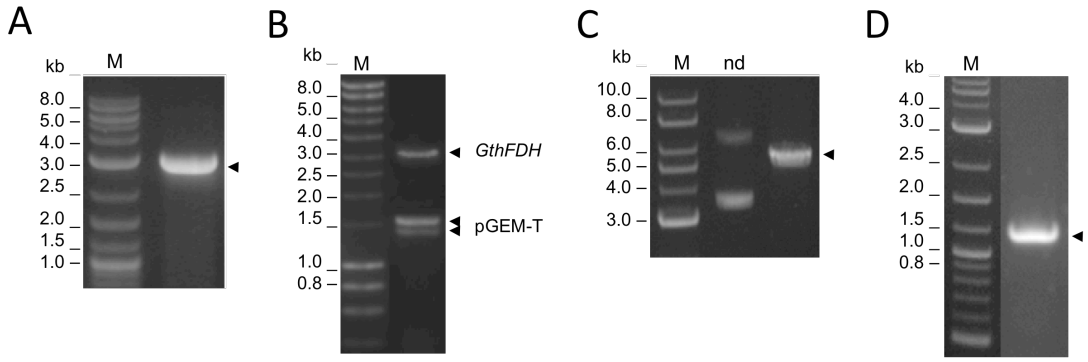


Figure 4.8. 1% agarose gel electrophoresis analysis of pET28a-Histhr:GthFDHa. (A) PCR of the *GthfdhA* (3 kb) gene as *Sac*II/*Xho*I insert; primers were SacII_GthfdhA_F *vs*. XhoI_GthfdhA_•R; (B) pGEM-T restriction analysis and excision of the insert (*Sac*II/*Xho*I and *Bsa*I); (C) *Sac*II/*Xho*I restriction of pET28a (5.2 kb) (nd: non-digested); (D) PCR analysis of pET28a-Histhr:GthFDHa construct; primers were pT7_F *vs* GthfdhA_4R (exp. 1.2 kb). M: 1 kb DNA ladder (kb). Phusion™ was used as DNA polymerase in PCR.

4.3.1.4 pTYB11 constructs

Bmmdh•, *Bmact*• and *Bmfalddh*• were first PCR-amplified from *B. methanolicus* PB1 genomic DNA as *Sap*I/*Xho*I fragments, while *GthfdhA*• was PCR-amplified from *G. thermoglucosidasius* genomic DNA (Figure 4.9-A). The inserts were gel-purified, cloned into pGEM®-T Easy and transformed into *E. coli* JM109 cells. White colonies were screened for the presence of insert-containing plasmid by restriction digestion, and the insert sequenced (Figure 4.9-B). The inserts were then *Sap*I/*Xho*I excised from pGEM®-T Easy (with the vector being further digested with *Bsa*I in the case of *GthfdhA*), ligated into pTYB11 and transformed into JM109 cells.

The intein-tag constructs pTYB11 *Bmmdh*•, pTYB11 *Bmact*•, pTYB11 *Bmfalddh*• and pTYB11 *GthfdhA*• were successfully generated.

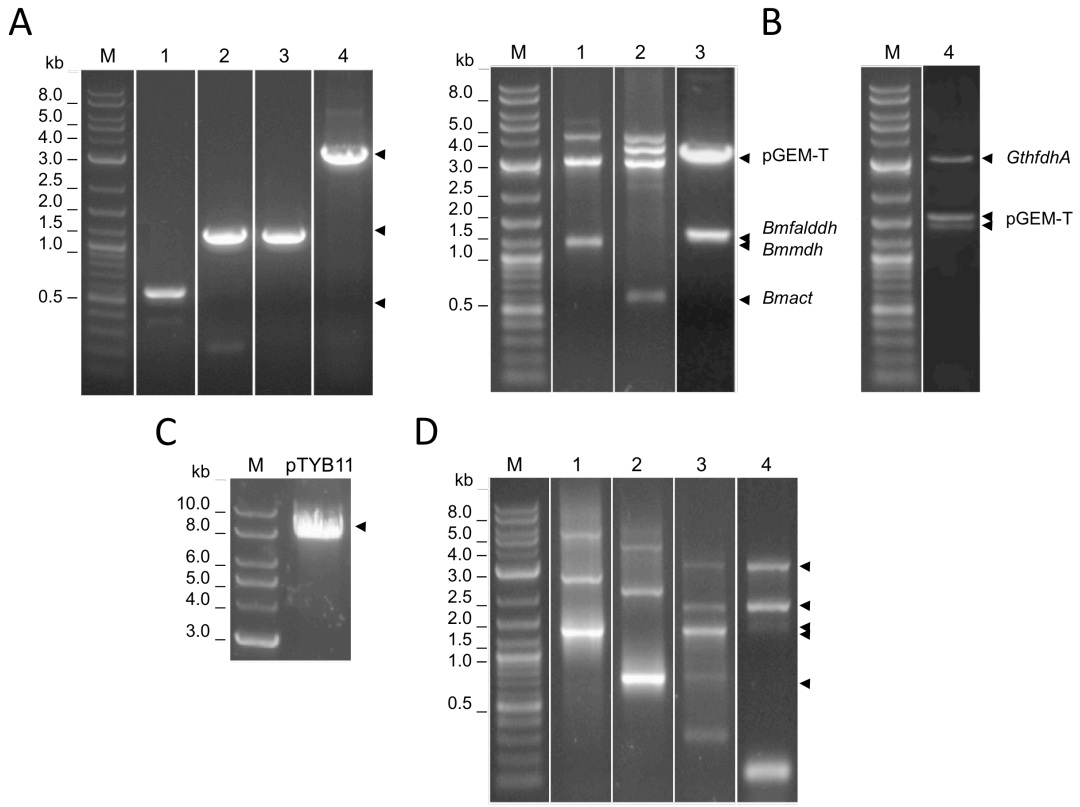


Figure 4.9. 1% agarose gel electrophoresis analysis of the pTYB11 constructs. Lanes: 1- *Bmmdh*; 2- *Bmact*; 3—*Bmfalddh*; 4- *GthfdhA*. (A) PCR of the genes as *Sap*I/*Xho*I inserts; primers were SapI_F *vs*. Xho_•R; (B) pGEM-T restriction analysis and excision of the inserts (*Sap*I/*Xho*I, with *Bsa*I for *GthfdhA*); (C) *Sap*I/*Xho*I restriction of pTYB11 (7.4 kb); (D) PCR analysis of pTYB11s construct; primers were pT7_F *vs* •R (except for pTYB11 *GthfdhA*•, primers were pT7F vs. GthfdhA_7R (2 kb) and vs. GthfdhA_•R (3 kb)). M: 1 kb DNA ladder (kb). Phusion™ was used as DNA polymerase in PCR.

4.3.2 Methanol dehydrogenase

4.3.2.1 Expression of BmMDH in *E. coli* (pET vector)

Once the presence of the pET expression vector was confirmed in *E. coli* JM109 by growth on kanamycin- (pET28a) or ampicillin- (pET19b) supplemented media, and the presence of the correct insert confirmed by PCR and restriction digest analysis, the constructs were transformed into the *E. coli* expression host.

First, constructs pET28a *Bmmdh*:His• and pET28a *pBMmdh*:His•, as well as pET28a *Bmact*:His• and pET19b *Bmact*• were successfully transformed into the *E. coli* BL21(DE3)pLysS expression strain. Cell pellets of induced cultures were fractionated and the levels of over-expression of the genes were assessed via SDS-PAGE analysis of the cell pellet, and of the soluble and insoluble fractions (Figure 4.10).

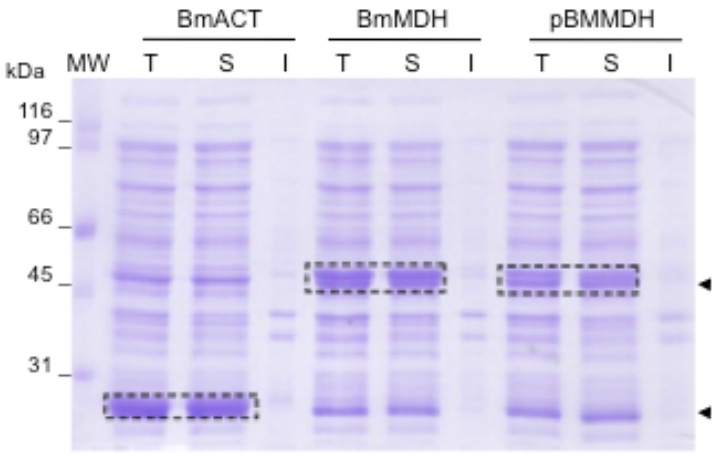


Figure 4.10. SDS-PAGE of *E. coli* BL21(DE3)pLysS pET28a *Bmact*:His•, pET28a *Bmmdh*:His• and *pBMmdh*:His• after overnight IPTG-induced expression. T: Total cell extract; S: Soluble lysate; I: Insoluble fraction. MW: Molecular weight markers (kDa). BmACT, BmMDH and pBMMDH are boxed. Coomassie blue staining.

Significant soluble expression of the methanol dehydrogenase from *B. methanolicus* PB1 was achieved for both pET28a *Bmmdh*:His• and pET28a *pBMmdh*:His• following an overnight incubation and induction with 1 mM IPTG. The M_r of the protein's monomeric subunit was calculated to be *ca.* 49 kDa by measuring the distance travelled by the individual band compared to that travelled by the size markers running alongside (Fig 4.10). Similarly, strong expression of the activator BmACT was achieved under the same conditions, with the protein subunit M_r calculated to be 27 kDa.

The M_r values of the proteins were estimated by co-electrophoresis of the protein samples with broad-range markers (Bio-Rad) (Figure 4.11).

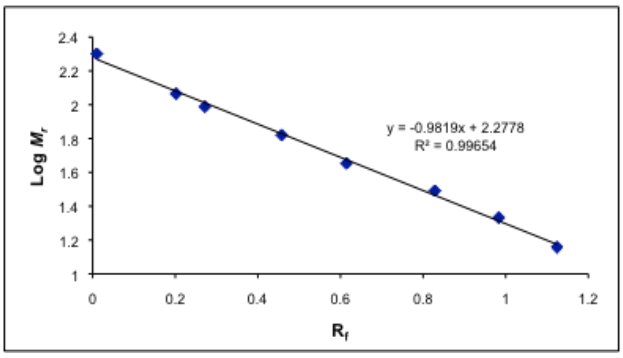


Figure 4.11. Calibration curve for SDS-PAGE broad-range markers (Bio-Rad). For each individual weight marker, the log M_r was plotted against the R_f value, defined as the distance travelled by the protein marker over the distance travelled by the migration front. A line of best fit was applied. Weight markers were: myosin 200 kDa, β-galactosidase 116.25 kDa, phosphorylase b 97.4 kDa, BSA 66.2 kDa, ovalbumin 45 kDa, carbonic anhydrase 31 kDa, soybean trypsin inhibitor 21.5 kDa, lysosyme 14.4 kDa (and aprotinin 6.5 kDa, not plotted).

When aliquots of the *E. coli* cultures were taken immediately prior to induction (0 h) and then at 4 h, 6 h and 18 h (over-night) post induction, a time-lapse expression profile was observed for both variants of the enzyme, genomic and plasmidic (Figure 4.12).

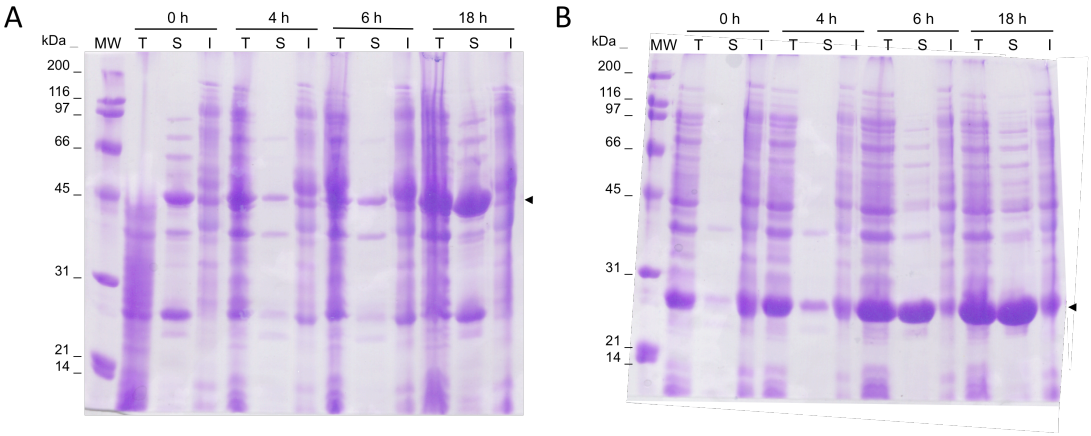


Figure 4.12. SDS-PAGE of the time course over-expression of (A) BmMDH and (B) BmACT in *E. coli* (pET28a constructs). T: Total extract; S: Soluble lysate; I: Insoluble fraction. MW: Molecular weight markers (kDa). Coomassie blue staining.

Expression of the recombinant BmMDH and BmACT was observed from 4 to 6 hours post-induction, but was best after overnight-induced growth. However, some leaky expression of the target protein could be observed pre-induction (t_0).

No differences in expression pattern were observed between the recombinant genomic and plasmidic BmMDH.

4.3.2.2 Assaying the recombinant *E. coli* pET28a *Bmmdh*:His•

The soluble fraction after lysis of the cells produced by the IPTG-induced *E. coli* BL21(DE3) pLysS pET28a *Bmmdh*:His• was assayed for methanol dehydrogenase activity. In phosphate buffer (pH 8.5) no activity was detected with methanol at 50 °C in the presence of 5 µM $MgSO_4$, 5 µM $ZnSO_4$ and 2 mM NAD^+. Also, no activity was detected when the assay was performed under the same conditions but in glycine buffer (0.2M glycine/KOH, pH 9.5).

Supplementing the assay with an aliquot of the soluble fraction from IPTG-induced *E. coli* BL21(DE3)pLysS pET28a *Bmact*:His• cell paste did not produce any measureable activity.

No activity was found with $NADP^+$ as the co-factor, nor with ethanol or butanol instead of methanol as the substrate.

When the assay was performed in the reverse direction (formaldehyde reductase), *i.e.* with NADH as the co-factor and formaldehyde as substrate, no measurable activity was detected.

However, NADH oxidase activity was recorded when the soluble fraction of the cell paste was assayed in phosphate buffer (pH 8.5) with NADH only (Figure 4.13). Without cell extract, NADH (0.2 mM) gives an A_{340nm} of *ca*. 1.1.

Untransformed *E. coli* cell paste was assayed for intrinsic methanol dehydrogenase activity, or alcohol dehydrogenase activity with methanol. The soluble fraction was assayed with methanol and NAD^+ and no activity was measured. When assayed with formaldehyde and NADH, the decrease in absorbance observed corresponded to the NADH oxidase. Replacing NAD(H) with NADP(H) did not change this observation.

The NADH oxidase activity observed would appear to mask the MDH activity in the *E. coli* cell extract, prompting for the purification of the BmMDH enzyme.

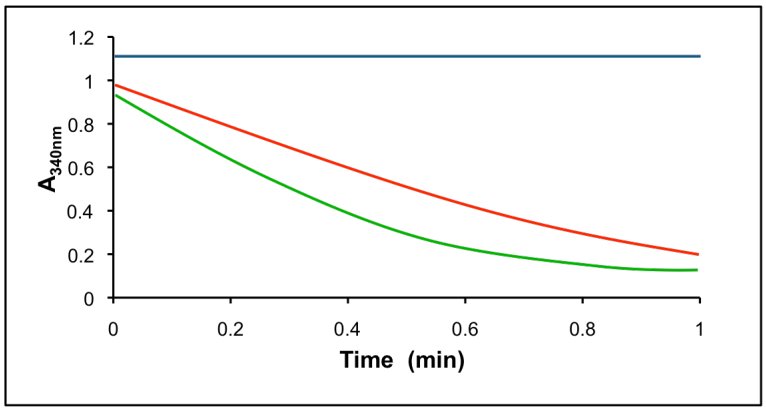


Figure 4.13. NADH oxidase assay on *E.coli* cell paste soluble fraction. In phosphate buffer (pH 8.5) with 0.2 mM NADH and no cell extract (blue/top), *E.coli* pET28a BmMDH:His cell extract (red/middle), *E.coli* pET28a BmACT:His cell extract (green/bottom).

4.3.2.3 Metal-affinity chromatography and purification of BmMDH

Since the genes *Bmmdh* and *Bmact* were cloned into the vector pET28a as N-terminal fusion to the histidine tag, the *E. coli* over-expressed proteins could be purified by metal-affinity chromatography.

Soluble cell extracts fraction of IPTG-induced cultures of *E. coli* BL21(DE3) pLysS pET28a *Bmmdh*:His• and *E. coli* BL21(DE3) pLysS pET28a *Bmact*:His• were loaded onto nickel-charged columns. Following passage through the columns, the recombinant His-tagged BmMDH and BmACT proteins were purified to near-homogeneity. The His-tagged BmMDH was eluted with 7% and 14% His-elute buffer, while the His-tagged BmACT was eluted with 14% and 40% His-elute buffer (percentage given as fraction of His-elute buffer in His-binding buffer) (Figure 4.14).

The two versions of the methanol dehydrogenase, MDH and pBMMDH, were expressed and eluted in the same way. For each recombinant protein, the semi-purified fractions were pooled.

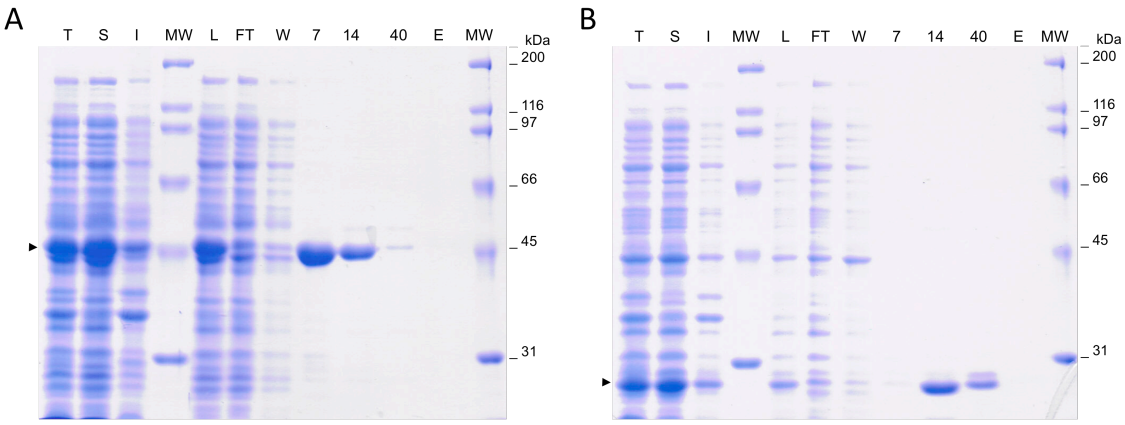


Figure 4.14. SDS-PAGE of Ni^{2+}-purification of (A) BmMDH:His , and (B) BmACT:His. pET28a constructs in *E. coli* BL21(DE3)pLysS. T: Total cell extract; S: Soluble lysate; I: Insoluble fraction; L: Loaded fraction; FT: Flow-through; W: Wash; 7: 7% His-elute buffer; 14: 14% His-elute buffer; 40: 40% His-elute buffer; E: 100% His-elute buffer. MW: Molecular weight markers (kDa). Coomassie blue staining.

Assaying the His-purified BmMDH:His and BmACT:His showed that even though no NADH oxidase was recorded in either sample, the His-purified BmMDH did not display activity with methanol and NAD^+ nor with formaldehyde and NADH, with or without the addition of its activator BmACT:His. The BmACT:His protein alone showed no activity either.

The metal-affinity purification procedure for the His-tagged proteins proved successful in that it removed the NADH oxidase, and as shown on the SDS-PAGE, the lanes corresponding to the eluted proteins contained almost exclusively the MDH and ACT proteins respectively. Assay buffer and conditions for the methanol dehydrogenase and formaldehyde reductase assays were optimised to the best of current knowledge, with both Mg and Zn present, and pH and temperature adjusted. The Ni^{2+}-purified plasmidic MDH also showed no measurable enzymatic activity.

BmMDH and BmACT co-purify

A non-tagged version of the activator, cloned in pET19b *Bmact*•, was successfully co-transformed into *E. coli* BL21(DE3)pLysS pET28a *Bmmdh*:His•. Selection of the co-transformants was carried out on LB supplemented with chloramphenicol, kanamycin and ampicillin.

In an attempt to determine whether the BmMDH and BmACT proteins interacted with each other, a cell extract from an IPTG-induced culture of *E. coli* BL21(DE3)pLysS pET28a *Bmmdh*:His• pET19b *Bmact*• was loaded onto a nickel-charged metal-affinity column. Co-elution of the proteins would signify that the non-tagged BmACT interacted with the His-tagged BmMDH and that the complex had adsorbed to the column and was eluted as a single entity.

SDS-PAGE analysis of the eluted fractions revealed that the co-expressed His-tagged BmMDH and the BmACT proteins co-eluted (Figure 4.15).

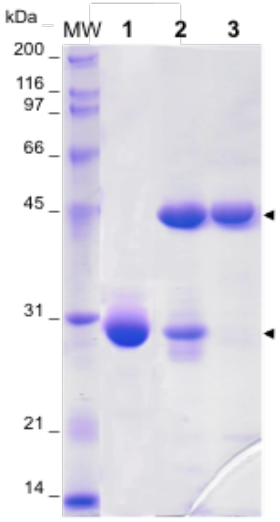


Figure 4.15. SDS-PAGE of the co-elution of BmACT with BmMDH-His. Lanes: 1- Ni^{2+}-purified BmACT from pET28a His:BmACT; 2- co-elution of Ni^{2+}-purified BmMDH-His (from pET28a BmMDH:His) and BmACT (pET19b BmACT, non His-tagged); 3- Ni^{2+}-purified BmMDH from pET28a BmMDH:His; MW: Molecular weight markers (kDa). Coomassie blue staining.

The soluble cell extracts of separate IPTG-induced cultures of *E. coli* BL21(DE3) pLysS pET28a *Bmmdh*:His• and *E. coli* BL21(DE3) pLysS pET19b *Bmact*• were also combined and loaded onto a nickel-charged metal-affinity column. SDS-PAGE analysis confirmed that, although separately expressed, the His-tagged BmMDH and the non-tagged BmACT proteins co-eluted.

The recombinantly-expressed methanol dehydrogenase and its activator from *B. methanolicus* seem to interact with each other. However, assay of the co-eluted

BmMDH:His and BmACT for methanol dehydrogenase activity or formaldehyde reductase activity revealed no enzymatic activity.

4.3.2.4 Assessing the multimeric assembly of BmMDH

The purified BmMDH:His and BmACT:His proteins were concentrated before loading onto a Hiload 16/60 Superdex 200 gel filtration column. Up to 4 mL of the purified protein were loaded onto an Amicon Ultra-10K NMWL centrifugal filter device (Millipore) and centrifuged down to 0.75 mL. The gel filtration column was loaded with 0.5 mL of concentrated protein (BmMDH:His and BmACT:His) and elution (running buffer 50 mM Tris, pH 8) recorded on an ÄKTA (following the A_{280nm}). Protein-containing fractions were collected and stored on ice. The K_{av} for the protein was calculated and its M_r extrapolated from the low- or high-M_r markers calibration curve. The gel filtration columns were initially calibrated with the standard protein solutions for high and low MW, and calibration curves generated (Figure 4.16).

With BmMDH:His, fractions A1 to D3 were collected. From the elution profile, it seems that BmMDH:His eluted in one major peak, at an elution volume (V_e) of 43.02 mL (> 500 mAU), with a shoulder at *ca.* 40 mL, and a small secondary peak at V_e = 50.36 mL (< 50 mAU) (Figure 4.17-A). SDS-PAGE analysis of fractions A9 to A13 revealed a strong clean single band of M_r 43 kDa, especially in fractions A11 to A13, confirming the presence of the recombinant protein under the main peak eluted (Figure 4.19-D). The 43 kDa band could be seen in fractions A9 and A10 albeit much less intensely.

In the case of BmACT:His, fractions A1 to D12 were collected. From its elution profile, BmACT:His eluted in one major peak at a V_e of 57 mL (> 1,200 mAU) (fractions B9 to B13) (Figure 4.17-B). However, prior to the main peak four small peaks corresponding to eluted proteins can be observed: at V_e of 22.97 mL (A3), 36.21 mL (A12), 41.52 mL (B1) and 47.49 mL (B4) (all *ca.* 100 mAU). SDS-PAGE analysis of fractions B9 to B13 revealed a strong clean single band of M_r 27 kDa in all these fractions, and especially B10 to B13, confirming the presence of the recombinant protein under the main peak (Figure 4.17-D). Furthermore, SDS-PAGE analysis allowed for the dismissal of fractions A3, A12, B1 and B4.

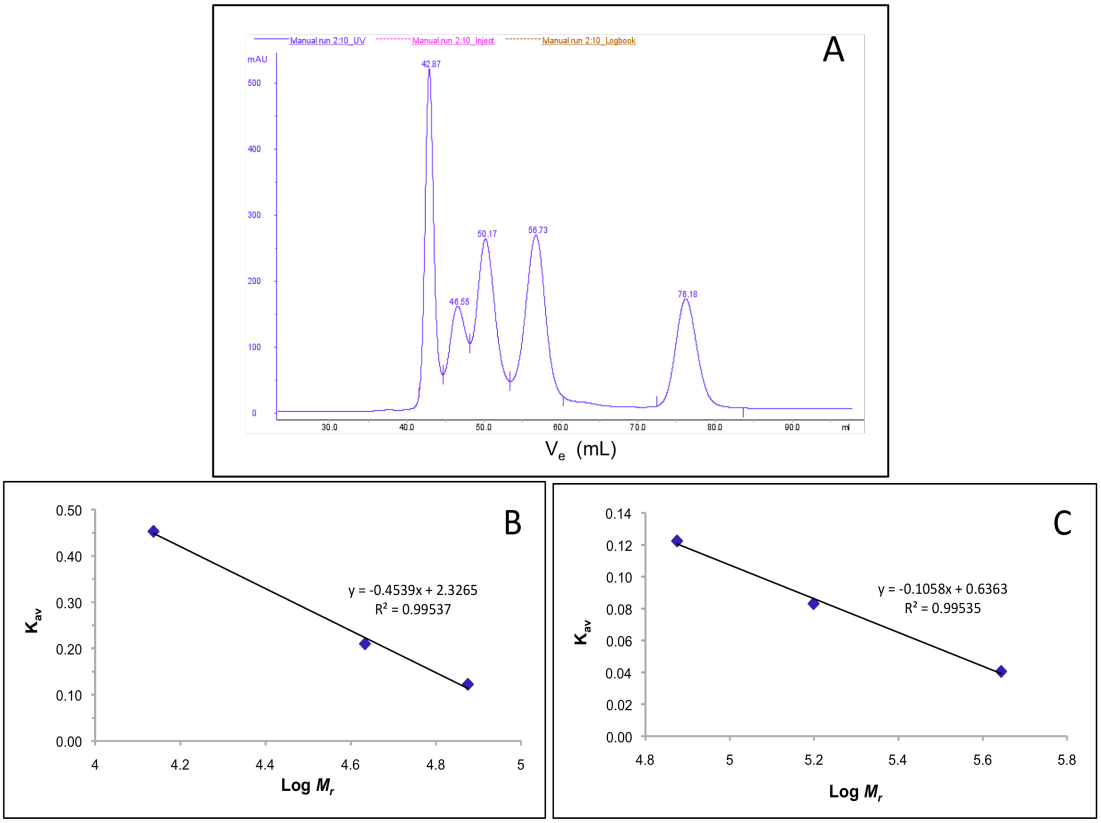


Figure 4.16. Calibration of the Hiload 16/60 Superdex 200 gel filtration column with molecular weight standards. (A) Gel filtration profile of MW standards on Hiload column. Standard curve for (B) Low M_r standards and (C) Low M_r standards. M_r standards (GE healthcare) were: ribonuclease M_r 13,700; ovalbumine M_r 43,000; conalbumine M_r 75,000; aldolase M_r 158,000; ferritin M_r 440,000.

With pBMMDH:His fractions A1 to B12 were collected. The elution profile showed several small peaks (<40 mAU) spread across fractions A5 to B5. Peaks were annotated at the following elution volumes (V_e): 7.13 mL (A5-A6), 8.73 mL (A7-A8), 11.72 mL (A9-A11), 16.23 mL (A15), 19.09 mL (B3), 20.11 mL (B4) and 21.19 mL (B5) (Figure 4.19-C). SDS-PAGE analysis of fractions A5 through to B6 showed that little protein was over-expressed by *E. coli*, especially compared to the BmMDH:His obtained from over-expression in the same strain of *E. coli* under the same conditions (Figure 4.17-E). The main product appeared in fraction A8, with a calculated M_r 42 kDa, corresponding to the recombinant protein.

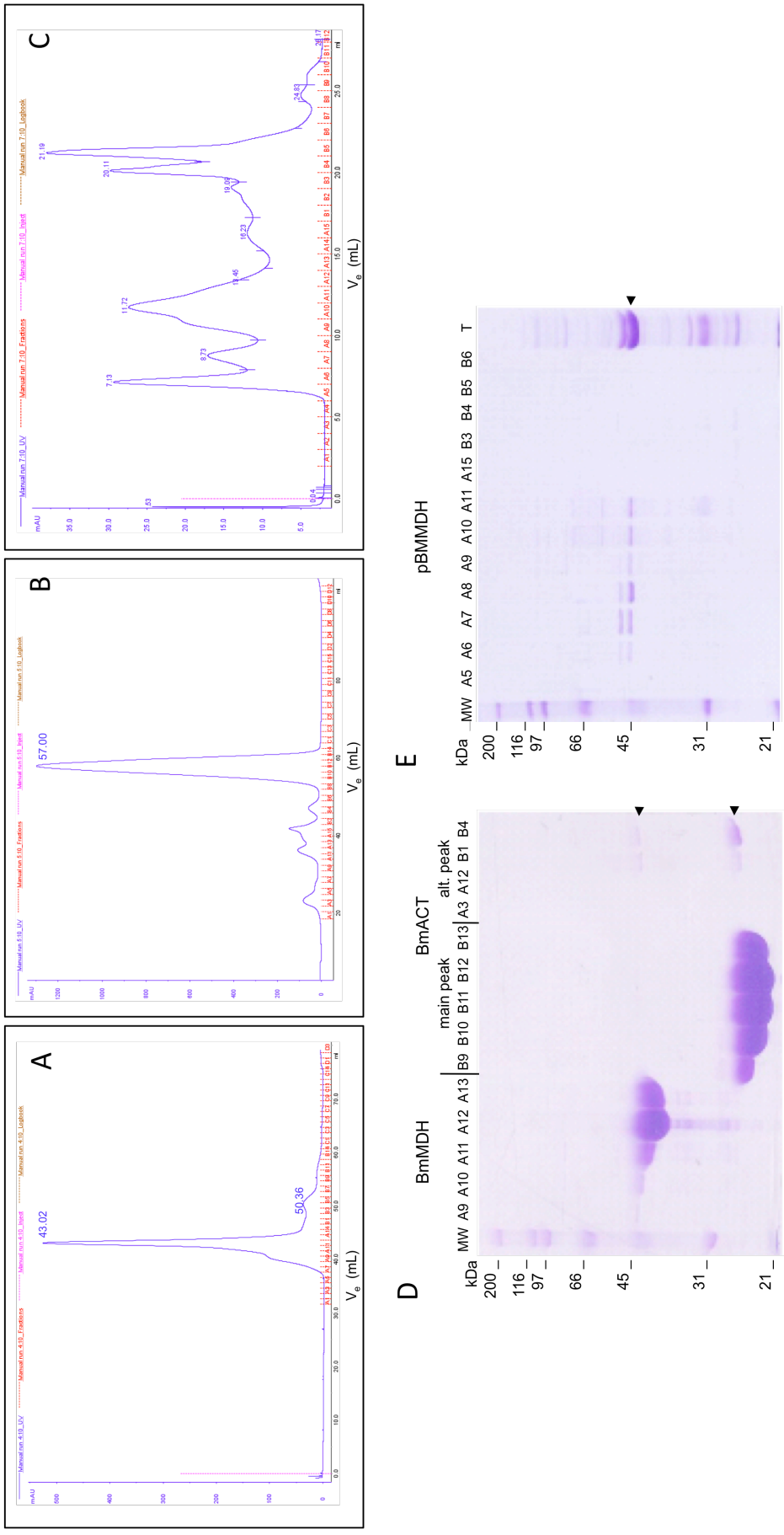


Figure 4.17. Gel filtration profiles for A- BmMDH, B- BmACT and C- pBMMDH (V_e in mL), and corresponding SDS-PAGE analysis for the collected fractions D- BmMDH and BmACT, and E- pBMMDH (MW: Molecular weight markers (kDa)).

The V_e values for the main peak fractions were used to calculate the K_{av} and thus the M_r of the native proteins by extrapolating from the calibration plots for low and high-MW markers (Figure 4.16). This was then used to calculate the number of subunits present. The results are presented in Table 4.3.

Table 4.3. Summary of the gel filtration data for BmMDH:His and BmACT:His.

Sample (fractions)	V_e (mL)	K_{av}	Log M_r	M_r (kDa) (native)	M_r (kDa) (subunit)	Nb of subunits
BmMDH:His (A11-13)	43.02	0.04	-5.61	410	43	9.5
BmACT:His (B10-13)	57.00	0.21	-4.65	48	27	1.7
pBMMDH:His (A6)	7.13	0.00	-5.64	440	42	10.5
(A8)	8.73	0.10	-5.32	210	42	5.0
(A10)	11.72	0.28	-4.73	54	42	1.3

The number of sub-units in the native protein was calculated by dividing the M_r of the native protein, as determined by the K_{av} after gel filtration, by the M_r of the sub-unit as determined under SDS-PAGE analysis. With a sub-unit number of 9.5, BmMDH:His appeared to assemble in a homodecamer. BmACT:His, with a sub-unit number of 1.7, assembled as a homodimer. However, pBMMDH:His seemed to exist under various states of assembly: a homodecamer in fraction A6, a homo-pentamer in fraction A8 and a monomer in fraction A10.

Assays for methanol dehydrogenase and formaldehyde reductase activity were carried out on the purified fraction of BmMDH:His (A11-A13) and BmACT:His (B10-B12). Even though no NADH oxidase activity was detected in either sample, no MDH activity could be measured, with or without the presence of activator, for both the His-tagged BmMDH and the His-tagged pBMMDH purified enzymes. As assembled, purifed (affinity chromatography followed by gel filtration), the proteins were expected to be active. Thus, a test was set up on a commercially-available alcohol dehydrogenase (ADH) exhibiting activity with methanol in order to confirm the suitability of the assay.

4.3.2.5 Horse liver ADH activity with methanol

The methanol dehydrogenase assay, as described in Chapter 2 (2.2.6.2) was carried out, replacing the recombinant enzyme extract with a commercially-available alcohol dehydrogenase displaying activity with methanol, in this case the alcohol dehydrogenase from horse liver (E.C. 1.1.1.1) (Sigma-Aldrich). The assay was performed at 37 °C, the core temperature of horse, and the reaction was started with 500 µM methanol. Enzyme activity was detected and was shown to be proportional to the amount of enzyme used (Figure 4.18) (an initial rate of 0.82 min^{-1} was measured with 0.024 U of horse liver ADH and 0.46 min^{-1} with 0.012 U).

The quoted specific activity for the stock enzyme was 1 $\mu mol.min^{-1}.mg^{-1}$. Here, the measured rates convert to a specific activity of 0.5 $\mu mol.min^{-1}.mg^{-1}$; these are in reasonable agreement given the enzyme stock had been in use for some years in the laboratory. This also shows that the assay itself is sensitive and can detect NAD^+ reducing enzyme activity to the level of enzymatic activity in the sample.

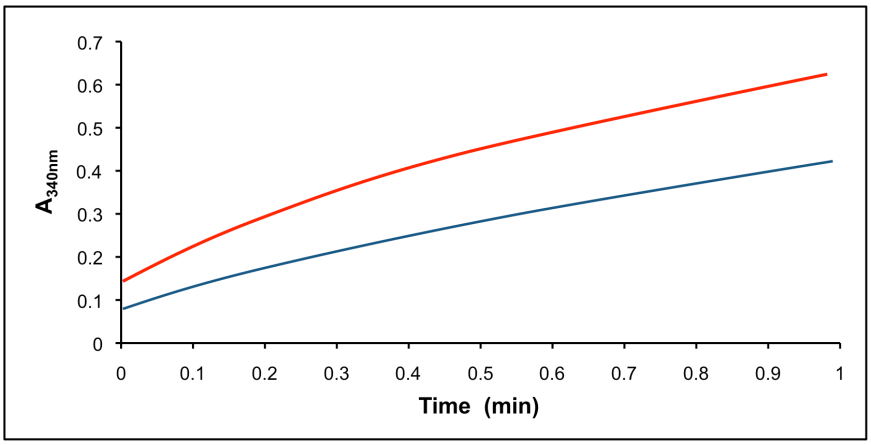


Figure 4.18. Dehydrogenase assay on horse liver ADH against methanol. The enzyme was assayed in phosphate buffer (pH 8.5), containing 5 µM $MgSO_4$ and 5 µM $ZnSO_4$, 2 mM NAD^+ at 37 °C. The assay was started with 0.5 mM methanol, and the A_{340nm} followed over 1 min. Traces (smoothed) are shown for 0.024 U (red/top) and 0.012 U (blue/bottom) horse liver ADH.

When *E. coli* cell extract was added, NADH oxidase activity partly though not wholly masked the NAD reduction by the ADH. However, it may prove enough to have masked the activity of the BmMDH, and the absence of detectable activity with the recombinant enzyme is therefore unlikely to be due to the assay method.

4.3.2.6 Spectrophotometric scan of BmMDH

The active site of the BmMDH enzyme is thought to be containing a tightly but not covalently-bound NADH (Arfman *et al.*, 1997). A spectrophotometric scan across the 250 to 400 nm region would show a distinctive cofactor peak at 340 nm.

The fractions of gel-purified proteins with the highest protein concentration were pooled: fractions A12 and A13 of BmMDH and fractions B10-B12 of BmACT (see 4.3.2.4, Figure 4.17-D).

Initial scanning experiments of BmMDH showed no peak in absorbance at 340 nm. When 5 mM DHA was added to the sample in the cuvette, the scan did not show a peak in absorbance at 340 nm neither (Figure 4.19). Spiking the protein aliquot with 0.072 mM NADH provided a positive control to this assay, showing a clear peak at 340 nm.

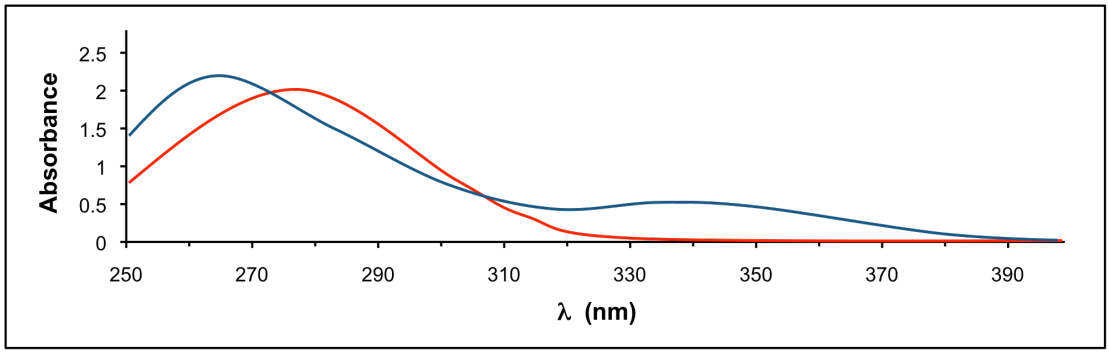


Figure 4.19. Scan (smoothed) of the DHA-treated (5mM) His-purified BmMDH enzyme over the 250-400nm rang. DHA-treated BmMDH (red) and spiked with 0.072 mM NADH (blue).

The ability of 5 mM DHA to reduce NAD^+ was confirmed by incubating 2 mM NAD^+ with 5 mM DHA under methanol dehydrogenase assay buffer conditions. Measured over 10 min, the increase of A_{340nm} confirmed the reduction of NAD^+ to NADH. Total reduction, *i.e.* reaching an $A_{340nm} \approx 1.2$, was achieved after overnight incubation.

Spiking the protein aliquot with 0.072 mM NADH provided a positive control to this assay, showing a clear peak at 340 nm.

4.3.2.7 *E. coli* KRX-expressed BmMDH(pET28a Histhr:) - a functional formaldehyde reductase

Once the presence of the pET28a Histhr:*Bmmdh*• and His:*Bmact*• constructs were confirmed in *E. coli* JM109 by growth on LB-Kan and PCR analysis, the constructs were transformed into the *E. coli* expression strain ArcticExpress and KRX.

Induced cultures were harvested and levels of over-expression were assessed by SDS-PAGE analysis of the fractionated respective cell pellets (Figure 4.20).

Soluble expression was achieved in KRX for the methanol dehydrogenase and its activator from *B. methanolicus* PB1 following induction with 1 mM IPTG and 0.1% rhamnose. The M_r values of the subunits were calculated to be 43 kDa and 27 kDa, respectively (as previously demonstrated (4.3.2.1)).

Since the KRX strain produced consistent quantities of cell pellets and soluble enzymes in the cell extracts, further focus was directed to this strain in preference to ArcticExpress™.

Prior to performing the enzyme assays, cell paste of induced overnight cultures of KRX pET28a Histhr:BmMDH• and KRX pET28a Histhr:BmACT• were harvested, lysed, and the soluble fractions loaded onto nickel-charged metal affinity column. The His:BmMDH and His:BmACT proteins were purified, with BmMDH eluting in the 7% eluting buffer fractions, and BmACT eluting in the 14% fractions.

When this KRX-expressed, His-purified BmMDH enzyme was assayed with methanol for substrate and NAD^+ as co-factor, it did not exhibit activity even with the addition of the activator BmACT. However, BmMDH was active against formaldehyde (NADH as co-factor). Typical assays were carried out in a final volume of 1 mL, in triplicate at 50 °C, in 50 mM HEPES buffer (pH 8.5, 5 µM $MgSO_4$, 5 µM $ZnSO_4$) and 0.2 mM NADH. Each assay contained 5 µg of enzyme and was started with the addition of formaldehyde. All further assays and parameters mentioned thereafter refer to the enzyme being assayed in the formaldehyde reductase orientation, *i.e.* with formaldehyde and NADH.

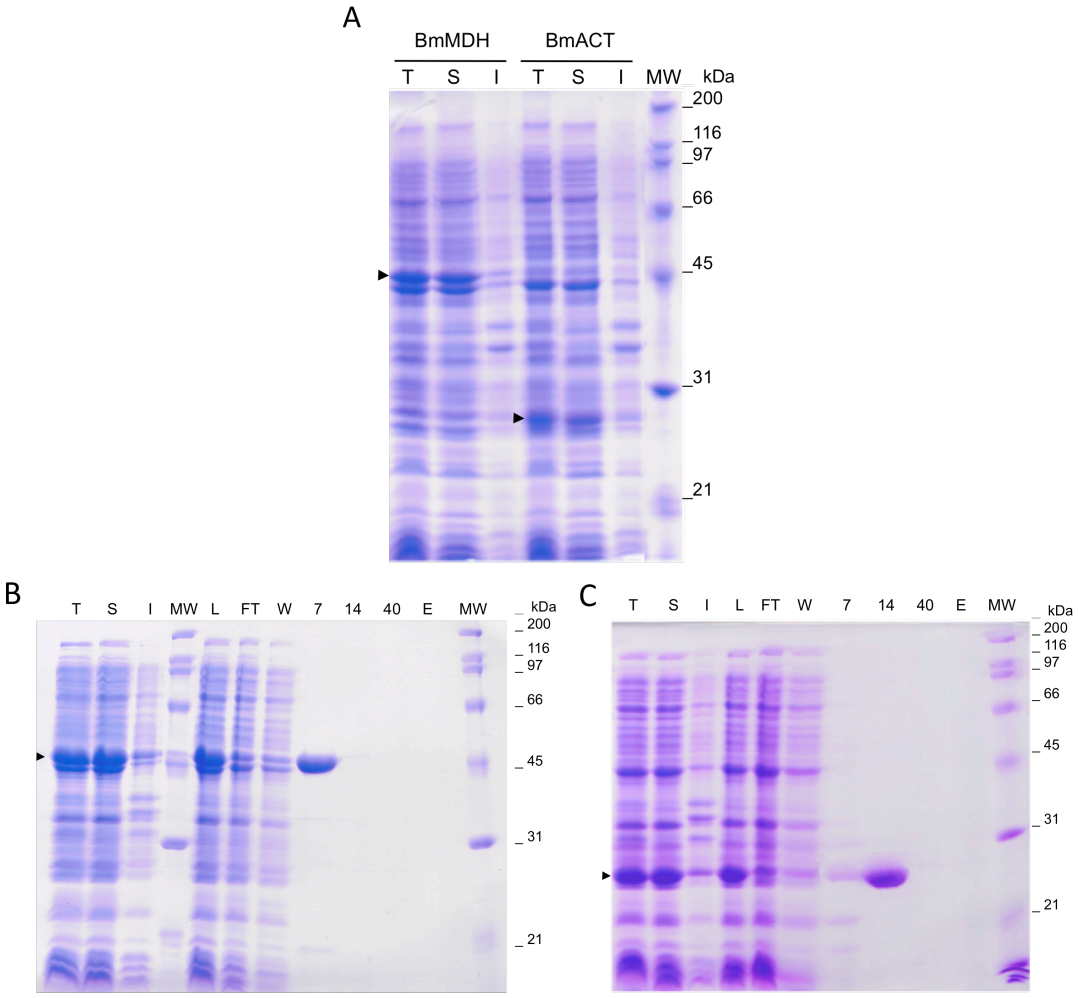


Figure 4.20. SDS-PAGE analysis of the proteins expressed in *E. coli* KRX pET28a Histhr:BmMDH• and pET28a Histhr:BmACT• (A), and Ni^{2+} purified proteins (B and C respectively). T: Total extract; S: soluble lysate; I: Insoluble fraction; L: loaded fraction; FT: Flow-through; W: Wash; 7: 7% His-elute; 14: 14% His-elute; 40: 40% His-elute; E: 100% His-elute; MW: Molecular weight markers (kDa). Coomassie blue staining.

4.3.2.8 Kinetic parameters of BmMDH: formaldehyde reductase

The protein concentration of the KRX-produced BmMDH was measured at 0.59 $mg.mL^{-1}$.

The proportionality of the amount of enzyme on the reaction rate was confirmed.

The effect of the concentration of formaldehyde substrate on specific activity of the enzyme BmMDH was investigated. The experimental data appear to fit the Michaelis-Menten equation (Figure 4.21).

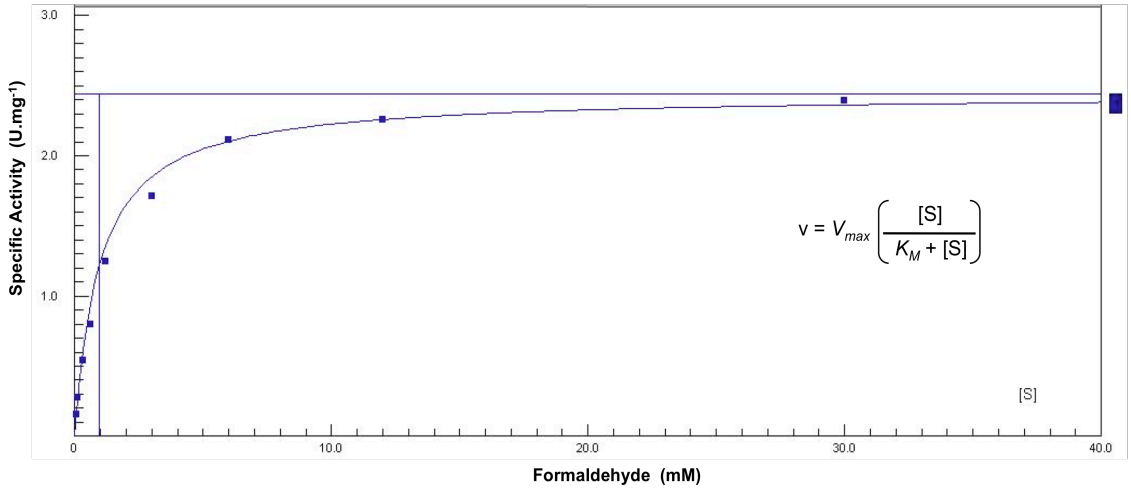


Figure 4.21. Dependence of BmMDH velocity on concentration of formaldehyde.. Michaelis-Menten equation (inset).

Values of V_{max} and K_M were determined by the direct linear plot of Eisenthal and Cornish-Bowden (1974) (1). However, to determine whether the data are a good fit to the Michaelis-Menten equation, a linear transformation was first performed on them. The preferred transformation is that of Hanes (2), where [S]/v is a linear function of [S]:

$$v = V_{max}\left(\frac{[S]}{K_M + [S]}\right) \quad (1) \qquad \text{to give} \qquad \frac{[S]}{v} = \left(\frac{1}{V_{max}}\right)[S] + \left(\frac{K_M}{V_{max}}\right) \quad (2)$$

In this linear transformation, the trendline has a gradient of $1/V_{max}$, and an estimate of K_M is given at the x-intercept ($-K_M$) while the y-intercept gives K_M/V_{max}. Furthermore, as both components of the equation contain the substrate concentration, the imprecision stays constant over the data range (*i.e.* as [S] increases).

A Hanes plot was fitted with the experimental data (Figure 4.22).

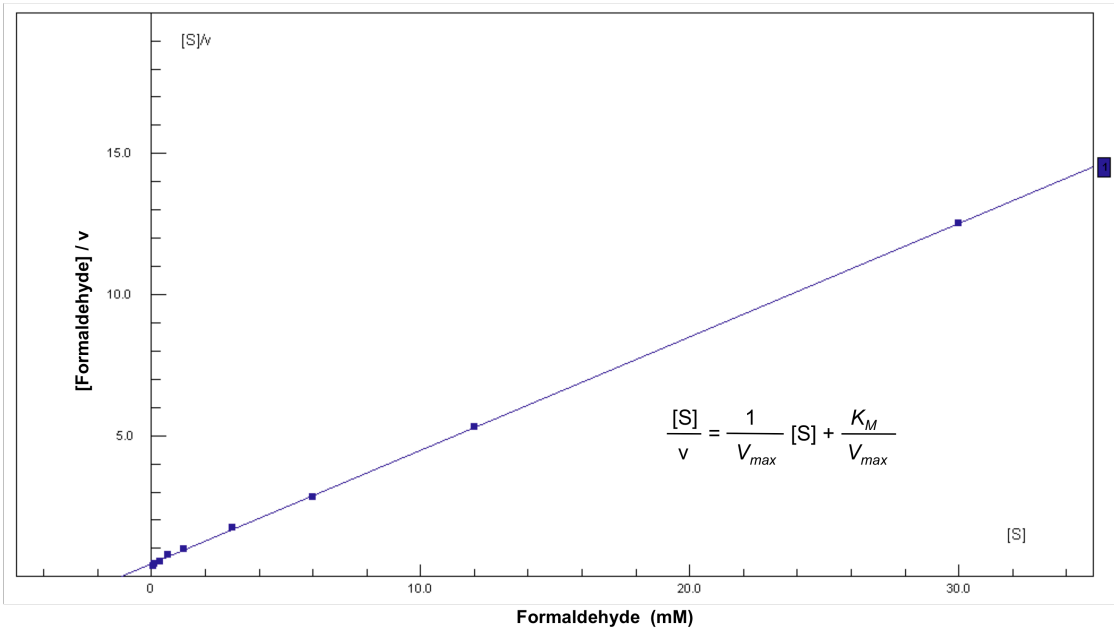


Figure 4.22. Hanes plot for BmMDH against formaldehyde. Hanes equation, inset ([S] in mM, v in U.mg^{-1}).

The recombinant BmMDH is catalytically active with formaldehyde and NADH as co-factor at 50 °C. The calculated K_M for formaldehyde is of 1.2 (± 0.2) mM and V_{max} of 2.5 (± 0.1) U.mg^{-1}. The addition of the activator BmACT did not have any effect on the kinetics of the enzyme.

4.3.2.9 Temperature profile and thermal inactivation

The activity of BmMDH with formaldehyde was assayed at various temperatures (Figure 4.23). The activity of BmMDH increased steadily with the assay temperature. The temperature optimum for the enzyme appears to lie near 85 °C to 90 °C. Past 90 °C, the activity dropped drastically. The Arrhenius analysis gave an E_a = 37.3 (± 3.9) kJ.mol^{-1}.

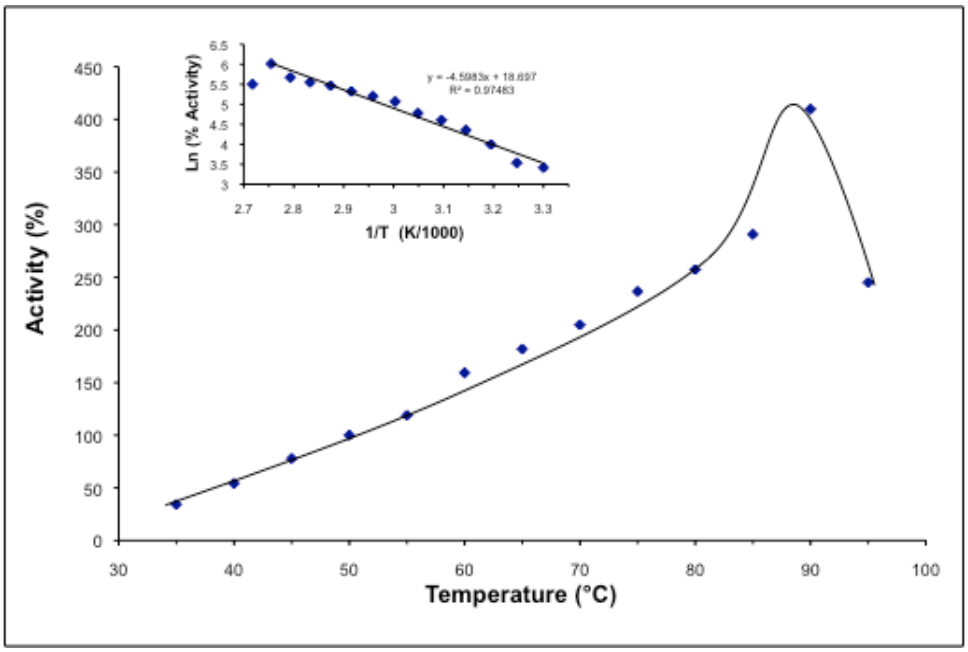


Figure 4.23. Thermal activity of the recombinant BmMDH with formaldehyde. Activity is expressed as a percentage of the activity of the enzyme at 50 °C (100%). Inset, Arrhenius plot of the data.

The thermostability of BmMDH was measured by incubation at various temperatures for up to 60 min (Figure 4.24).

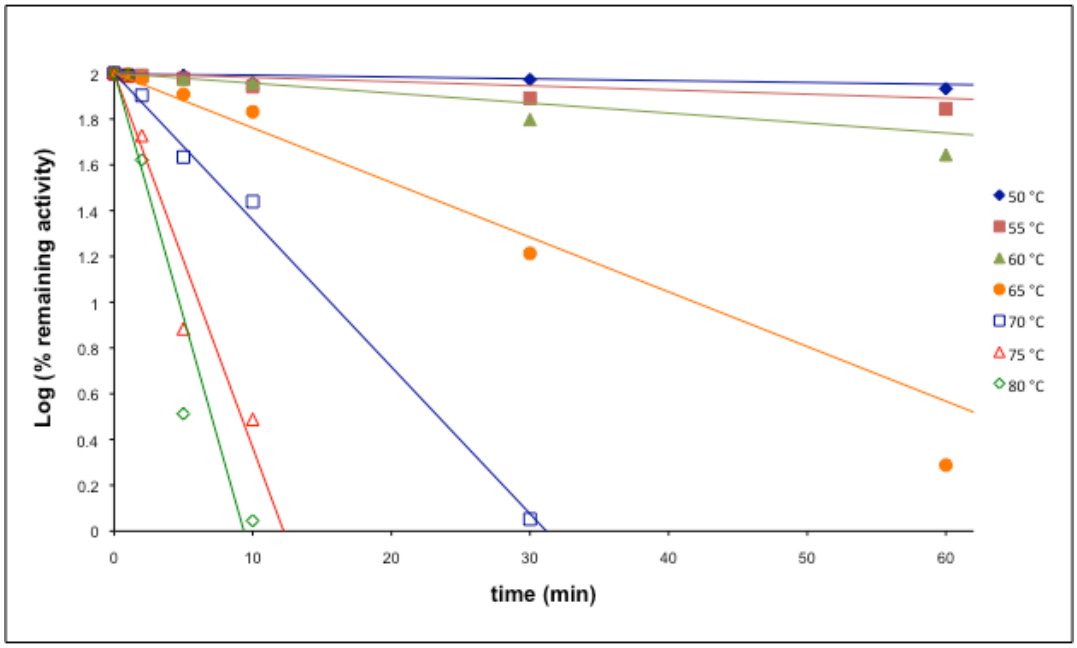


Figure 4.24. Thermal inactivation of His-purified BmMDH. Activity was measured at 50 °C after incubation at the stated temperatures for a given time.

While the enzyme retained 90% of its activity after one hour at 50 °C, the loss of activity was sharp past 60 °C. Estimates of half-life ($t_{1/2}$) are 8 h (476 min) at 50 °C, 3 h (187 min) at 55 °C, 68 min at 60 °C, 10 min at 65 °C. Past 70 °C, the protein is hardly active after more than a few minutes incubation. The addition of the activator BmACT did not enhance or impair the stability of the BmMDH.

SDS-PAGE analysis of the heat-incubated BmMDH indicated that thermally-induced precipitation was co-incident with the loss of enzyme activity inactivation (Figure 4.25).

Figure 4.25. SDS-PAGE analysis of heat-inactivation of (A) His-purified BmMDH• and (B) BmACT•. Samples were incubated for 10 min at a given temperature, from 50 to 95 °C. MW: Molecular weight marker (kDa); NT: non-heat treated sample; heat-treated samples (10 min) at given temperature (°C). Coomassie blue staining.

After 10 min incubation at 50 °C, 55 °C or 60 °C, the protein subunits appear to be stable. After 10 min at 65 °C, some degradation has occurred while after 10 min at 70 °C (and even more so at 75 °C), only traces of the protein can be seen on the SDS-PAGE, yet the MDH still registered 25% of its original activity. After treatment for 10 min at 80 °C and over, no protein could be seen on the gel, and very little activity if any could be measured with BmMDH. However, the observed degradation of the enzyme on the SDS-PAGE may not be accurate and the observation is more likely due to heat-precipitation of the protein, at least in the 70 to 85-90 °C region.

4.3.3 Formaldehyde dehydrogenase

4.3.3.1 Expression of BmFALDDH in *E. coli* (pET vector)

Once the pET vector containing the *Bmfalddh* gene was confirmed, transformed into *E. coli* JM109 cells by growth on kanamycin, and the insert itself confirmed by PCR analysis, the construct was transformed into the *E. coli* expression host.

The construct pET28a *Bmfalddh*:His• was successfully transformed into *E. coli* BL21(DE3)pLysS. The pET28a Histhr:*Bmfalddh*• construct was successfully transformed into the *E. coli* expression strains BL21(DE3)pLysS, Rosetta, ArcticExpress and KRX.

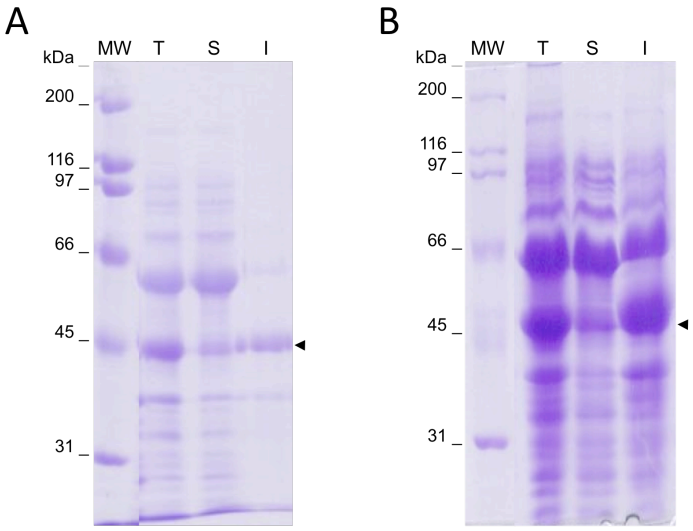


Figure 4.26. SDS-PAGE analysis of expression trial of pET28a Histhr:BmFALDDH in (A) *E. coli* Rosetta™ and (B) KRX. T: Total cell extract; S: Soluble fraction; I: Insoluble fraction; MW: Molecular weight markers (kDa). Coomassie blue staining.

The *E. coli* expression strain KRX performed better in expression tests than any of the others strain trialled, namely BL21(DE3)pLysS, Rosetta™ and ArcticExpress, by producing consistently-strong cell mass. A range of conditions of growth and induction patterns was also tested. However, no expression of formaldehyde dehydrogenase in a soluble form could be observed on the SDS-gels (Figure 4.26). The formaldehyde dehydrogenase appeared to be predominantly expressed in an insoluble form, with an M_r of ca. 46 kDa, although the soluble fraction was passed through a metal affinity column to purify any His-tagged soluble protein that was in

amounts too low to be visible on SDS-PAGE. SDS-PAGE of the eluate did not detect any protein corresponding to the formaldehyde dehydrogenase.

Re-solubilisation of the recombinant formaldehyde dehydrogenase expressed in the inclusion bodies was also attempted. They were first re-solubilised in 8 M urea, followed by purification on a nickel-charged affinity column. However, no binding between the His-tagged protein and the column's matrix occurred, with all the loaded material eluting through in the flow-through and wash fractions.

4.3.3.2 Assaying for formaldehyde dehydrogenase activity

Prior to assaying any recombinant activity, the intrinsic activity of the glutathione-dependent formaldehyde dehydrogenase from *E. coli* was measured. The assay was performed in HEPES buffer (pH 8.5) containing 5 mM glutathione and 0.5 mM NAD^+, and the assay was started with the addition of formaldehyde (6 mM for the standard assay). Using the soluble fraction of untransformed *E. coli* KRX cell paste, formaldehyde dehydrogenase activity was detected but only in the presence of gluthatione. Activity was proportional to the amount of enzyme, but do not seem to follow Michaelis-Menten kinetics (Figure 4.27).

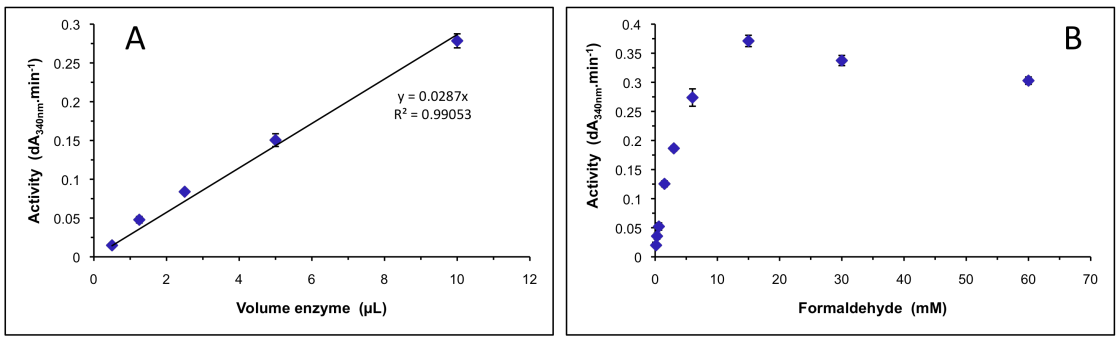


Figure 4.27. Formaldehyde dehydrogenase assay on *E. coli* KRX cell extract. Activity ($dA_{340nm}.min^{-1}$) as a function of (A) amount of enzyme (soluble fraction of cell extract), (B) formaldehyde concentration.

Then the His-purified fractions of the soluble lysate of *E. coli* KRX pET28a Histhr:BmFALDDH were assayed for formaldehyde dehydrogenase activity. However, with or without the addition of glutathione (5 mM), no such enzyme activity was detected in the His-purified eluates of the soluble fraction or the re-

solubilised inclusion bodies, with formaldehyde and NAD^+ or formate (sodium salt) and NADH (reverse direction).

The integrity of the assay was confirmed by assaying the commercial formaldehyde dehydrogenase from *Pseudomonas putida* (Sigma).

4.3.4 Formate dehydrogenase

4.3.4.1 Expression of GthFDH in *E. coli* (pET vector)

The pET vector containing the *Gthfdha* gene was confirmed transformed into *E. coli* JM109 cells by growth on kanamycin. The presence of the insert itself was confirmed by PCR analysis.

The construct pET28a Histhr:Gthfdh• was successfully transformed into *E. coli* expression strains Rosetta™, ArcticExpress and KRX.

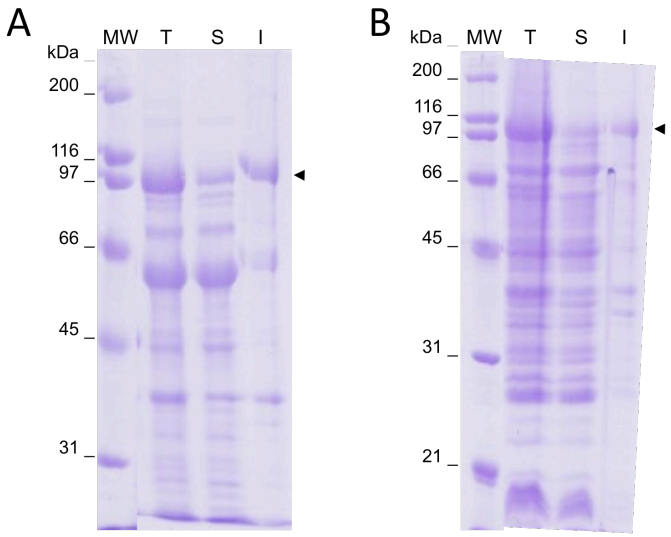


Figure 4.28. SDS-PAGE analysis of expression trials of pET28a Histhr:GthFDH in *E. coli* Rosetta™ (A) and KRX (B). T: Total cell extract; S: Soluble fraction; I: Insoluble fraction; MW: Molecular weight markers. Coomassie blue staining.

The *E. coli* expression strain KRX yielded more cell paste in expression trials than the other strains tested. However, no expression of a soluble recombinant formate dehydrogenase was observed in the range of growth conditions and induction patterns tested as portrayed in the SDS-PAGE analysis of the fractionated cell

extracts, with the over-expressed protein running at a M_r of *ca*. 105 kDa (Figure 4.28). The expected M_r of the formate dehydrogenase is 114 kDa.

The protein being His-tagged, the soluble fraction was loaded onto a nickel-charged metal affinity column. The fractions eluted in increasing imidazole concentrations did not show any purification of a recombinant FDH as no bands could be observed on SDS-PAGE.

The recombinant FDH that seemed to be expressed in an insoluble form was subjected to re-solubilisation in urea (8 M). There again, following loading onto a metal-affinity column, no nickel-purified protein could be observed when analysing the eluted fraction by SDS-PAGE.

4.3.4.2 Assaying for formate dehydrogenase activity

The nickel-purified fractions of the soluble cell lysate of *E. coli* KRX pET28a Histhr:GthFDH were assayed for formate dehydrogenase activity. In phosphate buffer, Glycine/KOH buffer or HEPES buffer, no formate dehydrogenase activity was detected. The assay conditions tested covered pH 7.0 to 9.5, and temperatures from 37 to 60 °C.

The *E. coli* cell paste was also tested for formate dehydrogenase activity. However, no formate dehydrogenase activity was detected.

The integrity of the assay was confirmed by assaying the commercial formate dehydrogenase from *Candida boidinii* (Sigma).

4.3.5 Intein-tagged constructs (pTYB11)

The *Bmmdh*, *pBMmdh*, *Bmact*, *Bmfaldh* and *Gthfdh* genes were successfully cloned into the pTYB11 vector as SapI/XhoI inserts, resulting in constructs pTYB11 *Bmmdh*, pTYB11 *pBMmdh*, pTYB11 *Bmact*, pTYB11 *Bmfalddh* and pTYB11 *GthfdhA*.

The constructs were transformed into the *E. coli* KRX expression strain, and recombinant protein expression was induced by the addition of 0.1% rhamnose and 1 mM IPTG.

The SceVMA1 intein-CDB tag protein adds 56 kDa to the molecular weight of the protein. Therefore, the M_r of the over-expressed intein-fusion recombinant proteins should be of *ca*. 95 kDa for BmMDH, *ca*. 83 kDa for BmACT, *ca*. 95 kDa for BmFALDDH fusion and *ca*. 170 kDa for GthFDHa.

SDS-PAGE analysis of the cell extracts (Figure 4.29-A) clearly showed M_r values of the over-expressed proteins that were greater than those of the His-tagged versions (see this Chapter - 4.3.2.1 (Fig. 4.10), 4.3.3.1 (Fig. 4.26), 4.3.4.1 (Fig. 4.28)). The four intein-fusion proteins of BmMDH, BmACT, BmFALDDH and GthFDHa had calculated M_r values of *ca*. 90 kDa, 77 kDa, 83 kDa and 160 kDa, respectively, all of which are slightly lower than the M_r expected. This is especially noticeable for BmFALDDH which migrated to an M_r in-between that of the intein fusion of BmMDH and BmACT.

The over-expressed fusion proteins also appeared to be found mainly concentrated in the insoluble fraction of their respective *E. coli* cell extracts. While some of the BmMDH and most of the BmACT intein fusion proteins appeared to have been expressed in a soluble form, no such observation could be made for the intein fusions of BmFALDDH andGthFDHa. The over-expression of pBMMDH as an intein fusion under the same conditions produced a very similar profile (Figure 4.29-B).

Some soluble expression of BmMDH and BmACT may have been occurring, and so the clarified lysates of these were loaded onto a chitin column. DTT-induced cleavage of the fusion partners did not seem to occur after overnight incubation at 4 °C, and the eluate fractions did not show the expected specific banding pattern corresponding to the release of intein (56 kDa) from the BmMDH (47 kDa) or BmACT (27 kDa) respectively. No cleavage seemed to occur and most of the soluble fraction loaded appear to have been retained on the column.

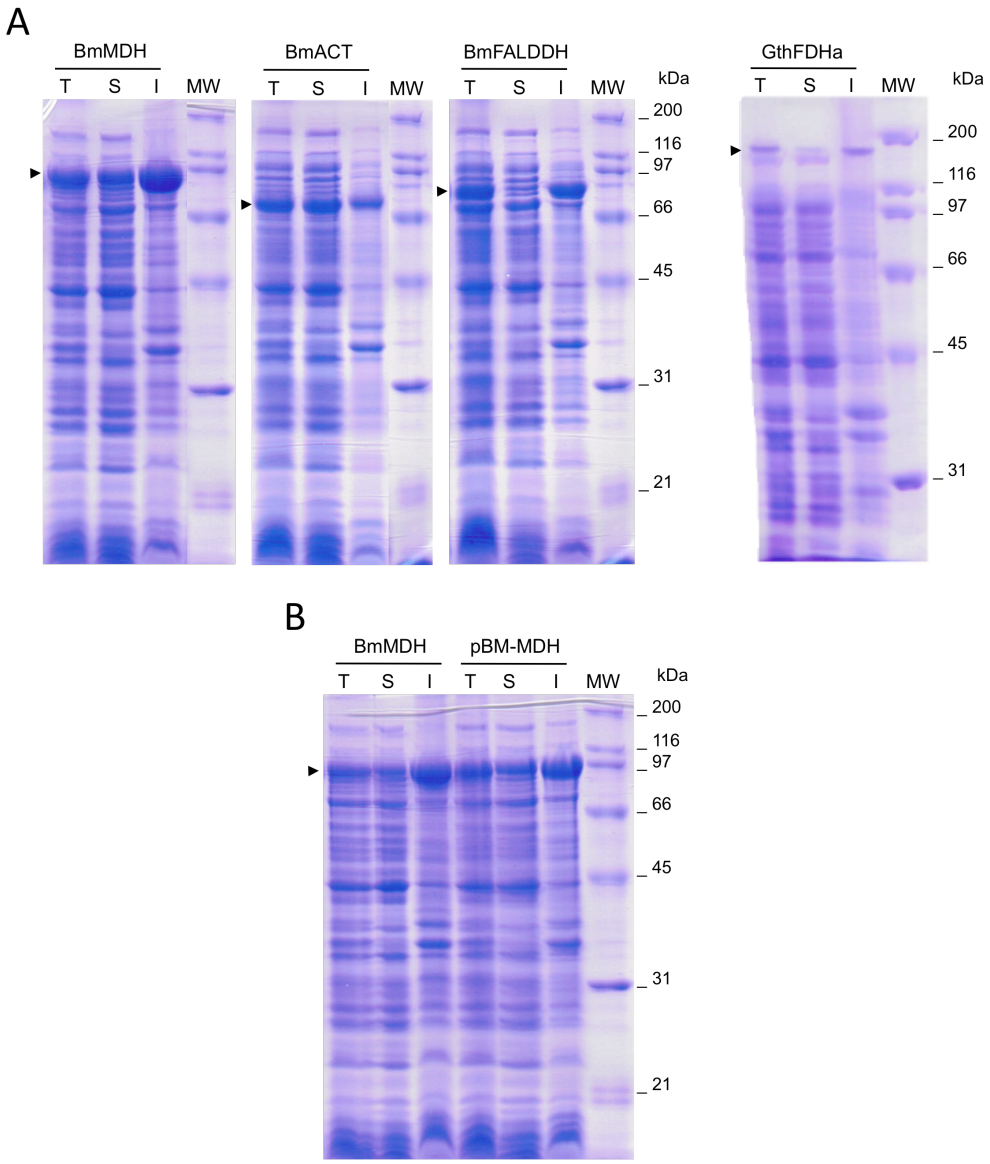


Figure 4.29. SDS-PAGE analysis of expression trials in *E. coli* KRX of (A) the intein-tagged BmMDH, BmACT, BmFALDDH and GthFDHa (pTYB11 constructs); (B) comparison of the expression of the intein-tagged BmMDH and pBMMDH. T: Total cell extract; S: Soluble fraction; I: Insoluble fraction; MW: Molecular weight markers (kDa). Coomassie blue staining.

4.4 Discussion

Heterologous expression in *E. coli* of the *Bacillus methanolicus* dehydrogenase enzymes involved in methanol dissimilation, with the exception of the formate dehydrogenase being from a *Geoacillus* background, was achieved but the recombinant enzymes did not behave as well as expected.

The cytoplasmic NAD-dependent methanol dehydrogenase in *B. methanolicus* had previously been expressed recombinantly in *E. coli*. A 2.5 kb fragment from *Sau*3AI-digested *B. methanolicus* C1 DNA was cloned into the vector pBS+ and the purified recombinant protein displayed a specific activity much lower than that of the native enzyme in *B. methanolicus* (3.5 $U.mg^{-1}$ and 19.6 $U.mg^{-1}$ respectively) (de Vries *et al*., 1992, Arfman *et al*., 1989). However, in this case, the *mdh* gene was likely to be under the control of its own promoter.

In this work, both the methanol dehydrogenase *mdh* and the activator *act* genes from *B. methanolicus* PB1 were cloned in the vector pET28a in an attempt to achieve soluble expression of active enzymes. Both MDH and ACT were expressed with C-terminal histidine tags as soluble proteins with observed M_r of 49 and 27 kDa, close to the reported M_r values of 43 and 27 kDa respectively (Arfman *et al.*, 1989, Arfman *et al*., 1991). The recombinant nickel-purified His-tagged MDH and ACT were also shown to have adopted their native conformations by gel filtration with M_r values of 410 and 48 kDa, close to the previously reported 430 and 50 kDa where the subunits arrangement is likely to have adopted the pairing of two sandwich pentagonal rings (Vonck *et al*., 1991, Arfman *et al*., 1991). However, no activity could be detected with the purified recombinant MDH.

In *B. methanolicus*, the affinities of MDH for its substrates are strongly enhanced by the activator, given that the requirement of the enzyme for NAD^+ and Mg^{2+} ions is fulfilled (Vonck *et al.*, 1991, Arfman *et al.*, 1991 and Kloosterman *et al.*, 2002). The involvement of magnesium ions is a peculiar feature of the MDH in *B. methanolicus* since all other family III NAD-dependent alcohol dehydrogenase characterised thus far require zinc or iron (Conway and Ingram, 1989; Drewke and Ciriacy, 1988). Furthermore, the MDH in *B. methanolicus* is the first example of a naturally-

occuring NAD-dependent ADH containing a bound co-factor NAD(H) (Arfman *et al.*, 1997), with it being vital for the activity of the enzyme. Arfman *et al.* (1997) performed a spectrophotometric scan on the purified MDH to demonstrate the presence of this tightly but not covalently bound NAD(H) in the active site of the enzyme subunit. A shoulder in the 300-350 nm region indicated the presence of a chromophore that disappeared with the addition of formaldehyde and which was restored with the addition of methanol. In taking in methanol, the MDH-bound cofactor is reduced and then oxidized with the release of formaldehyde. With no such shoulder peak observed in the current work, there is no oxidized co-factor bound to the enzyme that is accessible to DHA, while spiking the reaction with NADH restored the absorbance peak.

The activator facilitates the re-oxidation of the bound NADH cofactor. While general activity of the enzymatic complex as a whole necessitates exogenous NAD^+ co-enzyme, it is not interchanged with the bound co-factor moiety during catalysis. The bound NAD^+ acts as an electron acceptor and the exogenous NAD re-oxidises the NADH cofactor. With ACT facilitating the re-oxidation of the bound cofactor, ACT is indirectly changing the kinetics of MDH.

MDH recombinantly expressed as an N-terminal fusion to a Histidine tag in the *E. coli* KRX displayed activity but only in the formaldehyde reductase direction. Kinetics parameters were calculated (V_{max} 2.5 $U.mg^{-1}$, K_M 1.22 mM) and were comparable to that reported by Arfman (V_{max} 1.22 $U.mg^{-1}$, K_M 2 mM). ACT only affects the forward reaction. The temperature optimum for the enzyme was estimated to be 85-90 °C. This is perhaps surprising given that the growth optimum temperature for *B. methanolicus* is 45-55 °C and that its growth temperature range lies between 35 and 65 °C. While it is not uncommon for enzymes to have temperature optima for activity above their growth temperature range (Lee *et al.*, 2007), it is counter-balanced by the thermal stability of the enzyme, as in this case the half-life at 50 °C of 8 h for the recombinant MDH declined sharply past 60 °C. It could be that the presence of substrate (in the temperature optima assays) has a protective effect on the enzyme. However, while the enzyme assays were performed over 1 min, the temperature inactivation of the enzyme was followed over

longer periods of time. Thermostabilities are to be strictly compared over the same time-scale as the assay.

The recombinant expression of the formaldehyde and formate dehydrogenases in *E. coli* resulted in insoluble proteins and despite efforts to return the proteins to solution, re-folding proved unsuccessful and no activity could be recorded. It has been shown that culturing expressing *E. coli* strains at elevated or downshifted temperatures could help in the production of soluble and correctly-folded proteins (Koma *et al*., 2006, Trimpin and Brizzard, 2009), an approach also employed by large fusion protein partners (Evans *et al*., 1999). It was with that knowledge that the pTYB11 constructs were generated, with the self-cleavable 56 kDa intein fusion. Although being recombinantly expressed at the right size - a 56 kDa upward shift in M_r-, all MDH, ACT, FALDDH and FDH happened to express as insoluble inclusion bodies.

A codon quality analysis was run on the nucleotide sequences of the three dehydrogenase genes. Codon usage is biased depending on the organism and genes of a *Bacillus* origin may be less effectively transcribed in an *E. coli* environment (Sharp *et al*., 1988; Mozner *et al*., 1999). Codon quality distribution plots showed that the genes used rare codons for *E. coli* with a high frequency. Later recombinant expression of the dehydrogenases was performed in the *E. coli* strain KRX which, amongst other features, carries the plasmid pRARE2 encoding for rare tRNAs, supplementing *E. coli*'s own distribution. Yet, the codon adaptation index (CAI), a simple and effective measure of the codon bias, was calculated for each of the dehydrogenases and they scored at 0.65 (*Bmmdh*), 0.64 (*Bmact*), 0.52 (*Bmfalddh*) and 0.58 (*GthfdhA*). The particularly low scores of the formaldehyde and formate dehydrogenase genes may in part explain their lower over-expression in *E. coli* compared to that of the methanol dehydrogenase and its activator. Codon-optimised versions of the gene sequences were simulated (CAI 0.97-0.98).

An alternative to heterologous recombinant expression is to express the protein of interest in a homologous (or semi-homologous) host. While, as discussed earlier, the advantages of heterologous expression and particularly in *E. coli* are numerous, specific requirements or properties of the gene and protein of interest can have a decisive influence on the expression system of choice. Homologous expression systems have been developed for a number of hosts and conditions. Halophilic or (hyper)thermophile proteins benefit from being expressed in alternative homologous systems, such as in the case of *Sulfolobus solfataricus* (e.g.) (Albers *et al*., 2006; She *et al*., 2009). Expression of the three dehydrogenase in a (genus-)homologous system will be described and discussed in the next chapter.

Chapter 5- Homologous Expression (*B. subtilis*)

5.1 Introduction

High-level production of recombinant proteins is a pre-requisite for their subsequent purification and characterisation. As described in the previous chapter, most expression systems for recombinant proteins are based on *Escherichia coli* as a host strain because, in part, of the availability of a wide variety of vector plasmids (e.g. pUC- and pET-vectors series). However, problems can arise during heterologous gene expression and purification of the recombinant protein: low expression rates, improper folding, formation of inclusion bodies and problems of toxicity of the recombinant gene product in the heterologous host (Baneyx, 1999). These can necessitate a change from a heterologous to a homologous host for expression of recombinant gene proteins.

Gram-positive bacteria are well known for their contributions to the agricultural, food biotechnology and pharmaceutical industries (Harwood, 1992; Schallmey *et al*., 2004; Westers *et al*., 2004). Amongst these is *Bacillus subtilis*, a non-pathenogenic GRAS (Generally Regarded As Safe) microorganism. *B. subtilis* is also the best-studied Gram-positive bacterium, benefiting from a large body of literature regarding its transcription, translation, protein folding and secretion machinery, genetic manipulation and large-scale fermentation. It also has a rapid growth rate.

B. subtilis has the ability to become naturally genetically competent. Natural competence is the ability of bacteria to take up exogenous DNA and incorporate it into the genome. This has opened the organism to genetic manipulation, allowing cloning, mutant generation and gene mapping in *B. subtilis*. (Anagnostopoulos and Spizizen, 1961; Kunst and Rapoport, 1995). The competence of *B. subtilis* to accept foreign DNA and factors influencing it has been studied extensively (Young, 1967; Bott and Wilson, 1968, Ayad and Barker, 1969). Generally, competence develops in

B. subtilis as a global response to the onset of stationary phase, accompanied by the expression of late gene products that are required for the binding, processing and uptake of transforming DNA such as regulatory genes involved in the process, including transcription factors and regulatory information (Dubnau, 1991a and 1991b). However, natural isolates of *B. subtilis* are difficult to transform due to a much-reduced natural level of competence. The low or unnatural competence shown by several strains has led to the development of other strategies involving phage transduction (Yasbin and Young, 1974), protoplast fusion (Chang and Cohen, 1979) and the versatile electroporation method (Xue *et al*., 1999; Ito and Nagane, 2001). Overall, the different transformation efficiencies range from 10^3-10^7 transformants per µg DNA (Hauser and Karamata, 1994; Martinez *et al*., 1999; Xue *et al*., 1999; Romero *et al*., 2006; Duitman *et al*., 2007). Wittke *et al*. (2002) even developed a transformation under non-selective conditions, albeit at low frequency, in chocolate-milk using plasmid pMG36enpr as a model DNA.

Most convenient vector plasmids for *B. subtilis* expression have been derived from natural plasmids detected in *Staphylococcus aureus* such as pUB110 (Gryczan *et al*., 1978; Keggins *et al*., 1978) and pC194 (Ehrlich, 1977; Gryczan *et al*., 1978; Gryczan and Dubnau, 1978). The pC194 vector has been used in the electroporation of *B. cereus*, with the use of wall-weakening agents (Turgeon *et al*., 2006), of *B. polyxyma* in the presence of 0.25 M sucrose (Mallonee and Speckman, 1989), and the transformation of vegetative cells of *B. thuringiensis* in the presence of 30% sucrose (Heierson *et al*., 1987). Protoplast transformation has also been developed for the transformation of undomesticated strains of *B. subtilis* (Romero *et al*., 2006). Of interest here, the thermostable alcohol dehydrogenase (ADH-T) gene (*adhT*) from *B. stearothermophilus* was cloned in the vector plasmid pTB524 and transformed in *B. subtilis* according to the protocol of Imanaka *et al*. (1981) (Sakoda and Imanaka, 1992).

Other Gram-positive expression systems have also been developed in *Lactococcus lactis* (de Ruyter *et al*., 1996) and *Lactobacillus delbrueckii* subsp. *lactis* and subsp. *bulgaricus* (Maassen, 1999; Serror *et al*., 2002). Shuttle vectors for ectopic insertion of transgenes into the *B. subtilis* chromosome have also been developed (Middleton and Hofmeister, 2004).

The increase in the development of expression systems in particular progressed in parallel with inducible gene-expression systems. The *E. coli lac* repressor and operator, equipped with appropriate promoter and RBS, were first used in conjunction with the inducible P_{spac} promoter (Yansura and Henner, 1984). Many other systems were implemented in *B. subtilis*. The T7 system has been adapted, inserting the T7 RNA polmerase *rpo*T7 into the *B. subtilis* chromosome under control of a xylose-inducible promoter, and with a plasmid containing the T7 promoter driving expression of the target gene (Conrad *et al*., 1996; de Vos *et al*., 1997).

While the pUB110 and other vectors mentioned earlier replicate in *B. subtilis*, addition of recombinant DNA often confers structural and segregational instability. In *B. subtilis* these plasmids replicate as rolling circles producing single-stranded DNA as an intermediate, and short direct repeats within this ssDNA may lead to the deletion of one of the two repeats and the intervening DNA.

Versatile expression vectors for *Bacillus subtilis* have been created, such as pBSMuL (Brockmeier *et al*., 2006), improved versions of pUB110 (Wu and Wong, 1999) and the vector pHT01 used in this study (Nguyen *et al*., 2007). The *B. subtilis-E. coli* shuttle vector, pHT01, is based on the *E. coli-B. subtilis* shuttle vector pMTLBS72, carrying the replication region and *bla* gene (for ampicillin resistance) of pBR322 (*E. coli*) and the replication region and chloramphenicol resistance gene of pBS72 (*B. subtilis*) (Titok *et al*., 2003). Efficient termination is provided by the insertion of the *trpA* transcriptional terminator sequence immediately downstream of the recombinant gene (at the *Aat*II site) (Nguyen *et al*., 2005). IPTG-inducible expression of the recombinant gene is driven by the strong P_{groES} promoter preceding the *groE* operon of *B. subtilis* and the *lac* operator of *E. coli* (Phan *et al*., 2006). *Bacillus subtilis* can be used as an alternative protein factory to *Escherichia coli*.

Cue *et al*. (1996, 1997) characterised a restriction modification system and achieved genetic transformation of *Bacillus methanolicus* protoplasts with the construction of the plasmid pDQ503, an integrative shuttle vector based on the pHP13 plasmid (Haima *et al*., 1987). The pTB1.9 plasmid, a minimal pBM19 containing its replicon *repB* and replication origin, has been transformed into *B. methanolicus* but with very poor efficiency (Brautaset *et al*., 2004). Hypothesising that increasing the number of

copies of the *B. methanolicus* chromosomal genes hexulose phosphate synthase *hps* and phosphohexuloisomerase *phi*, needed in the initial phase of fixation of formaldehyde into the RuMP cycle, would elevate the activities of HPS and PHI in the cells and lead to a higher formaldehyde assimilation rate synonymous with a higher tolerance for methanol, these two genes were cloned into the shuttle vector pHP13 and the plasmid electroporated into *B. methanolicus* competent cells (Jakobsen *et al*., 2006). Yet, transformation efficiencies were poor (Brautaset, personal communication).

5.2 Materials and Methods

The pHT01 vector was purchased from Mobitec Gmbh (Göttingen, Germany). *Bacillus subtilis* 168 (*trp*$^-$ or *ind*$^-$) (DSM402) was obtained from DSMZ (http://www.dsmz.de/index.htm).

DNA manipulations are described in Chapter 2- General Materials & Methods. PCR amplification of the *Bmmdh*, *pBMmdh*, *Bmfalddh* and *GthfdhA* genes was performed as previously described (2.2.2), with restriction sites added at the 5' ends of the forward and reverse primers to allow complementary cloning in pHT01.

5.2.1 Cultivating *B. subtilis*

Bacillus subtilis 168 was routinely grown and maintained on TS medium (broth or agar plates) at 28 °C, with shaking (200 rpm) for the broth cultures.

5.2.2 Cloning into pHT01

The *E. coli-B. subtilis* shuttle vector pHT01 (Figure 5.1) is based on the shuttle vector pMTLBS72 (Nguyen *et al*., 2005). It encodes ampicillin (carbenicillin)

resistance (Ap^r) in *E. coli*, and chloramphenicol resistance (Cm^r) in a *Bacillus* host. The transgene is driven by the σ^A-dependent promoter preceding the *groE* operon in *B. subtilis* with the addition of the *lac* operator (*lacO*) for IPTG-induction.

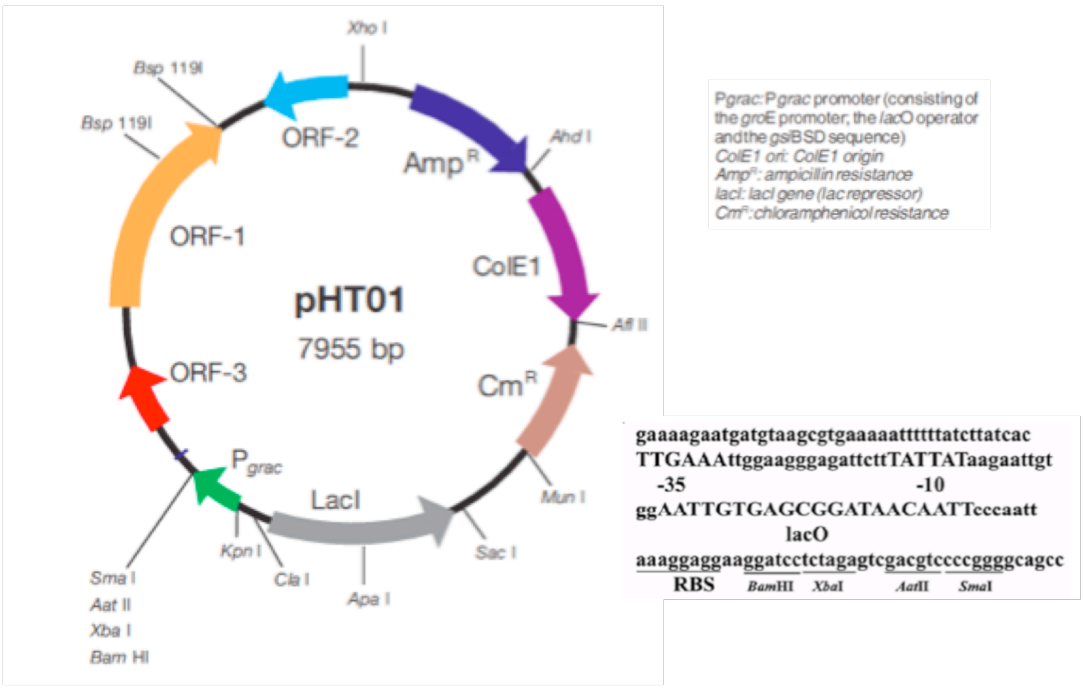


Figure 5.1. Vector map of pHT01. (Adapted from Mobitec).

5.2.2.1 Cloning strategy

Genes were cloned in pHT01 as *Bam*HI/*Xma*I inserts since these restriction-site recognition sequences do not exist in any of the four dehydrogenase genes and can therefore be added at the extremities. The genes were first PCR-amplified using gene-specific primer pairs with a *Bam*HI recognition sequence added at the 5' end of the forward primers and an *Xma*I recognition sequence at the 5' end of the reverse primers (Table 5.1). The stop codon was included in the reverse primer sequences. PCR was performed as previously described, with Phusion™ as the DNA polymerase. PCR products were separated on a 1% agarose gel, and the appropriate size fragment gel-extracted, A-tailed, cloned into pGEM®-T Easy and their sequence confirmed (with primers pT7_F and pSP6_F). Correct inserts were *Bam*HI/*Xma*I-excised and purified, and ligated with T4 DNA ligase into a *Xma*I/*Bam*HI-digested, alkaline phosphatase-treated pHT01 vector. The ligation reaction was then transformed into *E. coli* JM109 cells. Plasmids displaying the expected pattern for

the insert in PCR or restriction digest screen were further sequenced (with primers pHTseq_F and pHTseq_R and/or a gene-specific primer) before being transformed into *B. subtilis* 168 cells.

Table 5.1. Primers sequences for Bmdh, pBMmdh, Bmact, Bmfalddh and GthfdhA. Restriction sites added at the 5' ends are underlined. Translational start and stop signals are in bold.

Primer name	Sequence	Tm (°C)
BamHI_Bmmdh_F	CGCGGATCC**ATG**ACAAACTTTTTCATTCC	71.3
XmaI_Bmmdh_•R	CCCCCCGGG**TTA**CAGAGCGTTTTTGATG	70.5
BamHI_pBMmdh_F	CGCGGATCC**ATG**ACAACAAACTTTTTCATTCC	72.1
XmaI_pBMmdh_•R	CCCCCCGGG**TTA**CATAGCGTTTTTGATG	73.3
BamHI_Bmact_F	CGCGGATCC**ATG**GGAAAATTATTGAGC	71.4
XmaI_Bmact_•R	CCCCCCGGG**TCA**TTTATGTTTGAGAGC	72.3
BamHI_Bmfalddh_F	CGCGGATCC**ATG**AAAGCTGTAACATATCAAGGACC	75.0
XmaI_Bmfalddh_•R	CCCCCCGGG**TCA**CGGTTTAAAAACGACTTTGATGC	78.3
BamHI_Gthfdh_F	CGCGGATCC**ATG**CTAAAAAACTACTCCACC	71.2
XmaI_Gthfdh_•R	CCCCCCGGG**TTA**GCCACGCGGCTCTGCCTCC	74.8
pHTseq_F	CTTATCACTTGAAATTGGAAGGGAG	64.6
pHTseq_R	CAGTTGCAGACAAAGATCTCATGG	66.8

5.2.2.2 Constructing pHT:His and pHT:Intein

Since the vector pHT01 does not carry a fusion tag, two mutated versions of the vector were produced, echoing the pET28a (N-term His tag) and pTYB11 (intein tag) vectors.

pHT:His

The pHT:His vector was produced by site-directed mutagenesis. Two homologous primers, pHTHisthr_sdmF and pHTHisthr_sdmR were synthesised. These consisted of the 22 bp of pHT01 sequence (underlined) up to the *Bam*HI site (not included), the His-tag and thrombin cleavage site as of pET28a, and 21 bp of pHT sequence from the *Bam*HI site (underlined) (*Bam*HI in bold).

pHTHisthr_sdmF

5' CAATTCCCAATTAAAGGAGGAACATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGC**GGATCC**CGTCTAGAGTCGACG

pHTHisthr_sdmR

5' CGTCGACTCTAGACG**GGATCC**GCTGCCGCGCGGCACCAGGCCGCTGCTGTGATGATGATGATGATGTTCCTCCTTTAATTGGGAATTG

A PCR was set up (as previously described in Chapter 2) using pHT01 as template, the primers pHTHisthr_sdmF and sdmR, and Phusion™ as the DNA polymerase. The cycling parameters were altered to the following: 98 °C for 1 min, 18 cycles [98 °C 15 s, 60 °C 15 s, 68 °C 5 min], followed by a final step at 68 °C for 10min. The DNA was then digested with 10 U *Dpn*I for 1 h at 37 °C. *Dpn*I digests the parental methylated and hemimethylated supercoiled dsDNA. JM109 *E. coli* were transformed with 2 µL of *Dpn*I-treated DNA as previously described.

Transformants were routinely PCR-screened using primers pHTseq_F and pHTseq_R (see Table 5.1 above).

pHT:intein

The intein tag from the pTYB11 vector was PCR-amplified using primers 5'SceIntein_BamHI F *vs*. 3'SceIntein_XbaI R, which includes a *Sap*I restriction site.

5'SceIntein_BamHI F 5' CGGGATCCTGCTTGCCAAGGGTACCAATG

3'SceIntein_XbaI R 5' CGTCTAGAGCTCTTCTGTTCTGTACAACAACCTGG

The 1552 bp fragment was digested with *Bam*HI/*Xba*I, gel purified, and ligated into the *Bam*HI/*Xba*I-digested, alkaline phosphatase-treated pHT01. Ligation products were transformed into *E. coli* JM109 cells. Transformants were PCR-screened using primers pHTseq_F *vs*. SceIntein_R and SceIntein_F *vs*. pHTseq_R.

Cloning of the intein fusion partner was as a *Sap*I/*Xma*I insert, as previously described in Chapter 4 (4.2.2).

5.2.3 *Bacillus* transformation

Fresh cultures of *B. subtilis* were prepared for each transformation, following a protocol adapted from Bolhius *et al*. (1999). *Bacillus subtilis* was grown in GCHE medium (Kunst and Rapoport, 1995) at 28 °C with shaking to an $OD_{600} \approx 1.0$. Then 1-2 µg DNA was added, and growth was continued for 4 h, during which time the cells will express the antibiotic resistance gene. The cells were then plated onto TS agar plates supplemented with chloramphenicol (5 $\mu g.mL^{-1}$).

A two-step protocol was also employed whereby *B. subtilis* was first grown in GCHE medium as above. At the transition between the exponential and stationary phases (t_0), the culture was diluted in an equal volume of GE medium (GCHE without casein hydrolysate) and incubated for a further 1 h. DNA was added as described above.

5.2.4 Expression in *B. subtilis*

A fresh culture of a *B. subtilis* transformant was grown on TS supplemented with chloramphenicol (5 $\mu g.mL^{-1}$) at 28 °C, with shaking, for 4 h or until an OD_{600} of approximately 0.6 was achieved. IPTG (1 mM) was added and the culture further incubated overnight.

The cultures were harvested by centrifugation at 3,800*g*, the supernatant discarded and the pellet resuspended in one-tenth volume of phosphate buffer (50 mM HEPES, pH 8.0). Cells were lysed (see Chapter 2 - 2.2.5.2) and the cell debris and insoluble fraction pelleted (13,000*g*, 20 min). The supernatant (soluble fraction) was transferred to a fresh tube and the pellet (insoluble fraction) was resuspended in the same volume of 8 M urea.

Samples of each fraction (total, soluble and insoluble) were analysed on SDS-PAGE.

5.2.5 Enzyme assays

Enzyme assays were performed as previously described (see Chapter 2 - 2.2.6).

5.3 Results

5.3.1 Amplification and cloning of the dehydrogenase genes in the *E. coli-B. subtilis* shuttle vector pHT01

With the intention of cloning into pHT01 as *Bam*HI/*Xma*I inserts, each of the three dehydrogenase genes along with the activator were first PCR-amplified from their respective DNA sources with the addition of a *Bam*HI site at the 5' end and an *Xma*I site at the 3' end.

Bmmdh, *Bmact*, and *Bmfalddh* were PCR-amplified from *B. methanolicus* PB1 genomic DNA as *Bam*HI/*Xma*I fragments. *pBMmdh* was PCR-amplified from the pBM19 plasmid from *B. methanolicus* PB1, while *GthfdhA* was PCR-amplified from *G. thermoglucosidasius* genomic DNA, both as a BamHI/XmaI fragment. The PCR products of 1,162 bp, 574 bp, 1,150 bp, 1,165 bp and 3,067 bp respectively were gel-purified, cloned into pGEM®-T easy, transformed into *E. coli* JM109, and white colonies screened for insert-containing pHT plasmids. Inserts were then *Bam*HI/*Xma*I-excised from the pGEM®-T Easy, ligated into pHT01 and transformed into JM109 cells. Colonies were PCR-screened for the presence of the insert, confirming the successful production of constructs pHT01 *Bmmdh*•, pHT01 *Bmact*•, pHT01 *pBMmdh*•, pHT01 *Bmfalddh*• and pHT01 *GthfdhA*•.

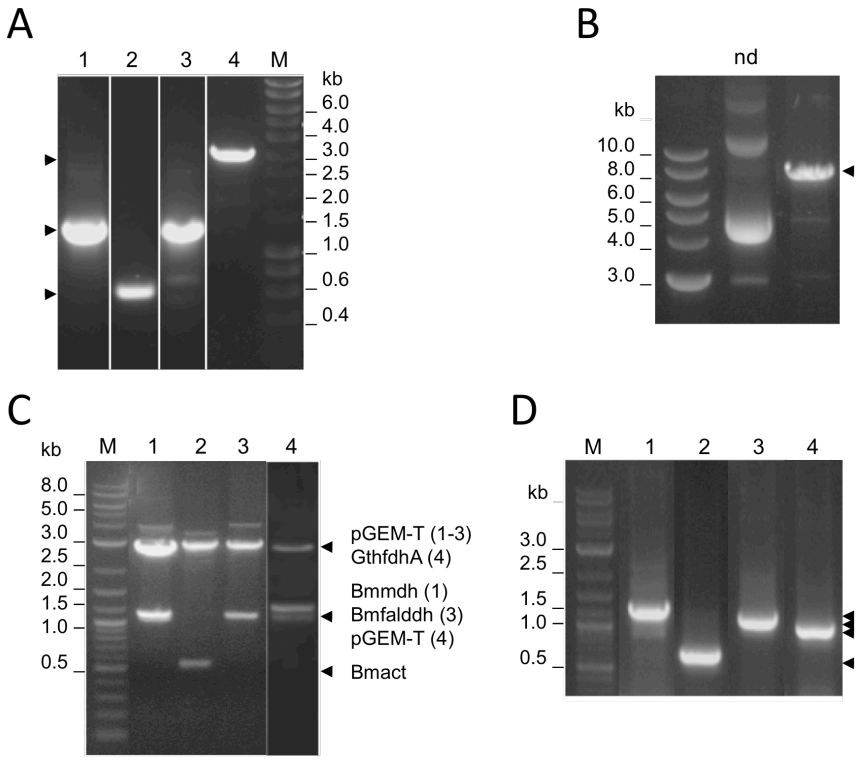


Figure 5.2. pHT01 *Bmmdh*, pHT01 *Bmact*, pHT01 *Bmfalddh* and pHT01 *GthfdhA*. 1% agarose gel electrophoresis analysis. (A) BamHI_F *vs*. XmaI_R PCR of *Bmmdh* (1,162 bp), *Bmact* (574 bp), *Bmfalddh* (1,150 bp), *GthfdhA* (3 kb). (B) *Bam*HI/*Xma*I restriction of pHT01 (7.9 kb) (nd: non digested). (C) pGEM®-T Easy restriction analysis and excision of the inserts (*Bam*HI/*Xma*I) (and *Bsa*I for pGEM®-T Easy *GthfdhA*). (D) PCR analysis of pHT01 constructs pHT01 *Bmmdh*, pHT01 *Bmact*, pHT01 *Bmfalddh* and pHT01 *GthfdhA*; primers were pHTseq_F *vs*. Bmmdh_R, Bmact_R, Bmfalddh_R and GthfdhA_4R respectively. Lane 1: Bmmdh; lane 2: Bmact; lane 3: Bmfalddh; lane 4: GthfdhA; M: 1 kb DNA ladder. Phusion™ was used as DNA polymerase.

5.3.2 pHT:Histhr and pHT:intein constructs

5.3.2.1 Making the pHT:Histhr and pHT:intein vectors, pHT01 derivatives

A 6x His-tag was successfully inserted in the pHT01 vector by site-directed mutagenesis with primers pHTHisthr_sdmF *vs*. pHTHisthr_sdmR, creating the vector pHT:Histhr, so that *Bam*HI/*Xma*I inserts could be cloned in and expressed as fusions to the N-terminal His-tag.

The intein tag from the vector pTYB11 (*Sce* VMA1 intein) was PCR-amplified as a *Bam*HI/*Xba*I insert, using primers 5'SceIntein_BamHI F vs. 3'Sceintein_XbaI R. The 1,552 bp fragment was digested with *Bam*HI and *Xba*I, gel-purifed, ligated into pHT01 and transformed into *E. coli* JM109 cells. Colonies were PCR-screened (pHTseq_F *vs*. SceIntein_R and SceIntein_F *vs*. pHTseq_R) and a specimen showing the correct size of insert was kept.

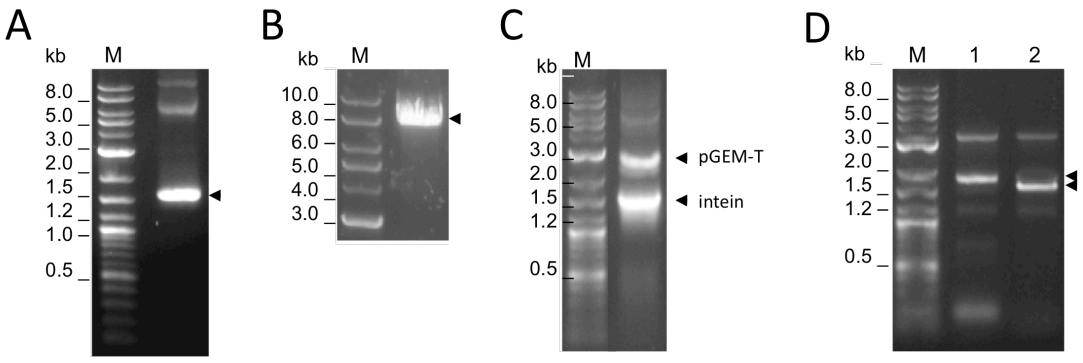


Figure 5.3. pHT:intein. 1% agarose gel electrophoresis analysis. (A) PCR of the *Sce* VMA1 intein gene (1,152 bp). (B) pHT01 BamHI/XbaI restriction (7.9 kb). (D) PCR screen of pHT:intein; primers were lane 1: pHTseq_F *vs*. SceIntein_R and lane 2: SceIntein_F *vs*. pHTseq_R. M: 1 kb DNA ladder. Phusion™ was used as DNA polymerase.

The new pHT:intein vector allows for the insertion of genes as *Sap*I/*Xma*I N-terminal fusions.

5.3.2.2 pHT:Histhr and pHT:intein contructs

pHT:Histhr constructs

All four genes of interest were PCR-amplified as *Bam*HI/*Xma*I fragments, cloned into pGEM®-T Easy and transformed into *E. coli* JM109 cells (see this chapter - 5.3.1). The correctly-excised inserts were ligated into *Xma*I/*Bam*HI pHT:Histhr (8 kb) and transformed into JM109 cells. Following sequence checks, the following plasmid constructs were successfully generated: pHT:Histhr:Bmmdh•, pHT:Histhr:Bmact•, pHT:Histhr:Bmfalddh• and pHT:Histhr:GthfdhA•.

pHT:Intein constructs

All four genes of interest (*Bmmdh*, *pBMmdh*, *Bmact*, *Bmfalddh* and *GthfdhA*) were PCR-amplified from their respective DNA sources as *Sap*I/*Xma*I fragments, gel-purified, cloned into pGEM®-T Easy and transformed into *E. coli* JM109 cells. White colonies were screened and inserts were excised by restriction digest with *Sap*I and *Xma*I. The gel-purified inserts were ligated into the *Xho*I/*Sap*I pHT:intein vector, and transformed in *E. coli* JM109 cells.

The following vectors were successfully generated: pHT:Intein:Bmmdh•, pHT:Intein:Bmact•, pHT:Intein:Bmfalddh• and pHT:Intein:GthfdhA•.

5.3.3 Homologous expression in *B. subtilis* (pHT01)

The four pHT01 constructs, pHT01 Bmmdh•, pHT01 Bmact•, pHT01 Bmfalddh• and pHT01 GthfdhA• were successfully transformed into *B. subtilis* 168, as confirmed by growth on chloramphenicol (5 $\mu g.mL^{-1}$) and PCR checks.

In protein expression trails, cultures were induced with IPTG (1 mM). Following overnight growth, the cell pellets were fractionated and analysed by SDS-PAGE (Figure 5.4). For comparison and as a control, a culture of *B. subtilis* transformed with the empty pHT01 vector, induced with IPTG, was also analysed.

The MDH and activator were expressed strongly and in a soluble form in *B. subtilis*. The M_r of the MDH and ACT subunits was calculated at 46 and 27 kDa respectively. The FALDDH did not appear to be expressed as no clear specific band could be observed at the expected M_r of 45 kDa. The FDH seemed to have been successfully expressed in *B. subtilis* with an extra band appearing in the total extract fraction at a calculated $M_r \approx 105$ kDa. However, this enzyme appeared to segregate mainly in the insoluble fraction.

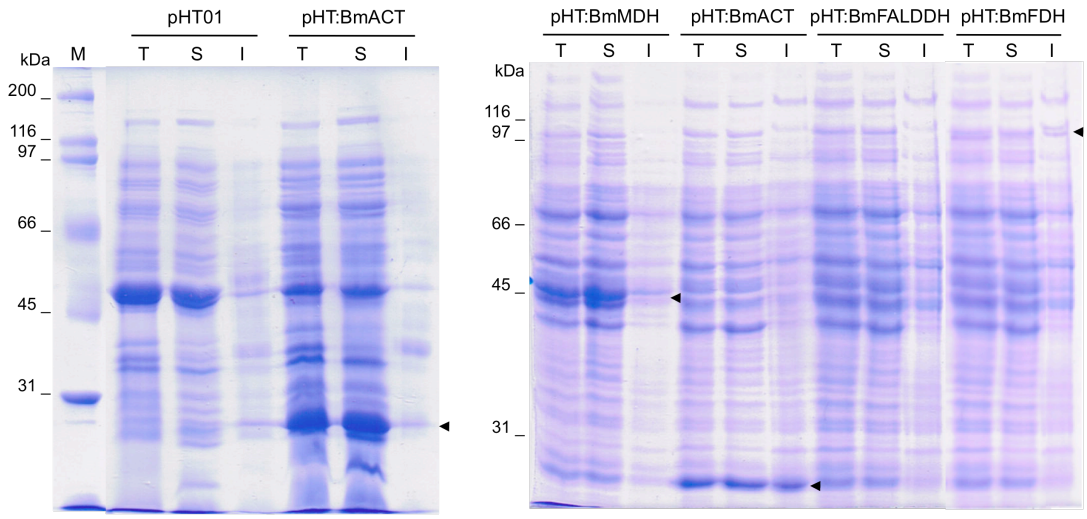


Figure 5.4. SDS-PAGE analysis of the cell lysates of *B. subtilis* pHT01 constructs (pHT01, pHT:BmMDH, pHT:BmACT, pHT:BmFALDDH and pHT:GthFDHa). T: Total extract; S: soluble lysate; I: insoluble fraction. MW: Molecular weight markers. Coomassie blue staining.

From the pHT:Histhr constructs, his-tagged variants of the MDH and ACT were also over-expressed in *B. subtilis*. The soluble lysates of the IPTG-induced *B. subtilis* pHT:Histhr:Bmmdh• and pHT:Histhr:Bmact• cultures were loaded onto nickel-charged columns. The recombinant His-tagged MDH and ACT were eluted separately from the rest of the soluble proteins, mainly in the 14 - 40% and 7 – 14% His-elute buffer fractions respectively (Figure 5.5).

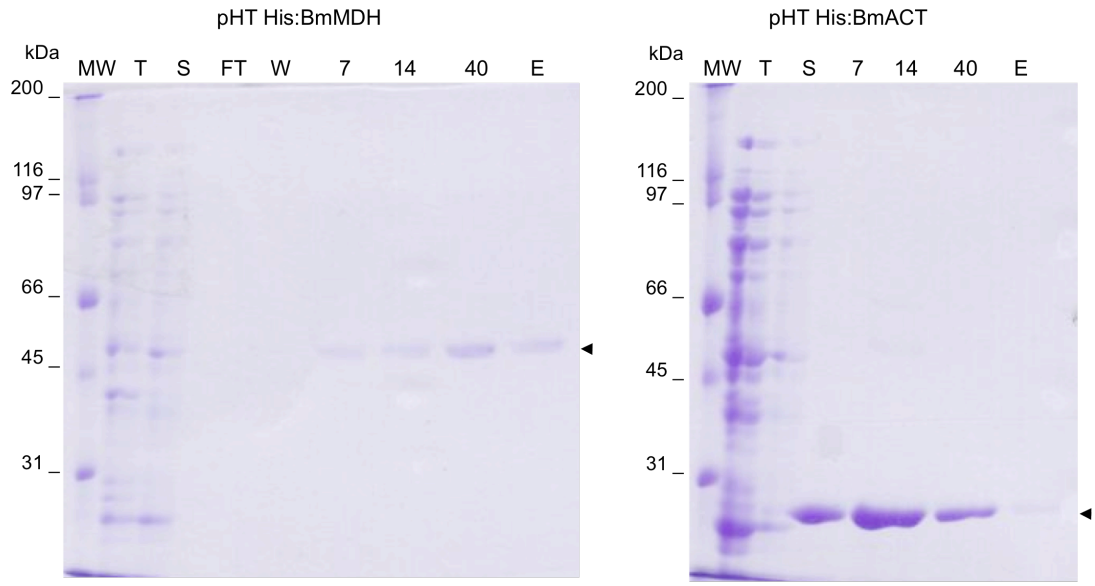


Figure 5.5. SDS-PAGE analysis of the Ni^{2+} purification proteins expressed in *B. subtilis* pHT01 His:BmMDH• and pHT01 His:BmACT•. T: Total extract; S: soluble lysate (loaded fraction); FT: flow-through; W: wash; 7: 7% His-elute; 14: 14% His-elute; 40: 40% His-elute; E: 100% His-elute; MW: molecular weight markers (kDa). Coomassie blue staining.

5.3.4. Kinetic parameters of *Bacillus*-expressed BmMDH (formaldehyde reductase)

First, the soluble lysate of fractionated IPTG-induced overnight cultures of *B. subtilis* pHT01 Bmmdh• was assayed for MDH activity. No enzymatic activity was detected in the assay, run in HEPES buffer (50 mM, pH 8.5, with 5 µM $MgSO_4$ and $ZnSO_4$) and 2 mM NAD^+, and started with the addition of 6 mM methanol. The addition of soluble lysate from a *B. subtilis* pHT01 Bmact• IPTG-induced culture did not increase MDH activity so that it could be not measured under these conditions.

As previously mentioned (see Chapter 4 - 4.3.2.7), activity of the MDH was also assayed in the reverse direction (formaldehyde reductase), with NADH (0.2 mM) in place of NAD^+ and the assay started with formaldehyde. The intrinsic *Bacillus* NADH oxidase activity was subtracted from the overall dA_{340}. Formaldehyde reductase activity was measured in the soluble lysate of *B. subtilis* pHT01 Bmmdh•

IPTG-induced cultures. The activity was proportional to the amount of enzyme (soluble cell extract) in the assay (Figure 5.6).

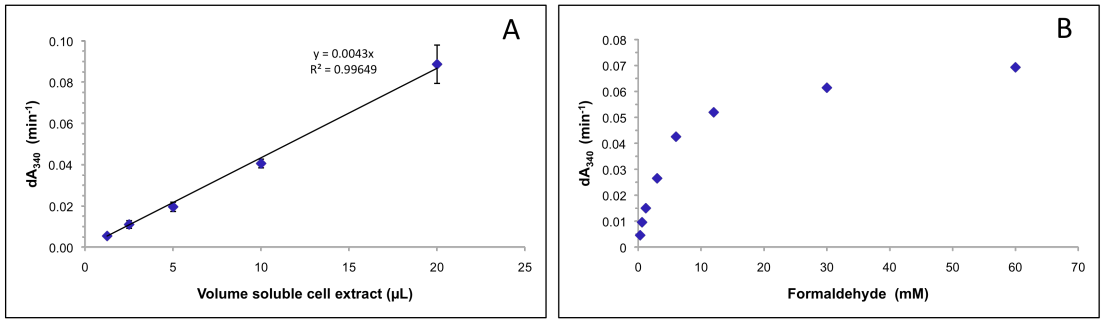


Figure 5.6. Formaldehyde reductase assay on *B. subtilis* pHT01 BmMDH cell extract (soluble lysate). Enzyme activity ($dA_{340}.min^{-1}$) as a function of (A) amount of enzyme (soluble fraction of cell extract) and (B) formaldehyde concentration.

The *B. subtilis*-expressed His-tagged BmMDH, was assayed for enzymatic activity. With methanol (substrate) and NAD^+ (co-factor), and in the presence of Mg^{2+} and Zn^{2+} ions, no MDH activity was detected. However, activity was measured in the reverse direction, with formaldehyde and NADH.

The proportionality of the activity to the amount of enzyme in the assay was confirmed. The effect of substrate (formaldehyde) concentration on the specific activity of the enzyme was studied (Figure 5.7).

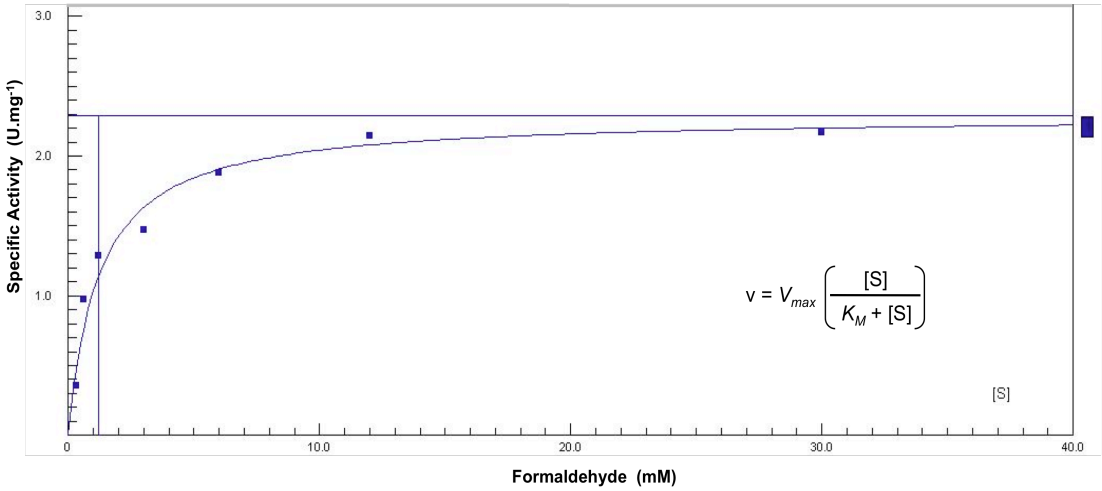


Figure 5.7. Dependence of velocity on substrate concentration of the *B. subtilis*-expressed His-tagged BmMDH against formaldehyde. Michaelis-Menten equation (inset).

Plotting the data according to the transformation of Hanes (Figure 5.8) indicated that the enzyme obeyed Michaelis-Menten kinetics. The kinetic parameters, V_{max} and K_M, were then determined by the Direct-Linear method of Eisenthal and Cornish-Bowden (1974).

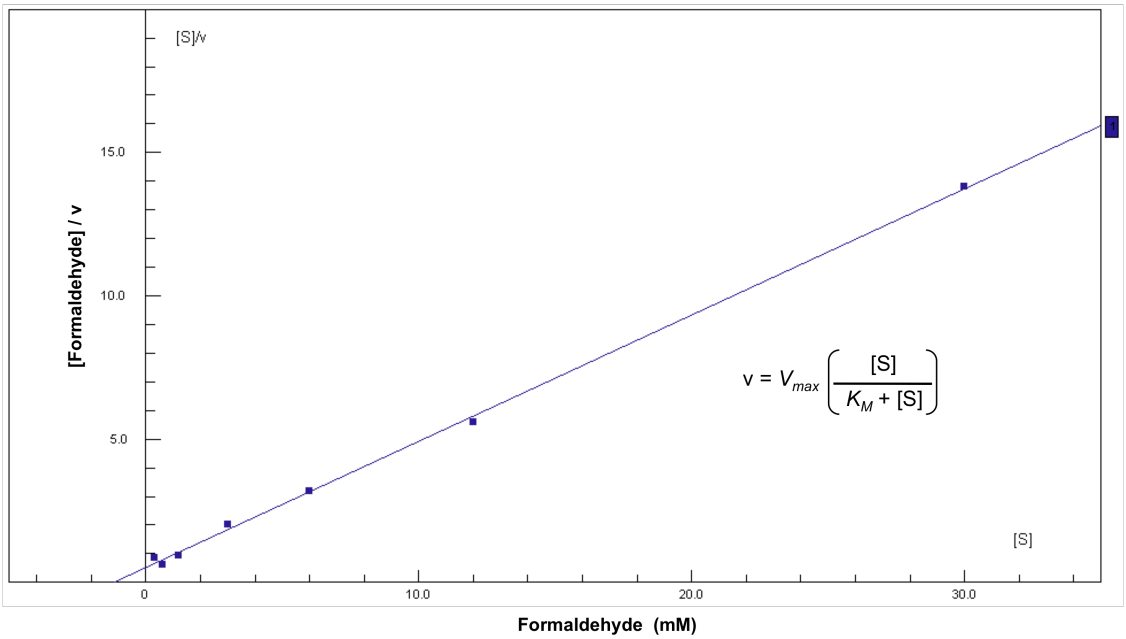


Figure 5.8. Hanes plot for the *B. subtilis*-expressed His-tagged BmMDH against formaldehyde. Hanes equation (inset) ([S] in mM, v in $U.mg^{-1}$).

In *B. subtilis*, the recombinantly-expressed His-tagged BmMDH is catalytically active against formaldehyde at 50 °C, with NADH as co-factor. The V_{max} of the enzyme was calculated as 2.27 (± 0.2) $U.mg^{-1}$ and the K_M for formaldehyde as 1.24 (± 0.1) mM. The addition of BmACT did not alter the kinetics.

The temperature profile of BmMDH was also generated by assaying the *B. subtilis*-expressed Ni^{2+}-purified enzyme at a range of temperatures from 25 to 85 °C (Figure 5.9). The temperature optimum for the *B. subtilis*-expressed enzyme lies at 75 °C. The Arrhenius analysis gave an E_a = 36.1 (± 5.0) $kJ.mol^{-1}$.

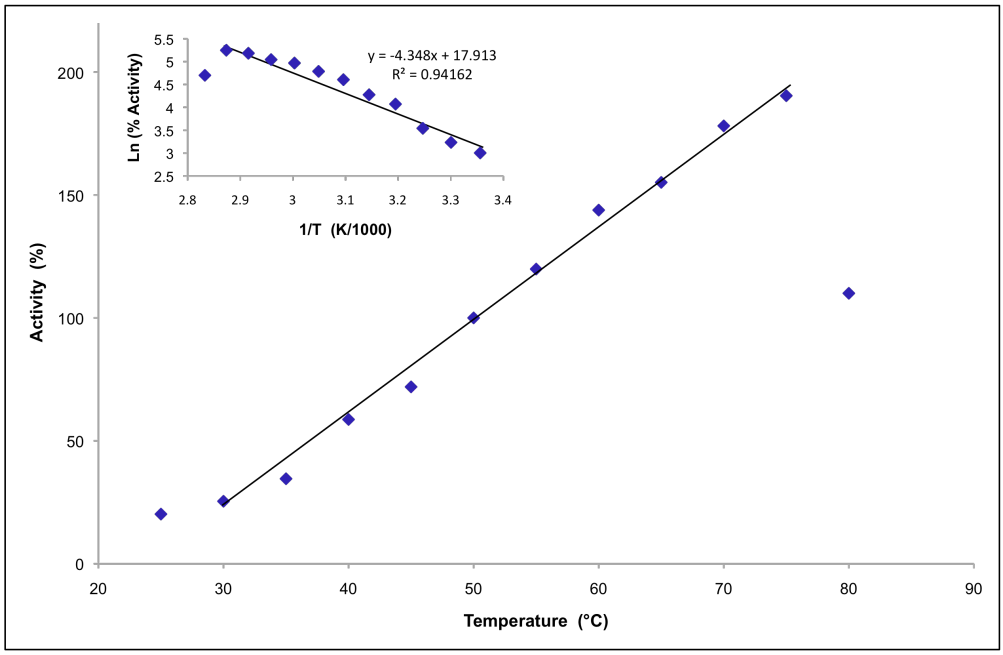


Figure 5.9. Thermal activity of the recombinant *B. subtilis*-expressed His-tagged BmMDH against formaldehyde. Activity is expressed as a percentage of the activity of the enzyme at 50 °C (100%). Inset, Arrhenius plot of the data.

The *Bacillus*-expressed BmFALDDH and GthFDH were assayed but no enzymatic activity could be detected. The respective soluble lysate from the IPTG-induced *B. subtilis* cultures harbouring the constructs pHT01 *Bmfalddh*• and pHT01 *GthfdhA*• did not exhibit activity with their respective substrates (formaldehyde and NAD^+ or formate and NADH for FALDDH activity, and formate and NAD^+ for FDH activity). Metal-affinity purification of the corresponding His-tagged constructs did not result in purification of soluble protein. Neither FALDDH nor FDH enzymatic activities could be detected in the respective column eluates.

5.4 Discussion

Expression of the three dehydrogenases involved in the dissimilation of methanol in *Bacillus methanolicus* in the homologous host system *Bacillus subtilis* has not been completely successful. While the methanol dehydrogenase MDH and its activator ACT have been successfully expressed in a soluble and active form, the same cannot

be said regarding the formaldehyde dehydrogenase and formate dehydrogenase which have not demonstrated strong over-expression, if any, for a product that remained at least partially insoluble.

There are numerous Gram-positive-based expression systems to chose from and initial directions were towards a *Lactobacillus* system (de Ruyter *et al*., 1996; Maassen, 1999; Serror *et al*., 2002). Although this would not have constituted homologous expression of the genes of interest, this was a Gram-positve expression system for Gram-positive genes. However, since three of the genes of interest here are from *Bacillus*, the fourth from a *Geobacillus*, a *Bacillus*-based expression system was finally decided upon. The shuttle vector pHT01 was also chosen as it is one of the few commercially-developed systems for this host. Numerous *Bacillus* transformation protocols were tried from high-osmolarity based electroporation (Xue *et al*., 1999) to electroporating protoplasts (Romero *et al*., 2006), but all with poor transformation efficiency and no satisfactory transformants were obtained despite various attempts and thought improvements. Discussions with Dr. Brautaset led to discarding the method described for the transformation of *B. methanolicus* by Cue *et al*. (1997) and Brautaset *et al*. (2003), and the direct transformation system described by Bolhuis *et al*. (1999) was finally adopted after discussions with the author. Successful transformation of *B. subtilis* was thus achieved with good transformation efficiencies.

As observed during the heterologous expression, only the MDH and ACT proteins expressed solubly and intensely. The FALDDH, and even more so FDH, were weakly over-expressed and appeared to be overwhelmingly found in the insoluble fraction.

Furthermore, and as observed previously, MDH could only be assayed in the reverse direction (formaldehyde reductase) as not activity was detected with methanol and NAD. The reported effect of the activator on the kinetics of the reaction in the forward direction was not observed either.

However, in the reverse direction (i.e. with formaldehyde and NADH), the kinetic parameters measured for the *Bacillus*-expressed MDH were similar to those calculated for the *E. coli*-expressed enzyme, and comparable to that previously

reported by Arfman *et al.* (1997). The temperature optimum of the *Bacillus*-expressed MDH was lower than that estimated for the *E. coli*-expressed enzyme at *ca*. 75 °C.

The construction of the pHT:His vector, bearing an N-terminal His-tag for the protein of interest had an error of conception embedded within: the design of the pHTHisthr forward and reverse primers (see p 143) included the highly conserved *Bacillus* RBS motif but did not contain a start codon (ATG) and therefore no expression of any target protein could be possible. Instead the primer pair should have been designed as presented below (*Bam*HI in bold, start codon as **ATG** followed by an 8xHis and the thrombin recognition sequence from pET28a; sequence homologous to pHT01 is underlined)

pHTHisthr_sdmF

5' CAATTCCCAATTAAAGGAGGAAGG**ATG**CATCACCATCACCATCACCATCACAGCAGCGGCCTGGTGCCGCGC GGCAGC**GGATCC**CCGTCTAGAGTCGACG

pHTHisthr_sdmR

5' CGTCGACTCTAGACG**GGATCC**GCTGCCGCGCGGCACCAGGCCGCTGCTGTGGTAGTGATGGTGATGGTGATG CATCCTTCCTCCTTTAATTGGGAATTG

It is worthy noting that Mobitec's own version of a His-tag pHT vector, pHT08, also lacks a start codon. It is likely that the error incurred in designing and constructing pTH:Histhr came from taking pHT08 as a model.

The recent development of a versatile vector for expression in *Geobacillus* spp. could prove a useful alternative (Taylor *et al*., 2008). It is based on the plasmid pUCG18, which uses kanamycin for selection (the most thermo-tolerant, commonly-used antibiotic with resistance encoded by a thermostable kan nucleotidyltransferase), and possesses the *ori* of pBST1 *G. stearothermophilus* cloned into *E.coli* pUC18. Plasmid pUCG18 replicated in both organisms with a transformation efficiency of 10^4 transformants per µg DNA, was stable at 68 °C, and demonstration of expression was performed with the expression of the pyruvate decarboxylase (PDC) from *Z. palmae* at 45 °C. Thompson *et al*. (2008) later described the heterologous expression of PDC from *Zymomonas mobilis* in *G. thermoglucosidasius*.

Chapter 6
General Discussion and suggestions for future work

The concept of a cell-free enzymatic pathway for biohydrogen production has the potential to be adapted and tailored to a variety of substrates and applications.

In this study, methanol was the substrate in a proof-of-concept system for the enzymatic production of biohydrogen. The genes encoding for the three dehydrogenase enzymes involved in methanol dissimilation to hydrogen and CO_2 were cloned from two thermophilic organisms, namely *Bacillus methanolicus* and *Geobacillus thermoglucosidasius*, fulfilling a primary objective of this study. However, specific enzymatic activity could not be detected for all three individual enzymes using *in vitro* assays. The methanol dehydrogenase from *B. methanolicus* (BmMDH) exhibited activity only in the reverse direction, *i.e.* acting as a formaldehyde reductase, with similar kinetic parameters calculated for both the heterologously and homologously expressed enzymes. In principle, all chemical reactions are reversible, and an enzyme will accelerate the forward or reverse reactions equally since the K_{eq} is a constant. The Haldane relationship (1) highlights the interdependence of the kinetic parameters V_{max} and K_M of the forward (f) and backwards (b) reactions, so that K_{eq} stays constant (Haldane, 1930).

$$K_{eq} = \left[\frac{V_{maxf} \cdot K_{Mb}}{V_{maxb} \cdot K_{Mf}} \right] \qquad (1)$$

The BmMDH enzymatic rate and kinetic parameters were calculated for the reverse (backwards) direction of the reaction, with formaldehyde and NADH, but no activity could be observed for the forward direction with methanol and NAD^+. However, the high NAD to NADH ratio *in vivo* may favour the oxidation (dehydrogenation) of methanol to formaldehyde (San *et al*., 2002). Also, the equilibrium may be pulled towards formaldehyde *in vivo*, through the subsequent reactions catalysed by formaldehyde dehydrogenase and formate dehydrogenase. It should be noted that the

enzymes were assayed *in vitro* under optimal conditions (particularly of pH), which may not reflect the status *in vivo*, although the methanol dehydrogenase assay was validated in the forwards direction with commercial alcohol dehydrogenase from horse liver (Sigma).

The recombinant formaldehyde dehydrogenase and formate dehydrogenase could not be obtained in a soluble form and attempts at resolubilising the enzymes proved unsuccessful. The formaldehyde and formate dehydrogenase assays were validated by assaying commercial versions of the enzymes (Sigma).

The construction of a synthetic pathway was attempted, by coupling the Ni^{2+}-purified methanol dehydrogenase from *B. methanolicus* with the commercial formaldehyde dehydrogenase from *Pseudomonas putida* (Sigma). It was hoped the formaldehyde dehydrogenase would pull the methanol dehydrogenase reaction in the forward direction by catalysing the oxidation of formaldehyde formed from methanol by BmMDH into formate in the presence of cofactor NAD^+. However, this did not occur. Replacing the BmMDH with the commercial ADH from horse liver (Sigma) in this scheme, with methanol as primary substrate and the cofactor NAD^+, saw an increased initial velocity of the coupled reaction compared to the reaction with the ADH alone. Further expanding this synthetic pathway with the formate dehydrogenase from *Candida boidinii* (Sigma) did not show a further increase in NAD^+ reduction to NADH. However, only preliminary assays were performed, and assay conditions may have had to be customised to accommodate these three enzymes to function in concert.

Therefore, much work remains. An initial priority is to express soluble recombinant formaldehyde and formate dehydrogenases (BmFALDDH and GthFDHa), and to purify and characterise them. Identifying a suitable formate dehydrogenase in *B. methanolicus* (BmFDH) would add to the homogeneity of the scheme as all three dehydrogenases would then come from a single source. This would also provide evidence for a linear pathway for formaldehyde dissimilation in *B. methanolicus*. Numerous expression constructs have been generated here for the three dehydrogenases, but not all have been tested. Furthermore, other combinations of vector, expression tag, expression system and expression strain could be envisaged.

Particularly, expressing the BmFALDDH and GthFDHa as secreted proteins (e.g. using the pHT43 expression vector in a *Bacillus* expression system) may allow for the purification of soluble proteins.

With all three dehydrogenases characterised individually, the cascade could then be set up and biohydrogen evolution from methanol by an enzymatic pathway estimated.

As introduced previously, Woodward *et al*. (1996 and 2000) and more recently Zhang *et al.* (2007) designed multi-enzyme synthetic pathways for biohydrogen production from glucose and starch respectively (see 1.2.2.4). A step towards the utilisation of even more complex carbohydrate substrates for conversion to hydrogen was taken when Ye *et al*. (2009) who further developed the synthetic pathway envisaged by Zhang *et al*. (2007) to accommodate a cellobiose or cellodextrin substrate (Figure 6.1). Cellobiose is a major product of the enzymatic hydrolysis of cellulosic materials, while cellodextrins are the product of the acid hydrolysis of cellulose or biomass. With these substrates, the first segment of the synthetic pathway was modified to a cellobiose phosphorylase (CBP) or a cellodextrin phosphorylase (CDP) upstream of the phosphoglucomutase (Ye *et al*., 2009). The overall yield of hydrogen production was 11.2 mol/mol anhydroglucose unit of cellobiose in solution (cellobiose being a polymer of anhydroglucose), *i.e.* 93% of the theroretical yield and comparable to that reported by Woodward *et al.* (2000b) from glucose.

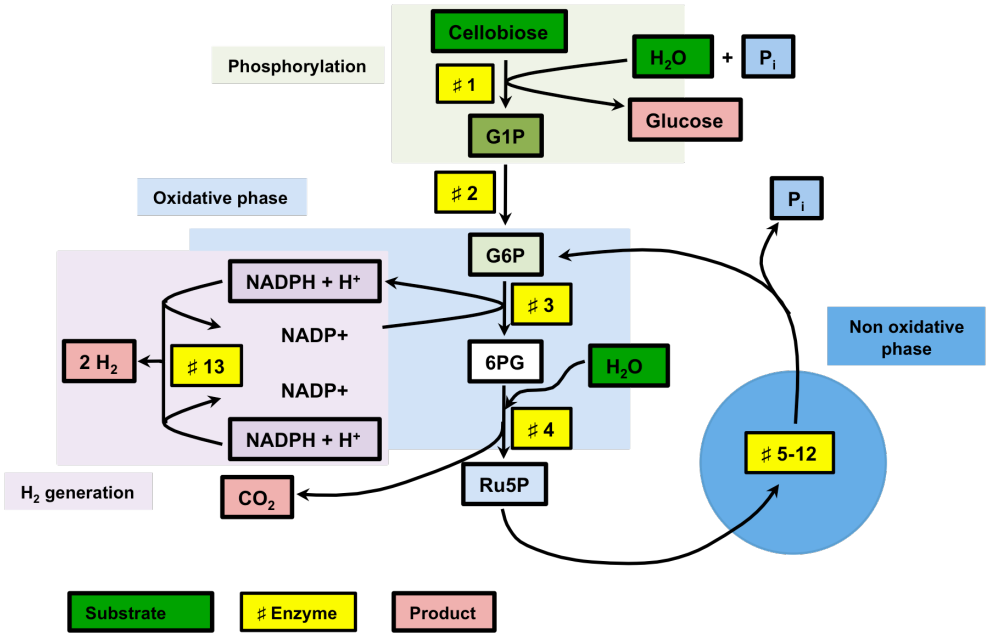


Figure 6.1. Hydrogen production from cellodextrin and water by a synthetic enzymatic pathway. Abbrev.: G1P, glucose-1-phosphate; G6P, glucose-6-phosphate; 6PG, 6-phosphogluconate; Ru5P, ribulose-5-phosphate; P_i, inorganic phosphate. The enzymes are: #1 cellobiose phosphorylase (cellodextrin phosphorylase); #2 phosphoglucomutase; #3 G6P dehydrogenase; #4 6PG dehydrogenase; #5 phosphoribose isomerase; #6 Ru5P epimerase; #7 transaldolase; #8 transketolase; #9 triose phosphate isomerase; #10 aldolase; #11 phosphoglucose isomerase; #12 fructose-1-6-bisphosphatase; #13 hydrogenase. (Adapted from Ye *et al*., 2009).

Metabolic engineering could provide the means for improvement of biohydrogen production, whether it be through genetically engineering microorganisms with enhanced hydrogen evolution capabilities, or creating synthetic *de novo* microorganisms with tailored genetics as is the vision driving research at the J. Craig Venter™ Institute. However, complex cellular responses to genetic perturbations may hinder development of such tailored, synthetic or not, microorganisms. It is in this context that cell-free, enzymatic systems could generate an economic advantage. However, one major obstacle could be the cost and life span of the enzymes. The use of enzymes from (hyper)thermophilic sources may help lower production costs while the development of improved purification techniques, the use of immobilisation systems, and other foreseeable system improvements may participate in making these systems more competitive.

As for hydrogen itself, it is an extremely efficient energy carrier, with the physical properties of being a fluid. And the reason why hydrogen could be the fuel of tomorrow is precisely because it is the fuel of today.

References

Abreu, A. A., Danko, A. S., Costa, J. C., Ferreira, E. C., Alves, M. M. (2009). Inoculum type response to different pHs on biohydrogen production from L-arabinose, a component of hemicellulosic biopolymers. *Int. J. Hydrogen Energy* **34**(4):1744-51

Adams, M. W. W., Holden, J. F., Menon, A. L., Schut, G. J., Grunden, A. M., Hou, C., Hutchins, A. M., Jenney, F. E., Jr., Kim, C., Ma, K., Pan, G., Roy, R., Sapra, R., Story, S. V., Verhagen, M. F. J. M. (2001). Key role for sulfur in peptide metabolism and in regulation of three hydrogenases in the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol*. **183**(2):716-24

Adeosun, E. K., Smith, T. J., Holberg, A.-M., Velarde, G., Ford, R., Dalton, H. (2004). Formaldehyde dehydrogenase preparations from *Methylococcus capsulatus* (Bath) comprise methanol dehydrogenase and methylene tetrahydromethanopterin dehydrogenase. *Microbiology* **150**(3):707-13

Akutsu, Y., Lee, D.-Y., Chi, Y.-Z., Li, Y.-Y., Harada, H., Yu, H.-Q. (2009). Thermophilic fermentative hydrogen production from starch-wastewater with bio-granules. *Int. J. Hydrogen Energy* **34**(12)5061-71

Ali Hassan, M., Shirai, Y., Kusubayashi, N., Ismail Abdul Karim, M., Nakanishi, K., Hasimoto, K. (1997). The production of polyhydroxyalkanoate from anaerobically treated palm oil mill effluent by *Rhodobacter sphaeroides*. *J. Ferm. Bioeng.* **83**(5):485-8

Anagnostopoulos, C. and Spizizen, J. (1961). Requirements for transformation in *Bacillus subtilis*. *J. Bacteriol*. **81**(5):741-6

Andreadeli, A., Platis, D., Tishkov, V., Popov, V., Labrou, N. E. (2008). Structure-guided alteration of coenzyme specificity of formate dehydrogenase by saturation mutagenesis to enable efficient utilization of NADP+. *FEBS J*. **275**(15):3859-69

Albers, S. V., Jonuscheit, M., Dinkelaker, S., urich, T., Kletzin, A., Tampe, R., Driessen, A. J. M., Schleper, C. (2006). Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Appl. Envir. Microbiol*. **72**(1):102-11

Anthony, C. (1982). The biochemistry of methylotrophs. Academic Press, Inc. (London) Ltd., London. 350p.

Anthony, C. (1986). The bacterial oxidation of methane and methanol. *Adv. Microbiol. Physiol.* **27**:113-210

Arfman, N., Watling, E. M., Clement, W., van Oosterwijk, R. J., de Vries, G. E., Harder, W., Attwood, M. M., Dijkhuizen, L. (1989). Methanol metabolism in thermotolerant methylotrophic *Bacillus* strains involving a novel catabolic NAD-dependent methanol dehydrogenase as a key enzyme. *Arch. Microbiol.* **152**(3):280-8

Arfman, N. and Dijkhuizen, L. (1990). Methanol dehydrogenase from the thermotolerant methyloytroph *Bacillus methanolicus* C1. *Meth. Enzymol*. **188**:223-6

Arfman, N., Vanbeeumen, J., de Vries, G. E., Harder, W., Dijkhuizen, L. (1991). Purification and characterization of an activator protein for methanol dehydrogenase from thermotolerant *Bacillus* spp. *J. Biol. Chem.* **266**(6):3955-60

Arfman, N., de Vries, K. J., Moezelaar, H. R., Attwood, M. M., Robinson, G. K., van Geel, M., Dijkhuizen, L. (1992a). Environmental regulation of alcohol metabolism in thermotolerant methylotrophic *Bacillus* strains. *Arch. Microbiol*. **157**(3):272-8

Arfman, N., Dijkhuizen, L., Kirchhof, G., Ludwig, W., Schleifer, K. H., Bulygina, E. S., Chumakov, K. M., Govorukhina, N. I., Trotsenko, Y. A., White, D., Sharp, R.J. (1992b). *Bacillus methanolicus* sp. nov., a new species of thermotolerant, methanol-utilizing, endospore-forming bacteria. *Int. J. Syst. Bacteriol.* **42**(3):439-45

Arfman, N., Hektor, H. J., Bystrykh, L. V.,Govorukhina, N. I., Dijkhuizen, L., Frank, J. (1997). Properties of an NAD(H)-containing methanol dehydrogenase and its activator protein from *Bacillus methanolicus*. *Eur. J. Biochem.* **244**(2):426-33

Argun, H., Kargi, F., Kapdan, I. K. (2009). Hydrogen production by combined dark and light fermentation of ground wheat solution. *Int. J. Hydrogen Energy* **34**(10):4305-11

Asada, Y., Tokumoto, M., Aihara, Y., Oku, M., Ishimi, K., Wakayama, T., Miyake, T., Tomiyama, M., Kohno, H. (2006). Hydrogen production by co-cultures of Lactobacillus and a photosynthetic bacterium, *Rhodobacter sphaeroides* RV. *Int. J. Hydrogen Energy* **31**(11):1509-13

Asano, Y., Sekigawa, T., Inukai, H., Nakazawa, A. (1988). Purification and properties of formate dehydrogenase from Moraxella sp. strain C-1. *J. Bacteriol*. **170**(7):3189-93

Attwood, M. M., Arfman, N., Weusthuis, R. A., Dijkhuizen, L. (1992). Purification and characterisation of an NAD(+)-linked formaldehyde dehydrogenase from the facultative RuMP cycle methylotroph Arthrobacter P1. *J. Microbiol. (Antonie van Leeuwenhoek)* **62**(3):201-7

Ayad, S. R. and Barker, G. R. (1969). The integration of donor and recipient deoxyribonucleic acid during transformation of Bacillus subtilis. *Biochem J*. **113**:167-174

Ball, M. and Wietschel, M. (2009). The future of hydrogen – opportunities and challenges. *Int. J. Hydrogen Energy* **34**(2):615-627.

Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol*. **10**(5):411-21

Barber, R. D., Rott, M. A., Donohue, T. J. (1996). Characterization of a glutathione-dependent formaldehyde dehydrogenase from Rhodobacter sphaeroides. *J. Bacteriol*. **178**(5):1386-93

Barber, R. D. and Donohue, T. J. (1998). Function of a glutathione-dependent formaldehyde dehydrogenase in Rhodobacter sphaeroides formaldehyde oxidation and assimilation. *Biochemistry* **37**(2):530-7

Berks, B. C., Palmer, T., Sargent, F. (2003). The tat protein translocation pathway and its role in microbial plysiology. *Adv. Microb. Physiol*. **47**:187-254

Bessman, M. J., Frick, D. N., Ohandley, S. F. (1996). The MutT proteins or "Nudix" hydrolases, a family of versatile, widely distributed, "house-cleaning" enzymes. *J. Biol. Chem.* **271**(41):25059-62

Bisaillon, A., Turcot, J., Hallenbeck, P. C. (2006). The effect of nutrient limitation on hydrogen production by batch cultures of *Escherichia coli*. *Int. J. of Hydrogen Energy* **31**(11):1504-8

Bodrossy, L., Holmes, E. M., Holmes, A. J., Kovàcs, K. L., Murrell, J. C. (1997). Analysis of 16S rRNA and methane monooxygenase gene sequences reveals a novel group of thermotolerant and thermophilic methanotrophs, *Methylocaldum* gen. nov. *Arch. Microbiol*. **168**(6):493-503

Bolhuis, A., Sorokin, A., Azevedo, V., Ehrlich, S. D., Braun, P. G., deJong, A., Venema, G., Bron, S., vanDijl, J. M. (1996). *Bacillus subtilis* can modulate its capacity and specificity for protein secretion through temporally controlled expression of the *sipS* gene for signal peptidase I. *Mol. Microbiol.* **22**(4):605-18

Bolhuis, A., Tjalsma, H., Stephenson, K., Harwood, C. R., Venema, G., Bron, S., van Dijl, J. M. (1999). Different mechanisms for thermal inactivation of *Bacillus subtilis* signal peptidase mutants. *J. Biol. Chem.* **274**(22):15865-8

Borodin, V. B., Tsygankov, A. A., Rao, K. K., Hall, D. O. (2000). Hydrogen production by *Anabaena variabilis* PK84 under simulated outdoor conditions. *Biotechnol. Bioeng*. **69**(5):478-85

Bott, K. F. and Wilson, G. A. (1968). Metabolic and nutritional factors influencing development of competence for transfection of *Bacillus subtilis*. *Bact. Rev*. **32**(4):370

Bowman, J. P., McCammon, S. A., Skrratt, J. H. (1997). *Methylosphaera hansonii* gen. nov., sp. nov., a psychrophilic, group I methanotroph from Antarctic marine-salinity, meromictic lakes. *Microbiology* **143**(4):1451-9

Bradford, M. M. (1976). Rapid and sensitive method for quantitation of microgram quantities of protein untilizing principle of protein-dye binding. *Anal. Biochem*. **72**(1-2):248-54

Brautaset, T., Williams, M. D., Dillingham, R. D., Kaufmann, C., Bennaars, A., Crabbe, E., Flickinger, M. C. (2003). Role of *Bacillus methanolicus* citrate synthase gene, *citY*, in regulating the secretion of glutamate in L-lysine-secreting mutants. *Appl. Envir. Microbiol.* **67**(7):3986-95

Brautaset, T., Jakobsen, O. M., Flickinger, M. C., Valla, S., Ellingsen, T. E. (2004). Plasmid-dependent methylotrophy in thermotolerant *Bacillus methanolicus*. *J. Bacteriol*. **186**(5):1229-38

Brautaset, T., Jakobsen, O. M., Josefsen, K. D., Flickinger, M. C., Ellingsen, T. E. (2007). *Bacillus methanolicus*: a candidate for industrial production of amino acids from methanol at 50 degrees C. *Appl. Microbiol. Biotechnol*. **74**(1):22-34

Brockmeier, U., Wendorff, M., Eggert, T. (2006). Versatile expression and secretion vectors for *Bacillus subtilis*. *Curr. Microbiol.* **52(**2):143-8

Bukau, B. and Horwich, A. L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell* **92**(3):351-66

Bystrykh, L. V., Vonck, J., van Bruggen, E. F. J., van Breeumen, J., Samyn, B., Govorukhina, N. I., Arfman, N., Duine, J. A., Dijkhuizen, L. (1993a). Electron microscopic analysis and structural oxidoreductases from the gram-positive methylotrophic bacteria *Amycolatopsis methanolica* and *Mycobacterium gastri* MB19. *J. Bacteriol.* **175**(6):1814-22

Bystrykh, L. V., Govorukhina, N. I., van Ophem, P. W., Hektor, H. J., Dijkhuizen, L., Duine, J. A. (1993b). Formaldehyde dismutase activities in Gram-positive bacteria oxidizing methanol. *J. Gen. Microbiol.* **139**:1979-85

Bystrykh, L. V., Govorukhina, N. I., Dijkhuizen, L., Duine, J. A. (1997). Tetrazolium-dye-linked alcohol dehydrogenase of the methylotrophic actinomycete *Amycolatopsis methanolica* is a three-component complex. *Eur. J. Biochem*. **247**(1):280-7

Cao, G., Ren, N., Wang, A., Lee, D.-J., Guo, W., Liu, B., Feng, Y., Zhao, Q. (2009). Acid hydrolysis of corn stover for biohydrogen production using *Thermoanaerobacterium thermosaccharolyticum* W16. *Int. J. Hydrogen Energy* **34**(17):7182-8

Chang, S. and Cohen, S. N. (1979). High-frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. *Mol. Gen. Genet*. **168**(1):111-5

Chen, X., Sun, Y., Xiu, Z., Li, X., Zhang, D. (2006). Stochiometric analysis of biological hydrogen production by fermentative bacteria. *Int. J. Hydrogen Energy* **31**(4):539-49

Cheong, D. Y. and Hansen, C. L. (2006). Bacterial stress enrichment enhances anaerobic hydrogen production in cattle manure sludge. *Appl. Microbiol. Biotechnol*. **72**(4):635-43

Chistoserdova, L. V. and Lidstrom, M. E. (1994a). Genetics of the serine cycle in *Methylobacterium extorquens* AM1: identification of *sgaA* and *mtdA* and sequences of *sgaA*, *hprA*, and *mtdA*. *J. Bacteriol.* **176**(7):1957-68

Chistoserdova, L. V. and Lidstrom, M. E. (1994b). Genetics of the serine cycle in *Methylobacterium extorquens* AM1: cloning, sequence, mutation, and physiological effect of *glyA*, the gene for Serine hydroxymethyltransferase. *J. Bacteriol*. **176**(21):6759-63

Chistoserdova, L. V. and Lidstrom, M. E. (1996). Molecular characterisation of a chromosomal region involved in the oxidation of acetyl-coA to glyoxylate in the isocitrate-lyase-negative methylotroph *Methylobacterium extorquens* AM1. *Microbiology* **142**:1459-68

Chistoserdova, L. V. and Lidstrom, M. E. (1997). Molecular and mutational analysis of a DNA region separating two methylotrophy gene clusters in *Methylobacterium extorquens* AM1. *Microbiology* **143**:1729-36

Chistoserdova, L., Vorholt, J. A., Thauer, R. K., Lidstrom, M. E. (1998). C_1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea. *Science* **281**(5373):99-102

Chistoserdova, L., Laukel, M., Portais, J. C., Vorholt, J. A., Lidstrom, M. E. (2004). Multiple formate dehydrogenase enzymes in the facultative methylotroph *Methylobacterium extorquens* AM1 are dispensable for growth on methanol. *J. Bacteriol*. **186**(1):22-8

Chong, M.-L., Rahim, R. A., Shirai, Y., Hassan, M. A. (2009). Biohydrogen production by *Clostridium butyricum* EB6 from palm oil mill effluent. *Int. J. Hydrogen Energy* **34**(2):764-71

Chow, C. M. and RajBhandary, U. L. (1993). Developmental regulation of the gene for formate dehydrogenase in *Neurospora crassa*. *J. Bacteriol.* **175**(12):3703-9

Claassen, P. A. M., Budde, M. A. W., van Noorden, G. E., Hoekema, S., Hazewinkel, J. H. O., van Groenestijn, J. W., de Vrije, G. J. (2004). Biological hydrogen production from agro-food by-products. Total food: exploiting co-products, minimising waste. Institute of Food Research, Norwich.

Clark, D. P. (1989). The fermentation pathways of *Escherichia coli*. *FEMS Microbiol. Rev*. **63**(3):223-34

Colas des Francs-Small, C., Ambard-Bretteville, F., Small, I. D., Remy, R. (1993). Identification of a major soluble protein in mitochondria from nonphotosynthetic tissues as NAD-dependent formate dehydrogenase. *Plant Physiol*. **102**(4):1171-7

Conrad, B., Savchenko, R. S., Breves, R., Hofemeister, J. (1996). A T7 promoter-specific inducible promoter expression system for *Bacillus subtilis*. *Mol. Gen. Genetics* **250**(2):230-6

Conway, T. and Ingram, L. O. (1988). Similarity of *Escherichia coli* propanediol oxidoreductase (*fuc*O product) and an unusual alcohol dehydrogenase from *Zymomonas mobilis* and *Saccharomyces cerevisia*. *J. Bact* **171**(7):3754-9

Cooper, G., Kimmich, N., Belisle, W., Sarinana, J., Brabham, K., Garrel, L. (2001). Carbonaceous meteorites as a source of sugar-related organic compounds for the early Earth. *Nature* **414**(6866):879-83

Cue, D., Lam, H., Hanson, R. S., Flickinger, M. C. (1996). Characterization of a restriction-modification system of the thermotolerant methylotroph *Bacillus methanolicus*. *Appl. Envir. Microbiol*. **62**(3):1107-11

Cue, D., Lam, H., Dillingham, R. L., Hanson, R. S., Flickinger, M. C. (1997). Genetic manipulation of *Bacillus methanolicus*, a Gram-positive, thermotolerant methylotroph. *Appl. Envir. Microbiol.* **63**(4):1406-20

Davila-Vazquez, G., Cota-Navarro, C. B., Rosales-Colunga, L. M., de Leon-Rodriguez, A., Razo-Flores, E. (2009). Continuous biohydrogen production using cheese whey: Improving the hydrogen production rate. *Int. J. Hydrogen Energy* **34**(10):4296-304

d'Ippolito, G., Dipasquale, L., Vella, F. M., Romano, I., Gambacorta, A., Cutignano, A., Fontana, A. (2010). Hydrogen metabolism in the extreme thermophile *Thermotoga neapolitana*. *Int. J. Hydrogen Energy* **35**(6):2290-5

de Ruyter, P. G. G. A., Kuipers, O. P., de Vos, W. M. (1996). Controlled gene expression systems for *Lactobacillus lactis* with the food-grade inducer nisin. *Appl. Envir. Microbiol*. **62**(10):3662-7

de Vos. W. M., Kleerebezem, M., Kuipers, O. P. (1997). Expression systems for industrial Gram-positive bacteria with low guanine and cytosine content. *Curr. Opin. Biotechnol*. **8**:547-53

de Vries, G. E., Harms, N., Maurer, K., Papendrecht, A., Stouthamer, A. H. (1988). Physiological regulation of *Paracoccus denitrificans* methanol dehydrogenase synthesis and activity. *J. Bacteriol*. **170**(8):3731-7

de Vries, G. E., Kues, U., Stahl, U. (1990). Physiology and genetics of methylotrophic bacteria. *FEMS Microbiol. Rev.* **6**(1):57-101

de Vries, G. E., Arfman, N., Terpstra, P., Dijkhuizen, L. (1992). Cloning, expression, and sequence analysis of the *Bacillus methanolicus* C1 methanol dehydrogenase gene. *J. Bacteriol*. **174**(16):5346-53

Dedysh, S. N., Liesack, W., Khmelenina, V. N., Suzina, N. E., Trotsenko, Y. A., Semrau, J. D., Bares, A. M., panikov, N. S., Tiedje, J. M. (2000). *Methylocella palustris* gen. nov., sp. nov., a new methane-oxidising acidophilic bacterium from peat bogs, representing a novel subtype of serine pathway methanotrophs. *Int. J. Sys. Evol. Microbiol*. **50**(3):955-69

Dijkhuizen, L., Hansen, T. A., Harder, W. (1985). Methanol, a potential feedstock for biotechnology processes. *Trends Biotechnol*. **3**(10):262-7

Dijkhuizen, L., Arfman, N., Attwood, M. M., Brooke, A. G., Harder, W., Waitling, E. M. (1988). Isolation and initial characterisation of thermotolerant methylotrophic *Bacillus* strains. *FEMS Microbiol. Rev.* **52**(3):209-14

Dijkhuizen, L. and Arfman, N. (1990). Methanol metabolism in thermotolerant methylotrophic *Bacillus* species. *FEMS Microbiol. Rev*. **87**(3-4):215-9

Drewke, C. and Ciriacy, M. (1988). Overexpression, purification and properties of alcohol dehydrogenase-IV from *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* **950**:54-60

Dubnau, D. (1991a). The regulation of competence in *Bacillus subtilis*. *Mol. Microbiol*. **5**(1):11-8

Dubnau, D. (1991b). Genetic competence in *Bacillus subtilis*. *Microbiol. Rev.* **55**(3):395-424

Duine, J. A. (1999). Thiols in formaldehyde dissimilation and detoxification. *Biofactors* **10**(2-3):201-6

Duitman, E. H., Wyczawski, D., Boven, L. G., Venema, G., Kuipers, O. P., Hamoen, L. W. (2007). Novel methods for genetic transformation of natural *Bacillus subtilis* isolates used to study the regulation of the mycosubtilin and surfactin synthetases. *Appl. Envir. Microbiol*. **73**(11);3490-6

Dunn, C. A., O'Handley, S. F., Frick, D. N., Bessman, M. J. (1999). Studies on the ADP-ribose pyrophosphatase of the Nudix hydrolases and tentative identification of *trgB*, a gene associated with tellurite resistance. *J. Biol. Chem*. **274**(45):32318-24

Dutta, D., De, D., Chaudhuri, S., Bhattacharya, S. (2005). Hydrogen production by cyanobacteria. *Microb. Cell Fact*. **4**(36):1-11

Ehrlich, S. D. (1977). Replication and expression of plasmids from *Staphylococcus aureus* in *Bacillus subtilis*. *PNAS USA* **74**(4):1680-2

Eisenthal, R. and Cornish-Bowen, A. (1974). The direct linear plot. A new graphical procedure for estimating enzyme kinetic parameters. *Biochem. J.* **139**:715-20

Emptage, C. D. (2007). Nitroreductases for use in enzyme prodrug therapy. *PhD thesis*, University of Bath.

Eroğlu, E., Gündüz, U., Yücel, M., Türker, L., Eroğlu, I. (2004). Photobiological hydrogen production by using olive mill wastewater as a sole substrate source. *Int. J. Hydrogen Energy* **29**(2):163-71

Eroğlu, E., Eroğlu, I., Gündüz, U., Türker, L., Yücel, M. (2006). Biological hydrogen production from olive mill wastewater with two-stage processes. *Int. J. Hydrogen Energy* **31**(11):1527-35

Evans, T. C., Benner, J., Xu, M. Q. (1999). The *in vitro* ligation of bacterially expressed proteins using an intein from *Methanobacterium thermoautotrophicum*. *J. Biol. Chem*. **274**(7):3923-6

Ferrer, M., Chernikova, T. N., Timmis, K. N., Golyshin, P. N. (2004). Expression of a temperature-sensitive esterase in a novel chaperone-based *Escherichia coli* strain. *Appl. Envir. Microbiol*. **70**(8):4499-504

Goldberg, S. L., Cino, P. M., Patel, R. N., Nanduri, V. B., Jahnston, R. M., Johnston, R. (2004). USA patent US2004/0038237 AI of 26.02.2004

Goodwin, P. M. and Anthony, C. (1998). The biochemistry, physiology and genetics of PQQ and PQQ-containing enzymes. *Adv. Microb. Physiol*. **40**:1-80

Gryczan, T. J., Contente, S., Dubnau, D. (1978). Characterisation of *Staphilococcus aureus* plasmids introduced by transformation into *Bacillus subtilis*. *J. Bacteriol*. **134**(1):318-29

Gryczan, T. J. and Dubnau, D. (1978). Construction and properties of chimeric plasmids in *Bacillus subtilis*. *PNAS USA* **75**(3):1428-32

Guo, Y.-P., Fan, S.-Q., Fan, Y.-T., Pan, C.-M., Hou, H.-W. (2010). The preparation and application of crude cellulase for cellulose-hydrogen production by anaerobic fermentation. *Int. J. Hydrogen Energy* **35**(2):459-68

Gupta, S. K. and Ghosh, T. C. (2000). GUCG: a non-redundant codon usage database from complete genomics. *Curr. Science* **78**(1):28-9

Hagel, J. M. and Facchini, P. J. (2010). Biochemistry and occurence of O-demethylation in plant metabolism. *Frontiers in Physiol*. **1**:1-7

Haima, P., Bron, S., Venema, G. (1987). The effect of restriction on shotgun cloning and plasmid stability in *Bacillus subtilis* Marburg. *Mol. Gen. Genet*. **209**(2):335-42

Haldane, J. B. S. (1930). *In:* Enzymes, Longmans, London. pp30

Hallenbeck, P. C. (2005). Fundamentals of the fermentative production of hydrogen. *Water Sci. Technol.* **52**(1-2):21-9

Hanes, C. S. (1932). Studies on plant amylases: the effect of starch concentration upon the velocity of hydrolysis by the amylase of germinated barley. *Biochem. J*. **26**(5):1406-21

Hanson, R. S. and Hanson, T. E. (1996). Methanotrophic bacteria. *Microbiol. Rev*. **60**(2):439-471

Harms, N., Ras, J., Reijnders, W. N., van Spanning, R. J., Stouthamer, A. H. (1996). S-formylglutathione hydrolase of Paracoccus denitrificans is homologous to human esterase D: a universal pathway for formaldehyde detoxification. *J. Bact*. **178**(21):6296-9

Harwood, C. R. (1992). *Bacillus subtilis* and its relatives: molecular biology and industrial workhorses. *Trends Biotechnol*. **10**:247-56

Hauser, P. M. and Karamata, D. (1994). A rapid and simple method for *Bacillus subtilis* transformation on solid media. *Microbiology* **140**(7):1613-7

Heierson, A., Landen, R., Lovgren, A., Dalhammar, G., Boman, H. G. (1987). Transformation of vegetative cells of *Bacillus thuringiensis* by plasmid DNA. *J. Bacteriol*. **169**(3):1147-52

Hektor, H. J., Kloosterman, H., Dijkhuizen, L. (2000). Nicotinoprotein methanol dehydrogenase enzymes in Gram-positive methylotrophic bacteria. *J. Mol. Catalysis B: Enzymatic* **8**:103-9

Hektor, H. J., Kloosterman, H., Dijkhuizen, L. (2002). Identification of a magnesium-dependent NAD(P)(H)-binding domain in the nicotinoprotein methanol dehydrogenase from *Bacillus methanolicus*. *J. Biol. Chem*. **277**(49):46966-73

Hirt, W., papoutsakis, E., Krug, E., Lim, H. C., Tsao, G. T. (1978). Formaldehyde incorporation by a new methylotroph (L3). *Appl. Envir. Microbiol*. **36**(1):56-62

Hollenberg, C. P. and Janowicz, Z. (1989). Eur. Patent EP1987000110417, bulletin 89/03

Hsiao, C.-L., Chang, J.-J., Wu, J.-H., Chin, W.-C., Wen, F.-S., Huang, C.-C., Chen, C.-C., Lin, C.-Y. (2009). *Clostridium* strain co-cultures for biohydrogen production enhancement from condensed molasses fermentation solubles. *Int. J. Hydrogen Energy* **34**(17):7173-81

Huber, R., Langworthy, T. A., Konig, H., Thomm, M., Woese, C. R., Sleytr, U. B., Stetter, K. O. (1986). *Thermotoga maritima* sp Nov. represents a new genus of unique extremely thermophilic Eubacteria growing up to 90 °C. *Arch. Microbiol*. **144**(4):324-33

Hyatt, D., Chen, G.-L., LoCascio, P., Land, M., Larimer, F., Hauser, L. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**:119

Ingason, H. T., Ingnolfsson, H. P., Jensson, P. (2008). Optimizing site selection for hydrogen production in Iceland. *Int. J. Hydrogen Energy* **33**(14):3632-43

Imanaka, T., Tanaka, N., Tsunekawa, H., Aiba, S. (1981). Cloning of the genes for penicillase, *pen*P and *pen*I, of *Bacillus licheniformis* in some vector plasmids and their expression in *Escherichia coli*, *Bacillus subtilis* and *Bacillus licheniformis*. *J. Bacteriol*. **147**(3):776-86

Irvine, W. M. (1999). The composition of interstellar clouds. *Space Sci. Rev*. **90**:203-218

Ito, K., Takahashi, M., Yoshimoto, T., Tsuru, D. (1994). Cloning and high-level expression of the glutathione-independant formaldehyde dehydrogenase gene from *Pseudomonas putida*. *J. Bacteriol.* **176**(9):2483-91

Ito, M. and Nagane, M. (2001). Improvement of the electro-transformation efficiency of facultatively alkaliphilic *Bacillus pseudofirmus* OF4 by high osmolarity and glycine treatment. *Biosci. Biotechnol. Biochem.* **65**(12):2273-5

Ivanova, G., Rákhely, G., Kovács, K. L. (2009). Thermophilic biohydrogen production from energy plants by *Caldicellulosiruptor saccharolyticus* and comparison with related studies. *Int. J. Hydrogen Energy* **34**(9):3659-70

Jahn, A., Keuntje, B., Dörffler, M., Klipp, W., Qelze, J. (1994). Optimising photoelectrotrophic H_2 production by *Rhodobacter capsulatus* upon interposon mutagenesis in the hupL gene. *Appl. Microbiol. Biotechnol*. **40**(5):687-90

Jakobsen, O. M., Benichou, A., Flickinger, M. C., Valla, S., Ellingsen, T. E., Brautaset, T. (2006). Upregulated transcription of plasmid and chromosomal ribulose monophosphate pathway genes is critical for methanol assimilation rate and methanol tolerance in the methylotrophic bacterium *Bacillus methanolicus*. *J. Bacteriol*. **188**(8):3063-72

Jayalakshmi, S., Joseph, K., sukumaran, V. (2009). Bio hydrogen generation from kitchen waste in an inclined plug flow reactor. *Int. J. Hydrogen Energy* **34**(21):8854-8

Jeon, Y. J., Fong, J. C. N., Riyanti, E. I., Neilan, B. A., Rogers, P. L., Svenson, C. J. (2008). Heterologous expression of the alcohol dehydrogenase (adhI) gene from *Geobacillus thermoglucosidasius* strain M10EXG. *J. Biotechnol*. **135**(2):127-33

Kadar, Z., de Vrijek, T., van Noorden, G. E., Budde, M. A. W., Szengyel, Z., Reczey, K., Claassen, P. A. M. (2004). Yields from glucose, xylose, and paper sludge hydrolysate during hydrogen production by the extreme thermophile *Caldicellulosiruptor saccharolyticus*. *Appl. Biochem. Biotechnol.* **113**:497-508

Kato, N., Yurimoto, H., Thauer, R. K. (2006). The physiological role of the ribulose monophosphate pathway in bacteria and archaea. *Biosci. Biotechnol. Biochem*. **70**(1):10-21

Kato, T., Miyanaga, A., Haruki, M., Imanaka, T., Morikawa, M., Kanaya, S. (2001). Gene cloning of an alcohol dehydrogenase from thermophilic alkane-degrading *Bacillus thermoleovorans* B23. *J. Biosci. Bioeng.* **91**(1):100-2

Kato, T., Miyanaga, A., Kanaya, S., Morikawa, M. (2010). Gene cloning and characterisation of an aldehydes dehydrogenase from long-chain alkane-degrading *Geobacillus thermoleovorans* B23. *Extremophiles* **14**(1):33-9

Keggins, K. M., Lovett, P. S., Duvall, E. J. (1978). Molecular cloning of genetically active fragments of *Bacillus* DNA in *Bacillus subtilis* and properties of the vector plasmid pUB110. *PNAS USA* **75**(3):1423-7

Kerr, R. A. (2007). Scientists tell policymakers we're all warming the world. *Science* **315**(5813):754-7

Kerr, R. A. (2007). The looming oil crisis could arrive uncomfortably soon. *Science* **316**(5823):351

Kilstrup, M. and Kristiansen, K. N. (2000). Rapid genome walking: a simplified oligo-cassette mediated polymerase chain reaction using a single genome-specific primer. *Nucleic Acids Res.* **28**(11):E55

Kim, M.-S., Lee, T. J., Yoon, Y. S., Lee, I. G., Moon, K. W. (2001). Hydrogen production from food processing wastewater and sewage sludge by anaerobic dark fermentation combined with photofermentation. *In:* Miyake, J., Matsunaga, T., SanPietro, A. (Eds.) Biohydrogen II: an approach to environmentallly acceptable technology. Pergamon, Oxford. pp:263-72

Kim, M.-S., Baek, J. S., Lee, J. K. (2006). Comparison of H_2 accumulation by *Rhodobacter sphaeroides* KD131 and its uptake hydrogenase and PHB synthase deficient mutant. *Int. J. Hydrogen Energy* **31**(1):121-7

Kloosterman, H., Vrijbloed, J. W., Dijkhuizen, L. (2002). Molecular, biochemical, and functional characterization of a Nudix hydrolase protein that stimulates the activity of a nicotinoprotein alcohol dehydrogenase. *J. Biol. Chem.* **277**(38):34785-92

Koma, D., Sawai, T., Harayama, S., Kino, K. (2006). Overexpression of the genes from thermophiles in *Escherichia coli* by high-temperature cultivation. *Appl. Microbiol. Biotechnol*. **73**(1):172-80

Komives, C. F., Chueng, L. Y.-Y., Pluschkell, S. B., Flickinger, M. C. (2005). Growth of *Bacillus methanolicus* in seawater-based media. *J. Ind. Microbiol. Biotechnol*. **32**(2):61-6

Koonin, E. V. (1993). A highly-conserved sequence motif defining the family of MutT-related proteins from eubacteria, eucaryotes and viruses. *Nucleic Acids Res*. **21**(20):4847

Kosourov, S., Tsygankov, A., Seibert, M., Ghirardi, M. L. (2002). Sustained hydrogen photoproduction by *Chlamydomonas reinhardtii*: effects of culture parameters. *Biotechnol. Bioeng*. **78**(7):731-40

Kunst, F. and Rapoport, G. (1995). Salt stress is an environmental signal affecting degradative enzyme synthesis in *Bacillus subtilis*. *J. Bacteriol*. **177**(9):2403-7

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680-5

Lee, C. K., Daniel, R. M., Shepherd, C., Saul, D., Cary, S. C., Danson, M. J., Eisenthal, R., Peterson, M. E. (2007). Eurythermalism and the temperature dependence of enzyme activity. *FASEB J.* **21**(8):1934-41

Levin, D. B., Islam, R., cicek, N., Sparling, R. (2006). Hydrogen production by *Clostridium thermocellum* 27405 from cellulosic biomass substrates. *Int. J. Hydrogen Energy* **31**(11):1496-1503

Lidstrom, M. E. and Wopat, A. E. (1984). Plasmids in methylotrophic bacteria: isolation, characterisation and DNA hybridisation analysis. *Arch. Microbiol*. **140**(1):27-33

Liu, Z., Qiu, Z., Luo, Y., Mao, Z., Wang, C. (2010). Operation of first solar-hydrogen system in China. *Int. J. of Hydrogen Energy* **35**(7):2762-6

Lloyd, D., Ralphs, J. R., Harris, J. C. (2002). Hydrogen production in *Giardia intestinalis*, a eucaryote with no hydrogenosomes. *Trends Parasitol*. **18**(4):155-6

Lo, Y. C., Saratale, G. D., Chen, W. M., Bai, M. D., Chang, J. S. (2009). Isolation of cellulose-hydrolytic bacteria and applications of the cellulolytic enzymes for cellulosic biohydrogen production. *Enzyme and Microbial Technol*. **44**(6-7):417-25

Maassen, C. B. (1999). A rapid and safe plasmid isolation method for efficient engineering of recombinant *Lactobacilli* expressing immunogenic or tolerogenic epitopes for oral administration. *J. Immunol. Methods* **223**(1):131-6

Mallonee, D. H. and Speckman, R. A. (1989). Transformation of *Bacillus polyxyma* with plasmid DNA. *Appl. Envir. Microbiol*. **55**(10):2517-21

Marbán, G. and Valdés-Solís, T. (2007). Towards the hydrogen economy? *Int. J. Hydrogen Energy* **32**(12):1625-37

Marino-Marmolejo, E. N., De Leon-Rodriguez, A., de la Rosa, A. P. B., Santos, L. (2009). Heterologous expression and characterisation of an alcohol dehydrogenase from the Archeon *Thermoplasma acidophilum*. *Mol. Biotechnol*. **42**(1):61-7

Martinez, M. A., Dezar, C., Baigori, M., Sineriz, F. (1999). Simple method for plasmid-mediated transformation of different *Bacillus* species. *Biotechnol. Techniques* **13**;337-40

Marx, C. J., Miller, J. A., Chistoserdova, L., Lidstrom, M. E. (2004). Multiple formaldehyde oxidation/detoxification pathways in *Burkholderia fungorum* LB400. *J. Bacteriol*. **186**(7):2173-8

Masukawa, H., Mochimaru, M., Sakurai, H. (2002). Disruption of the uptake hydrogenase gene, but not of the bidirectional hydrogenase gene, leads to enhanced photobiological hydrogen production by the nitrogen-fixing cyanobacterium *Anabaena* sp. PCC 7120. *Appl. Microbiol. Biotechnol*. **58**(5):618-24

Mathews, M. B. and Vennesland, B. (1950). Enzymic oxidation of formic acid. *J. Biol. Chem*. **186**(2):667-82

Melis, A. and Melnicki, M. R. (2006). Integrated biological hydrogen production. *Int. J. Hydrogen Energy* **31**(11):1563-73

Merk, O. and Speit, G. (1998). Significance of formaldehyde-induced DNA-protein crosslinks for mutagenesis. *Envir. Mol. Mut*. **32**(3):260-8

Middleton, R. and Hofmeister, A. (2006). New shuttle vectors for ectopic insertion of genes into *Bacillus subtilis*. *Plasmid* **51**(3):238-45

Misset-Smits, M., van ophem, P. W., Sakuda, S., Duine, J. A. (1997). Mycothiol, 1-O-(2'-[N-acetyl-cysteinyl]amido-2'-deoxy-[α]-glucopyranosyl)-myo-inositol, is the factor of NAD/factor-dependent formaldehyde dehydrogenase. *FEBS Lett*. **409**(2):221-2

Miyake, J., Miyake, M., Asada, Y. (1999). Biotechnological hydrogen production: research for efficient light energy conversion. *J. Biotechnol*. **70**(1):89-101

Moszner, I., Rocha, E. P., Danchin, A. (1999). Codon usage and lateral gene transfer in *Bacillus subtilis*. *Curr. Opin. Microbiol.* **2**(5):524-8

Nanba, H., Takaoka, Y., Hasegawa, J. (2003). Purification and characterization of an alpha-haloketone-resistant formate dehydrogenase from *Thiobacillus* sp. strain KNK65MA, and cloning of the gene. *Biosci. Biotechnol. Biochem*. **67**(10):2145-53

Nguyen, H. D., Nguyen, Q. A., Ferreira, R. C., Ferreira, L. C. S., Tran, L. T., Schumann, W. (2005). Construction of plasmid-based expression vectors for *Bacillus subtilis* exhibiting full structural stability. *Plasmid* **54**(3):241-8

Nguyen, H. D., Phan, T. P. P, Schumann, W. (2007). Expression vectors for the rapid purification of recombinant proteins in *Bacillus subtilis*. *Curr. Microbiol.* **55**(2):89-93

O Brien, J. E., McKellar, M. G., Harvego, E. A., Stoots, C. M. (2010). High-temperature electrolysis for large-scale hydrogen and syngas production from nuclear energy - summary of system simulation and economic analyses. *Int. J. Hydrogen Energy* **35**(10):4808-19

Olah, G. (2005). Beyond oil and gas: the methanol economy. *Angew. Chem. Int. Ed. Engl.* **44**:2636-39

Orita, I., Yurimoto, H., Harai, R., Kawarabayasi, Y., Sakai, Y., Kato, N. (2005). The archaeon *Pyrococcus horikoshii* possesses a bifunctional enzyme for formaldehyde fixation via the ribulose monophosphate pathway. *J. Bacteriol*. **187**(11):3636-42

Öztürk, Y., Yücel, M., Daldal, F., Mandaci, S., Gündüz, U., Türker, L., Eroğlu, I. (2006). Hydrogen production by using *Rhodobacter capsulatus* mutants with genetically modified electron transfer chains. *Int. J. Hydrogen Energy* **31**(11):1545-52

Penfold, D. W., Forster, C. F., Macaskie, L. E. (2003). Increased hydrogen production by *Escherichia coli* strain HD701 in comparison with the wild-type parent strain MC4100. *Enzyme and Microbial Technol.* **33**:185-9

Penfold, D. W. and Macaskie, L. E. (2004). Production of H_2 from sucrose by *Escherichia coli* strains carrying the pUR400 plasmid, which encodes invertase activity. *Biotechnol. Lett.* **26**(24):1879-83

Penfold, D. W., Sargent, F., Macaskie, L. E. (2006). Inactivation of the *Escherichia coli* K-12 twin-arginine translocation system promotes increased hydrogen production. *FEMS Microbiol. Lett*. **262**(2):135-7

Phadtare, S., Alsina, J., Inouye, M. (1999). Cold-shock response and cold-shock proteins. *Curr. Opin. Microbiol.* **2**(2):175-80

Phan, T. P. P., Nguyen, H. D., Schumann, W. (2006). Novel plasmid-based expression vectors for intra- and extracellular production of recombinant proteins in *Bacillus subtilis*. *Protein Expr. Purif.* **46**(2):189-95

Pluschkell, S. B. and Flickinger, M. C. (2002). Dissimilation of [^{13}C]methanol by continuous cultures of *Bacillus methanolicus* MGA3 at 50 degrees C studied by ^{13}C NMR and isotope-ratio mass spectrometry. *Microbiology* **148**(10):3223-33

Popov, V. O. and Lamzin, V. S. (1994). NAD(+)-dependent formate dehydrogenase. *Biochem. J.* **301**(3):625-43

Quievryn, G. and Zhitkovich, A. (2000). Loss of DNA-protein crosslinks from formaldehyde-exposed cells occurs through spontaneous hydrolysis and an active repair process linked to proteosome function. *Carcinogenesis* **21**(8):1573-80

Ras, J., van Ophem, P. W., Reijnders, W. N. M., van Spanning, R. J. M., Duine, J. A., Stouthamer, A. H., Harms, N. (1995). Isolation, sequencing, and mutagenesis of the gene encoding NAD- and glutathione-dependent formaldehyde dehydrogenase (GD-FALDH) from *Paracoccus denitrificans*, in which GD-FALDH is essential for methylotrophic growth. *J. Bacteriol*. **177**(1):247-51

Redwood, M. D. and Macaskie, L. E. (2006). A two-stage, two-organism process for biohydrogen from glucose. *Int. J. Hydrogen Energy* **31**(11):1514-21

Redwood, M. D. and Macaskie, L. E. (2007). Efficient bioH_2 and PEM-FC catalyst. *In:* Proceedings of the 7th hydrogen power and theoretical engineering solutions international symposium (HyPoThSIS VII), Merida, Mexico.

Resch, A., Rosenstein, R., Nerz, C. (2005). Differential gene expression profiling of *Staphylococcus aureus* cultivated under biofilm and planktonic conditions. *Appl. Envir. Microbiol*. **71**(5):2663-76

Richardson, A., Landry, S.J., Georgopoulos, C. (1998). The ins and outs of a molecular chaperone machine. *Trends Biochem. Sci*. **23**(4):138-43

Rishi, A.S., Nelson, N.D., Goyal, A. (2004). Genome walking of large fragments: an improved method. *J. Biotechnol*. **111**(1):9-15

Rocha, E.P.C., Danchin, A., Viari, A. (1999). Translation in *Bacillus subtilis*: roles and trends of initiation and termination, insights from a genome analysis. *Nucleic Acids Res*. **27**(17):3567-76

Romero, D., Perez-Garcia, A., Veening, J.W., de Vicente, A., Kuipers, O.P. (2006). Transformation of undomesticated strains of *Bacillus subtilis* by protoplast electroporation. *J. Microbiol. Methods* **66**(3):556-9

Rosales-Colunga, L. M., Razo-Flores, E., Ordonez, L. G., Alastride-Mondragon, F., de Leon-Rodriguez, A. (2010). Hydrogen production by *Escherichia coli* ΔhycA ΔlacI using cheese whey as substrate. *Int. J. Hydrogen Energy* **35**(2):491-9

Rose, T. M., Schultz, E. R., Henikoff, J. G., Pietrokowski, S., McCallum, C. M., Henikoff, S. (1998). Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res*. **26**(7):1628-35

Sakai , Y., Mitsui, R., Katayama, Y., Yanase, H., Kato, N. (1999). Organization of the genes involved in the ribulose monophosphate pathway in an obligate methylotrophic bacterium, *Methylomonas aminofaciens* 77a. *FEMS Microbiol. Lett*. **176**(1):125-30

Sakoda, H. and Imanaka, T. (1992). Cloning and sequencing of the gene coding for alcohol dehydrogenase of *Bacillus stearothermophilus* and rational shift of the optimum pH. *J. Bacteriol*. **174**(4):1397-1402

Saleeba, J. A., Cobbett, C. S., Hynes, M. J. (1992). Characterisation of the Amda-regulated *acia* gene of *Aspergillus nidulans*. *Mol. Gen. Genet.* **235**(2-3):349-58

Sambrook, J. and Russell, D.W. (2001). Molecular cloning: a laboratory manual (Third edition). Cold Spring Harbour Laboratory Press, New York. 2, 344p

Sawers, R. G., Ballantine, S. P., Boxer, D. H. (1985). Differential expression of hydrogenase isoenzymes in *Escherichia coli* K-12: evidence for a third isoenzyme. *J. Bacteriol*. **164**(3):1324-31

Schallmey, M., Singh, A., Ward, O.P. (2004). Developments in the use of Bacillus subtilis species for industrial production. *Can J. Microbiol*. **50**(1):1-17.

Schendel, F. J., Bremmon, C. E., Flickinger, M. C., Guettler, M., Nanson, R. S. (1990). L-lysine production at 50 degrees C by mutants of a newly isolated and characterized methylotrophic *Bacillu*s sp. *Appl. Envir. Microbiol*. **56**(4):963-70

Schirwitz, K., Schmidt, A., Lamzin, V.S. (2007). High-resolution structures of formate dehydrogenase from *Candida boidinii*. *Protein Sci.* **16**(6):1146-56

Schmitz, R. P. and Diekert, G. (2003). Purification and properties of the formate dehydrogenase and characterization of the *fdhA* gene of *Sulfurospirillum multivorans*. *Arch. Microbiol.* **180**(6):394-401

Schut, G. J. and Adams, M. W. W. (2009). The iron hydrogenase of *Thermotoga maritima* utilises ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production. *J. Bacteriol*. **191**(13):4451-7

Serror, P., Sasaki, T., Ehrlich, S.D., Maguin, E. (2002). Electrotransformation of *Lactobacillus delbrueckii subsp bulgaricus* and *L. delbrueckii subsp lactis* with various plasmids. *Appl. Envir. Microbiol.* **68**(1):46-52

Sharp, P.M., Crowe, E., Higgins, D.G., Shields, D.C., Wolfe, K.H., Wright, F. (1988). Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster* and *Homo sapiens*; a review of the considerable within-species diversity. *Nucleic Acids Res*. **16**(17):8207-11

She, Q., Zhang, C., Deng, L., Peng, N., Chen, Z., Ling, Y. X. (2009). Genetic analyses in the hyperthermophilic archaeon *Sulfolobus islandicus*. *Biochem. Soc. Trans*. **37**(1):92-6

Shine, J. and Dalgarno, L. (1975). Determinant of cistron specificity in bacterial ribosomes. *Nature* **254**(5495):34-8

Shinoda, T., Satoh, T., Mineki, S., Lida, M., Tagushi, H. (2002). Cloning, nucleotide sequencing, and expression in *Escherichia coli* of the gene for formate dehydrogenase of *Paracoccus* sp. 12-A, a formate-assimilating bacterium. *Biosci. Biotechnol. Biochem.* **66**(2):271-6

Snedecor, B. and Cooney, C.L. (1974). Thermophilic mixed culture of bacteria utilising methanol for growth. *Appl. Microbiol*. **27**(6):1112-7

Steinfeld. A. (2002). Solar hydrogen production via to-step water-splitting thermochemical cycle based on Zn/ZnO redox reactions. *Int. J. Hydrogen Energy* **27**(6):611-9

Stoots, C.M., O'Brien, J.E., Condie, K.G., Hartvigsen, J.J. (2010). High-temperature electrolysis for large-scale hydrogen production from nuclear energy - Experimental investigations. *Int. J. Hydrogen Energy* **35**(10):4861-70

Studier, F.W., Rosenberg, A.H., Dunn, J.J., Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Meth. Enzymol*. **185**:60-89

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif*. **41**(1):207-34

Takács, M., Toth, A., Bogos, B., Varga, A., Rakhely, G., Kovacs, K. L. (2008). Formate hydrogenlyase in the hyperthermophilic archaeon, *Thermococcus litoralis*. *BMC Microbiol*. **8**(1):88

Takami, H., Takaki, Y., Chee, G.J., Nishi, S., Shimamura, S., Suzuki, H., Matsui, S., Uchiyama, I. (2004). Thermoadaptation trait revealed by the genome sequence of thermophilic *Geobacillus kaustophilus*. *Nucleic Acids Res*. **32**(21):6292-303

Tanaka, N., Kusakabe, Y., Ito, K., Yoshimoto, T., Nakamura, K.T. (2002). Crystal structure of formaldehyde dehydrogenase from *Pseudomonas putida*: the structural origin of the tightly bound cofactor in nicotinoprotein dehydrogenases. *J. Mol. Biol*. **324**(3):519-533

Tate, S. and Dalton, H. (1999). A low-molecular-mass protein from *Methylococcus capsulatus* (Bath) is responsible for the regulation of formaldehyde dehydrogenase activity in vitro. *Microbiology* **145**(1):159-67

Taylor, M. P., Esteban, C. D., Leak, D. J. (2008). Development of a versatile shuttle vector for gene expression in *Geobacillus* spp. *Plasmid* **60**(1):45-52

Taylor, M. P., Eley, K. L., Martin, S., Tuffin, M. I., Burton, S. G., Cowan, D. A. (2009). Thermophilic ethanologenesis: future prospects for second-generation bioethanol production. *Trends Biotechnol*. **27**(7):398-405

Thauer, R. K. (1977). Limitation of microbial H2 formation via fermentation. *In:* Schegel, H. G., Barnea, J. (Eds.) Microbial energy conversion. Pergamon, Oxford. pp 201-4

Thompson, A. H., Studholme, D. J., Green, E. M., Leak, D. J. (2008). Heterologous expression of pyruvate decarboxylase in *Geobacillus thermoglucosidasius*. *Biotechnol. Lett.* **30**(8):1359-1365

Thuy Le, A. T. and Schumann, W. (2007). A novel cold-inducible expression system for *Bacillus subtilis*. *Protein Expr. Purif.* **53**(2):264-9

Tishkov, V. I., Galkin, A. G., Yegorov, A. M. (1991). NAD-dependent formate dehydrogenase from methylotrophic bacterium *Pseudomonas* sp 101 – cloning, expression and study of gene structure. *Doklady Akademii Nauk. SSSR* **317**(3):745-8

Titok, M.A., Chapuis, J., Selezneva, Y.V., Lagodich, A.V., Prokulevich, V.A., Ehrlich, S.D., Jannière, L. (2003). *Bacillus subtilis* soil isolates: plasmid replicon analysis and construction of a new theta-replicating vector. *Plasmid* **49**(1):53-62

Tolga Balta, M., Dincer, I., Hepbasli, A. (2009). Thermodynamic assessment of geothermal energy use in hydrogen production. *Int. J. Hydrogen Energy* **34**(7):2925-39

Trimpin, S. and Brizzard, B. (2009). Analysis of insoluble proteins. *BioTechniques* **46**(5):321-6

Turgeon, N., Laflamme, C., Ho, J., Duchaine, C. (2006). Elaboration of an electroporation protocol for *Bacillus cereus* ATCC 14579. *J. Microbiol Methods* **67**(3):543-8

Ulleberg, Ø., Nakken, T., Eté, A. (2010). The wind/hydrogen demonstration system at Utsira in Norway: Evaluation of system performance using operational data and updated hydrogen energy system modeling tools. *Int. J. Hydrogen Energy* **35**(5):1841-52

van Ginkel, S. and Logan, B. E. (2005). Inhibition of biohydrogen production by undissociated acetic and butyric acids. *Envir. Sci. Technol*. **39**(23):9351-6

van Ophem, P. W., van Beeumen, J., Duine, J. A. (1992). NAD-linked, factor-dependent formaldehyde dehydrogenase or trimeric, Zinc-containing, long-chain alcohol-dehydrogenase from *Amycolatopsis methanolica*. *Eur. J. Biochem*. **206**(2):511-8

Vellanoweth, R. L. and Rabinowitz, J. C. (1992). The Influence of ribosome binding site elements on translational efficiency in *B. subtilis* and *E. coli in vivo*. *Mol. Microbiol*. **6**(9):1105-14

Vignais, P. M., Colbeau, A., Willison, J. C., Jouanneau, Y. (1985). Hydrogenase, nitrogenase, and hydrogen metabolism in the photosynthetic bacteria. *Adv. Microbial Physiol*. **26**:155-234

Vinals, C., Depiereux, E., Feytmans, E. (1993). Prediction of structurally conserved regions of D-specific hydroxy acid dehydrogenases by multiple alignment with formate dehydrogenase. *Biochem. Biophys. Res. Comm*. **192**(1):182-8

Vonck, J., Arfman, N., de Vries, G. E., van Beemen, J., van Bruggen, E. F., Dijkhuizen, L. (1991). Electron microscopic analysis and biochemical characterization of a novel methanol dehydrogenase from the thermotolerant *Bacillus* sp. C1. *J. Biol. Chem*. **266**(6):3949-54

Voskuil, M. I. and Chambliss, G. H. (1993). Rapid isolation and sequencing of purified plasmid DNA from *Bacillus subtilis*. *Appl. Envir. Microbiol*. **59**(4):1138-42

Waks, Z. and Silver, P. A. (2009). Engineering a synthetic dual-organism system for hydrogen production. *Appl. Environ. Microbiol*. **75**(7):1867-75

Wallman, P. H., Thorsness, C. B., Winter, J. D. (1998). Hydrogen production from wastes. *Energy* **23**(4):271-8

Wang, M. Y., Olson, B. H., Chang, J. S. (2008). Relationship among growth parameters for *Clostridium butyricum*, *hydA* gene expression, and biohydrogen production in a sucrose-supplemented batch reactor. *Appl. Microbiol. Biotechnol.* **78**(3):525-32

Wang, M. Y., Tsai, Y. L., Olson, B. H., Chang, J. S. (2008). Monitoring dark hydrogen fermentation performance of indigenous *Clostridium butyricum* by hydrogenase gene expression using RT-PCR and qPCR. *Int. J. Hydrogen Energy* **33**(18):4730-8

Westers, L., Westers, H., Quax, W. J. (2004). *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochim. Biophys. Acta* **1694**(1-3):299-310

Whittenbury, R. and Dalton, H. (1981). The methylotrophic bacteria. *In:* Starr, M. P., Stolp, H., Trüper, H. G., Balows, A. and Schlegel, H. G. (Eds.) The prokaryotes. Springer, New York, N.Y. pp 894-902

Wilson, G. A. and Bott, K. F. (1968). Nutritional factors influencing the development of competence in the *Bacillus subtilis* transformation system. *J. Bacteriol*. 95(4):1439-49.

Wittke, A. A., Lick, S., Heller, K. J. (2002). Transformation of *Bacillus subtilis* in chocolate milk: evidence for low frequency of establishment of cells transformed under non-selective conditions. *Syst. Appl. Microbiol.* **25**(4):478-82

Woodward, J., Mattingly, S. M., Danson, M., Hough, D., Ward, N., Adams, M. W. W. (1996). *In vitro* hydrogen production by glucose dehydrogenase and hydrogenase. *Nature Biotechnol*. **14**(7):872-4

Woodward, J., Cordray, K., Edmonston, R. J., Blanco-Rivera, M., Mattingly, S. M., Evans, B. R. (2000). Enzymatic hydrogen production: conversion of renewable resources for energy production. *Energy and Fuels* **14**:197-201

Woodward, J., Orr, M., Cordray, K., Greenbaum, E. (2000). Enzymatic production of biohydrogen. *Nature* **405**(6790):1014-5

Worin, N. A., Lissolo, T., Colbeau, A., Vignais, P. M. (1996). Increased H_2 photoproduction by *Rhodobacter capsulatus* strain deficient in uptake hydrogenase. *J. Mar. Biotechnol*. **4**:28-33

Wu, S. C. and Wong, S. L. (1999). Development of improved pUB110-based vectors for expression and secretion studies in *Bacillus subtilis*. *J. Biotechnol*. **72**:185-95

Xue, G. P., Johnson, J. S., Dalrymple, B. P. (1999). High osmolarity improves the electro-transformation efficiency of the gram-positive bacteria *Bacillus subtilis* and *Bacillus licheniformis*. *J. Microbiol. Methods* **34**:183-91

Yansura, D. G. and Henner, D. J. (1984). Use of the *Escherichia coli lac* repressor and operator to control gene-expression in *Bacillus subtilis*. *PNAS USA* **81**(2):439-43

Yasbin, R. E. and Young, F. E. (1974). Transduction in *Bacillus subtilis* by bacteriophage Spp1. *J. Virol.* **14**(6):1343-8

Ye, X. H., Wang, Y. R., Hopkins, R. C., Adams, M. W. W., Evans, B. R., Mielenz, J. R., Zhang, Y. H. P. (2009). Spontaneous High-Yield Production of Hydrogen from Cellulosic Materials and Water Catalyzed by Enzyme Cocktails. *Chem. Sus. Chem.* **2**(2);149-52

Yetis, M., Gündüz, U., Eroğlu, I., Yücel, M., Türker, L. (2000). Photoproduction of hydrogen from sugar refinery wastewater by *Rhodobacter sphaeroides* O.U. 001. *Int. J. Hydrogen Energy* **25**(11):1035-41

Yokoi, H., Maki, R., Hirose, J., Hayashi, S. (2002). Microbial production of hydrogen from starch-manufacturing wastes. Biomass and Bioenergy **22**(5):389-95

Yoon, J. H., Hae Shin, J., Kim, M. S., Jun Sim, S., Park, T. H. (2006). Evaluation of conversion efficiency of light to hydrogen energy by Anabaena variabilis. *Int. J. Hydrogen Energy* **31**(6):721-7

Yoshida, A., Nishimura, T., Kawaguchi, H., Inui, M., Yukawa, H. (2007). Efficient induction of formate hydrogen lyase of aerobically grown *Escherichia coli* in a three-step biohydrogen production process. *Appl. Microbiol. Biotechnol*. **74**(4):754-60

Young, F. E. 1967. Competence in *Bacillus subtilis* transformation system. *Nature* **213**(5078):773-5

Yurimoto, H., Kato, N., Sakai, Y. (2009). Genomic organization and biochemistry of the ribulose monophosphate pathway and its application in biotechnology. *Appl. Microbiol. Biotechnol.* **84**(3):407-16

Zahn, J. A., Bergmann, D. J., Boyd, J. M., Kunz, R. C., DiSpirito, A. A. (2001). Membrane-associated quinoprotein formaldehyde dehydrogenase from *Methylococcus capsulatus* (Bath). *J. Bacteriol*. **183**(23):6832-40

Zhang, Y., Olsen, D. R., Nguyen, K. B., Olson, P. S., Rhodes, E. T., Mascarenhas, D. (1998). Expression of eukaryotic proteins in soluble from *Escherichia coli*. *Prot. Expr. Purif*. **12**(2):159-65

Zhang, Z. and Gurr, S. J. (2000). Walking into the unknown: a 'step down' PCR-based technique leading to the direct sequence analysis of flanking genomic DNA. *Gene* **253**(2):145-50

Zhang, H.-Y. P., Evans, B. R., Mielenz, J. R., Hopkins, R. C., Adams, M. W. W. (2007). High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. *PLoS ONE* **2**(5):e456

Zhang, C., Ma, K., Zing, X. H. (2009). Regulation of hydrogen production by Enterobacter aerogenes by external NADH and NAD+. *Int. J. Hydrogen Energy* **34**(3):1226-32

Zhao, C. X., O-Thong, S., Karakashev, D., Angelidaki, I., Lu, W. J., Wang, H. T. (2009). High yield simultaneous hydrogen and ethanol production under extreme-thermophilic (70 °C) mixed culture environment. *Int. J. Hydrogen Energy* **34**(14):5657-65

arpa-e.gov US DoE Advanced Research Project Agency Energy

www.eere.energy.gov US DoE Energy Efficiency & Renewable Energy

www.hm-treasury.gov.uk/stern_review_report.htm Stern review

www.hydrogen.energy.gov US DoE, Hydrogen program

www.ipcc.ch Intergovernmental Panel on Climate Change

www.npl.co.uk/ National Physical Laboratory

unfccc.int/kyoto_protocol/items/2830.php Kyoto protocol

unfccc.int/meetings/cop_15/items/5257.php Copenhagen Accord

www.worldenergy.org World Energy Council

www.worldmapper.org worldmapper resource

APPENDICES

Methanol dehydrogenase, *Bacillus methanolicus* PB1

>BmMDH
MTNFFIPPASVIGRGAVKEVGTRLKQIGATKALIVTDAFLHGTGLSEEVAKNIREAGLDAVI
FPKAQPDPADTQVHEGVDIFKQEKCDALVSIGGGSSHDTAKAIGLVAANGGRINDYQGVNSV
EKPVVPVVAITTTAGTGSETTSLAVITDSARKVKMPVIDEKITPTVAIVDPELMVKKPAGLT
IATGMDALSHAIEAYVAKRATPVTDAFAIQAMKLINEYLPRAVANGEDIEAREAMAYAQYMA
GVAFNNGGLGLVHSISHQVGGVYKLQHGICNSVNMPHVCQFNLIARTERFAHIAELLGENVS
GLSTASAAERAIVALQRYNKNFGIPSGYAEMGVKEEDIELLANNAYQDVCTLDNPRVPTVQD
IAQIIKNAL*

>*Bmmdh*
atgacaaacttttcattccaccagctagcgtaattggacgcggcgctgtaaaagaagtagg
aacaagacttaagcaaattggagctacaaaagcacttatcgttacagatgcatttcttcatg
gcacaggtttgtcagaagaagttgctaaaaacattcgtgaagctggccttgatgctgtaatt
ttcccaaaagctcaaccagatccagcagatacacaagttcatgaaggcgtagatatattcaa
acaagaaaaatgtgatgcacttgtttctatcggtggaggtagctctcacgatacagcaaaag
caatcggtttagttgcagcaaacggcggaagaatcaacgactatcaaggtgtaaacagtgta
gaaaaaccggttgttccagtagttgcaatcactacaacagctggtactggtagtgaaacaac
atctcttgcggttattacagattctgcacgtaaagtaaaaatgccagttatcgatgagaaaa
ttacaccaactgtagcaattgttgacccagaattaatggtgaaaaaaccagctggattaaca
attgcaactggtatggatgcattatcccatgcaattgaagcatatgttgcaaaacgtgctac
accagttactgatgcgtttgcaattcaagcaatgaaactcattaatgaatacttaccacgtg
cggttgcaaatggagaagacatcgaagcacgtgaagcaatggcttatgcacaatacatggca
ggagtggcatttaacaacggaggtttaggattagtacactctatttctcaccaagtaggtgg
agtttacaagttacaacacggaatctgtaactcagttaatatgccacacgtttgccaattca
acttaattgctcgtactgaacgcttcgcacacattgctgagcttttaggcgagaatgtttct
ggcttaagcactgcatctgctgctgagagagcaattgtagcgcttcaacgctataacaaaaa
cttcggtatcccatctggctatgcagaaatgggcgtaaaagaagaggatatcgaattattag
cgaacaacgcgtaccaagacgtatgtactctagataacccacgtgttcctactgttcaagac
attgcacaaatcatcaaaaacgctctg**taa**

>pBMMDH
mttnffippasvigrgavkevgtrlkqigatkalivtdaflhgtglseevaknireagldav
ifpkaqpdpadtqvhegvdifkqekcdalvsigggsshdtakaiglvaanggrindyqgvns
vekpvvpvvaitttagtgsettslavitdsarkvkmpvidekitptvaivdpelmvkkpagl
tiatgmdalshaieayvakratpvtdafaiqamklineylpravangediearеamayaqym
agvafnngglglvhsishqvggvyklqhgicnsvnmphvcqfnliarterfahiaellgenv
sglstasaaeraivalqrynknfgipsgyaemgvkeediellannayqdvctldnprvptvq
diaqiiknam*

>*pBMmdh*
atgacaacaaacttttcattccaccagctagcgtaattggacgcggcgctgtaaaagaagt
aggaacaagacttaagcaaattggagctacaaaagcacttatcgttacagatgcatttcttc
atggcacaggtttgtcagaagaagttgctaaaaacattcgtgaagctggccttgatgctgta
attttcccaaaagctcaaccagatccagcagatacacaagttcatgaaggcgtagatatatt
caaacaagaaaaatgtgatgcacttgtttctatcggtggaggtagctctcacgatacagcaa
aagcaatcggtttagttgcagcaaacggcggaagaatcaacgactatcaaggtgtaaacagt

gtagaaaaaccggttgttccagtagttgcaatcactacaacagctggtactggtagtgaaac
aacatctcttgcggttattacagattctgcacgtaaagtaaaaatgccagttatcgatgaga
aaattacaccaactgtagcaattgttgacccagaattaatggtgaaaaaaccagctggatta
acaattgcaactggtatggatgcattatcccatgcaattgaagcatatgttgcaaaacgtgc
tacaccagttactgatgcgtttgcaattcaagcaatgaaactcattaatgaatacttaccac
gtgcggttgcaaatggagaagacatcgaagcacgtgaagcaatggcttatgcacaatacatg
gcaggagtggcatttaacaacggaggtttaggattagtacactctatttctcaccaagtagg
tggagtttacaaattacaacacggaatctgtaactcagttaatatgccacacgtttgccaat
tcaacttaattgctcgtactgaacgcttcgcacacattgctgagcttttaggcgagaatgtt
tctggcttaagcactgcatctgctgctgagagagcaattgtagcgcttcaacgctataacaa
aaacttcggtatcccatctggctatgcagaaatgggcgtaaaagaagaggatatcgaattat
tagcgaacaacgcgtaccaagacgtatgtactctagataacccacgtgttcctactgttcaa
gacattgcacaaatcatcaaaaacgctatg**taa**

Activator, *Bacillus methanolicus* PB1

>BmACT
MGKLFEEKTIKTEQIFSGRVVKLQVDDVELPNGQTSKREIVRHPGAVAVIAVTNENKIVMVE
QYRKPLEKSIVEIPAGKLEKGEDPRITALRELEEETGYQCEQMEWLISFATSPGFADEIIHL
YVAKGLSKKENAAGLDEDEFVDLIELTLEEAXQYIKEQRIYDSKTVIAVQYLQLQEALKHK*

>Bmact
atgggaaaattatttgaggaaaaaacgataaaaacagaacaaatttttcgggaagagttgt
caaattgcaagtagacgatgttgagctaccaaacggacaaacatcaaagcgggaaattgtta
ggcatcctggtgctgtagctgtaatcgcagtaacaaatgaaaataaaatcgtgatggtggaa
caataccggaagccactagaaaaatcaatagttgaaattcctgccggtaaactggaaaaagg
tgaagatccgcgtattaccgcattgcgcgagttagaagaggaaacgggatatcaatgcgaac
agatggaatggttgatctcttttgcgacctctcctggttttgcggatgaaattattcattta
tatgtagcaaaaggactgtccaaaaaggaaaatgcagcaggactggatgaggatgagtttgt
cgatcttattgaactgacacttgaggaagcgcttcaatatataaaagaacagagaatctacg
attcaaaaactgtaattgccgttcaatatttgcagcttcaagaggctctcaaacataaa**tga**

Formaldehyde dehydrogenase, *Bacillus methanolicus* PB1

>BmFALDDH
MKAVTYQGPKTLVVKEVEDAKIEKKDDVIVRITSTAICGSDLHIYQGNFPLPEGYIIGHEPM
GIVEEVGPEVTKVKKGDRVVVPFNVACGHCFYCEHDMTSQCDNANDNYDSGGYFGFGEKYGN
HPGGQAEYLKVPFGNFTPFVIPESCELEDESLLFLSDVLPTAWWSVENADVKKGDTVIVLGC
GPVGLMAQKFAWMKGAKRVIAVDYLDYRLDHAKRMNNVEVFNFTKNDDMGEYLKEITQGGAD
KVIDAVGMDGKKTPLEFVGQKLKLQGGTLSPLKIAAKAVRKNGTIQITGVYGGNYNNFELGA
LWIRNVTIKMGQAPVIPIMPKLFDKIMNKEFDPKEIITHKIPLNDASHGYKIFNDHQDNCIK
VVFKP*

>*Bmfalddh*
gtgcctttggaaatttcacccctttgtcatcccggaatcgtgtgaactggaagatgaatct
tgctctaaacctcgttcaaaaaaaatgcagaataaagttggtcaagaggaacatattgaata
tttagctcgtagttttcatgagagtcgattgccaagaaaacccacgccacctacaacggttc
ctgatgaggtggttagcatagttcttaatataagttttaatatacagcctgaaaatcttgag
agaataaaagaagaacatcgattttccatggcagctgagtaatattgtaggagatcttctaa
aggaggtgctctagaatgaaagctgtaacatatcaaggacctaaaaccttagtagtaaagga
agttgaagatgcaaaaatagaaaaaaaagatgatgttatcgtacgaataacatctacagcca
tatgcgggtcagaccttcacatctatcaaggaaactttccgttgcctgaaggttatattatt

ggacacgaaccgatggggatcgtggaagaagtcggccctgaagtaacaaaagttaagaaagg
ggaccgtgttgttgtcccgtttaatgttgcctgtggacattgcttttattgcgaacatgata
tgacgagccaatgtgataatgccaatgataattatgattcaggagggtattttggttttgga
gaaaagtacggaaatcatccgggaggccaggcggaatatttaaaagtgccttttggaaattt
cacccctttgtcatcccggaatcgtgtgaactggaagatgaatcccttctcttttatctg
atgtcttgcctactgcctggtggagtgtagaaaatgccgatgtaaaaaaaggggataccgtt
attgtgcttggctgtggacctgtcggtttaatggcgcaaaagtttgcctggatgaaaggtgc
gaaacgtgtaattgctgtcgattacttagattatcgcttagaccatgcaaaacgaatgaaca
atgttgaggtgtttaattttacaaaaaacgatgacatgggcgagtatttgaaggaaataact
caaggaggagcagacaaggtaattgatgccgtcggtatggatggaaaaaaaacacccctttga
attcgttgggcaaaaattaaagctgcaaggcggaaccctaagtcctttaaagattgctgcga
aagctgtaagaaaaaacggaacgattcaaattaccggtgtatatggaggcaactacaataat
tttgaattaggagccttatggataagaaatgttacaattaaaatgggacaggcaccggttat
cccaattatgccgaagctatttgataaaattatgaataaagaatttgatccaaaagaaatta
taacacacaaaatccctcttaatgatgcaagccatggctataagattttcaacgatcatcaa
gataattgcatagtcgtttttaaacct**taa**ggctggcgatggattcatacaccgggacattc
tcccggacatatttccttgttcagagatgaagatcgttccctaattgccggtgacgcatttg
ttacggtgaaacaagagtcgctttacaaggtagtgacgcaaattcaggaaatcagcggtccg
ccgagatattttacaaccgattgggaggcagcgaaggagtccgtcaaaaagcttgaagcttt
aaaaccgacagttgcaattaccgggcacggaatgccaatgtccggcggcgaacttacaaaag
gcctggaaaaattggcaagagactttgaacgtatagcaattccggattatgggcgctatgtt
tatcgaaagcattgattttcgaataactcaaaaaaatttattaaaacgtgtaaaaaaatgta
acaagttaatttcttcttgattggagtgtaaaaaatgaaagctgtaacatatcaaggaccta
aaaccttagtagtaaaggaagttgaagatgcaaaaatagaaaaaaaagatgatgttatcgta
cgaataacatctacagccatatgcgggtcagaccttcacatctatcaaggaaacaatcacta
gtgaattcgcggccgcctgcaggtcgaccatatgggagagctcccaacgcgttggatgcata
gcttgagtattctatagtgtcacctaaatagcttggcgtaatcatggtcatagctgtttcct
gtgtgaaattgttatccgctcacaattccacacaacatacgagccggaagcat

(the alternative putative start codon (atg) are indicated in blue. Putative RBS and TATA box are underlined)

Formate dehydrogenase, *Geobacillus thermoglucosidasius*

>GthFDHa
MLKNYSTMICIICIHGGVFILMIRRFGLSMNEQLITITINGTDYNAKQGMTILEIVNENGIP
HPQVCYTPELGAIQTCDTCIVEVDGKLVRACSTPAENGMNVELASARAKAAQKEAMDRILEN
HLLYCTVCDNNNGNCKLHNTAELMGIEHQSYPYRPKVDPSEVDMSHPFYRYDPNQCIACGQC
VEVCQNLQVNETLSIDWEAERPRVIWDNGVPINESSCVSCGQCVTVCPCNALMEKSMLGEAG
FMTGLDQEILNPMINFVKEVEPNYTSIFAVSEIEAAMRKQRIKKTKTVCTFCGVGCSFEVWT
KGRKILKIQPVSEAPVNAISTCVKGKFGWDFVNSEARLTKPLIRKGDTFVESTWEEALSLVA
EKLGEIKQKYGGNAIGFISSSKISNEENYLMQKLARQVFETNNVDNCSRYCQSPATDGLFRT
VGMGGDSGTIHDIASAGLVIIIGANPAEGHPVLATRVKRAHKLFGQKLIVADLRKNEMAERA
DLFIRPKQGTDQVWLMAVTKYIIDQGWHDEKFIRERVHFFDEFKQVLEKYTLDYAEEVTGVA
KEDLIRIAEMIHEADGTCVLWGMGVTQNTGGSDTSAAISNLLLATGNYGRPGAGAFPLRGHN
NVQGACDMGSLPGWLPGYQHITDDIARAKFEKAYGVRIDGKPGLDNIQMIEAAEQGKLKAMY
IVGEDMAFVDCNANHVQKVLSELDFLVVQDIFLSKTAQFADVVLPATPSLEKEGTFTNTERR
IQRFYQALEPLGDSKPDWWIIQEIAKRLGADWNYAGPKEIMDEIASLAPLFSQAQYENLEGW
NSLCWGSYDGAHTPILYKERFNFPDGKARFALADWVQPTEYPEEYDLLVNNGRLLEHFHEGN
LTYKSQGILKKFPEVFVEVSPELAKERGIKDGSLVRLESPFGHVKVRVLVTDRVKGKELFLP
MHSATNESAINILTGPATDHRTNTPAFKQTKVRMQVLEVEGTPPLPRTNPRFKKRNPRKGVE
VERKWRRPDYVPLTDKWKEAEPRG*

>*GthfdhA*

atgctaaaaaactactccaccatgatttgcataatctgcattcatggtggagtttttatttt
aatgataaggaggtttggtttaagcatgaacgaacagttgataacgattacaataaacggaa
cggattacaatgccaaacaaggaatgacgattttagaaatcgttaacgaaaacggcattccg
catccgcaagtttgctatactcctgaacttggagccattcagacctgtgatacgtgcatcgt
cgaggtagacggaaaattagtgcgcgcctgttcgacccctgcagaaaacgggatgaatgtcg
aactagcttcagcacgggcgaaagcagcgcaaaaagaagcaatggaccgcattttagaaaac
cacttattatattgtactgtatgtgacaacaataacgggaactgcaaactgcataatacggc
ggaattgatgggaattgagcaccagtcgtatccatatcgcccaaaagtggatccatcggaag
tcgatatgtcccatccgttctatcgttatgatccaaaccaatgcatcgcctgcggacaatgt
gtcgaagtatgtcaaaacttgcaagtaaacgaaacgttatcgatcgactgggaggcggaacg
tcctcgcgtcatttgggacaacggtgttccgatcaatgaatcttcctgtgtcagctgtggac
agtgcgtcaccgtttgtccatgcaatgcgttaatggaaaaatcgatgctaggcgaagccggg
tttatgactggattagaccaagaaatcttaaatccgatgatcaattttgtcaaagaagtgga
gccgaactacacgagcatttttgccgtttccgaaatcgaagcggcgatgcggaaacagcgca
tcaaaaagacgaaaacggtttgcacgttctgcggtgttggctgttcgtttgaagtctggaca
aaaggacggaagattttgaaaatccagcctgtctccgaagcgccggtcaatgcgatttccac
atgtgttaaaggaaaattcggctgggatttcgtcaacagcgaagcacgccttacaaaaccgc
tcatccgcaaaggtgatacgtttgtagagtcaacgtgggaagaagcgctttcgcttgttgcg
gaaaaacttggcgaaattaagcaaaaatacggcggcaatgccattggctttatttcatcatc
aaaaatatctaatgaagaaaactatttaatgcaaaaattggcccggcaagtgtttgaaacga
acaatgtcgacaactgttcacgctattgccaatctccggcaaccgacggattgttccgcaca
gtaggaatgggcggcgattccggaacgattcatgatatcgcatcagctgggcttgtcattat
tatcggcgctaacccagcagaaggtcaccctgttttagcgacccgcgtcaaacgcgcccaca
aactgttcggccaaaaactaattgttgccgacttgcgcaaaaacgaaatggcagaacgcgcc
gatctatttatcagaccaaaacaaggcactgaccaagtatggttgatggcggtcacgaaata
tatcatcgaccaaggatggcacgacgaaaaattcattcgcgaacgcgttcatttcttcgacg
aattcaaacaagtgcttgaaaaatatacgctcgattacgcagaagaagtaacaggagttgcg
aaagaagacttgattcgcatcgctgaaatgatccatgaagcggatggcacgtgcgtcctgtg
gggaatgggagtcacgcaaaacacgggaggaagcgatacatctgcagccatctcgaacttat
tgctcgcaaccggcaactacggccgcccaggcgccggagcgttcccgctgcgcggtcataac
aatgtgcaaggcgcttgtgatatgggctcgctccctggctggctcccaggataccaacatat
tacagatgacatcgcgcgggctaaatttgaaaaagcatacggtgtccgcatcgatggaaaac
cagggcttgacaacattcaaatgattgaggctgctgaacaaggaaaattaaaagcaatgtat
atcgtcggtgaagatatggcgtttgtcgactgcaacgccaatcatgtccaaaaagtgctttc
cgaattagacttcttagtagtacaagacattttcctatcgaaaacagctcagtttgctgatg
ttgttttgccagcaacaccaagcttagaaaaagaaggaacgtttaccaatacggaacgccgc
attcaacgttttaccaagcgcttgagccgcttggtgattcaaagccagactggtggattat
tcaagaaattgcaaaacggctgggcgctgattggaactacgccggtccaaaagaaattatgg
atgaaattgcaagcctcgcgccactattctcgcaagcgcaatatgaaaacttggaaggatgg
aacagcctctgctggggcagctatgacggtgcgcatacaccaattctttataaagaacgctt
caacttcccggatggaaaagcacgctttgcgcttgccgattgggtgcaaccgacagaatatc
cggaagaatatgacttgcttgtcaataacgggcggttgcttgaacatttccatgaaggaaat
ttaacatacaaatcgcaaggcattctgaaaaaattcccggaagtgttcgtcgaagtatcgcc
ggaacttgcgaaagaacgcggcatcaaagacggctcgctcgttcgtttagaatcgccattcg
gccacgtgaaagtgcgcgtgctcgtcacagaccgcgtcaaaggaaaagaattgttcctgcca
atgcactcggcgacgaacgaaagcgccattaacatccttaccggccctgcgacggaccatcg
caccaacacgccggcattcaaacaaacaaaagtacggatgcaagtgttagaagtcgagggca
cgcctccgcttccgcgcacaaatccgcgctttaaaaagcggaatccgcgtaaaggcgtcgaa
gtcgaacgaaaatggcggcgcccggattacgttcctctaacggataaatggaaggaggcaga
gccgcgtggc**taa**