University of Sheffield

D. A. G. Skibinski

Functional and regulatory analysis of the 12-gene *hyf* operon of *Escherichia coli*

A thesis submitted for the degree of Doctor of Philosophy

2001

Functional and regulatory analysis of the 12-gene *hyf* operon of *Escherichia coli*

David Alexander George Skibinski, BSc.
(University of Warwick)

A thesis submitted for the degree of Doctor of Philosophy

Department of Molecular Biology and Biotechnology,
University of Sheffield

August 2001

Functional and regulatory analysis of the 12-gene *hyf* operon of *Escherichia coli*

David Skibinski

Sequence analysis of the 55.8-56.0 min region of the *Escherichia coli* genome has revealed a 12-gene operon designated the *hyf* operon (*hyfABCDEFGHIR-focB*). The *hyf* operon encodes a putative ten-subunit hydrogenlyase complex (hydrogenase four or Hyf), a potential formate sensing σ^{54}-dependent transcriptional activator, HyfR (related to FhlA), and a possible formate transporter, FocB (related to FocA). It has been proposed that Hyf in conjunction with Fdh-H forms a second formate hydrogenlyase pathway (Fhl-2) in *Escherichia coli*, which unlike the *hyc* operon encoded Fhl pathway (Fhl-1) is a respiration-linked proton translocating Fhl complex.

Initial experiments directly investigated these proposals and were conducted with *hyf* and hydrogenase-1, -2 and -3 mutants grown under *hyf* operon optimal transcriptional activation conditions. Radiolabelling experiments with ^{63}Ni did not detect the proposed large subunit of hydrogenase-4, despite the detection of ^{63}Ni-associated polypeptides likely to correspond to the large subunits of hydrogenase-1, -2 and -3. Also, Fdh-H, hydrogenase and hydrogen production assays detected no activity attributable to the *hyf* operon. Immunoblotting experiments with anti-HycE and anti-Hyf sera did not detect Hyf polypeptides, suggesting that expression of the *hyf* operon was very low under optimal transcriptional activation conditions.

Transcriptional analysis of the *hyf* operon using a *hyfA-lacZ* transcriptional fusion showed that, like the *hyc* operon, the *hyf* operon is induced by formate at low pH via the formate sensing, σ^{54}-dependent transcriptional activator FhlA. The proposed transcriptional activator HyfR was also found to activate *hyf* operon transcription in a σ^{54}-dependent manner. However the co-effector(s) used by HyfR has yet to be identified.

Finally bioreactors were used to analyse the growth and metabolism of *hyf* mutants. However, no differences in growth and metabolism attributable to the *hyf*

operon were observed during anaerobic controlled batch cultivation and both aerobic and anaerobic glucose-limited chemostat cultivation.

PRESENTATIONS

Skibinski, D. A. G., Golby, P., Berks, B. C., Attwood, M. M., Guest, J. R. & Andrews, S. C. (2001). Hydrogenase-4 of *E. coli*. Oral presentation abstract at EC Workshop (COST) – Biodiversity of Hydrogenases. University of Reading.

Skibinski, D. A. G., Golby, P., Berks, B. C., Attwood, M. M., Guest, J. R. & Andrews, S. C. (2000). Hydrogenase-4 of *E. coli*. Oral presentation and poster abstract at 6th International Conference on the Molecular Biology of Hydrogenases. Potsdam, Berlin.

Skibinski, D. A. G., Golby, P., Berks, B. C., Attwood, M. M., Guest, J. R. & Andrews, S. C. (1998). Hydrogenase-4 of *E. coli*. Oral presentation at the Biochemical Society Meeting 665. University of Southampton.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr Margaret Attwood and Dr Simon Andrews (University of Reading) for their guidance over the last four years. I also thank Dr Paul Golby and Andy (Y. S. Chang) for their technical assistance and for making some of the stains and plasmids used in this study. Thanks also to Dr Jeff Green, Professor Robert Poole and Professor John Guest for their help and advice.

I would also like to thank Dr Frank Sargent (University of East Anglia) for his technical assistance, encouragement and for donating some of the strains used in this study.

I thank Professor August Böck (Universität München) for donating many of the strains and plasmids used in this study.

I acknowledge the BBSRC for it's funding.

Last, but by no means least, I thank my parents for their unwavering support, help and love.

CONTENTS

List of Figures

List of Tables

COMMON ABBREVIATIONS

ADH	Alcohol dehydrogenase
Al-DH	Aldehyde dehydrogenase
ATP	Adenosine triphosphate
CoA	Co-enzyme A
CTAB	Cetyl-trimethyl-ammonium bromide
°C	Degrees centigrade
DCIP	2,6-dichloroindolphenol
DMF	Dimethyl formamide
DMK	Demethylmenaquinone
DMSO	Dimethyl sulfoxide
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
e^-	Electron
EDTA	Disodium ethylene diamine tetra-acetate
Fdh	Formate dehydrogenase
FdhF	Formate dehydrogenase-H large subunit
Fdh-H	Formate dehydrogenase-H
Fhl	Formate hydrogenlyase
Fhl-1	*hyc* operon encoded formate hydrogenlyase complex
Fhl-2	*hyf* operon encoded formate hydrogenlyase complex
Fr	Fumarate reductase
FTN	Ferritin
g	Gravitational force (centrifugal force)
GMP	Guanine mononucleotide phosphate
GOD	Glucose oxidase
h	Hour
H^+	Proton
HPLC	High pressure liquid chromatography
Hyc	Hydrogenase-3 complex
Hyf	Hydrogenase-4 complex
IHF	Integration host factor

kD	Kilodaltons
l	Litre
min	Minute
μ	Micro
m	Milli
M	Molar
MGD	Molybdenum guanine dinucleotide cofactor
MK	Menaquinone
mol	Mole
MOPS	3-(N-morpholine) ethane sulphuric acid
n	Nano
NAD	Nicotinamide adenine dinucleotide (oxidised)
NADH	Nicotinamide adenine dinucleotide (reduced)
NBT	Nitro blue tetrazolium
Nr	Nitrate reductase
OD	Optical density
ONPG	O-nitrophenyl-β-D-galactosidase
OX	Oxidised
Pdh	Pyruvate dehydrogenase
PCR	Polymerase chain reaction
Pfl	Pyruvate formatelyase
pfu	Plaque forming units
POD	Peroxidase
Q	Quinone
rpm	Revolutions per minute
RED	Reduced
TEMED	N, N, N', N'-tetramethylethylenediamine
TMAO	Trimethylamine N-oxide
TYEP	Tryptone yeast extract phosphate
SDS	Sodium diethyl sulphate
U	Units
UQ	Ubiquinone
URS	Upstream regulatory sequence

V	Volts
vrs	Versus
v/v	Volume/volume
w/v	Weight/volume
X-gal	5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside

1. INTRODUCTION

1.1 Formate metabolism in *Escherichia coli*

In *Escherichia coli* pyruvate is generated during both aerobic and anaerobic growth, from the oxidation of hexoses (such as glucose) by the Embden-Meyerhof-Parnas pathway (Böck & Sawers, 1996). During aerobic growth, the oxidative decarboxylation of pyruvate by the pyruvate dehydrogenase (Pdh) complex generates acetyl CoA, CO_2 and one molecule of NADH. The reducing equivalents generated by this reaction are passed from NADH to O_2 via the electron transport respiratory chain generating a proton motive force which in turn is utilised by other membrane proteins to do 'work' (Guest *et al.*, 1989; Kaiser & Sawers, 1994; Böck & Sawers, 1996). During anaerobic growth however there is no exogenous electron acceptor able to reoxidise NADH and as a consequence the Pdh complex is functionally restricted to respiratory metabolism. Therefore during fermentative growth the nonoxidative cleavage of pyruvate to acetyl coenzyme A and formate by the pyruvate formate-lyase (Pfl) enzyme is of more benefit to *E. coli* (Fig. 1.1; Knappe & Sawers, 1990; Kessler & Knappe, 1996; Böck & Sawers, 1996). The acetyl coenzyme A produced has two alternative fates. It may either be reduced to ethanol, reoxidising two molecules of NADH, or be converted to acetate, generating a molecule of ATP. The initial part of this chapter however is concerned with the fate of the formate produced by the pyruvate formate lyase reaction.

1.1.1 Formate transport

Formate is exported from and imported into the cell by FocA (formate channel), a 31 kD integral membrane protein, functioning as a specific bi-directional formate channel (Suppmann & Sawers, 1994). The mechanisms by which FocA transports formate have not yet been identified, but it has been speculated to be via a proton symport-based mechanism, a theory supported by the observation that formate is imported into the cell at low external pH values (Rossmann *et al.*, 1991; Suppmann & Sawers, 1994). It has also been suggested that formate uptake by FocA is coupled to formate metabolism, since a *selC* mutant unable to synthesize any of the three

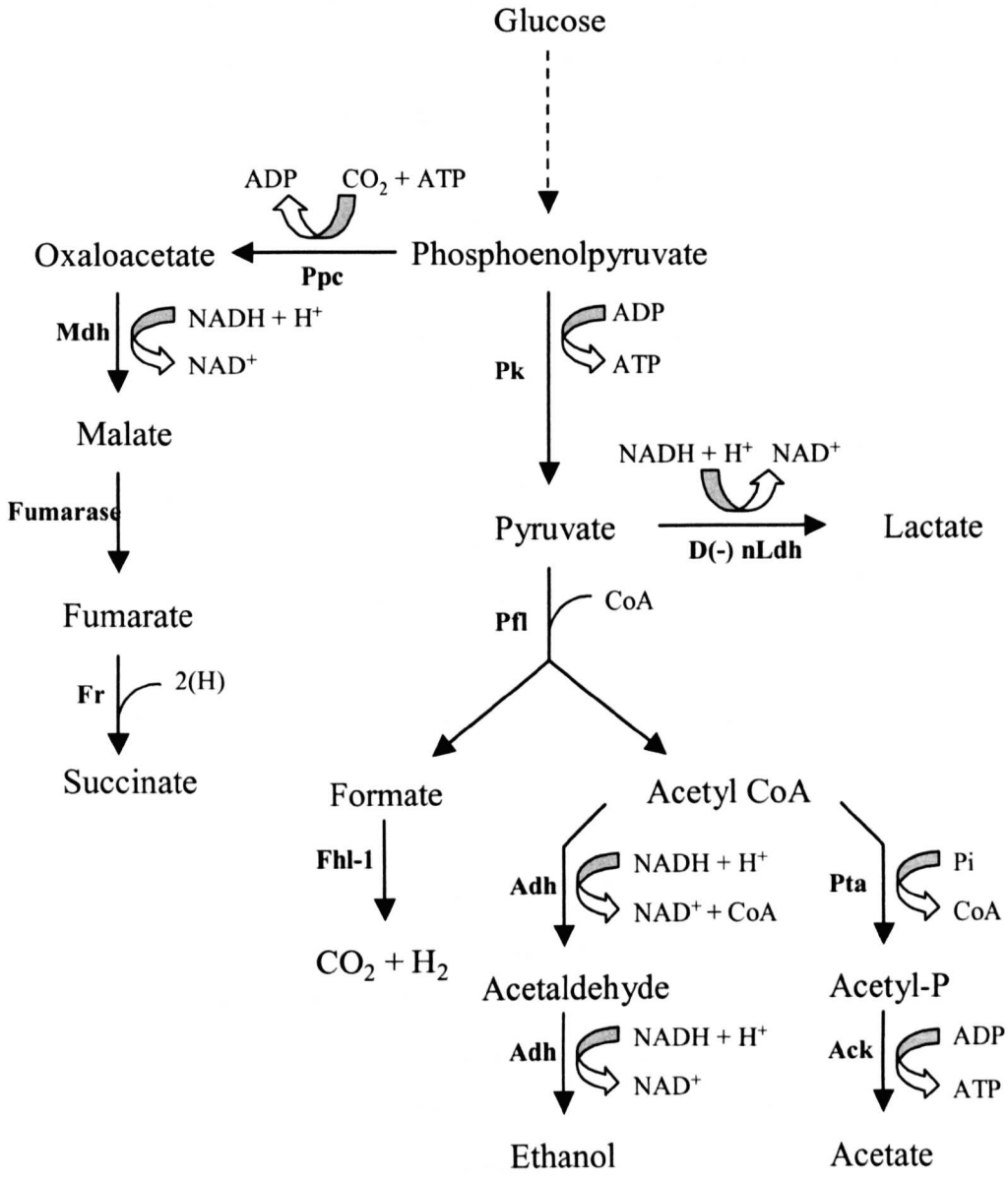


Fig. 1.1. The fermentation products of *E. coli*. Abbreviations: **Ack**, acetate kinase; **Adh**, alcohol dehydrogenase; **Fhl-1**, formate hydrogenlyase; **Fr**, fumarate reductase; **D-(-)-Ldh**, NAD^+-dependent D-(-)-lactate dehydrogenase; **Mdh**, malate dehydrogenase; **Ppc**, phosphoenolpyruvate carboxylase; **Pfl**, pyruvate formate lyase; **Pk**, pyruvate kinase; **Pta**, phosphotransacetylase; Pi, phosphate. The broken arrow signifies the oxidation of glucose to phosphoenolpyruvate by the Embden-Meyerhof-Parnas pathway.

selenium containing Fdh isoenzymes (Fdh-N, section 1.1.2.1; Fdh-O, section 1.1.2.5; Fdh-H, see section 1.1.5.1) does not re-import excreted formate (Suppmann & Sawers, 1994).

FocA is encoded by the *focA* gene, which forms an operon with the gene encoding the enzyme PFL (section 1.1; Table 1.1; Sawers & Böck, 1989; Suppmann & Sawers, 1994). Expression of the *focA-pfl* operon is induced during anaerobic growth with transcriptional levels increased 10-15 fold over levels observed during aerobic growth (Sawers & Böck, 1988). This anaerobic induction of expression is positively regulated by two global transcriptional regulators, ArcA (anaerobic respiration control) and Fnr (fumarate nitrate reduction) (Table 1.2; Sawers & Böck, 1988, 1989; Sawers & Suppmann, 1992).

The existence of at least one other system involved in transporting formate out of and into the cell has been proposed as formate can still exit and enter the cell in a *focA* mutant (Suppmann & Sawers, 1994). This second system, unlike *focA*, is thought to be synthesized in both aerobically and anaerobically grown cells (Suppmann & Sawers, 1994).

Formate has a low pKa (3.75) and at physiological pH little is present in the undissociated, membrane permeable form (Suppmann & Sawers, 1994).

1.1.2 The formate-nitrate respiratory pathway

The major pathway for the anaerobic respiration of the formate produced during fermentation is the formate-nitrate respiratory pathway (Enoch & Lester, 1974). Nitrate is the preferred electron acceptor for anaerobic respiration by *E. coli*, and during anaerobic growth the presence of nitrate in the growth medium induces the synthesis of the formate-nitrate respiratory chain and suppresses the reduction of other potential anaerobic electron acceptors, such as dimethyl sulfoxide (DMSO) and fumarate (Stewart, 1988). Anaerobic respiration of formate with these alternative electron acceptors is discussed later in this chapter (see section 1.1.3). The formate-nitrate respiratory pathway involves two membrane-bound multisubunit enzymes, formate dehydrogenase N (Fdh-N) and nitrate reductase A (Nr-A) (Enoch & Lester, 1974). It is thought that Fdh-N catalyses the oxidation of external formate to CO_2. The electrons produced by this oxidation are passed across the cytoplasmic membrane via quinone (ubiquinone or menaquinone) and Nr-A to intracellular

Gene/Operon	Linkage map position (min)	Enzyme/ Protein	Presumed function(s)	Position in text
arcA	0 min	ArcA	global aerobic transcriptional regulator	section 1.1.1
dcuSR	94 min	DcuS	*frdABCD* gene regulation (C4-carboxylate sensor)	section 1.1.3.2
		DcuR	*frdABCD* gene regulation (response regulator)	
dmsABC	20 min	DMSO reductase	reduction of DMSO	section 1.1.3.3
fdhD	88 min	FdhD	processing of Fdh-O	section 1.1.2.1
fdoGHI-fdhE	88 min	Fdh-O (FdoGHI) (also called Fdh-Z)	respiratory oxidation of formate; role in adaptation from aerobiosis to anaerobiosis	section 1.1.2.5
		FdhE	processing or assembly of Fdh-O	
fdnGHI	32 min	Fdh-N	respiratory oxidation of formate; component of formate-nitrate respiratory chain	section 1.1.2.1
fnr	29 min	Fnr	global anaerobic transcriptional regulator	section 1.1.1
focA-pfl	20 min	FocA	formate export and import	section 1.1.1
		Pfl	cleavage of pyruvate to acetyl CoA and formate	section 1.1, section 1.1.1
frdABCD	93 min	fumarate reductase	reduction of fumarate to succinate	section 1.1.3.2
modEF	17 min	ModE	molybdate-responsive transcriptional regulator	section 1.1.3.3
		ModF	unknown	
narGHJI	27 min	Nr-A (NarGHI)	respiratory reduction of nitrate; component of formate-nitrate respiratory chain	section 1.1.2.2
		NarJ	Nr-A assembly	
narK	27 min	NarK	nitrite export	section 1.1.2.4
narP	46 min	NarP	gene regulation (response regulator)	section 1.1.2.1
narQ	53 min	NarQ	gene regulation (nitrate and nitrite sensor)	section 1.1.2.1
narU	33 min	NarU	nitrite export?	section 1.1.2.6
narXL	27 min	NarX	gene regulation (nitrate and nitrite sensor)	section 1.1.2.1
		NarL	gene regulation (response regulator)	
narZYWV	33 min	Nr-Z (NarZYV)	respiratory reduction of nitrate during aerobic stress	section 1.1.2.6
		NarW	Nr-Z assembly?	section 1.1.2.6
nrfABCDEFG	93 min	NrfAB	respiratory reduction of nitrite to ammonium	section 1.1.3.1
		NrfCD	components of Nrf pathway	
		NrfEFG	assembly of NrfBD	
rpoS	59 min	RpoS (σ^s)	RNA polymerase σ^s subunit; stationary phase regulator	section 1.1.2.6
torCAD	22 min	TMAO reductase	reduction of TMAO	section 1.1.3.4
torR	22 min	TorR	*torCAD* gene regulation (response regulator)	section 1.1.3.4
torS	22 min	TorS	*torCAD* gene regulation (sensor)	section 1.1.3.4
torT	22 min	TorT	*torCAD* gene regulation (periplasmic inducer-binding protein)	section 1.1.3.4

Table 1.1 (Page 4). The main genes involved in respiratory formate metabolism in *E. coli*. The reactions catalysed by DMSO reductase, fumarate reductase, Nr-A, Nr-Z, NrfAB and TMAO reductase may all be coupled to the respiratory oxidation of formate by either Fdh-N or Fdh-O. Gene regulation described in text and summarised in **Table 1.2**. The Fhl-1 complex and hydrogenases-1 and -2 also play a role in formate respiration but are described elsewhere in this chapter (section 1.1.5 and section 1.1.6, respectively).

Gene/Operon	Linkage map position (min)	Anaerobiosis		Nitrate/nitrite		Position in text
		Fnr	ArcA	NarL	NarP	
dmsABC	20 min	+	O	-/ND	ND/ND	section 1.1.3.3
fdhD	88 min	O	ND	O/O	O/O	section 1.1.2.1
fdoGHI-fdhE	88 min	O	O	O/O	O/O	section 1.1.2.5
fdnGHI	32 min	+	O	+/+	+/+	section 1.1.2.1
focA-pfl	20 min	+	+	O/O	O/O	section 1.1.1
frdABCD[a]	93 min	+	O	-/O	O/O	section 1.1.3.2
narGHJI	27 min	+	O	+/+	O/O	section 1.1.2.2
narK	27 min	+	O	+/+	O/O	section 1.1.2.4
narU	33 min	O	O	O/O	O/O	section 1.1.2.6
narZYWV	33 min	O	O	O/O	O/O	section 1.1.2.6
nrfABCDEFG	93 min	+	ND?	+ & -/+ & -	+ & -/+ & -	section 1.1.3.1
torCAD[b]	22 min	O	O	O/O	O/O	section 1.1.3.4

Table 1.2. Anaerobic, nitrate and nitrite regulation of the main genes involved in respiratory formate metabolism in *E. coli*. Key: +, positive regulation (activation); -, negative regulation (repression or inhibition); O, not regulated by indicated regulatory protein; ND, not investigated. Based on 'Table 11' presented in Gennis & Stewart, 1996.

[a]Activation of the *frdABCD* operon also mediated by DcuSR in response to C_4-dicarboxylates (section 1.1.3.2).

[b]Activation of the *torCAD* operon mediated by TorR (section 1.1.3.4).

nitrate reducing it to nitrite (Enoch & Lester, 1974; Berg *et al.*, 1991; Sawers, 1994; Suppmann & Sawers, 1994; Berks *et al.*, 1995; Berks, 1996). This movement of electrons is coupled to proton translocation which in turn may be utilised by other membrane proteins to do 'work', such as ATP synthesis, solute transport or flagella rotation (Jones, 1980a; Gennis & Stewart, 1996).

1.1.2.1 Formate dehydrogenase-N (Fdh-N)

Subunit structure

Fdh-N is composed of three subunits α (FdnG), β (FdnH) and γ (FdnI) with molecular weights of 110 kD, 32 kD and 20 kD respectively (Enoch & Lester, 1975). The large α (FdnG) subunit is a selenomolybdoprotein containing the catalytic site of formate oxidation (Enoch & Lester, 1975; Berg *et al.*, 1991). Molybdenum is associated with this α (FdnG) subunit in the form of a molybdopterin guanine dinucleotide cofactor which catalyses formate oxidation (Axley *et al.*, 1990; Berks *et al.*, 1995). Selenium is incorporated into the protein in the form of selenocysteine (Berg *et al.*, 1991). The α (FdnG) subunit also possesses an iron sulphur cluster (Axley *et al.*, 1990; Brenton *et al.*, 1994; Berks *et al.*, 1995). The β (FdnH) subunit binds four iron sulphur clusters and is thought to mediate electron transfer between α (FdnG) and γ (FdnI) (Enoch & Lester, 1975; Berg *et al.*, 1991). The γ (FdnI) subunit is a cytochrome *b* possessing two *b*-type haems that transfer electrons from β (FdnH) to quinone (ubiquinone or menaquinone) (Enoch & Lester, 1975; Berks *et al.*, 1995; Ballard & Ferguson, 1988).

Cytoplasmic membrane topology of Fdh-N

Fdh-N has been shown to be a membrane bound enzyme complex (Giordano *et al.*, 1983). Initial sequence analysis of the *fdnGHI* operon (see genes encoding Fdh-N and their regulation) suggested that the α (FdnG) subunit and the majority of the β (FdnH) subunit are located in the cytoplasm, whilst the γ subunit is an integral membrane protein spanning the cytoplasmic membrane (Berg *et al.*, 1991). Berg and co-workers (1991) also speculated that the carboxy-terminal tail of the β (FdnH) subunit extends across the cytoplasmic membrane into the periplasmic space and together with the γ (FdnI) subunit may serve to anchor the enzyme complex to the

cytoplasmic membrane. Subsequent sequence analysis, however, speculated a periplasmic orientation for the active site of Fdh-N given the presence of a RRXFXK motif conserved in the N-terminal part of precursor polypeptides of periplasmic proteins binding redox co-factors (Berks *et al.*, 1995). A periplasmic orientation for the active site of Fdh-N fits our current understanding of formate metabolism in *E. coli* (Sawers, 1994; Suppmann & Sawers, 1994; Berks *et al.*, 1995). However a recent study by Benoit and co-workers (1998) has determined a cytoplasmic location for Fdh-O (an isoenzyme of Fdh-N in *E. coli*, section 1.1.2.5) and speculates that the α and β subunits of Fdh-N are also located in the cytoplasm.

Genes encoding Fdh-N and their regulation

The *fdnGHI* operon, located at 32 min on the *E. coli* genetic map, encodes the α (FdnG), β (FdnH) and γ (FdnI) subunits of Fdh-N (Table 1.1; Berg & Stewart., 1990; Berg *et al.*, 1991). Expression of the *fdnGHI* operon requires anaerobiosis, with the additional presence of nitrate necessary for optimal transcriptional levels (Berg & Stewart, 1990).

The anaerobic induction of *fdnGHI* operon expression is regulated by FNR (Table 1.2; Berg & Stewart, 1990).

The nitrate-mediated induction of *fdnGHI* operon expression is regulated by a system consisting of dual-interacting two-component regulatory systems with homologous membrane bound sensor proteins (NarX and NarQ) and homologous DNA binding response regulators (NarL and NarP) (Table. 1.2; Rabin & Stewart, 1993). It is proposed that upon sensing nitrate either NarX or NarQ is able to transfer a phosphate group to either NarL or NarP which consequentially effects a conformational change in these proteins, converting them to their transcriptionally active forms (Stewart, 1993). NarL mediates most of the nitrate induction of the *fdnGHI* operon, although NarP can weakly induce expression (Fig. 1.2; Rabin & Stewart, 1993). NarX, NarL, NarP and NarQ are encoded by the *narXL* operon and the *narP* and *narQ* genes located at 27 min, 46 min and 53 min on the *E. coli* genetic map respectively (Table 1.1). Expression of the *fdhGHI* operon is also weakly induced by nitrite (Berg & Stewart, 1990; Rabin & Stewart, 1993). It is proposed that upon sensing nitrite the NarQ protein phosphorylates NarL, which mediates this nitrite induction. This nitrite dependent phosphorylation of NarL is inhibited by

NarX (Berg & Stewart, 1990; Rabin & Stewart, 1993, Stewart, 1993). In the presence of nitrate, the Nar regulatory proteins also induce the synthesis of other formate-nitrate respiratory chain components (see section 1.1.2.2; section 1.1.2.4, section 1.1.2.5, section 1.1.2.6) and repress the synthesis of other anaerobic respiratory chain components (see section 1.1.3) (Table 1.2).

Fdh-N accessory genes

Genetic studies have revealed two genes, *fdhE* and *fdhD*, required for the synthesis of active Fdh-N (Mandrand-Berthelot *et al.*, 1988). The *fdhE* gene is located at 88 min on the *E. coli* genetic map and is co-transcribed with the *fdoGHI* genes (see section 1.1.2.). The *fdhD* gene is immediately upstream of and divergently transcribed from the *fdoGHI-fdhE* operon (Table 1.1; Plunkett *et al.*, 1993). Evidence suggests that FdhE may have a function in the processing or assembly of the Fdh-N isoenzyme (perhaps also Fdh-O below, section 1.1.2.5), and that FdhD may function as a protease required for the maturation of Fdh-N (perhaps also Fdh-O and Fdh-H below, section 1.1.2.5 and section 1.1.5.1, respectively) (Paveglio *et al.*, 1988; Sawers, 1994; Mandrand-Bethelot *et al.*, 1988). Expression of *fdhD* is decreased about threefold by aerobiosis but is unaffected by nitrate or nitrite (Schlindwein *et al.*, 1990). The anaerobic-mediated increase in *fdhD* expression is independent of FNR (Schlindwein *et al.*, 1990).

1.1.2.2 Nitrate Reductase-A (Nr-A)

Nr-A has a similar subunit organisation to Fdh-N, and is composed of a large molybdoprotein containing the catalytic site of nitrate reduction (α, NarG), an electron transfer subunit (β, NarH) containing four iron sulphur clusters and a cytochrome *b* (γ, NarI) (Enoch & Lester, 1975; Choudhry & MacGregor, 1983a; Choudhry & MacGregor, 1983b; Blasco *et al.*, 1989; Berks *et al.*, 1995). The α (NarG) and β (NarH) subunits are homologues of the equivalent Fdh-N subunits (Berks *et al.*, 1995).

Both the α (NarG) and β (NarH) subunits are peripheral membrane proteins anchored to the cytoplasmic side of the cell membrane by the γ subunit (NarI), an integral membrane protein (Ballard & Ferguson, 1988; Stewart, 1988; Berks *et al.*, 1995; Magalon *et al.*, 1997).

Nr-A is encoded by the *narGHJI* operon located at 27 min on the *E. coli* genetic map (Table 1.1; Bonnefoy-Orth *et al.*, 1981; Stewart & MacGregor, 1982; Edwards *et al.*, 1983; Sodergren & DeMoss, 1988; Blasco *et al.*, 1989). The *narJ* gene is thought to encode a chaperone (NarJ) that binds to the α (NarG) subunit during Nr-A assembly holding it in an appropriate conformation for molybdate co-factor insertion to occur (Blasco *et al.*, 1998). As observed for the *fdnGHI* operon, anaerobiosis is required for *narGHJI* operon expression with the additional presence of nitrate necessary for optimal transcriptional levels. This anaerobic and nitrate mediated induction of *narGHJI* operon expression is mediated by the Fnr and NarL proteins respectively (Table 1.2) (Stewart, 1982).

1.1.2.3 Proposed model for proton translocation by the formate-nitrate respiratory chain

The orientations of the α (FdnG), β (FdnH) and γ (FdnI) subunits of Fdh-N with respect to the cytoplasmic membrane have not been clearly determined, complicating attempts to explain the mechanism of proton translocation by the formate-nitrate respiration chain (topology discussed above, section 1.1.2.1; Gennis & Stewart, 1996; Benoit *et al.*, 1998). A periplasmic orientation for the α (NarG) and β (NarH) subunits has been speculated from sequence analysis. The current view therefore is that the active site of Fdh-N is orientated towards the periplasm and a model (proposed by Berks *et al.*, 1995) for the formate-nitrate respiration chain based on this arrangement is described below. This model is coherent with our current understanding of formate and nitrate metabolism (Rowe *et al.*, 1994; Sawers, 1994; Suppmann & Sawers, 1994), but it should be noted that subsequent topological studies have identified a cytoplasmic location for the α and β subunits of Fdh-O (section 1.1.2.5) and speculated a similar location for these subunits in Fdh-N (Benoit *et al.*, 1998).

Formate oxidation with nitrate as an electron acceptor supports transmembrane proton translocation with an estimated ratio of $H^+/2e^-=4$ (Jones, 1980a; Jones *et al.*, 1980).

It has been proposed that Fdh-N and Nr-A each generate a proton motive force by catalysing two half-reactions on opposite sides of the cytoplasmic membrane (a scalar mechanism; Jones, 1980a; Jones *et al.*, 1980). One half reaction

consumes protons from the cytoplasm (the reduction of quinone to quinol by Fdh-N; the reduction of nitrate to nitrite by Nr-A), and the second delivers protons to the periplasm (the oxidation of formate to CO_2 by Fdh-N; the oxidation of quinol to quinone by Nr-A). A schematic model for the formate-nitrate respiratory chain based on the models of membrane bound formate dehydrogenase and membrane bound nitrate reductase proposed by Berks and co-workers (1995) is shown in Fig. 1.2. Formate is oxidised by the active site of the α (FdnG) subunit of Fdh-N on the periplasmic side of the membrane. The two protons produced by this oxidation remain in the periplasm, however the two electrons are transferred to the γ (FdnI; cytochrome *b*) subunit via the iron sulphur clusters of the α (NarG) and β (NarH) subunits. The γ (FdnI; cytochrome *b*) subunit contains two *b*-type haem groups, which provide a transmembrane electron transport pathway transferring electrons to the cytoplasmic side of the membrane (Ballad & Ferguson, 1988; Berks *et al.*, 1995). These electrons together with cytoplasmic protons reduce quinone to quinol. Reoxidisation of this quinol by the γ (NarI; cytochrome *b*) subunit of Nr-A on the periplasmic side of the membrane releases these two protons into the periplasm (Jones & Garland, 1977; Morpeth & Boxer, 1985; Berks *et al.*, 1995). However the two electrons are transferred across the membrane by the γ (NarI; cytochrome *b*) subunit of Nr-A (Magalon *et al.*, 1997). At the cytoplasmic side of the membrane the electrons are transferred from the γ (NarI; cytochrome *b*) subunit to the active site of the α (NarG) subunit of Nr-A via the iron-sulphur clusters of the β (NarH) and α (NarG) subunits (Blasco *et al.*, 1990; Augier *et al.*, 1993a & 1993b; Breton *et al.*, 1994; Berks *et al.*, 1995). The active site of the α (NarG) subunit catalyses the reduction of nitrate by the transferred electrons and two cytoplasmic protons forming nitrite plus water.

1.1.2.4 Nitrate/nitrite transport

A system for the uptake of nitrate has not yet been identified in *E. coli,* although it has been speculated to be via a proton symport-based mechanism (Rowe *et al.*, 1994).

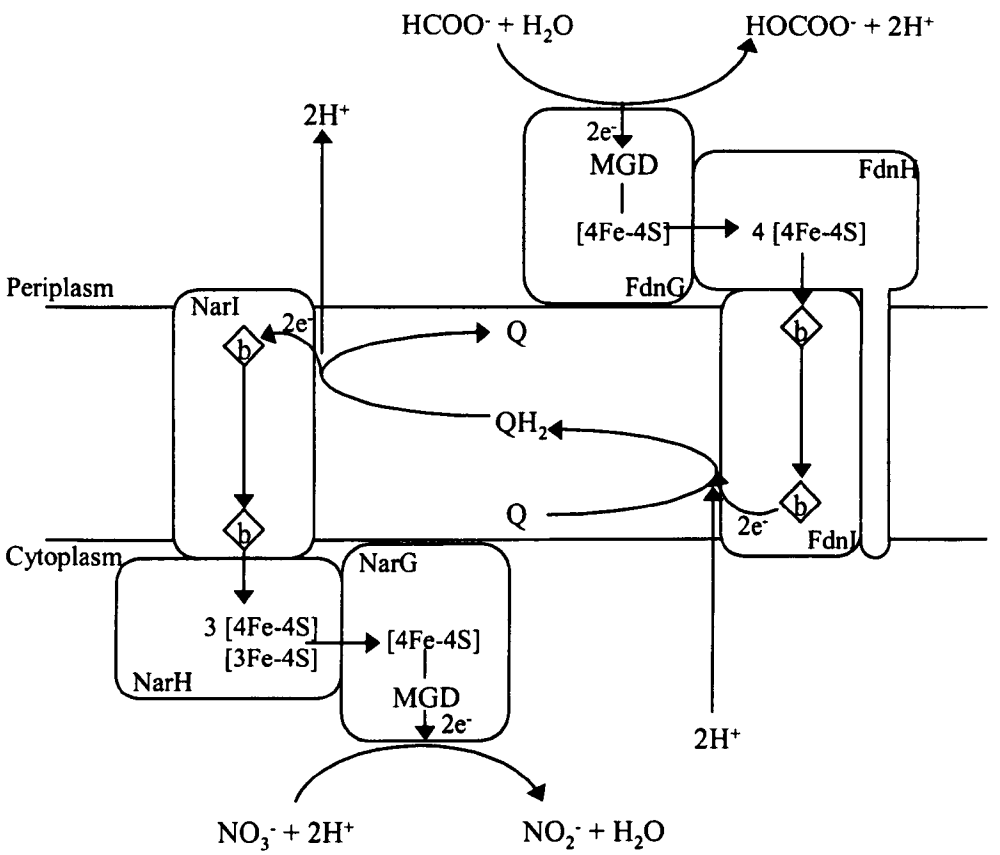


Fig. 1.2. Schematic depiction of proposed subunit and cofactor arrangement of Fdh-N and Nr-A and the proposed mechanism for proton translocation by the formate-nitrate respiratory chain. Based on the schematic depictions presented in Berks *et al.* (1995). FdnG, FdnH and FdnI are, respectively, the α, β and γ subunits of Fdh-N, and NarG, NarH and NarI are, respectively, the α, β and γ subunits of Nr-A. MGD represents the GMP conjugate of the molybdopterin cofactor. Q is quinone (ubiquinone or menaquinone).

Nitrite (the product of nitrate reduction by Nr-A) is exported from the cell by NarK, which serves to prevent the intracellular concentration of nitrite rising to toxic levels during anaerobic nitrate respiration (Rowe *et al.*, 1994). NarK is encoded by the *narK* gene located immediately upstream of the *narGHJI* operon (Table 1.1; Stewart & McGregor, 1982). Expression of *narK* is positively regulated by Fnr and NarL (Table 1.2; Kolesnikow *et al.*, 1992).

1.1.2.5 Formate dehydrogenase-O (Fdh-O)

Another formate dehydrogenase, distinct from Fdh-N has been identified which together with nitrate reductase-Z (Nr-Z; see section 1.1.2.6) forms a second formate-nitrate respiratory pathway in *E. coli* (Sawers *et al.*, 1991; Pommier *et al.*, 1992). It has also been speculated that Fdh-O (also called Fdh-Z) can catalyse the oxidation of formate using oxygen as a terminal electron acceptor (Sawers *et al.*, 1991; Sawers, 1994; Gennis & Stewart, 1996).

From sequence analysis it has been predicted that Fdh-O is composed of three subunits, α (FdoG), β (FdoH) and γ (FdoI), which exhibit a considerable degree of identity with those of Fdh-N (Berg *et al.*, 1991; Plunkett *et al.*, 1993). The α (FdoG) subunit is a selenomolybdoprotein containing the catalytic site of formate oxidation, the *β* (FdoH) subunit is an electron transfer unit harbouring four iron-sulphur clusters and the γ (FdoI) subunit is a cytochrome *b*. The α (FdoG) and β (FdoH) subunits of Fdh-O are located in the cytoplasm and anchored to the membrane by the γ (FdoI) subunit (Benoit *et al.*, 1998).

Fdh-O is encoded by the *fdoGHI* operon located at 88 min on the *E. coli* genetic map (Table 1.1; Plunkett *et al.*, 1993). The *fdoGHI* operon is expressed during both aerobic and anaerobic conditions with transcriptional levels slightly induced by nitrate during anaerobic growth (Abaibou *et al.*, 1995). Expression of the *fdoGHI* operon is not regulated by Fnr, ArcA, NarL or NarP (Table 1.2; Abaibou *et al.*, 1995).

It has been proposed that this second formate nitrate-respiratory pathway serves to ensure rapid adaptation to anaerobiosis during a sudden shift from aerobic to anaerobic conditions in the presence of nitrate (Abaibou *et al.*, 1995).

1.1.2.6 Nitrate Reductase-Z (Nr-Z)

Nr-Z is composed of three subunits, α (NarZ), β (NarY) and γ (NarV). Biochemical and DNA sequence analysis has revealed a high degree of similarity between these subunits and the corresponding subunits of Nr-A (Blasco *et al.*, 1990; Iobbi-Nivol *et al.*, 1990). Nr-Z is thought to possess the same membrane topology as Nr-A, with the α (NarZ) and β (NarY) subunits located in the cytoplasm and the γ (NarV) subunit anchored to the membrane (Berks *et al.*, 1995).

The α (NarZ), β (NarY) and γ (NarV) subunits of Nr-Z are encoded by the *narZYWV* operon located at 33 min on the *E. coli* genetic map (Bonnevoy *et al*; 1987; Blasco *et al.*, 1990). Another gene, *narU*, is located immediately upstream of the *narZ* gene and may be the first gene in a *narUZYWV* operon (Bonnefoy & DeMoss, 1994). The *narU* gene encodes a homologue of the nitrite extrusion protein NarK (section 1.1.2.4), and the *narW* gene encodes a homologue of the Nr-A assembly protein NarJ (section 1.1.2.2) (Bonnefoy & DeMoss, 1994). During exponential growth the *narZYWV* operon is expressed constitutively and is indifferent to the presence of either oxygen or nitrate (Table 1.2; Bonnefoy & DeMoss, 1994). However transcriptional activation of the operon is highly growth phase dependent (expression increased during stationary phase) and is highly dependent on the stationary phase regulator RpoS (σ^s) (Chang *et al.*, 1999), encoded by *rpoS* at 59 min (Hengge-Aronis, 1996). It has been recently proposed that Nr-Z serves to allow the cell to harness nitrate efficiently as an alternative electron acceptor under aerobic, stress-associated conditions (Chang *et al.*, 1999).

1.1.3 Respiratory formate oxidation by Fdh-N and Fdh-O with electron acceptors other than nitrate

In the absence of nitrate, *E. coli* is able to use nitrite, fumarate, DMSO, TMAO and O_2 as electron acceptors for respiratory formate oxidation.

1.1.3.1 Nitrite

Nitrite, the product of nitrate reduction by the formate-nitrate respiratory chain (section 1.1.2), is reduced to ammonium by either of two main routes (Stewart, 1993; Gennis & Stewart, 1996).

Nitrite present in the cytoplasm (section 1.1.2.4) is reduced by the soluble NADH-nitrite reductase (Page *et al.*, 1990). This enzyme serves no direct respiratory function and acts to regenerate NAD^+ and detoxify nitrite in the cytoplasm.

Nitrite exported from the cell is reduced by the formate-dependent respiratory nitrite reductase (NrfAB, nitrate reduction by formate) located on the periplasmic surface of the cytoplasmic membrane (Darwin *et al.*, 1993a). This reduction of nitrite can be coupled to the oxidation of formate by any of the three formate dehydrogenases of *E. coli* (Fdh-N, section 1.1.2.1; Fdh-O, section 1.1.2.5; Fdh-H, section 1.1.5) (Darwin *et al.*, 1993b). Nitrate reduction by NrfAB results in the generation of a proton motive force, although the mechanism for net proton translocation is unclear (Motteram *et al.*, 1981). NrfAB is encoded by the *nrfABCDEFG* operon located at 93 min on the *E. coli* genetic map (Table 1.1; Darwin *et al.*, 1993a; Hussain *et al.*, 1994). The *nrfC* and *nrfD* genes encode further components of the Nrf pathway, and the *nrfB, nrfF* and *nrfG* genes encode proteins involved in the assembly of NrfAB (Hussain *et al.*, 1994; Eaves *et al.*, 1998). Expression of the *nrfABCDEFG* operon is induced by Fnr in response to anaerobiosis, and NarL and NarP in response to low levels of nitrate and nitrite (<1 mM and < 2 mM, respectively) (Darwin *et al.*, 1993a; Hussain *et al.*, 1994; Wang & Gunsalus, 2000). However NarL represses the operon in response to higher levels of nitrate and nitrite (Table. 1.2; Rabin & Stewart, 1993; Tyson *et al.*, 1994; Wang & Gunsalus, 2000).

1.1.3.2 Fumarate

The reduction of fumarate by the cytoplasmically orientated membrane bound fumarate reductase (Lemire *et al.*, 1983) can be coupled to the oxidation of formate to generate a proton motive force. However fumarate reductase itself is unable to catalyse net transmembrane proton translocation (Jones 1980a; Gennis & Stewart, 1996). Hydrogen is also an effective electron donor for fumarate reduction (section 1.1.5) (Yamamoto *et al.*, 1977). Fumarate reductase is encoded by the *frdABCD* operon, located at 93 min on the *E. coli* genetic map (Table 1.1; Lambden & Guest, 1976; Jones & Gunalus, 1985). Expression of the *frdABCD* operon is induced by Fnr in response to anaerobiosis and repressed by NarL in response to nitrate (Table. 1.2; Jones & Gunalus, 1985, 1987; Iuchi & Lin, 1987). The FrdABCD operon is also induced (22 fold) by the DcuS-DcuR two component sensor regulator system in

response to the presence of C4-dicarboxylates (aspartate, fumarate, malate and succinate) (Table 1.2; Zientz *et al.*, 1998; Golby *et al.*, 1999).

1.1.3.3 Dimethyl sulfoxide (DMSO)

The reduction of DMSO by DMSO reductase, located on the cytoplasmic surface of the cell membrane (Sambasisarao *et al.*, 1990), can be coupled to the oxidation of formate (Rothery & Weiner, 1993). This DMSO reduction by DMSO reductase results in the generation of a proton motive force probably by a scalar mechanism (section 1.1.2.3) (Bilous & Weiner, 1985; Gennis & Stewart, 1996). DMSO reductase is also able to couple TMAO reduction with the respiratory oxidation of formate (Sambasivarao & Weiner, 1991). DMSO reductase is encoded by the *dmsABC* operon located at 20 min on the *E. coli* genetic map (Table 1.1; Bilous *et al.*, 1988; Bilous & Weiner, 1988). Expression of the *dmsABC* operon is induced by FNR in response to anaerobiosis and repressed by NarL in response to nitrate (Table. 1.2; Cotter & Gunsalus, 1989). The molybdate-responsive transcription factor, ModE, is also required for optimal expression of the *dmsABC* operon in response to anaerobiosis and for repression in response to nitrate (Table 1.1; McNicholas *et al.*, 1998).

1.1.3.4 Trimethylamine *N*-oxide (TMAO)

The reduction of TMAO by TMAO reductase (TorACD), located on the periplasmic surface of the cytoplasmic membrane (Silvestro *et al.*, 1989), can be coupled to the oxidation of formate to generate a proton motive force. However, it is not established whether TMAO reductase itself is able to catalyse net transmembrane proton translocation (Takagi *et al.*, 1981; Gennis & Stewart, 1986). TMAO reductase is encoded by the *torCAD*, located at 22 min on the *E. coli* genetic map (Table 1.1; Mejean *et al.*, 1994). Expression of the *torCAD* operon is induced by anaerobiosis and TMAO (or related compounds such as DMSO, tetramethylene sulfoxide and pyridine N-oxide) (Pascal *et al.*, 1984). Fnr and ArcA do not mediate the anaerobic induction of the *torCAD* operon and the regulator is still unknown (Simon *et al.*, 1994). The TMAO-mediated induction of expression is regulated by the TorS-TorR two-component regulatory system (Simon *et al.*, 1994; Jourlin *et al.*, 1996a). The sensor protein TorS detects TMAO and phosphorylates the response regulator TorR which, in turn, activates transcription of *torCAD*. Moreover, TorS is

able to interact with a periplasmic inducer-binding protein called TorT, which is essential for *tor* operon induction (Joulin *et al.*, 1996b).

A third TMAO reductase of *E. coli* (including DMSO reductase, section 1.1.3.3) has been identified recently, encoded by the *torYZ* operon (Gon *et al.*, 2000). This operon is expressed constitutively at very low levels and as yet a specific role for this reductase in *E. coli* respiration has not been elucidated (Gon *et al.*, 2000).

1.1.3.5 Oxygen (O_2)

Fdh-O (section 1.1.2.5) is the only Fdh isoenzyme synthesised during aerobic growth conditions. It has been speculated that in an aerobic environment the electrons generated by the oxidation of formate by Fdh-O can enter the quinone pool, be transferred to one of the terminal cytochrome oxidases and ultimately oxygen with consequent energy conservation (Sawers, 1994).

1.1.4 Quinones in *E. coli*

Quinones mediate the transfer of electrons between protein components of respiratory chains (Gennis & Stewart, 1996). *E. coli* synthesises three types of quinones: a benzoquinone, ubiquinone (UQ), and two napthoquinones, menaquinone (MK) and demethylmenaquinone (DMK) (Gennis & Stewart, 1996). All three quinones are dissolved within the lipid bilayer of the cytoplasmic membrane (Gennis & Stewart, 1996). The oxidised quinone species (ubiquinone, menaquinone and methylmenaquionone) is reduced to form the quinol species (ubiquinol, menaquinol and demethylmenaquinol) by the transfer of two electrons from the protein components of respiratory chains (Gennis & Stewart, 1996). Studies using knockout mutants have revealed a general pattern for quinone function whereby UQ is used for oxygen respiration, both UQ and MK are used for nitrate respiration and both MK and DMK are used for anaerobic respiration with acceptors other than nitrate (Gennis & Stewart, 1996). The cells of aerobic cultures contain about four to five times more UQ than MK plus DMK, whereas anaerobic cells contain about one third as much UQ as MK plus DMK (Wallace & Young, 1997; Wissenbach *et* al., 1992; Gennis & Stewart, 1996).

1.1.5 The formate hydrogenlyase pathway

The formate hydrogenlyase pathway catalyses the disproportionation of formate to carbon dioxide and hydrogen and has a role in pH homeostasis during fermentative growth (Böck & Sawers, 1996).

Fermentative growth by *E. coli* on hexoses, such as glucose, results in the production of a number of fermentation products including ethanol, acetate, formate and succinate (Fig. 1.1; Clark, 1989). These fermentation products are excreted from the cell (Böck & Sawers, 1996). Acetate, formate, lactate and succinate are acidic and as the extracellular concentration of these fermentation products increases the pH of the medium decreases. To counteract this pH drop formate is re-imported into the cell by FocA (see section 1.1.1) and broken down to carbon dioxide and hydrogen by the formate hydrogenlyase pathway. The hydrogen produced can either diffuse away from the cell or be re-oxidised by hydrogenases-1 and –2 (see section 1.1.6) (Böck & Sawers, 1996). The formate hydrogenlyase pathway constitutes a multiprotein complex (Fhl-1) located on the inner aspect of the cytoplasmic membrane (for topology see section 1.1.5.2) (Sauter *et al.*, 1992).

1.1.5.1 Subunit structure of the Fhl-1 complex

Fhl-1 is composed of a formate dehydrogenase (Fdh-H), the large and small subunits of a [Ni-Fe] hydrogenase (HycE and HyfG respectively), two electron carriers (HycB and HycF) and two membrane-spanning proteins (HycC and HycD) (Peck & Gest, 1957; Zinoni *et al.*, 1986; Bohm *et al.*, 1990; Sauter *et al.*, 1992; Sawers, 1994).

Fdh-H is composed of a single polypeptide with a molecular weight of 79 kD. It is a peripheral membrane protein which catalyses the oxidation of formate to carbon dioxide with the release of a proton and two electrons (Cox *et al.*, 1981; Pecher *et al.*, 1985; Zinoni *et al.*, 1986; Sawers *et al.*, 1986). The crystal structure of Fdh-H has been solved, and Fdh-H was shown to contain selenocysteine, molybdenum, two molybdopterin guanine dinucleotide cofactors, and an Fe_4S_4 cluster at the active site of formate oxidation (Boyington *et al.*, 1997). Boyington and co-workers (1997) suggested a reaction mechanism for Fdh-H that involved the selenocysteine residue in proton abstraction from formate, and the molybdenum, molybdopterin and Fe_4S_4 cluster in electron transfer.

HycE and HyfG are the large and small subunits, respectively, of a third [Ni-Fe] hydrogenase of *E. coli* which catalyses the reduction of protons to hydrogen (designated the name, hydrogenase-3, from the chronological order of its discovery) (Sawers *et al.*, 1985; Bohm *et al.*, 1990; Sauter *et al.*, 1992). Hydrogenases-1 and –2 are described later in this chapter (see section 1.1.6). The large, 65-kD, HycE subunit is a peripheral membrane protein containing the [Ni-Fe] metal centre that constitutes the active site of proton reduction (Bohm *et al.*, 1990; Rossmann *et al.*, 1994). The crystal structure of the [Ni-Fe] hydrogenase from *Desulfovibrio gigas* has been solved, showing that the nickel is co-ordinated by four cysteine residues, which are strictly conserved also among [Ni-Fe] hydrogenases from other organisms including HycE of *E. coli* (Volbeda *et al.*, 1995). Two cysteine-derived thiolates form a bridge to a second metal, iron, which is additionally co-ordinated by three non-protein derived ligands (Volbeda *et al.*, 1995). These ligands were identified as one cyanide (CN) and two carbon monoxide (CO) molecules (Happe *et al.*, 1997). The small, 28-kD, HycG subunit is tightly attached to the membrane and functions in electron transfer within the Fhl-1 complex (Bohm *et al.*, 1990; Sauter *et al.*, 1992). HycE and HycG have homology to the subunits of NADH-ubiquinone-oxidoreductases (Nuo) (see section 1.2.1; Bohm *et al.*, 1990).

The 22 kD HycB and 20 kD HycF subunits are membrane associated ferrodoxin-like electron transport proteins that are thought to constitute the two electron carriers of the Fhl-1 complex (Peck & Gest, 1957; Sauter *et al.*, 1992). HycB is probably the small (β) subunit of Fdh-H (see section 1.1.2.1; Sauter *et al.*, 1992).

The 64 kD HycC and 33 kD HycD subunits are both membrane-spanning proteins and, like HycE and HycG, have homology to the subunits of NADH-ubiquinone-oxidoreductase (Nuo) (section 1.2.1; Bohm *et al.*, 1990). Weiss and co-workers (1991a) proposed a possible quinone binding site for HycD and suggested that HycC and HycD could couple formate oxidation to quinone reduction.

1.1.5.2 Cytoplasmic membrane topology of Fhl-1

The current view for the membrane topology of Fhl-1 proposes that the complex is located on the cytoplasmic surface of the cell membrane (Sauter *et al.*, 1992). A schematic model of the putative topology of the Fhl-1 complex and the possible path

of electron flow, based on the model proposed by Sauter and co-workers (1992), is shown in Fig. 1.3.

1.1.5.3 Genes required for the synthesis of Fhl-1 (The formate regulon)

The genes required for the synthesis of Fhl-1 are organised in four transcriptional units, namely the *hyc* and *hyp* operons, the *fdhF* gene, and the *hydA* locus containing the *hydN* and *hypF* genes (Pecher *et al.*, 1985; Bohm *et al.*, 1990; Lutz *et al.*, 1991; Tomiyama *et al.*, 1991; Sawers, 1994; Maier *et al.*, 1996). Together these genes form the formate regulon (Fig 1.4).

The *hyc* operon (*hycABCDEFGHI*), located at 58 min on the *E. coli* genetic map, encodes structural components of Fhl-1 (HycBCDEFG; see section 1.1.5.2), a negative transcriptional regulator of the formate regulon (HycA) and a 17-kD protease (HycI) required for the maturation of the large subunit of hydrogenase-3 (HycE) (Table 1.3; Bohm *et al.*, 1990; Sauter *et al.*, 1992; Rossmann *et al.*, 1995). The *hycH* gene encodes a 16-kD polypeptide of unknown function (Bohm *et al.*, 1990).

The *hyp* (hydrogenase pleiotropic) operon (*hypABCDEfhlA*) is located immediately upstream of and divergently transcribed from the *hyc* operon (Fig. 1.4). The products of the *hyp* genes are involved in hydrogenase maturation (see section 1.1.5.5) and the *fhlA* gene encodes a σ^{54} dependent transcriptional activator of the formate regulon (FhlA) (Table 1.3; Schlensog & Böck, 1990; Lutz *et al.*, 1991).

The *fdhF* gene is located at 93 min on the *E. coli* genetic map and encodes Fdh-H (Table 1.3; Pecher *et al.*, 1985).

The *hydA* locus is positioned approximately four to five kilobases upstream of the *hyc* operon. The product of *hydN* appears to have some role in electron flow from or to Fdh-H, and the product of *hypF* is involved in hydrogenase maturation (see section 1.1.5.5) (Table 1.3; Tomiyama *et al.*, 1991; Sawers, 1994; Maier *et al.*, 1996).

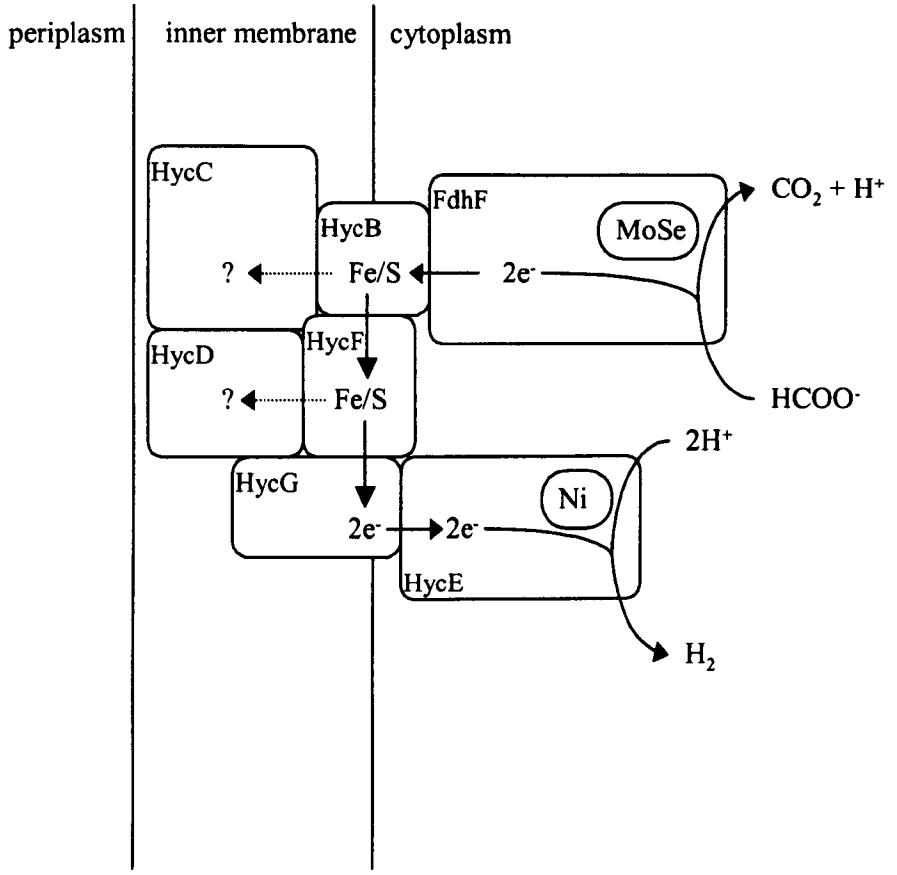


Fig. 1.3. Diagram of proposed subunit and cofactor arrangement of the Fhl-1 complex of *E. coli* and the proposed path of electron flow from formate to H^+.
Based on the schematic depiction presented in Sauter *et al.* (1992).
The trace elements involved in the catalytic reaction of Fdh-H (MoSe for molybdenum and selenium) and HycE (Ni for nickel) are indicated. Broken arrows show possible electron flow to HycC and HycD, which may then transfer electrons to the quinone pool (section 1.1.5.6).

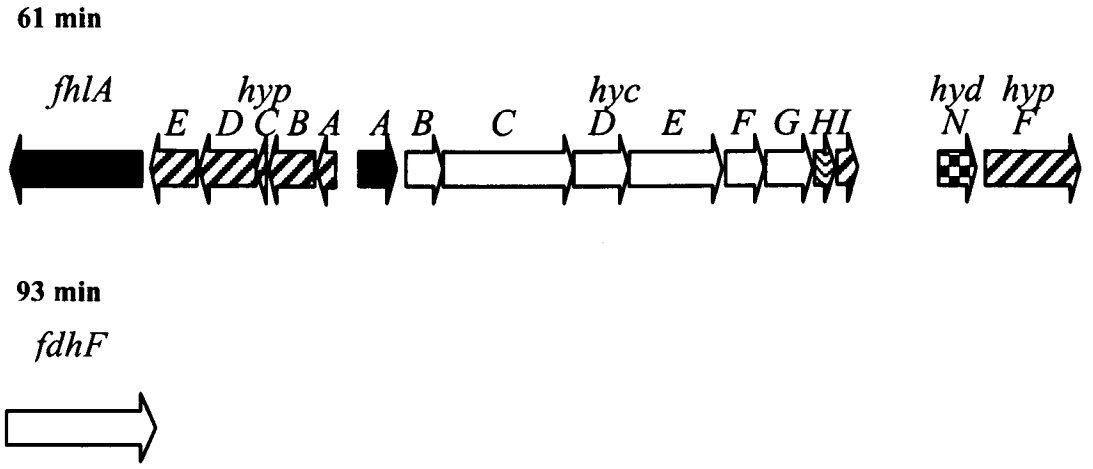


Fig. 1.4. Genetic organisation of the formate regulon at 58 min and 93 min on the *E. coli* chromosome. Key: white shading, genes encode structural components of Fhl-1; black , genes encode transcriptional regulators of the formate regulon; stripes, gene product involved hydrogenase maturation; checks, gene product involved in electron flow from or to Fdh-H; waves, function unknown.

Gene/Operon	Linkage map position (min)	Enzyme/ Protein	Presumed function(s)	Position in text
arcA	0 min	ArcA	global aerobic transcriptional regulator	section 1.1.1; section 1.1.6.3
appY	13 min	AppY	global transcriptional regulator	section 1.1.6.3
fdhF	93 min	Fdh-H	component of Fhl-1; breakdown of formate to CO_2 and H_2	section 1.1.5.3
fnr	29 min	Fnr	global anaerobic transcriptional regulator	section 1.1.5.4; section 1.1.6.3
ihfA	37 min	IHF	transcriptional activation of the *hyc* and *hyp* operons	section 1.1.5.4
ihfB	20 min			
hycABCDEFGHI	61 min	HycA	negative transcriptional regulator of the formate regulon	section 1.1.5.3, section 1.1.5.4
		HycBCDEFG	hydrogenase-3; components of Fhl-1; breakdown of formate to CO_2 and H_2	section 1.1.5.1; section 1.1.5.3
		HycH	function unknown	section 1.1.5.3
		HycI	specific processing protease of HycE	section 1.1.5.3; section 1.1.5.5
hyaABCDEF	22 min	HyaABC	hydrogenase-1; hydrogen uptake during anaerobic stress	section 1.1.6.1; section 1.1.6.3
		HyaD	specific processing protease of HyaB	section 1.1.6.3; section 1.1.6.4
		HyaE	maturation of hydrogenase-1	
		HyaF	maturation of hydrogenase-1	
hybOABCDEFG	65 min	HybOABC	hydrogenase-2; respiratory hydrogen uptake	section 1.1.6.1; section 1.1.6.3
		HybD	specific processing protease of HybC	section 1.1.6.3; section 1.1.6.4
		HybE	maturation of hydrogenase-2	
		HybF	maturation of hydrogenase-2	
		HybG	maturation of hydrogenases-1 & -2; chaperone like protein	
hydN	61 min	HydN	electron flow from or to Fdh-H	section 1.1.5.3
hypABCDEfhlA	61 min	HypA	nickel incorporation into hydrogenase-3	section 1.1.5.3; section 1.1.5.5
		HypB	nickel insertion into hydrogenases-1, -2 & -3	
		HypC	maturation of hydrogenases-3; chaperone like protein	
		HypD	nickel incorporation into hydrogenases-1, -2 & -3	
		HypE	CN/CO delivery to hydrogenases-1, -2 & -3	
		FhlA	positive transcriptional regulator of the formate regulon	section 1.1.5.3; section 1.1.5.4
hypF	61 min	HypF	CN/CO delivery to hydrogenases-1, -2 & -3	section 1.1.5.3
modEF	17 min	ModE	molybdate-responsive transcriptional regulator	section 1.1.3.3; section 1.1.5.4; section 1.1.7.2
		ModF	unknown	section 1.1.5.4
moeAB	18 min	MoeA	activation of molybdenum	section 1.1.5.4; section 1.1.7.2
		MoeB	synthesis of molybdopterin	section 1.1.7.2
narP	46 min	NarP	gene regulation (response regulator)	section 1.1.2.1 section 1.1.6.3

Gene/Operon (cont.)	Linkage map position (min) (cont.)	Enzyme/ Protein (cont.)	Presumed function(s) (cont.)	Position in text (cont.)
narXL	27 min	NarX	gene regulation (nitrate and nitrite sensor)	section 1.1.2.1; section 1.1.6.3
		NarL	gene regulation (response regulator)	
ntrA (*rpoN, glnF*)	72 min	NtrA (σ^{54})	RNA polymerase σ^{54} subunit	section 1.1.5.4
oxyS	90 min	OxyS	a small RNA; represses FhlA synthesis	section 1.1.5.4
rpoS	59 min	RpoS (σ^{s})	RNA polymerase σ^{s} subunit; stationary phase regulator	section 1.1.2.6

Table 1.3 (Page 23 and 24). The main genes required for Fhl-1, hydrogenase-1 and hydrogenase-2 synthesis in *E. coli*. Gene regulation described in text and summarised in **Fig. 1.5**.

1.1.5.4 Regulation of the formate regulon

Optimal expression of the formate regulon requires anaerobiosis, the absence of nitrate (or other compounds which may act as electron acceptors for the respiratory oxidation of formate, see section 1.1.2 and section 1.1.3), the presence of formate, the presence of molybdate (see ModE and MoeA) and an acidic pH (Peck & Gest, 1957; Wimpenny & Cole, 1967; Pecher *et al.*, 1983; Birkmann *et al.*, 1987a; Schlensog *et al.*, 1989). All these factors act at the level of transcription and all do not act singly but are channelled via a single signal, namely the intracellular concentration of formate (Birkmann *et al.*, 1987a, 1989; Schlensog *et al.*, 1989; Schlensog & Böck, 1990; Rossmann *et al.*, 1991; Lutz *et al.*, 1990, 1991). Formate only accumulates in the cell during anaerobiosis, in the absence of nitrate and at acidic pH because it is not produced during aerobiosis (Pfl not synthesised, see section 1.1), it is oxidised by the formate-nitrate respiratory pathway in the presence of nitrate (see section 1.1.2) and it is not imported into the cell at non-acidic pH (section 1.1.1; Rossmann *et al.*, 1991). This formate-mediated induction of expression is regulated by FhlA, a σ^{54}-dependent transcriptional activator (Maupin & Shanmugam, 1990; Schlensog & Böck, 1990).

FhlA

FhlA, like other regulators of the σ^{54} family, is composed of three distinct domains; an N-terminal A domain involved in the regulation of the activity of the protein, a central C domain involved in ATP binding and hydrolysis, and interaction with σ^{54} associated RNA polymerase, and a short C-terminal D domain responsible for DNA binding (Shingler, 1996). Transcription of the genes of the formate regulon occurs from -24/-12 promoters by the σ^{54} associated RNA polymerase. FhlA is activated by the binding of formate to its N-terminal A domain. Once activated FhlA is able to bind to upstream activating sequences (UAS) located about 100 bp upstream of corresponding regulon transcriptional start sites and activate transcription by interacting with σ^{54} associated RNA polymerase (Birkmann *et al.*, 1987b; Birkmann & Böck, 1989; Lutz *et al.*, 1990; Schlensog & Böck, 1990; Schlensog *et al.*, 1994; Korsa & Böck, 1997). Each UAS consists of two binding sites separated by a spacer region (Leonhartsberger *et al.*, 2000). FhlA binds to each UAS as a tetramer (Leonhartsberger *et al.*, 2000). ATP hydrolysis by the σ^{54}-dependent family of

transcriptional regulators is required for the conversion of a closed complex of RNA polymerase with DNA to the open form necessary for initiation of transcription (Weiss *et al.*, 1991b). FhlA has been shown to hydrolyse ATP (Hopper & Böck, 1995; Hopper *et al*, 1996).

The *fhlA* gene is transcribed from three different promoters. It is transcribed constitutively from a weak promoter located within the intergenic region between *hypE* and *fhlA*, and this level is enhanced during anaerobiosis through the activity of an Fnr-dependent promoter located within *hypA* (Lutz *et al.*, 1990; Lutz *et al.*, 1991). This Fnr promoter also plays a role during anaerobic respiratory growth conditions ensuring that sufficient levels of the Hyp proteins are present for the maturation of catalytically active hydrogenase isoenzymes (Lutz *et al.*, 1991). Transcription of the *fhlA* gene is further enhanced in the presence of formate from the σ^{54}-dependent promoter located upstream of the *hypA* gene (Lutz *et al.*, 1990). The RNA polymerase σ^{54} subunit is encoded by the *ntrA* gene (also known as *rpoN* or *glnF*) located at 72 min on the *E. coli* genetic map (Table1.3; Hirschmann *et al.*, 1985; Hunt & Magasanik, 1985).

HycA also plays a role in transcriptional control, negatively regulating transcription from the formate regulon (Sauter *et* al., 1992). It is not known whether HycA acts by directly interacting with FhlA, or by preventing binding of FhlA to the upstream activator sequence (Sauter *et al.*, 1992).

Translation of FhlA is repressed by OxyS, a small untranslated RNA that is induced in response to oxidative stress and acts as a global regulator affecting the expression of multiple genes. OxyS RNA represses *fhlA* through base-pairing with two short sequences on *fhlA* messenger RNA, one which overlaps the ribosome binding site and the other resides further downstream, within the coding region of *fhlA* (Altuvia *et al.*, 1998; Argaman & Altuvia, 2000). OxyS is encoded by the *oxyS* gene located at 90 min on the *E. coli* genetic map (Table 1.3; Altuvia *et al.*, 1997). Transcriptional regulation of the formate regulon is summarised in Fig. 1.5.

ModE and MoeA

FhlA regulates the molybdate-mediated induction of the *fdhF* and *hyc* operons via two independent proteins: ModE and MoeA.

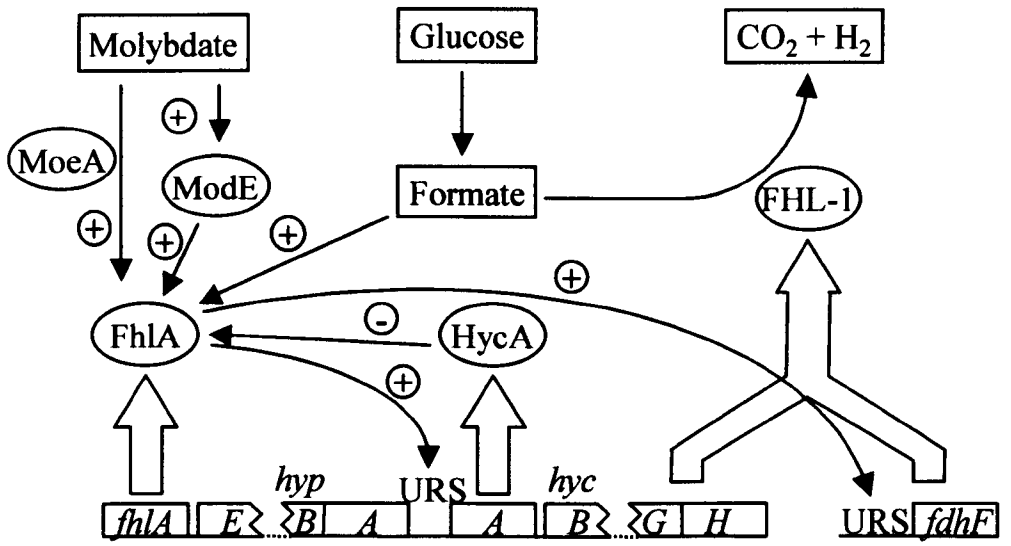


Fig. 1.5. A model depicting the transcriptional regulation of the formate regulon by FhlA. Abbreviations: URS, upstream regulatory sequence; +, activation; -, inhibition. Transcription of the *hypABCDEfhlA* operon also positively regulated by Fnr and IHF (see text, section 1.1.5.4). Transcription of the *hycABCDEFGH* operon is also positively regulated by IHF (see text, section 1.1.5.4).

ModE serves as a molybdate sensor and upon binding molybdenum, binds to the *hyc* promoter DNA upstream of the FhlA binding site and acts as a secondary activator of transcription (Grunden & Shanmugam, 1997; Self *et al.*, 1999). It is suggested that ModE-molybdate interacts with FhlA-formate during transcriptional activation of the *hyc* operon (Self *et al.*, 1999). ModE is encoded by the *modEF* operon located at 17 min on the *E. coli* genetic map (Table 1.3; Grunden *et al.*, 1996).

MoeA catalyses the production of a form of activated molybdenum species, which is used in the synthesis of the molybdenum co-factor present in Fdh-H (Hasona *et al.*, 1998a). The mechanism by which MoeA regulates transcription is unknown but it has been proposed that FhlA also binds this activated molybdenum species, besides formate, and activates transcription to optimum levels (Hasona *et al.*, 1998b; Self & Shanmugam, 2000). MoeA is encoded by the *moeAB* operon located at 18 min on the *E. coli* genetic map (Table 1.3; Nohno *et al.*, 1998).

IHF

The integration host factor (IHF) is also a regulator of the *hyc* and *hyp* operons (but not the *fdhF* gene) (Hopper *et al.*, 1994). It has been suggested that IHF functions by binding between the *hyp* and *hyc* promoters and facilitates nucleoprotein complex formation by bending adjacent DNA (Hopper *et al.*, 1994). IHF is encoded by the *hip* (also known as *himD*) and *himA* genes, located at 20 min and 37 min on the *E. coli* genetic map, respectively (Table 1.3; Friedmann, 1988).

1.1.5.5 Fhl-1 accessory genes (the *hyp* genes)

The Hyp proteins, encoded by the *hypABCDEfhlA* operon and the *hypF* gene, together with HycI, encoded by the *hyc* operon, are involved in the synthesis and insertion of the [NiFe] metal centre and the maturation of the large subunit of hydrogenase-3 (HycE) (Lutz *et al.*, 1991; Tomiyama *et al.*, 1991; Rossmann *et al.*, 1995; Maier *et al.*, 1996). HypB, HypD, HypE and HypF are involved in the maturation of hydrogenases-1, -2 and –3 of *E. coli*, whereas HypA and HycI are solely involved in the maturation of hydrogenase-3 (Lutz *et al.*, 1991). HypC is involved in the maturation of hydrogenases–1 and –3 (Casalot & Rousset).

HypB is a GTPase required for the formation of the [Ni-Fe] metal centre in the HycE precursor protein (pre-HycE) (Maier *et al.*, 1993; 1995). It is proposed that

HypB acts as a nickel donating system, in which GTP hydrolysis is thought to be involved in releasing HypB (or another nickel-binding protein) from the pre-HycE after the metal has been released (Maier *et al.*, 1993; 1995).

HypC is thought to be a specific chaperone-type protein which is able to form a stable complex with pre-HycE during the maturation process, thus keeping the protein in a folding state amenable to nickel insertion (Drapal & Böck, 1998; Magalon & Böck, 2000a). An additional or alternative function of HypC could also reside in the prevention of the association of the small subunit of hydrogenase-3 (HycG) to pre-HycE during the maturation process (Magalon & Böck, 2000b).

HypF is thought to interact with pre-HycE and synthesize and/or insert the cyanide (CN) and carbon monoxide (CO) ligands in the active site cavity of the large subunit (Casalot & Rousset, 2001). HypE is a possible partner of HypF in this maturation step (Casalot & Rousset, 2001).

The specific roles of the other Hyp proteins, HypA and HypD, are currently unknown.

HycI is a specific processing protease that removes a 32-amino acid peptide from the C-terminus of pre-HycE (Rossmann *et al.*, 1994; 1995).

A model has been proposed outlining the sequence of the steps involved in the maturation of hydrogenase-3 from *E. coli* (Magalon & Böck, 2000a,b). Firstly, HypC interacts with the HycE precursor enabling the incorporation of the diatomic ligands (CN and CO), iron and nickel into pre-HycE. After metal insertion the association between HypC and the C-terminus of the HycE precursor in cleaved, rendering pre-HycE a substrate for C-terminal endoproteolytic cleavage by HycI. This cleavage triggers a conformational change in HycE resulting in completion of the metal centre assembly and maturation of HycE (Magalon & Böck, 2000a). Association of the mature large subunit (HycE) with the small subunit (HycG) is now able to occur and is required for docking of hydrogenase-3 to the cytoplasmic membrane (Magalon & Böck, 2000b).

1.1.5.6 Respiratory role of Fhl-1

As well as its role in pH homeostasis, Fhl-1 also plays a role in anaerobic respiration producing hydrogen, which may be oxidised by the energy-conserving hydrogenase-1 and –2 isoenzymes (see section 1.1.6). Additionally to this it has been suggested that Fdh-H is able to divert electron flow from formate, possibly through the integral

membrane proteins HycC and HycD, to anaerobic reductases (Weiss *et al.*, 1991a; Sawers, 1994; Nandi & Sengupta, 1996).

1.1.6 Hydrogenase-1 and –2

Hydrogenases-1 and –2 are both membrane bound uptake hydrogenases catalysing the oxidation of dihydrogen to protons, and donating the resulting electrons to anaerobic reductases (such as fumarate reductases, section 1.1.3.3, and Nr-A, section 1.1.2.2) via quinone (ubiquinone, menaquinone or demethylmenaquinone) (Ballantine & Boxer, 1986; Sawers *et al.*, 1995; 1996). This oxidation of hydrogen supports transmembrane proton translocation (see section 1.1.6.5; Jones, 1980b).

The function of hydrogenase-1 remains to be resolved definitively, but it has been proposed to be involved in recycling the hydrogen produced by Fhl-1, since its synthesis showed strong correlation with that of Fhl-1 (Sawers *et al.*, 1985, Böck and Sawers, 1996). Hydrogenase-1 has more recently been suggested to be involved in response to stress maintaining the proton potential of the cytoplasmic membrane in an energy conserving manner (King & Przybyla, 1999).

Hydrogenase-2 has a respiratory function allowing cells to gain energy from the oxidation of molecular hydrogen during growth on non-fermentable carbon sources such as fumarate (Ballantine & Boxer, 1985; Sawers *et al.*, 1985; Menon *et al.*, 1994).

1.1.6.1 Subunit structures of hydrogenase-1 and -2

The hydrogenase-1 and –2 protein complexes are composed of the large (HyaB and HybC, respectively) and small (HyaA and HybO, respectively) subunits of the hydrogenase and a third subunit (HyaC and HybB, respectively) thought to be a cytochrome *b* (Ballantine & Boxer, 1986; Sawers & Boxer, 1986; Menon *et al.*, 1990; 1991; 1994; Dross *et al.*, 1992; Pryzbyla *et al.*, 1992; Sargent *et al.*, 1998).

The large (HyaB and HybC) and small (HyaA and HybO) subunits of both hydrogenases-1 and –2 show extensive similarities with the respective polypeptides of hydrogenase-3 of *E. coli* and hydrogenases from other organisms (see Subunit structure of the Fhl-1 complex, section 1.1.5.1; Böck & Sawers, 1996).

HyaC and HybB are homologues, and both are homologous to the third subunit of the *Wolinella succinogenes* hydrogenase (Przybyla *et al.*, 1991; Dross *et al.*, 1992). This third subunit is thought to mediate electron transfer from the small hydrogenase subunit (HyaA & HybA) to quinone, and is required for integrity of hydrogenase in the cytoplasmic membrane (Menon *et al.*, 1991; Dross *et al.*, 1992).

The hydrogenase-2 protein complex also possesses another subunit, HybA, that is proposed to serve as an electron acceptor from the small subunit, HybO (Sargent *et al.*, 1998).

1.1.6.2 Cytoplasmic membrane topologies of hydrogenases-1 and –2

In contrast to hydrogenase-3, hydrogenases-1 and –2 are integral membrane proteins (Ballantine & Boxer, 1986; Sawers & Boxer, 1986). Hydrogenase-2 has been shown to be attached to the periplasmic side of the cytoplasmic membrane and a similar localisation is proposed, although not resolved definitively, for hydrogenase-1 (Rhode *et al.*, 1989; Rodrigue *et al.*, 1996). Berks and co-workers (1995) stated that membrane bound [Ni-Fe] hydrogenases have a similar subunit organisation to the membrane-bound formate dehydrogenases and membrane-bound nitrate reductases (see Fdh-N, section 1.1.2.1; Nr-A, section 1.1.2.2). This organisation consists of two peripheral membrane subunits, α (HyaB and HybC) and β (HyaA and HybO) and a third integral membrane subunit, γ (HyaC and HybB) (Fig. 1.6; Berks *et al.*, 1995). The HybA subunit of hydrogenase-2 is also thought to be located at the periplasmic face of the cytoplasmic membrane (Sargent *et al.*, 1998).

1.1.6.3 Genes encoding hydrogenases-1 and-2, and their regulation

The *hyaABCDEF* operon, located at 22 min on the *E. coli* genetic map, encodes the structural components of the hydrogenase-1 complex (HyaABC), and three proteins required for hydrogenase-1 maturation HyaD, HyaE and HyaF (see section 1.1.6.4; Table 1.3; Menon *et al.*, 1990; 1991). Expression of the *hyaABCDEF* operon is induced by anaerobiosis and repressed by nitrate (Bronsted & Atlung, 1994; Richard *et al.*, 1999). The anaerobic induction of expression is positively regulated by two independent transcriptional regulators, ArcA and AppY (Bronsted & Atlung, 1994; Richard *et al.*, 1999). AppY appears to be activated in response to a metabolite from fermentation, however the identity of this effector is still unknown (Atlung *et al.*,

1997). AppY is encoded by the *appY* gene located at 13 min on the *E.* coli genetic map (Atlung *et al.*, 1989). The nitrate-mediated repression of *hyaABCDEF* operon expression is regulated by NarL and NarP (see section 1.1.2.1; Richard *et al.*, 1999). Transcriptional activation of the *hyaABCDEF* operon is also growth phase dependent (expression increases during stationary phase) and dependent on the stationary phase regulator RpoS (σ^s) (see section 1.1.2.6; Atlung *et al.*, 1997). Expression of the *hyaABCDEF* operon was also increased under acidic conditions (King & Przybyla, 1999).

The *hybOABCDEFG* operon, located at 65 min on the *E. coli* genetic map, encodes the structural components of the hydrogenase-2 complex (HybOABC), and four proteins required for hydrogenase maturation (HybD, HybE, HybF and HybG) (see section 1.1.6.4; Table 1.3; Menon *et al.*, 1994; Sargent *et al.*, 1998). Expression of the *hybOABCDEFG* operon, like the *hyaABCDEF* operon, is induced by anaerobiosis and repressed by nitrate (Richard *et al.*, 1999). The system responsible for controlling this anaerobic induction of expression is yet to be identified. The *hybOABCDEFG* operon is repressed by the transcriptional regulator, ArcA (Richard *et al.*, 1999). The nitrate-mediated repression of *hybOABCDEFG* operon expression is regulated by NarL and NarP (Richard *et al.*, 1999). In contrast to the *hyaABCDEF* operon, expression of the *hybOABCDEFG* operon is increased under alkaline conditions (King & Przybyla, 1999).

Fnr also induces expression of the *hyaABCDEF* and *hybOABCDEFG* operons, however it is yet to be determined whether it controls expression in a direct or indirect manner (Richard *et al.*, 1999).

1.1.6.4 Genes involved in the maturation of hydrogenase-1 and -2

The non-structural genes, *hyaD, hyaE* and *hyaF*, of the *hyaABCDEF* operon, together with the *hybF* and *hybG* genes of the *hybOABCDEFG* operon and the *hyp* genes (except *hypA*; see section 1.1.5.5) are involved in the maturation of the large subunit of hydrogenase-1 (HyaB) (Menon *et al.*, 1991; 1994; Jacobi *et al.*, 1992; Pryzybyla *et al.*, 1992; Casalot & Rousset, 2001). HyaD is a specific endopeptidase required to cleave the C-terminal extension of the precursor of the large subunit of hydrogenase-1 (HyaB) after metal centre assembly (Menon *et al.*, 1991). HybG, a homologue of HypC, is a chaperone-like protein, which forms a stable complex with

the HyaB precursor protein during the maturation process (Blokesch *et al.*, 2001). HypC is able to carry out the role of HybG in the maturation of hydrogenase-1 (see section 1.1.5.5; Blokesch *et al.*, 2001). The roles of the other Hyp proteins in the maturation of hydrogenase-1 are the same as described for the maturation of hydrogenase-3 (see section 1.1.5.5). The specific roles of HyaE, HyaF and HybF are currently unknown.

The non-structural genes, *hybD, hybE, hybF* and *hybG*, of the *hyaOABCDEFG* operon together with the *hyp* genes (except *hypA* and *hypC*; see section 1.1.5.5) are involved in the maturation of the large subunit of hydrogenase-2 (HybC) (Jacobi *et al.*, 1992; Przybyla *et al.*, 1992; Menon *et al.*, 1994; Casalot & Rousset, 2001). HybD is the endopeptidase required to cleave the C-terminal extension from HybC (Fritsche *et al.*, 1999). The roles of the HybG and the Hyp proteins in hydrogenase-2 large subunit maturation have been described previously in this chapter (see earlier in this section and section 1.1.5.5). The specific roles of HybE and HybF are currently unknown.

1.1.6.5 Proposed model for proton translocation by the membrane bound [Ni-Fe] uptake hydrogenases

Hydrogenase oxidation supports transmembrane proton translocation with an estimated H^+/e^- ratio of 2 and 4 with fumarate and nitrate as electron acceptors, respectively (fumarate reductase, section 1.1.3.2; Nr-A, section 1.1.2.2; Jones, 1980b). Hydrogenases-1 and –2 are thought to generate this proton motive force by a scalar mechanism (as described for the formate-nitrate respiratory chain, section 1.1.2.3; Gennis & Stewart, 1996). A schematic model for proton translocation by hydrogenase-1 and –2, proposed by Berks and co-workers (1995), is shown in Fig. 1.6.

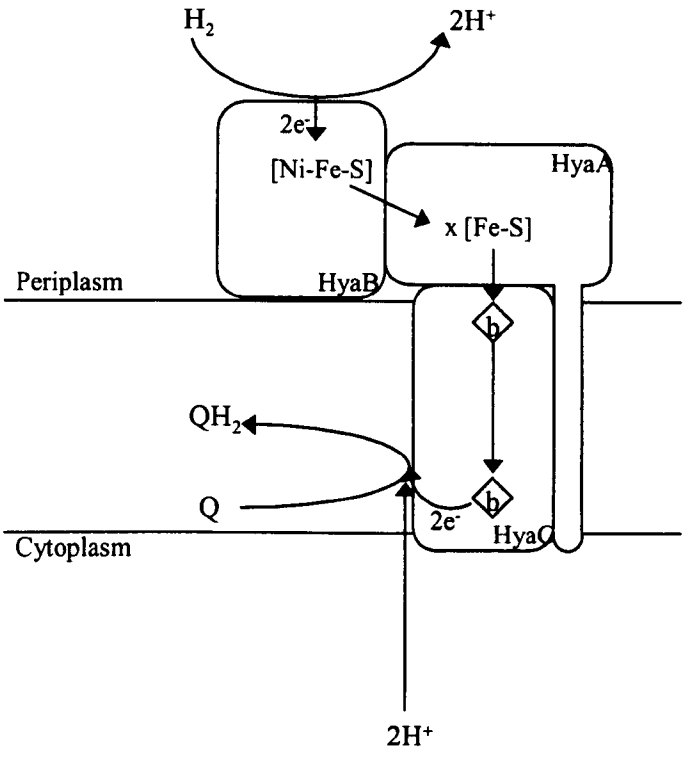


Fig. 1.6. Diagram of proposed subunit and cofactor arrangement and the proposed mechanism for proton translocation by the membrane bound [Ni-Fe] hydrogenase-1. Subunit, cofactor arrangement and mechanism of proton translocation proposed to be the same for hydrogenase-2, except for an additional subunit HybA that is proposed to serve as an electron acceptor from the small β subunit, HybO. Based on the schematic depictions presented in Berks *et al.* (1995). HyaB, HyaA and HyaC are, respectively, the α, β and γ subunits of hydrogenase-1. Q is quinone (ubiquinone, menaquinone or dimethylmenaquinone).

1.1.7 Auxiliary systems required for formate dehydrogenase and hydrogenase enzyme formation

The formation of the formate dehydrogenase and hydrogenase enzymes of *E. coli* depends on the availability, the uptake, and the incorporation of a number of metals (Böck & Sawers, 1996). These include selenium and molybdenum, which are required for formate dehydrogenase synthesis, and nickel, which is required for hydrogenase synthesis.

1.1.7.1 Selenium uptake and incorporation

Selenium is incorporated into the large α subunit of the three Fdh isoenzymes of *E. coli* (Fdh-N, section 1.1.2.1; Fdh-O, section 1.1.2.5; Fdh-H, section 1.1.5.1) in the form of a single selenocysteine residue (Böck & Sawers, 1996).

The mechanism of selenium uptake, as the selenite ion, into the cell has not been identified, but a high affinity selenite-specific transport system has been proposed from the observation that selenocysteine incorporation into Fdh-H is saturated at 100 nM (Zinoni *et al.*, 1987; Sawers, 1994).

The products of the *selAB* operon (located at 80 min on the *E. coli* genetic map), the *sel*C gene (82 min) and the *selD* gene (38 min) are required for the biosynthesis and incorporation of the selenocysteine residue into the α subunit of formate dehydrogenase (Table 1.4; Leinfelder *et al.*, 1988a). The *selC* gene encodes a unique tRNA, which is aminoacylated with serine by the enzyme seryl-tRNA synthetase (Leinfelder *et al.*, 1998b). This serine residue is converted to selenocysteine on the tRNA through the action of selenophosphate synthetase (SelD) and selenocysteine synthase (SelA) (Leinfelder *et al.*, 1990; Forchhammer *et al.*, 1991ab). SelB is a special elongation factor which specifically binds to selenocystyl-tRNA and delivers it to the translating ribosome where the tRNA decodes the UGA codon effecting the cotranslational insertion of selenocysteine into the polypeptide chain (Forchhammer *et al.*, 1989). The *sel* genes are expressed constitutively in *E. coli* (Sawers *et al.*, 1991).

1.1.7.2 Molybdenum uptake and incorporation

Molybdenum is an essential component of several enzymes in *E. coli* and is present associated with molybdopterin as the co-factor molybdopterin guanine dinucleotide

(MGD) (Rajagopalan, 1996). These enzymes include Fdh-N (section 1.1.2.1), Fdh-O (section 1.1.2.5), Nr-A (section 1.1.2.2), DMSO reductase (section 1.1.3.3), TMAO reductase (section 1.1.3.4) and Fdh-H (section 1.1.5.1) (Earhart, 1996).

Uptake of molybdenum, as the molybdate ion, MoO_4^{2-}, is carried out by the ModABCD high affinity molybdate-specific transporter. ModABCD is an ABC-transporter, and consists of a molybdate-specific periplasmic binding protein (ModA), an integral membrane channel-forming protein (ModB) and an ATP-binding protein (ModC) (Mauplin-Furlow *et al.*, 1995). A second integral membrane protein, ModD, is also a component of the transporter (Earhart, 1996). ModABCD is encoded by the *modABCD* operon located at 17 min on the *E. coli* genetic map (Table 1.4; Johann & Hinton, 1987; Shanmugam *et al.*, 1992). Expression of the *modABCD* is repressed by ModE in response to high concentrations of molybdate (ModE described above, see section 1.1.5.4; Walkenhorst *et al.*, 1995; Grunden *et al.*, 1996). When present in high concentrations, molybdate is transported by another, yet to be identified, transport system (Earhart, 1996).

Additional to the genes required for molybdate transport, biosynthesis of the molybdenum cofactor, molybdopterin guanine dinucleotide (MGD), requires the products of four transcriptional units, namely the *moaABCDE* (located at 17 min on the *E.* coli genetic map) operon, the *moeAB* operon (18 min), the *mogA* gene (0 min) and the *mobAB* operon (87 min) (Table 1.4). The products of the *moaABCDE* operon, together with MoeB, are involved in the synthesis of the molybdopterin component of the molybdenum cofactor, whilst MobA and MobB are required for the synthesis of the cofactor MGD from molybdopterin (Pitterle & Rajagoplan, 1993; Plunkett *et al.*, 1993; Palmer *et al.*, 1994; Rajagopalan, 1996). MogA and MoeA are thought to be involved in the processing and activation of molybdenum, before it is inserted into the cofactor (MoeA described previously, section 1.1.5.4; Rajagopalan, 1996; Hasona *et al.*, 1998).

1.1.7.3 Nickel uptake and incorporation

Nickel is required as an essential cofactor for the anaerobic biosynthesis and maturation of hydrogenases-1, -2 and –3 from *E. coli* (Rodrigue *et al.*, 1996).

Since in natural environments, nickel ions are usually present only in trace amounts, *E. coli* possesses an ATP-dependent nickel permease (NikABCDE), which has a very high affinity for nickel (Navarro *et al.*, 1993; De Pina *et al.*, 1995). This

Operon/ gene	Linkage map position (min)	Enzyme/ Protein	Presumed function(s)	Position in text
corA	86 min	CorA	magnesium transport	section 1.1.7.3
fis	74 min	Fis	growth phase transcriptional regulator	section 1.1.7.3
fnr	29 min	Fnr	global anaerobic transcriptional regulator	section 1.1.5.4; section 1.1.7.3
moaABCDE	17 min	MoaA	synthesis of molybdopterin	section 1.1.7.2
		MoaB		
		MoaC		
		MoaD		
		MoaE		
mobAB	87 min	MobA	synthesis of MGD from molybdopterin	section 1.1.7.2
		MobB		
modABCD	17 min	ModABCD	high affinity specific transporter of molybdate	section 1.1.7.2
modEF	17 min	ModE	molybdate-responsive transcriptional regulator	section 1.1.3.3; section 1.1.5.4; section 1.1.7.2
		ModF	unknown	section 1.1.5.4
moeAB	18 min	MoeA	activation of molybdenum	section 1.1.5.4; section 1.1.7.2
		MoeB	synthesis of molybdopterin	section 1.1.7.2
nikABCDE	77 min	NikABCDE	high affinity specific transporter of nickel	section 1.1.7.3
nikR	77min	NikR	nickel responsive transcriptional regulator	section 1.1.7.3
selAB	80 min	SelA	selenocysteine synthase	section 1.1.7.1
		SelB	selenocystyl-tRNA^{Sec}-specific elongation factor	
selC	82 min	SelC	tRNA^{Sec}	
selD	38 min	SelD	selenophosphate synthetase	

Table 1.4. The main genes required for the uptake of selenium, molybdenum or nickel, and for the incorporation of these metals into the formate dehydrogenase and hydrogenase enzymes of *E. coli*. Genes involved in the incorporation of nickel into hydrogenases-1, -2 and -3 of *E. coli* were described previously in this chapter (see section 1.1.5.5 and section 1.1.6.4).

NikABCDE nickel transport system belongs to a class of transport systems termed the ABC-transporters, and consists of a nickel sensing periplasmic binding protein (NikA), two integral membrane pore-forming proteins (NikB and NikC) and two ATP-binding proteins (NikD and NikE) responsible for coupling energy to the transport of nickel (Navarro *et al.*, 1993). NikABCDE is encoded by the *nikABCDE* operon located at 77 min on the *E. coli* genetic map (Table 1.4; Navarro *et al.*, 1993). Expression of the *nikABCDE* operon is induced by Fnr in response to anaerobiosis and repressed by NikR, a nickel-responsive regulator, in response to high intracellular concentrations of nickel (Wu & Mandrand-Berthelot, 1986; Wu *et al.*, 1989; De Pina *et al.*, 1999). It is important that the transport of nickel ions is suppressed when their concentration in the environment is high, as nickel ions can also be toxic to cells, by binding non-specifically to biomolecules or by displacing other metals from their native binding sites. NikR, is encoded by the *nikR* gene located immediately downstream of the *nikABCDE* operon (De Pina *et al.*, 1999). When present in high concentrations (relative to magnesium ions), nickel ions can also be taken up from the medium by a magnesium transport system encoded by the *corA* gene located at 86 min on the *E. coli* genetic map (Table 1.4; Park *et al.*, 1976).

The incorporation of nickel into hydrogenases-1, -2 and –3 is described elsewhere in this chapter (hydrogenase-3, section 1.1.5.5; hydrogenases-2 and –3, section 1.1.6.4).

1.2 Respiratory complex I of *E. coli*

Respiratory complex I of *E. coli* (also called NADH:ubiquinone oxidoreductase) catalyses the oxidation of NADH, and couples the transfer of the resulting electrons to ubiquinone with the translocation of protons across the cytoplasmic membrane (Gennis & Stewart, 1996). Complex I is considered to pump four protons across the membrane for every molecule of NADH oxidised (Gennis & Stewart, 1996).

1.2.1 Subunit structure and membrane topology of complex I

Complex I was originally believed to consist of 14 subunits, however the recent discovery that two of the genes encoding the complex were fused to form a single gene (see section 1.2.2), led to the conclusion that it consists of 13 subunits in *E. coli* (Weidner *et al.*, 1993; Braun *et al.*, 1998). These subunits are listed and their details summarised in Table 1.5. A topological model has been proposed for complex I, with the subunits grouped into three subcomplexes referred to as the peripheral, connecting, and membrane fragments (Fig. 1.7; Friedrich *et al.*, 1995; Leif *et al.*, 1995).

The peripheral fragment is comprised of the NuoE, NuoF and NuoG subunits and exhibits the NADH dehydrogenase activity that oxidises NADH to NAD^+ (Friedrich *et al.*, 1993; Leif *et al.*, 1995). NuoF is thought to possess the flavin mononucleotide cofactor that catalyses NADH oxidation (Gennis & Stewart, 1996).

The connecting fragment is comprised of the NuoB, NuoCD and NuoI subunits (Friedrich *et al.*, 1993; Leif *et al.*, 1995).

The membrane fragment is comprised of the NuoA, NuoH, NuoJ, NuoK, NuoL, NuoM and NuoN fragments and catalyses the reduction of ubiquinone (Friedrich *et al.*, 1993; Leif *et al.*, 1995). The NuoH subunit has been tentatively assigned the ubiquinone binding site (Weidner *et al.*, 1993; Gennis & Stewart, 1996).

The NuoB, NuoCD, NuoH, NuoI and NuoL subunits are evolutionarily related to the HycG, HycE, HycD, HycF and HycC subunits of *E. coli* Fhl-1, respectively (Fhl-1 subunits described in section 1.1.5.1; Bohm *et al.*, 1990; Fearnly & Walker, 1992). Gennis and Stewart (1996) proposed that these Fhl-1 related Nuo subunits form a discrete functional and structural unit within complex I (Fig. 1.7).

Nuo subunit	Molecular mass (kD)[a]	Presumed function(s)	Presumed cofactor(s)
NuoA	16	Unknown (3 transmembrane helices)	None known
NuoB	25	Unknown	[4Fe-4S] or [2Fe-2S] cluster
NuoCD	67	Unknown	None known
NuoE	19	Unknown	[2Fe-2S] cluster
NuoF	50	NADH binding	FMN [4Fe-2S] cluster
NuoG	91	Unknown	Two [2Fe-2S] clusters One or two [4Fe-2S] cluster
NuoH	36	Q binding (8 transmembrane helices)	None known
NuoI	20	Unknown	Two [4Fe-2S] clusters
NuoJ	20	Unknown (5 transmembrane helices)	None known
NuoK	11	Unknown (3 transmembrane helices)	None known
NuoL	66	Unknown (13 transmembrane helices)	None known
NuoM	51	Unknown (10 transmembrane helices)	None known
NuoN	52	Unknown (12 transmembrane helices)	None known

Table 1.5. Products of the *nuoABCDEFGHIJKLMN* operon encoding respiratory complex I in *E. coli*. Based on 'Table 4' presented in Gennis & Stewart (1996). [a]Approximate molecular mass based on DNA sequence analysis. Abbreviations: FMN, flavin mononucleotide cofactor.

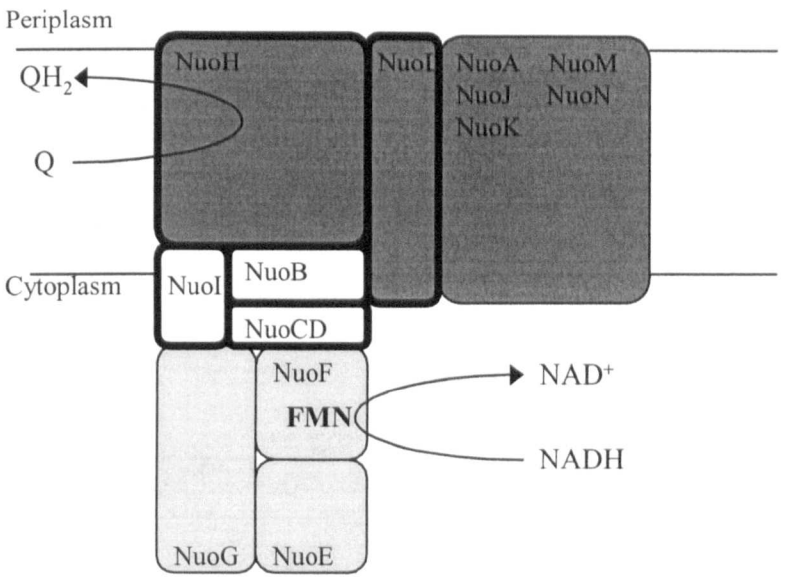


Fig. 1.7. Diagram of the subunit arrangements of respiratory complex I of *E. coli*. Abbreviations: FMN, flavin mononucleotide cofactor. The relative sizes of the subunits are not to scale. Based on the schematic depictions presented in Gennis & Stewart (1996) (Figure 6) and Falk-Krzensinski and Wolfe (1998) (Fig. 1). *E. coli* complex I is comprised of three distinct fragments: the peripheral (light grey), connecting (white), and membrane (dark grey) fragments. The subunits bordered with a thick line indicates groupings of subunits that are evolutionarily related to *E. coli* Fhl-1 (NuoB, NuoCD, NuoH, NuoI, NuoL).

1.2.2 Genes encoding complex I and their regulation

The *nuoABCDEFGHIJKLMN* operon, located at 51 min on the *E. coli* genetic map, encodes the structural components of complex I (Weidner *et al.*, 1993). As mentioned, the genes *nuoC* and *nuoD* are fused to form one gene *nuoCD* leading to a complex of 13 subunits (Braun *et al.*, 1998). Transcriptional activation of the operon is growth phase dependent, with expression increased during early exponential phase and decreased during late exponential phase and stationary phase (Wackwitz *et al.*, 1999). This growth phase dependent expression is mediated by the growth phase responsive regulator Fis located at 74 min on the *E. coli* genetic map (Johnson *et al.*, 1988; Wackwitz *et al.*, 1999). Induction of expression at early exponential phase, ensures that complex I is synthesised when large amounts of ATP are required (Wackwitz *et al.*, 1999).

1.3 The *hyf* operon of *E. coli*

Sequence analysis of the 55.8 min region of the *E. coli* genome revealed a putative 12-gene *hyfABCDEFGHIJRfocB* operon (Andrews *et al.*, 1997). Sequence similarities strongly suggested that the *hyfG* and *hyfI* genes encode the large and small subunits respectively, of a fourth [Ni-Fe] hydrogenase in *E. coli*, and therefore the operon was designated *hyf* (hydrogenase four) (Andrews *et al.*, 1997).

1.3.1 Nucleotide sequence of the *hyf* operon

Andrews and co-workers (1997) sequenced the 55.8 min region of the *E. coli* genome and using a combination of codon-usage analysis and Orf searches they revealed a total of 12 contiguous genes designated *hyfA-J*, *hyfR* and *focB* (Fig. 1.8; complete sequence of the *hyf* operon in Appendix I). These genes possess codon usages similar to those of very weakly expressed *E. coli* structural genes, suggesting the products of these genes are likely to be of very low abundance. A potential σ^{54} dependent promoter is located upstream of the *hyf* operon and approximately 110 bp downstream of a putative FhlA (and/or HyfR; see section 1.3.2) binding site (Appendix I). Each gene of the operon is preceded by an appropriately positioned Shine-Dalgarno sequence. A potential stem-loop structure, which could serve as a transcription terminator of the *hyf* operon, is located 4 bp downstream of the *focB* gene. The *hyfA-J* genes appear to be translationally coupled, as each stop codon (except *hyfJ*) overlaps the translation-initiation region of the gene immediately downstream. The *hyfJ-hyfR* and *hyfR-focB* intergenic regions are 29 and 21 bp respectively, suggesting that the translation of *hyfR* and *focB* is not tightly coupled to that of the upstream genes. Preliminary expression experiments have shown that a *hyfA-lacZ* translational fusion is expressed during anaerobic growth on fermentable carbon sources indicating that the *hyf* promoter is active and that *hyf* is likely to be a functional operon (Andrews *et al.*, 1997).

1.3.2 Translation products of the *hyf* operon

The main features of the *hyf* encoded polypeptides are shown in Table 1.6.

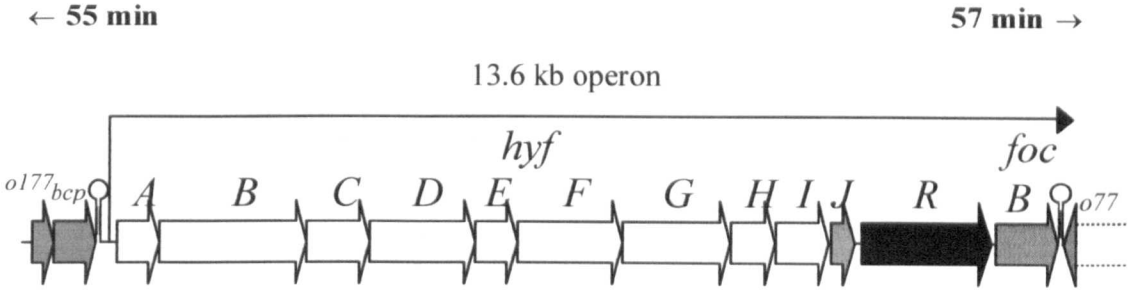


Fig. 1.8. Genetic organisation of the *hyf* opeorn on the *E. coli* chromosome. Key: white shading, genes encode structural components of Fhl-2; black, gene product involved in transcriptional regulation; stripes, gene product involved in formate transport; waves, function unknown; spots, neighbouring genes.

Hyf subunit (no. of residues)	Molecular mass (kD)[a]	Features(s)	Amino acid sequence identity with *hyf* products (%)		
			E. coli Hyc	*E. coli* Nuo	*E. coli*
HyfA (205)	22	16Fe ferredoxin, 4[4Fe-4S]	HycB, 50	-	-
HyfB (672)	73	Polytopic membrane protein, 16 transmembrane helices	HycC, 37	NuoM, 22	-
HyfC (315)	34	Polytopic membrane protein, 8 transmembrane helices	HycD, 51	NuoH, 22	-
HyfD (479)	52	Polytopic membrane protein, 14 transmembrane helices	HycC, 17	NuoL, 33	-
HyfE (216)	23	Polytopic membrane protein, 7 transmembrane helices	-	-	-
HyfF (526)	57	Polytopic membrane protein, 14 transmembrane helices	HycC, 22	NuoM, 21	-
HyfG (555)	63	Hydrogenase large subunit, with [Ni-Fe(CO)$(CN)_2$] cluster	HycE, 73	NuoCD	-
HyfH (181)	20	Ferredoxin, with 2[4Fe-4S] + [Fe-S]/Fe-Cys_4 ?	HycF, 44	NuoI, 25	-
HyfI (252)	28	Hydrogenase small subunit, with [4Fe-4S]	HycG, 63	NuoB, 27	-
HyfJ (137)	16	Unknown function	HycH, 47	-	-
HyfR (670)	75	Formate-sensing gene regulator	-	-	FhlA, 46
FocB (282)	31	Formate transporter, 6 transmembrane helices	-	-	FocA, 50

Table 1.6. Products of the *hyfABCDEFGHIJR-focB* operon encoding a putative proton-translocating formate hydrogenlyase system. Based on 'Table 1 & Table 2' presented in Andrews *et al.* (1997). [a]Approximate molecular mass based on DNA sequence analysis.

HyfA, B, C, G, H and I

HyfA, B, C, G, H and I more closely resemble the *hyc* operon encoded Fhl-1 subunits, HycB, C, D, E, F and G, respectively, than any other polypeptides (Table 1.6; Fhl-1 subunits described in section 1.1.5.1) (Andrews *et al.*, 1997). HyfG and HyfI are particularly closely related (73% and 63% identity) to the large and small subunits of hydrogenase-3 (HycE and HycG, respectively) (Andrews *et al.*, 1997). These findings suggest that the *hyf* operon, together with Fdh-H, form a second formate hydrogen lyase complex (Fhl-2) in *E. coli* (Andrews *et al.*, 1997). A second formate hydrogenlyase complex containing Fdh-H would explain the otherwise surprising observation that *fdhF* (located at 92 min on the *E. coli* genome) is not part of the *hyc* operon (61 min) (see section 1.1.5.3).

The products of two genes of the *hyc* operon, the negative transcriptional regulator HycA and the protease HycI, have no counterparts amongst the products of the *hyf* operon (Andrews *et al.*, 1997). The lack of a HycI-like protease is unexpected as homologues are encoded by the *hya* and *hyb* operons (Menon *et al.*, 1990; 1991). The primary structure of the HyfG subunit (hydrogenase-4 large subunit) indicates that it is processed, and therefore it is presumed that a processing protease for this subunit is encoded by a gene outside the *hyf* operon (Andrews *et al.*, 1997). In the complete *E. coli* genome the only detectable HycI homologues are HyaD and HybD encoded by the hydrogenase-1 and –2 operons, respectively. In view of the close similarity between HyfG and HycE it would appear likely that both are processed by HycI (Andrews *et al.*, 1997). This might explain why no hydrogenase-4 dependent benzylviologen dependent hydrogenase activity was detected by Sauter and co-workers (1992) in a Δ*hya* Δ*hyb* Δ*hycB-H* hydrogenase triple mutant. The Δ*hycB-H* mutation of the hydrogenase triple mutant contains a *camR* cassette, which could exert a polar effect on *hycI* expression.

HyfD, E and F

Five of the six *hyc* encoded structural proteins resemble complex I subunits (section 1.2.1) and these similarities extend to the corresponding *hyf* encoded subunits (Table 1.6) (Andrews *et al.*, 1997). However unlike the *hyc* encoded Fhl-1, the *hyf* operon encodes three additional integral membrane subunits (HyfD, E and F), of which the two largest (HyfD and HyfF) have counterparts in complex I (Table 1.6) (Andrews *et*

al., 1997). Andrews and co-workers (1997) proposed that these additional subunits confer a function associated with the respective subunits of complex I that is not found in Fhl-1, namely proton translocation. The basic organisation of the *hyc* structural genes is conserved in the *hyf* operon apart from the genes encoding these three extra subunits (*hyfDEF*), which are inserted as a block immediately after *hyfC* (Fig. 1.8).

HyfJ

HyfJ is a homologue of HycH, and like HycH, the function of HyfJ is unknown (Table 1.6; HycH described in section 1.1.5.3) (Andrews *et al.*, 1997).

HyfR

HyfR is closely related to the σ^{54}-dependent transcriptional activator of the formate regulon, FhlA (FhlA described in section 1.1.5.4) (Andrews *et al.*, 1997). HyfR and FhlA possess good similarity in their central C domains and C-terminal D domains suggesting that σ^{54} interaction and ATP binding/hydrolysis is similar in both molecules and that they recognise similar DNA-binding sites (domains of σ^{54} dependent regulators described in section 1.1.5.4). However HyfR and FhlA are less similar in their regulatory N-terminal A domains, with similarity restricted to two sub-domains involved in formate interaction. The remaining regions of the A domain of HyfR are only 13-14% identical to the corresponding regions of FhlA. Interestingly one of these less similar regions of HyfR possess a cysteine-rich segment which could serve as a binding site for a metal cofactor or iron sulphur cluster, endowing the HyfR protein with the ability to respond to changes in redox and/or oxygen status as for the Fnr transcriptional regulator protein.

The high degree of similarity between HyfR and FhlA suggests that HyfR regulates transcription of the *hyf* operon (and possibly other genes) in response to formate, and therefore *E. coli* could possess two formate-responsive σ^{54}-dependent transcriptional regulators. A lack of functional complementarity between HyfR and FhlA is apparent from the observation that *fhlA* mutants exhibit no residual formate-dependent regulation of the formate regulon and it is likely that both molecules would respond to different ranges of formate concentration, have slightly different

UAS recognition specificities or differ in their expression control or interaction with HycA (Schlensog *et al.*, 1989).

FocB

FocB most closely resembles the *E. coli* formate transporter FocA (section 1.1.1) (Andrews *et al.*, 1997). It would therefore appear that FocB is a second *E. coli* formate transporter, in which case it could account for the residual formate transport activity observed in FocA mutants (Suppmann & Sawers, 1994). The physiological role of FocB is presumably to import exogenous formate to provide substrate for the *hyf* complex.

The presence of the *hyfR* and *focB* genes reinforces the view that the products of the *hyf* operon function in formate metabolism.

1.3.3 A functional model for Fhl-2

As discussed above, it is proposed that the *hyf* operon, together with *fdhF*, encodes a second formate hydrogenlyase complex (Fhl-2) in *E. coli*, which unlike the *hyc* encoded Fhl-1 complex is proton translocating. A functional model was proposed by Andrews and co-workers (1997) whereby the free energy available from the formate hydrogenlyase reaction,

$$HCO_2^- + \quad H_2O \quad \rightarrow \quad HCO_3^- + \quad H_2,$$

was only sufficient to support proton translocation when the H_2 produced is removed by respiratory metabolism, either by *E. coli* itself or by other organisms in the environment. Suitably favourable conditions have been detected in ecosystems where interspecies hydrogen transfer is occurring, e.g. in the rumen (Thauer *et al.*, 1993). This Fhl-2 reaction is, however, unlikely to support transmembrane translocation with a H^+/e^- ratio of greater than 1 (Andrews *et al.*, 1997). It was also proposed by Andrews and co-workers (1997) that the hydrogen produced by Fhl-2 is oxidised in *E. coli* by hydrogenase-1 (see section 1.1.6). This is consistent with the proposed hydrogen-cycling role of hydrogenase-1 and would provide a physiological rationale for the observed induction of hydrogenase-1 during fermentative growth on formate (Sawers *et al.*, 1985; Sawers *et al.*, 1986; Bronsted & Atlung, 1994).

1.3.4 Structural model for Fhl-2

A schematic structural model for Fhl-2 proposed by Andrews and co-workers (1997) and based on the structural features of Fhl-1 and complex I is presented in Fig. 1.9.

1.3.5 Other enzyme complexes structurally related to the Hyf complex

Hydrogenase-3 (Hyc) and hydrogenase-4 (Hyf) of *E. coli* are members of a group of membrane-bound, multisubunit [Ni-Fe]-hydrogenase with significant sequence similarities to subunits of complex I. Other members of this group that show high similarity to the Hyf complex include, a CO-induced hydrogenase from *Rhodospirillum rubrum*, the Ech hydrogenase from *Methanosarcina barkeri* and the Mbh hydrogenase from *Pyrococcus furiosus* (Andrews *et al.*, 1997; Meuer *et al.*, 1999; Silva *et al.*, 2000). It is proposed that these [Ni-Fe] hydrogenases are proton translocating (Bott and Thauer, 1989; Kerby *et al.*, 1995; Silva *et al.*, 2000).

Furthermore the Hyf complex closely resembles the products of a partial gene cluster in *M. tuberculosis* that encode an electron transfer complex (Andrews *et al.*, 1997). In fact the Orf3 of *M. tuberculosis* is the only known HyfE homologue. It is unlikely that this *M. tuberculosis* enzyme complex is a hydrogenase because the HyfG homologue (Orf5) does not contain ligands for a hydrogenase metallocluster.

The Na^+/H^+ antiporter system of *Bacillus subtilis* also shows high similarity to the Hyf complex (Ito *et al.*, 1999). Hiramatsu and co-workers (1998) proposed that this Na^+/H^+ antiporter is a novel multisubunit secondary transporter that is energised by a proton motive force.

No close homologues of the *hyfABCDEFGHIJR-focB* operon are present in the genomes of *E. coli*'s sibling species *Salmonella enterica* serovars Typhimurium, Typhi and Paratyphi A and the genome of the close outgroup *Klebsiella pneumoniae* (McClelland *et al.*, 2000).

The aims of the experimental programme reported and discussed in the remainder of this thesis were to investigate the proposals for Hyf function, structure and regulation described in this chapter.

Initial experiments were established to obtain evidence for a fourth [NiFe] hydrogenase in *E. coli* encoded by the *hyf* operon. These experiments included Immunoblotting and ^{63}Ni incorporation experiments to detect *hyf* gene products, and

hydrogenase and formate dehydrogenase assays to detect enzyme activity attributable to the *hyf* operon (Chapter 3).

An extensive study into the regulation of the *hyf* operon was also undertaken. A λ*hyfA-lacZ* transcriptional fusion strain was used to study *hyf* expression under different growth conditions and in different mutant backgrounds (Chapter 4).

Finally bioreactors were used to compare the growth of wildtype and *hyf* mutant strains during anaerobic batch cultivation and both aerobic and anaerobic glucose-limited continuous cultivation (Chapter 5).

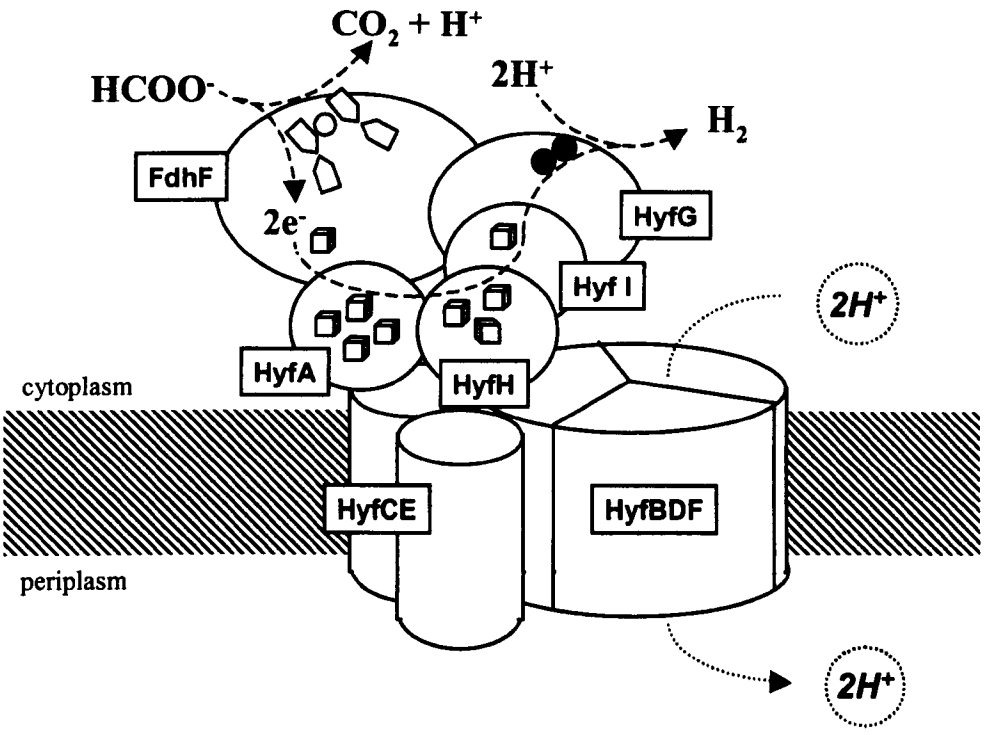


Fig. 1.9. Schematic model for the arrangement of the subunits of the Fdh-F: Hyf complex (Fhl-2) based on the Hyc model of Sauter *et al.* (1992), known structural features of complex I (Weiss *et al.*, 1991; Walker, 1992; Friedrich *et al.*, 1995), the crystal structure of Fdh-F (Boyington *et al.*, 1997) and deductions from Andrews *et al.*, 1997. Originally presented in Andrews *et al.*, 1997 (Fig. 5). Cubes represent putative iron-sulphur clusters. Circles represent other redox active metal centers: a *bis*-(molybdopterin guanine dinucleotide)-liganded molybdenum atom in Fdh-H (open circles); a [Ni-Fe(CO)$(CN)_2$] cluster in HyfG (closed circles).

2. MATERIALS AND METHODS

2.1 Bacterial strains, plasmids and bacteriophage

The bacterial strains used in this study are listed in Table 2.1. The plasmids are listed in Table 2.2.

2.2 Oligonucleotide primers

The oligonucleotide primers used for PCR are listed in Table 2.3.

2.3 Storage and maintenance of bacterial strains, plasmids and bacteriophage

Standard microbiological techniques, as described in Miller (1972), were used when manipulating bacterial strains. The optimum growth temperature (37 °C) was used in this study unless otherwise stated. Agar plates were stored at 4 °C.

For storage, bacterial strains were grown overnight in M9 minimal medium (section 2.5.3.1) supplemented with glucose (0.4 % w/v), centrifuged (12000 x g, 2 min, room temperature), resuspended in 1 % w/v M9 minimal salts containing glycerol (20 % v/v) and stored at –70 °C. Bacterial strains required for immediate use were sub-cultured onto L agar (section 2.5.2.1), M9 minimal agar or Standard minimal agar (section 2.5.3.2) plates from glycerol stocks.

Plasmids were stored at –70 °C in TE buffer (section 2.12.1). Plasmids were maintained within a bacterial host by sub-culturing onto L agar plates or M9 minimal agar containing the appropriate antibiotic.

Bacteriophage were stored at 4 °C in L broth supplemented with 2.5 mM $CaCl_2$, 10 mM $MgSO_4$ and approximately 10 % v/v chloroform.

Table 2.1 The strains of *E. coli* K12 used in this study.

Strains	Genotype	Source
BN450	MC4100, Δ(*ntr*D208::Tn*10*) Δ(*srl-recA*)306::Tn*10*	Birkmann *et al.* (1987)
DS5	MC4100, λ*hyfA-lacZ bla*	P. Golby & S. C. Andrews, University of Reading (unpublished)
DS6	MC4100, λ*hyfA-lacZ bla*, Δ*hyfR*::*spc*	This work
DS7	MC4100, λ*hyfA-lacZ bla*, *fhlA*::λ *plac*Mu53 *kan*	This work
DS8	MC4100, *fhlA*::λ *plac*Mu53 *kan*	This work
DS9	MC4100, λ*hyfA-lacZ bla*, Δ*hycA*	This work
DS10	MC4100, λ*hyfA-lacZ bla*, Δ*hycB-H*::*cat*	This work
DS11	MC4100, λ*hyfA-lacZ bla*, Δ(*ntr*Δ208::Tn*10*)	This work
FTD22	MC4100, Δ*hyaB*	F. Sargent, University of East Anglia (unpublished)
FTD67	MC4100, Δ*hybC*	F. Sargent, University of East Anglia (unpublished)
FTD147	MC4100, Δ*hyaB*, Δ*hybC*, Δ*hycE*	F. Sargent, University of East Anglia (unpublished)
HD700	MC4100, Δ*hycA-H*	Sauter *et al.* (1992)
HD701	MC4100, Δ*hycA*	Sauter *et al.* (1992)
HD705	MC4100, Δ*hycE*	Sauter *et al.* (1992)
HD709	MC4100, Δ*hycI*	Sauter *et al.* (1992)
HDJ123	Hfr (PO1 of Hfr Hayes), Δ(*gpt-lac*)5, *relA1*, *spoT1*, *thi-1*, Δ*hya*::*kan*, Δ*hyb*::*kan*, Δ*hycB-H*::*cat*	Sauter *et al.* (1992)
JRG3615	MC4100, Δ*hyfA-B*::*spc*	Y. S. Chang, P. Golby & S. C. Andrews, University of Reading (unpublished)
JRG3618	MC4100, Δ*hyfR*::*spc*	Y. S. Chang, P. Golby & S. C. Andrews, University of Reading (unpublished)
JRG3621	MC4100, Δ*hyfB-R*::*spc*	Y. S. Chang, P. Golby & S. C. Andrews, University of Reading (unpublished)
JRG3933	HD705, Δ*hyfB-R*::*spc*	This work
JRG3934	JRG3621, Δ*hycB-H*::*cat*	This work
MC10613	MC4100, λ*hycB'-lacZ*$^+$ (transcriptional fusion)	A. Böck, Universität München (unpublished)
MC4100	F$^-$, *araD139*, Δ*(argF-lac)U169*, *ptsF25*, *relA1*, *flbB5301*, *rpsL150*, λ$^-$	Casadaban & Cohen (1979)
M9S	MC4100, *fdhF*::Mu *d*(Ap^R*lac*)ts	Pecher *et al.* (1983)
SV83	MC4100, *fhlA*::λ *plac*Mu53 *kan*	Schlensog *et al.* (1989)

Table 2.2 Plasmids used in this study.

Plasmid	Properties	Source
pACYC184	*cat, tet*	Chang & Cohen (1978)
pGS1020	pSU18, *hyfA-focB*	P. Golby & S. C. Andrews, University of Reading (unpublished)
pGS1087	pSU18, *hyfR*	P. Golby & S. C. Andrews, University of Reading (unpublished)
pGS1037	pMAK705, Δ*hyfA-B*::*spc*	P. Golby & S. C. Andrews, University of Reading (unpublished)
pGS1038	pMAK705, Δ*hyfB-R*::*spc*	P. Golby & S. C. Andrews, University of Reading (unpublished)
pGS1039	pMAK705, Δ*hyfR*::*spc*	P. Golby & S. C. Andrews, University of Reading (unpublished)
pSH9	pACYC184, *fhlA*	A. Böck, Universität München (unpublished)
pSU18	*lacZ, cat*	Bartolome *et al.* (1991)

Table 2.3 Oligonucleotide primers used for PCR.
The position column of the table refers to the position of the base in the relevant sequence in the Appendix (Appendix I & II) that anneals to the 5' base of the primer. Primer *hyfRB*-1F is an exception; the terminal eight bases at the 5' end of this primer are not complementary to the region of the *E. coli* chromosome being amplified, so the 5' base of the complementary region of this primer is underlined.

Primer	Sequence (5'→3')	Appendix	Position (bp)
hyfBR-F1	GTTCTCACTAAGTCGTGTGGAAGCCGTAGC	I	3114
hyfBR-R1	GCAAAAGCAGAGGACAACACCTCGCGAACC	I	13978
hyfR-F1	GATGGCTATGTCAGACGAGGCGATGTTT	I	11954
hyfR-R1	TTCAGACTGTTACCACGAGTCAACAGTACC	I	13705
hyfRB-1F	CCGAATTCCTGGCTGAAAGAACACGC	I	982
lacZ	AGGCGATTAAGTTGGGTAACGCCAGGGTTTTCC	-	-
hyc-F	ATTACATCGCACAGCGGCATCGTCGC	II	242
hyc-R	TATCGTGGCGTCGACAATCAGCAGTCG	II	7798
hycA-L	TGGGAAATAAGCGAGAAAGC	II	220
hycA-R	TTTCGACACTCATCGACACG	II	740

2.4 Purity of strains

The purity of strains used in this study was checked regularly. Individual colonies from agar plates used in experiments were streaked out onto L agar plates (section 2.5.2.1) containing the appropriate antibiotics and incubated aerobically overnight. The presence or absence of growth was recorded.

For batch or continuous culture work carried out in bioreactors the purity of the strain in the bioreactor was checked regularly. A sample of culture was taken and streaked onto L agar and standard minimal agar plates (section 2.5.3.2) and incubated aerobically overnight. The morphologies of the resulting colonies were checked for homogeneity. Also 10 of these colonies were streaked onto L agar plates containing the appropriate antibiotics and incubated aerobically overnight. The presence or absence of growth was recorded. Failure of one colony to possess the desired phenotype, lead to the abandonment of the experiment.

2.5 Growth media

2.5.1 Media preparation

All solutions were prepared in distilled water unless otherwise stated. All thermostable solutions were sterilised by autoclaving (121 °C, 124 kPa, 15 min). Thermolabile solutions were sterilised by filtration through sterile 0.2 μm filters. All solutions were stored at room temperature unless otherwise stated.

2.5.2 Complex media

2.5.2.1 L broth (Lennox, 1955)

	$g\ l^{-1}$
Tryptone	10
Yeast extract	5
NaCl	5

Sterilised by autoclaving.

This medium was used for the growth of bacterial strains unless otherwise stated.

L agar was prepared by adding agar (1.5 % w/v) to the broth before autoclaving. After autoclaving, the medium was cooled (50 °C) before any sterile supplements were added.

Soft-top L agar was prepared by adding agar (0.6 % w/v) to L broth before autoclaving. After autoclaving, the medium was cooled (50 °C) before the addition of sterile $MgSO_4$ (5 mM), $CaCl_2$ (2.5 mM) and any further supplements.

2.5.2.2 TYEP (Begg *et al.*, 1977)

	pH 6.0 l^{-1}	pH 6.6 l^{-1}	pH 7.0 l^{-1}	pH 8.0 l^{-1}
Tryptone	10 g	10 g	10 g	10 g
Yeast extract	5 g	5 g	5 g	5 g
K_2HPO_4 (0.2 M)	61.5 ml	187.5 ml	305 ml	473.5 ml
KH_2PO_4 (0.2 M)	438.5 ml	312.5 ml	195 ml	26.5 ml

Sterilised by autoclaving.

TYEP was made to a pH of 6.6 unless otherwise stated.

2.5.3 Minimal media

2.5.3.1 M9 minimal medium

	l^{-1}
M9 minimal salts (Sigma)	10 g

Autoclaved and allowed to cool (50 °C) before the addition of the following autoclaved supplements;

	l^{-1}
$CaCl_2$ (0.1 M)	2 ml
$MgSO_4$ (1 M)	2 ml
Vitamin B_1 (1 % w/v)	1 ml
Glucose (20 % w/v)	20 ml

M9 minimal agar was prepared by adding agar (3 % w/v) to distilled water and autoclaving. After autoclaving the agar was cooled (50 °C) and an equal volume of x2 strength M9 minimal medium added.

2.5.3.2 Standard minimal medium for bioreactor work

This medium was used solely for work reported in chapter five of this thesis and is referred to simply as standard minimal medium in chapter five.

	l^{-1}
$(NH_4)_2SO_4$	2.0 g
K_2HPO_4	1.0 g
NaH_2PO_4	1.0 g
$MgSO_4 \cdot 7H_2O$	0.2 g
Vishniac's trace elements	2.0 ml

Sterilised by autoclaving (121 °C, 124 kPa, 50 min).

This medium was used for controlled batch and continuous culture work carried out in bioreactors. For batch growths, the pH of the medium was adjusted to 7.0 with KOH (5 M) prior to autoclaving.

Standard minimal agar was prepared by adding agar (1.5 % w/v) to the medium before autoclaving. After autoclaving the agar was cooled (50 °C) before sterile (by autoclaving) glucose (20 mM) or any further sterile supplements were added.

Vishniac's trace element solution (Vishniac and Santer, 1957)

	$g\ l^{-1}$
Na_2EDTA	50
$ZnSO_4 \cdot 7H_2O$	22
$CaCl_2 \cdot 6H_2O$	5.54
$MnCl \cdot 4H_2O$	5.06
$FeSO_4 \cdot 6H_2O$	4.99
$(NH_4)_6Mo_7O_{24} \cdot 4H_2O$	1.10
$CuSO_4.5H_2O$	1.57
$CoCl_2.6H_2O$	1.61

The salts were dissolved in the order shown above. The pH of the solution was adjusted to 6.0 with NaOH (5 M) in between the addition of each salt.

2.5.4 Carbon substrates used for growth

All carbon substrates except ethanol were sterilised by autoclaving and stored as sterile stock solutions. Ethanol was sterilised by filtration through sterile 0.2 μm filters. The concentrations used varied and are indicated for each experiment.

Substrate	Stock concentration
glucose	20 % w/v
sodium formate (pH 7.0)	1 M
sodium acetate	1 M
sodium lactate	1 M
gluconate	1 M
glucuronate	1 M
sorbitol	1 M
ethanol	95% (v/v)

2.5.5 Growth supplements

Growth supplements were filter sterilised and stored as sterile stock solutions. Actual concentrations used in experiments will be in the text associated with individual experiments.

Supplement	Stock concentration	Storage temperature
X-gal (in DMF)	20 g ml^{-1}	-20°C
tryptone	10 % w/v	room temperature
yeast extract	10 % w/v	room temperature

2.5.6 Antibiotic stock solutions

Antibiotic stock solutions were filter sterilised and stored as sterile stock solutions. Actual concentrations used in experiments are indicated in the text associated with individual experiments.

Antibiotic	Stock concentration	Storage temperature
chloramphenicol (prepared in methanol)	10 mg ml^{-1}	-20 °C
spectinomycin	10 mg ml^{-1}	-20 °C
kanamycin	10 mg ml^{-1}	-20 °C
ampicillin	100 mg ml^{-1}	-20 °C

2.6 Growth conditions

Unless otherwise stated, the growth of bacterial cultures was monitored by measuring the OD at 650 nm. OD was measured using a Thermo Unicam Heλios β spectrophotometer.

For the work presented in chapter five, OD at 430 nm was used to estimate the cell density of bacterial cultures and monitor their growth in the bioreactor. To investigate whether the relationship between OD at 430nm (measured using a Thermo Unicam Heλios β spectrophotometer) and cell density is linear the OD of an overnight culture was measured over a range of dilutions (Appendix III). OD (430 nm) has a linear relationship with cell density up to an OD (430 nm) of approximately 0.6. Therefore in chapter five bacterial culture samples were diluted 10 fold before OD (430 nm) was measured and OD (430nm) measurements above 0.6 were not used to estimate bacterial cell density or biomass.

OD at 430nm was also used to estimate the biomass (dry weight; g l^{-1}) of bacterial cultures in the bioreactors. Overnight cultures were harvested by centrifugation and the cells resuspended in different volumes of standard minimal medium for bioreactor work (section 2.5.3.2). The OD (430 nm) and dry weight (section 2.9.5) of the bacterial cultures was measured and the results used to generate a calibration curve (Appendix III).

2.6.1 Batch culture growths

2.6.1.1 Aerobic batch culture – 6″ test tubes

Aerobic growths in 6″ test tubes were used as an inoculation source for experiments. The tubes were filled with 2 - 5 ml of the appropriate growth medium and stoppered with cotton wool. The tubes were autoclaved either before or after the addition of the growth medium. The sterile growth medium was inoculated with a colony picked from an agar plate. Incubation was carried out overnight in a shaking incubator (200 rpm) or in a growth room on a rocking rack.

2.6.1.2 Aerobic batch culture – shake flasks

For the aerobic growth of 50 - 100 ml of bacterial culture, 250 ml conical flasks stoppered with cotton wool were used. The flasks were autoclaved either before or after the addition of the growth medium. The sterile growth medium was inoculated with 1 ml of an overnight culture or a colony picked from an agar plate. Incubation was carried out in a shaking incubator (250 rpm) or in a growth room on a shaking rack (250 rpm).

2.6.1.3 Anaerobic batch culture – 8ml Bijoux

Anaerobic batch growths were carried out in 8 ml Bijoux with rubber sealed screw top caps. The bijoux were autoclaved prior to being completely filled with inoculated growth medium. The sterile growth medium (100 ml) was supplemented as necessary, and inoculated (1 ml of an overnight culture), before being aliquoted into the bijoux. When a sample was required the contents of a single Bijoux were harvested. The bijoux were incubated in a water bath.

2.6.1.4 Controlled batch culture – fermenters

Controlled batch growth was carried out in vessels (LH Fermentation 500 series) controlled for pH, temperature and dissolved oxygen tension. The pH was monitored using a pH electrode (Ingold type 472, 170mm) connected to a pH controller (LH Fermentation 500). The pH was maintained at either 6.5 or 7.5 by titration with 2 M KOH fed to the bioreactor by the pH controller. The temperature was maintained at 37 °C by a thermostat and heater element connected to a temperature controller (LH Fermentation 503). The dissolved oxygen tension of the medium was measured using a galvanic type oxygen electrode (constructed in the laboratory) connected to an oxygen monitor (LH Fermentation 507).

The fermenter was filled with 900 ml Standard minimal medium (section 2.5.3.2) prior to autoclaving. After autoclaving the medium was supplemented with 20 mM glucose, 5 μM nickel chloride, 1.6 μM ferric citrate, 1 μM sodium molybdate and 1 μM sodium selenite.

Before inoculation the fermenter was flushed with sterile air (30 l hr^{-1}, impeller speed 500 rpm) and the dissolved O_2 tension set to 100 %. Once inoculated,

airflow was stopped to the fermenter, the impeller speed reduced to 150 rpm and the O_2 tension allowed to drop to 0 % with culture growth.

Bacterial strains to be cultivated were initially grown on standard minimal agar plates. Two 250 ml shake flasks each containing 50 mM standard minimal medium were each inoculated with three colonies from these plates and grown for 20-22 hrs. The volume of inoculum from each flask required to give an initial OD at 430 nm of 0.2 was calculated using the following equation:

$$X = \frac{V}{(A_1 \div A_2) - 1}.$$

X = volume of inoculum required to give desired initial OD at 430 nm A_2 (l).

V = volume of medium in the bioreactor (l).

A_1 = OD at 430 nm of inoculum (A).

A_2 = OD at 430 nm desired after the addition of inoculum (A).

Cell density was monitored, by measuring the OD at 430 nm. Sampling was carried out using a tap located at the base of the bioreactor. Any fermentation gases produced were left to accumulate in the bioreactor headspace and pressure build up was allowed to escape via a tube running from the bioreactor to an inverted water filled measuring cylinder (gas trap).

2.6.2 Controlled continuous culture – chemostats

Controlled continuous culture growths were carried out in vessels (LH Fermentation 500 series) controlled for pH, temperature and dissolved oxygen tension (section 2.6.1.4). Working volumes of 690, 700 and 785 ml were used in this study.

The chemostat was filled with 450 ml Standard minimal medium prior to autoclaving and subsequently supplemented with 20 mM glucose, 5 μM nickel chloride, 1 μM sodium molybdate and 1 μM sodium selenate.

Before inoculation the oxygen electrode was calibrated to 100 % as described in section 2.6.1.4. For aerobic continuous culture the O_2 tension was maintained above 30 % by adjustment of the air flow rate or impeller speed (LH fermentation 502D agitator). For anaerobic continuous culture, airflow was stopped to the chemostat immediately after inoculation and the O_2 tension allowed to drop to 0 %

with culture growth. When the OD at 430 nm reached approximately 0.5 or greater the culture was sparged with N_2 (flow rate and impeller speed in text associated with individual experiments).

Inoculation of controlled continuous cultures was as described for controlled batch cultures (section 2.6.1.4).

Cell density was monitored by measuring the OD at 430nm and a steady state was assumed to have been reached when the OD remained constant for five culture volumes. Sampling was carried out by temporarily closing exhaust gas flow, and allowing the build up in pressure inside the chemostat to force culture into a sampling tube.

2.7 Molecular biology techniques

2.7.1 Purification of genomic DNA

2.7.1.1 Promega Wizard™ genomic DNA Purification Kit

The bacterial strain was grown overnight and 1 ml harvested by centrifugation (12000 x g, 2 min, room temperature). The cells were resuspended in 600 µl Nuclei Lysis Solution and incubated at 80 °C for 5 min. After incubation, 3 µl of RNase (section 2.12.2) was added and the sample mixed by inversion. After mixing, the sample was incubated on ice for 5 min, centrifuged (12000 x g, 3 min, room temperature) and the supernatant removed to a clean microfuge tube containing 600 µl isopropanol. The sample was gently mixed by inversion until the threadlike strands of DNA became visible. The DNA was harvested by centrifugation (12000 x g, 2 min, room temperature) and washed in 600 µl ethanol (70 % v/v). The DNA was again harvested by centrifugation, the microfuge tube drained and allowed to air dry for 10 -15 min. The DNA was dissolved in 100 µl DNA Rehydration Solution by incubation at 65 °C for 1 h. The DNA was stored at -20°C.

2.7.1.2 Small scale preparation of genomic DNA (Ausubel *et al.*, 1989)

The bacterial strain was grown overnight and 1 ml harvested by centrifugation (12000 x g, 2 min, room temperature). The cells were resuspended in 567 µl TE

(50mM Tris, 20 mM EDTA), 30 µl SDS (10 % w/v), 3 µl Proteinase K (20 mg ml^{-1}) and 5 µl RNase (section 2.12.2). After incubation at 37 °C for 1 h, 100 µl of NaCl (5 M) were added and the sample mixed by inversion. CTAB-Mix (80 µl) were added, the sample mixed again and incubated at 65 °C for 10 min. After incubation, 700 µl of phenol/chloroform (1:1) was added, the sample vortexed and then centrifuged (12000 x g, 15 min, room temperature). The aqueous phase was removed to a microfuge tube containing 700 µl phenol/chloroform (1:1), vortexed, centrifuged and the aqueous phase removed to a microfuge tube containing 700 µl ethanol (100 % w/v). The sample was gently mixed by inversion until the threadlike strands of DNA became visible. The DNA was harvested by centrifugation (12000 x g, 15 min, room temperature) and washed in 700 µl ethanol (100 % v/v). The DNA was again harvested by centrifugation and washed in 700 µl ethanol (70 % v/v). After centrifugation, the microfuge tube was drained and dried under vacuum. The DNA was dissolved in 100 µl TE (section 2.12.1) and stored at –20°C.

CTAB Mix

	$g\ l^{-1}$
NaCl	10
Cetyl-trimethyl-ammonium bromide (CTAB)	100

After the addition of NaCl, the CTAB was added slowly, while heating. The CTAB-Mix was stored at 4 °C. The CTAB-Mix may precipitate at low temperatures and therefore was heated to 50 °C for a few min before use.

2.7.2 Ethanol precipitation of DNA

The ethanol precipitation of DNA in solution was achieved by the addition of 0.1 volumes of sodium acetate (3 M) and 3 volumes of 100 % v/v ethanol. After incubation for 20 min at –20 °C, the DNA was harvested by centrifugation (12000 x g, 15 min, room temperature) and washed with 70 % v/v ethanol. The DNA was again harvested by centrifugation, the microfuge tube drained and allowed to air dry for 10 - 15 min. The DNA was dissolved in TE (section 2.12.1) and stored at –20 °C.

2.7.3 PCR amplification of chromosomal DNA

The PCR was used in this study to verify the identity of new mutants created by P1 transduction. The oligonucleotides used in PCR are listed in table 2.3. A thermocycler (Progene Techne) was used for all PCR reactions. Some PCR reactions were 'hot started' by heating the reaction mixture to denaturing temperature in the thermocycler before the addition of the polymerase enzyme.

2.7.3.1 Amplification of *hyfB-R*

Primers *hyfBR*-F1 and *hyfBR*-R1 were used to amplify a 10864 bp DNA fragment from the *hyfB-R* region of the *hyf* operon (Table 2.3).

Amplification was performed using the Promega Expand™ Long Template PCR system. The PCR reaction mixture consisted of the following components:

Master mix 1:	
dNTP, 10 mM	10 µl
upstream primer (5 µM)	1 µl
downstream primer (5 µM)	1 µl
template DNA	0.5 µl
sterile distilled H_2O	12.5 µl

Master mix 2:	
10 x PCR buffer with $MgCl_2$ (promega buffer 2)	5 µl
enzyme mix (Taq/Pwo; 3.5 units/µl)	0.75 µl
sterile distilled H_2O	19.25 µl

Master mix 1 and master mix 2 were mixed together before cycling. The following temperature cycle profile was used:

1 x	Denature	94 °C	2 mins
10 x	Denature	94 °C	10 sec
	Anneal	60.5 °C	30 sec
	Elongate	68 °C	6 mins
20 x	Denature	94 °C	10 sec
	Anneal	60.5 °C	30 sec
	Elongate	68 °C	6 mins (+ 20 sec cycle elongation for each consecutive cycle)
1 x	Elongate	68 °C	7 mins

2.7.3.2 Amplification of *hyfR* and *hyfA-lacZ*

PCR primers *hyfR*-F1 and *hyfR*-R1 were used to amplify a 1751 bp region of DNA from the *hyfR* gene. In strains carrying the Δ*hyfR* mutation, the region of DNA amplified with these primers was 2190 bp long. Primers *hyfRB*-1F and *lacZ*-R were used to amplify a region of DNA from the *hyfA-lacZ* gene fusion.

The PCR reaction mixture consisted of the following components:

dNTP, 10 mM	1 µl
upstream primer (5 µM)	1 µl
downstream primer (5 µM)	1 µl
template DNA	0.5 µl
10 x Taq buffer	5 µl
sterile distilled H_2O	41 µl
Taq polymerase (5 units/µl) *	0.5 µl

*Hot start

The following temperature cycle profile was used:

35 x	Denature	94 °C	30 sec
	Anneal	56.5 °C	1 min
	Extend	74 °C	3 min 30 sec
			(+ 10 sec cycle elongation
			for each consecutive cycle)

2.7.3.3 Amplification of *hycB-H*

PCR primers *hyc*-F and *hyc*-R were used to amplify a 7556 bp fragment of DNA from the *hycB-H* region of the *hyc* operon. Amplification was performed using the Promega Expand™ Long Template PCR system. The PCR reaction components and the cycle profile used was identical to that previously described in section 2.7.3.1.

However, in a Δ*hycB-H* mutant the region amplified by these primers was only 1931 bp long so a different PCR protocol was sometimes used.

The PCR reaction mixture was as described in section 2.7.3.2.

The following temperature cycle profile was used:

1 x	Denature	95 °C	45 sec
35 x	Denature	95 °C	45 sec
	Anneal	63.2 °C	45 sec
	Extend	72 °C	3 min 30 sec
			(+ 10 sec cycle elongation
			for each consecutive cycle)

2.7.3.4 Amplification of *hycA*

PCR primers *hycA-F* and *hycA-R* were used to amplify a 521 bp fragment of DNA from the *hycA* gene. In strains carrying the Δ*hycA* mutation, the region of DNA amplified with these primers was 243 bp long.

The PCR reaction mixture consisted of the following components:

dNTP, 10 mM	1 µl
upstream primer (5 µM)	1 µl
downstream primer (5 µM)	1 µl
template DNA	0.75 µl
10 x Taq buffer	5 µl
$MgCl_2$ (50 mM)	1.5 µl
sterile distilled H_2O	39 µl
Taq polymerase (5 units/µl) *	0.75 µl

*Hot start

The following temperature cycle profile was used:

1 x	Denature	95°C	45 sec
35 x	Denature	95 °C	45 sec
	Anneal	51.3 °C	1 min
	Extend	72 °C	2 min 10 sec

2.7.4 Agarose gel electrophoresis

Agarose gels (0.7 % w/v) were used to detect and separate DNA fragments either amplified by PCR or cut in restriction enzyme digests. Agarose (0.7 g) and 100 ml TBE (x 0.5) were heated together in a 250 ml conical flask using a microwave (full power, 90 sec). The flask was allowed to cool on the bench for 15 min and incubated in a 50 °C water bath for a further 15 min. After incubation, the contents of the flask were poured into a taped BioRad electrophoresis plate, a comb inserted and the gel left to set for approximately 1 h. Once set the comb and tape were removed and the gel placed into an electrophoresis tank filled with TBE (x 0.5). Loading buffer (3 µl) and sterile H_2O (15 µl) were added to 2 µl of the sample and loaded onto the gel. The DNA was fractionated by electrophoresis at 70 V for 60 - 90 min.

The gel was stained either by the addition of 1 µl ethidium bromide (10 mg ml^{-1}) to the gel before pouring or by incubation of the gel in an ethidium bromide

solution (0.5-1.0 μg l^{-1} in x 0.5 TBE) for 15 – 30 min. DNA in the gel was visualised in a UV-transilluminator.

Size markers were run alongside samples so that fragment sizes could be estimated. The marker used is described in the text associated with individual experiments. The correct sizes of the fragments amplified by PCR were verified by plotting a calibration curve (fragment size (log_{10}) verses mobility).

x5 TBE

	$g\ l^{-1}$
Tris	54
boric acid	27.5
EDTA	4.65

Agarose gel loading buffer

xylene cyanol	25 mg
glycerol	3 ml
bromophenol blue	25 mg
EDTA	146 mg
distilled water	7 ml

Stored at –20°C.

2.7.5 Restriction enzyme digests

Restriction enzyme digests were used to verify the identity of plasmid preparations. Enzyme digest reactions consisted of the following:

plasmid preparation	5 μl
x 10 restriction enzyme buffer	2 μl
sterile distilled H_20	12 μl
restriction enzyme (10 units/μl)	1 μl

The reaction mixture was incubated at 37 °C for approximately 1 h.

The specific restriction enzyme used is indicated in the text associated with individual experiments.

2.7.6 Plasmid preparations

Medium scale plasmid preparations were carried out using the Promega Wizard® *Plus* Midiprep Kit. The appropriate transformant was grown in 100 ml L-broth with the required antibiotic and was harvested by centrifugation (3200 x g, 2 min, room temperature). The cells were resuspended in 3 ml Cell Resuspension Solution and lysed 3 ml Cell Lysis Solution. The sample was neutralised with 3 ml Neutralisation Solution and any cell debris removed by centrifugation (12000 x g, 15 min, 4 °C). After centrifugation, 10 ml DNA Purification Resin were added to the sample and loaded into the loading reservoir of a Midiprep Column. A vacuum was applied to the column to draw the sample through it. The column was washed twice, by drawing 15 ml Column Wash Solution through it. The loading reservoir was separated from the Midicolumn and the column transferred to a microfuge tube. Any residual Column Wash Solution was removed by centrifugation (12000 x g, 2 min, room temperature) and 300 µl of water was added to the column. The DNA was eluted from the column by centrifugation (12000 x g, 20 sec, room temperature) and any resin fines removed from the eluted DNA by centrifugation (12000 x g, 5 min, room temperature). The plasmid DNA was stored at –20 °C.

2.7.7 Transformation

2.7.7.1 Preparation of competent cells (Hanahan, 1985)

L broth (50 ml, section 2.5.2.1) was inoculated with 0.5 ml of an overnight culture and incubated at 37 °C with shaking (250 rpm) until the culture reached an OD at 650 nm of approximately 0.5. The cells were harvested by centrifugation (3200 x g, 15 min, room temperature) and resuspended in 8 ml ice cold TF-1. After incubation on ice for 15 min the cells were harvested by centrifugation as before and resuspended in 4 ml ice cold TF-2. The competent cells were aliquoted into microfuge tubes and stored at –70 °C.

TF-1

RbCl	12 g
(or KCl)	(or 7.4 g)
CH_3COOK (1 M, pH 7.5)	30 ml
$CaCl_2.2H_2O$	1.5 g
glycerol	150 g

The above components were made up to a final volume of 950 ml and the pH adjusted to pH 6.4 with 0.2 M acetic acid. After autoclaving the following were added:

$MnCl_2.4H_2O$ (1 M)	50 ml

TF-2

	g
RbCl	1.2
(or KCl)	(or 0.74)
$CaCl_2.2H_2O$	11
glycerol	150

The above components were made up to a final volume of 980 ml, autoclaved and the following added:

MOPS (0.5 M, pH 6.8 adjusted with KOH)	20 ml

2.7.7.2 Transformation reaction

Various dilutions of plasmid DNA were added to 0.2 ml competent cells and incubated on ice for 40 min. After incubation, the cells were heat shocked at 42 °C for 1 min and placed on ice for 5 min. Prewarmed L broth (0.7 ml, section 2.5.2.1) was added and the cells incubated at 37 °C for 1 h to allow resistance genes to be expressed. The cells were centrifuged (12000 x g, 2 min, room temperature) and 0.7 ml of the supernatant removed. The cells were resuspended in the remaining

supernatant and spread on L agar plates supplemented with the appropriate antibiotics. The L agar plates were incubated overnight at 37 °C.

2.7.8 P1 phage transduction

All strains produced in this study were generated by P1 phage transduction.

2.7.8.1 Production of P1 phage lysate

A previously titred P1 phage lysate was diluted in P1 dilution fluid (section 2.12.3) to approximately 10^6, 10^7 and 10^8 pfu ml^{-1}. Each dilution (0.1 ml) was added to 0.1 ml of an overnight culture of the donor strain. After incubation at 37 °C for 15 min, 2.5 ml prewarmed (50 °C) L-agar soft top (supplemented with 2.5 mM $CaCl_2$ and 10 mM $MgSO_4$, section 2.5.2.1) were added, briefly mixed and poured onto a prewarmed (50 °C) L agar plate (supplemented with 2.5 mM $CaCl_2$ and 5 mM $MgSO_4$). The L agar soft top was left to set and the plates incubated at 37 °C overnight. After overnight incubation, 3.5 ml L broth (supplemented with 2.5 mM $CaCl_2$ and 10 mM $MgSO_4$) were added to any plates exhibiting confluent lysis and these plates left on a level surface at 4 °C for 6 h. After 6 h incubation, the L broth was transferred to a glass universal, 1 ml chloroform added and any cell debris removed by centrifugation (3200 x g, 5 min, room temperature). The resulting phage lysate was stored at 4 °C.

2.7.8.2 Preparation of plating cells

L broth (50 ml, supplemented with 2.5 mM $CaCl_2$ and 10 mM $MgSO_4$, section 2.5.2.1) was inoculated with 0.5 ml overnight culture and incubated at 37 °C with shaking (250 rpm) until the culture reached an OD at 650 nm of approximately 0.5. The cells were harvested by centrifugation (3200 x g, 5 min, room temperature) and resuspended in 5 ml $CaCl_2$ (10 mM) and 5 ml $MgSO_4$ (20 mM). The plating cells were stored at 4 °C.

2.7.8.3 Titration of P1 phage lysates

A x 10 dilution series of the P1 lysate was set up in P1 dilution fluid (section 2.12.3) and 0.1 ml of each dilution added to 0.1 ml of plating cells. After incubation at 37 °C for 15 min, 2.5 ml prewarmed (50 °C) L agar soft-top (supplemented with 2.5 mM

$CaCl_2$ and 10 mM $MgSO_4$, section 2.5.2.1) were added, briefly mixed and poured onto a prewarmed (50 °C) L-agar plate (supplemented with 2.5 mM $CaCl_2$ and 5 mM $MgSO_4$). The L agar soft top was left to set and the plates incubated at 37 °C overnight. After overnight incubation, the plaques were counted and the titre of the lysate calculated.

2.7.8.4 P1 phage transduction

Plating cells (0.1 ml) were infected with approximately 4 x 10^6, 4 x 10^7 and 4 x 10^8 phage and incubated at 37 °C for 20 min. After incubation the cells were harvested, resuspended in 1 ml L broth (supplemented with 0.5 M trisodium citrate, section 2.5.2.1) and incubated at 37 °C for 1 h to allow any resistance gene to be expressed. The cells were centrifuged (12000 x g, 2 min, room temperature) and 0.9 ml of the supernatant removed. The cells were resuspended in the remaining supernatant and spread on L agar plates supplemented with the appropriate antibiotic. The plates were incubated overnight at 37 °C and transductants counted.

2.8 Protein techniques

2.8.1 Polyacrylamide gel electrophoresis

Polyacrylamide gels were used to separate proteins later visualised by Coomassie blue staining or autoradiography. Alternatively fractionated proteins were electroblotted and visualised by immunostaining. The polyacrylamide gels used in the ^{63}Ni incorporation experiments were different to those described here. The differences are described in (section 2.8.4).

Polyacrylamide resolving gels (12 %) were used and consisted of the following components:

resolving gel buffer (1.5 M Tris-Cl, pH 8.8)	5 ml
acrylamide (30 % w/v)	6.7 ml
distilled water	4.7 ml
SDS (10 % w/v)	200 µl
$(NH_4)_2S_2O_8$ (10 % w/v)	70 µl
TEMED	15 µl

Immediately after the addition of the TEMED, the gel was poured into a Protean mini gel assembly on a pouring rack. Water saturated butanol (300 µl) was pipetted onto the surface of the gel and the gel was left to set for 30 min. Once set, the butanol was removed and the top of the gel washed with distilled water. The stacking gels used consisted of the following components:

stacking gel buffer (0.5 M Tris-Cl, pH 6.8)	2.5 ml
acrylamide (30 % w/v)	1.5 ml
distilled water	5.9 ml
SDS (10 % w/v)	100 µl
$(NH_4)_2S_2O_8$ (10 % w/v)	35 µl
TEMED	15 µl

Immediately after the addition of the TEMED, the gel was poured into the gel assembly on top of the resolving gel and a comb inserted. Once set, the comb was removed and the wells rinsed with distilled water. The gel assembly was placed into an electrophoresis tank and the tank filled with gel running buffer.

Whole cell homogenates for electrophoretic analysis were prepared by growing the required bacterial strain to stationary phase. At stationary phase, 0.5 OD units of the bacterial culture were taken and the cells were harvested by centrifugation (3200 x g, 5 min, room temperature). The cells were resuspended in 100 µl SDS-digestion buffer and boiled at 100 °C for 3 min. Once cool, 10 µl of the sample was again centrifuged, the supernatant loaded onto the gel and the proteins in the sample fractionated by electrophoresis at 30 mA for 45 min.

Gel running buffer (pH 8.3)

	$g\ l^{-1}$
tris	3.05
glycine	14.4
SDS	1

SDS digestion buffer

resolving gel buffer (1.5 M Tris-Cl, pH 8.8)	2.0 ml
glycerol	0.7 ml
SDS (10 % w/v)	2.3 ml
bromophenol blue	10 mg
β-mercaptoethanol	0.5 ml

2.8.2 Coomassie blue gel staining

After electrophoresis the polyacrylamide gel (section 2.8.1) was placed in a staining dish and approximately 100 ml of Coomassie blue gel stain added. After gentle shaking for 1 h, the gel stain was removed, the gel rinsed with water and approximately 100 ml destain. The destain was renewed at regular intervals until background staining of the gel had been sufficiently reduced. The gel was viewed on a light box and photographed.

Destain

methanol	300 ml
acetic acid	100 ml
distilled water	600 ml

Coomassie blue gel stain

Coomassie blue R-250	2 g
Destain	1 l

2.8.3 Western blotting

After electrophoresis, the polyacrylamide gel (section 2.8.1) was soaked in electroblot buffer for 20 min, together with a similar sized piece of nitrocellulose membrane (Schleicher and Schuell-Protan BA83) and four similar sized pieces of filter paper (Whatman 3MM). The gel and blotting membrane were sandwiched together between the filter paper in an electrotransfer cassette and the cassette placed into an electroblotting tank. The tank was filled with electroblot buffer and electrotransfer was performed at 4 °C at 50 V for 1 h. Sea blue markers were run alongside samples to ensure that electrotransfer had been successful and so that fragment sizes could be estimated. After electrotransfer the membrane was washed in 100 ml TBS for 5 min and incubated in TBS (supplemented with 1 % w/v BSA) for 30 min. After incubation the membrane was washed twice in 100 ml TTBS for 5 min and incubated in 100 ml primary antiserum solution for 2 hr. In this study the primary antiserum used were either anti-HycE (0.1 % v/v) or anti-FTN (0.1 % v/v). These antisera were raised against proteins in rabbits and supplied by A. Böck (Universität München) and S. C. Andrews (University of Reading). After incubation the membrane was again washed twice in 100 ml TTBS and incubated with 100 ml Sigma anti-rabbit alkaline phosphatase conjugate solution (0.003 % v/v, secondary antibody solution) for 2 h. The membrane was washed in 100 ml TBS for 5 min and incubated in 100 ml substrate solution. The membrane was rinsed in water to stop colour development and photographed. Both antibody solutions were made up in TBS containing 1 % w/v BSA.

Electroblot buffer

Tris	9.09 g
glycine	43.2 g
methanol	600 ml
distilled water	2.4 ml

TBS

	$g\ l^{-1}$
Tris	2.42
NaCl	58.44

The pH was adjusted to 7.5 with HCl.

TTBS

TBS	1 l
Tween-20	0.5 ml

Substrate solution

	l^{-1}
$MgCl_2.6H_2O$	0.102 g
tris	12.1 g
NBT (30 mg ml^{-1}, 70 % v/v DMF)	10 ml
Boehringer BCIP (50 mg ml^{-1})	3 ml

Made on day of use.

2.8.4 ^{63}Ni incorporation experiments

TYEP (pH 6.6, section 2.5.2.2) was supplemented with 0.4 % w/v glucose, 50 mM sodium formate and 5 µM ^{63}Ni ($^{63}NiCl_2$ stock solution: approx. 30 mM, specific activity 27.92 $mCi.ml^{-1}$). The medium (100 ml) was inoculated with 0.5 ml of an overnight culture and incubated anaerobically at 37 °C in 8 ml bijoux until stationary phase was reached. At stationary phase, 5 OD units of the bacterial culture were taken and the cells harvested by centrifugation (3200 x g, 5 min, room temperature). The cells were washed in 1 ml saline (section 2.12.4) and resuspended in 300 µl TrS buffer. EDTA (30 µl, section 2.12.5) was added and the sample incubated at room temperature for 5 min. After incubation, 30 µl lysozyme (10 mg ml^{-1}) were added and the sample was incubated at room temperature for 10 min and vortexed periodically. The sphaeroplasts were harvested by low speed centrifugation (3000 x

g, 30 sec, room temperature) and lysed with 300 µl TrM with regular vortexing for 10 min. Triton X-100 (33.3 µl, 1 % w/v) and 30 µl glycine (50 % w/v) were added to the sample. Nucleic acids were precipitated by adding 30 µl streptomycin sulphate (10 % w/v) and incubating the sample on ice for 30 min. Nucleic acids and cell debris were removed by centrifugation (12000 x g, 5 min, room temperature).

To separate ^{63}Ni binding proteins, a 5 % w/v polyacrylamide resolving gel (section 2.8.1) was used. Triton X-100 (0.1 %) was added to the resolving gel, stacking gel and gel running buffer in place of SDS.

Native PAGE loading buffer containing 0.1 % w/v Triton X-100 (10 µl) was added to 20 µl of each cell extract and loaded onto the gel. The ^{63}Ni containing proteins were separated by electrophoresis at 100 V for 1 h. The gel was dried at 70 °C under vacuum.

For autoradiography, Kodak BioMax MS film with a Kodax BioMax intensifying screen and an exposure time of 2 weeks at –70 °C was used. Exposures were carried out in a Hypercassette Autoradiography Cassette Standard Kodak developing and fixing reagents were used.

TrS

Tricine (1 M)	20 ml
sucrose	50 g
distilled water	200 ml

pH adjusted to 8.0 with HCl. Stored at –20°C.

TrM

Tricine (1 M)	20 ml
$MgSO_4.7H_2O$	0.05 g
distilled water	200 ml

pH adjusted to 7.5 with HCl. Stored at –20°C.

Native PAGE loading buffer with 0.1% Triton X-100

distilled water	5.5 ml
Tricine (1M, pH 8)	2 ml
glycerol	2.4 ml
Triton X-100 (10 % w/v)	100 µl
bromophenol blue	20 mg

Stored at –20°C.

2.8.5 Enzyme assays

For hydrogenase assays (section 2.8.5.4) the amount of protein was estimated. It was assumed that 10^9 cells yield approximately 150 µg protein and that an OD at 650 nm of 1 corresponds to 10^9 cells ml^{-1}.

2.8.5.1 β-Galactosidase assay

β-Galactosidase assays were performed by the method described by Phillips-Jones and co-workers (1989). O-nitrophenyl-β-D-galactosidase (ONPG) is cleaved by β-galactosidase to release the yellow coloured o-nitrophenol:

$$\text{ONPG} \xrightarrow{\beta\text{-galactosidase}} \text{o-nitrophenol} + \text{galactose}$$

The bacterial strain to be assayed was grown in duplicate to late stationary phase and 0.5 OD units sampled from each culture throughout the growth curve. The cells were harvested by centrifugation (12000 x g, 2 min, room temperature), washed in 1 ml saline (section 2.12.4) and the cell pellet stored at –70 °C. Cells were thawed and resuspended in 300 µl TS buffer. EDTA (30µl, section 2.12.5) was added and the sample incubated at room temperature for 5 min. After incubation, 30 µl lysozyme (10 mg ml^{-1}) was added and the sample incubated at room temperature for 10 min, vortexing periodically. The sphaeroplasts were harvested by low speed centrifugation (3000 x g, 30 sec, room temperature) and lysed with 300 µl TMD with repeated vortexing for 10 min. Nucleic acids were precipitated by adding 30µl

streptomycin sulphate (10 % w/v) and incubating on ice for 30 min. Nucleic acids and cell debris were removed by centrifugation (12000 x g, 5 min, room temperature).

Each sample was assayed for β-galactosidase activity in duplicate. Cell extracts (50 µl) were mixed with 150 µl of ONPG solution in a flat bottomed microtitre plate and the rate of production of o-nitrophenol measured at 28 °C in a plate reader equipped with kinetic software using a 414 nm filter (10 nm band width). Normally, a Labsystems iEMS plate reader or a STECTRAmax 340 pc microplate reader were used in this study. TMD (45.5 µl) and 4.5 µl spectinomycin sulphate mixed with 150 µl ONPG solution was used as a blank for the reaction. One mmol o-nitrophenol in a volume of 200 ml has an OD at 414 nm of 10.6724 (as measured using the iEMS plate reader). The nitrophenol production rate (i.e. ONPG cleavage rate) and total protein concentration were used to calculate a specific β-galactosidase activity (mmol ONPG cleaved^{-1} min^{-1} mg protein^{-1}).

Protein concentrations were determined by the dye-binding method of Bradford (1976) using the Bio-Rad protein assay kit. The protein concentration of each sample was measured in duplicate. Bio-Rad reagent was diluted 1:4 v/v in water and filtered. Cell extracts (5 µl) were diluted with 15 µl distilled water and mixed with 180 µl of the diluted Bio-Rad reagent in a flat bottomed microtitre plate. After incubation at room temperature for 5 min, the OD at 600 nm was measured using a plate reader. A calibration curve was generated using γ globulin standards at 0, 10, 20, 50, 75, 100, 125 and 250 µg ml^{-1}. The standards were made up in TMD.

TS

Tris	2.43 g
sucrose	50 g
distilled water	200 ml

pH adjusted to 8.0 with HCl. Stored at –20°C.

TMD

Tris	2.422 g
$MgSO_4.7H_2O$	0.05 g
DTT	0.154 g
distilled water	200 ml

Stored at –20°C.

ONPG solution

Na_2HPO_4	1.704 g
NaH_2PO_4	1.248 g
KCl	0.15 g
MgSO4·7H2O	0.05 g
distilled water	200 ml

pH adjusted to 7.5 and the following added:

ONPG	0.18 g
DTT	0.1542 g

Divided into 20 ml aliquots, stored at –20°C and each aliquot discarded after use.

2.8.5.3 Formate dehydrogenase-H (Fdh-H) assay

The Fdh-H activity of whole cells made permeable with toluene was assayed by measuring formate dependent benzyl viologen reduction at 600 nm (Ballantine and Boxer, 1986; Axley *et al.*, 1990).

The bacterial strain to be assayed was grown anaerobically to stationary phase and 100 ml harvested by centrifugation (3200 x g, 5 min, room temperature). The cells were washed in 100 mM sodium phosphate buffer (supplemented with 10 mM EDTA; section 2.12.5) and resuspended in the same sodium phosphate buffer to a concentration of 250 mg cells ml^{-1}. The pH of the sodium phosphate buffer used in

Fdh-H assays was 6.6 or 7.6 depending on the growth conditions of the culture. Toluene (50 µl) was added and the sample incubated on ice for 5 min.

The assay mixture was as follows:

	Assay concentration	Stock concentration
sodium phosphate buffer	85 mM	100 mM
benzyl viologen	2 mM	40 mM
sodium formate	20 mM	1 M
ATP*	12.5 mM	200 mM
$MgSO_4$	4.25 mM	5 mM

* Added to assay mixture for assays at pH 7.6.

The assay mixture was titrated with a solution of sodium dithionite (10 mg ml^{-1} in 10 mM NaOH) until the absorbance at 600 nm was stable. The reaction was started by the addition of 10 - 20 µl of the whole cell suspension (made permeable with toluene). The increase in absorbance at 600 nm was measured as a function of time using distilled water as a blank. An extinction co-efficient value of 7400 M^{-1} cm^{-1} at 600 nm was assumed for reduced benzyl viologen (Ballantine and Boxer, 1986). The benzyl viologen reduction rate and total protein concentration were used to calculate a specific Fdh-H activity (µmol benzyl viologen reduced min^{-1} (mg protein^{-1})).

Protein concentrations in cell suspensions were determined using the Bio-Rad DC (Detergent Compatible) Protein Assay, based on the method of Lowry and co-workers (1951). The protein concentration of each sample was measured in duplicate. Cell suspensions were diluted 100 fold in 100 mM phosphate buffer (supplemented with 1 % w/v SDS) and boiled at 100 °C for 3 min. Diluted cell suspensions (5 µl) were mixed with 15 µl distilled water and 180 µl Bio-Rad DC reagent in flat-bottomed microtitre plates. After incubation at room temperature for 15 min the OD at 650 nm was measured using a STECTRAmax 340 pc microplate reader (Molecular Devices, Sunnyvale CA 94089, USA). A calibration curve was generated using γ globulin standards at 0, 0.5, 0.75, 1.0, 1.5 and 2.0 mg ml^{-1}. The

standards were made up in 100 mM phosphate buffer (supplemented with 1 % w/v SDS, pH 7.6).

2.8.5.4 Hydrogenase assays

The hydrogen uptake activity of whole cells permeabilised with toluene was assayed by measuring hydrogen dependent benzyl viologen reduction at 600 nm (Ballantine & Boxer, 1986).

The bacterial strain to be assayed was grown anaerobically to stationary phase and 2 – 2.5 OD units harvested by centrifugation (3200 x g, 5 min, room temperature). The cells were resuspended in 200 µl sodium phosphate buffer (100 mM, pH 7.0, degased) and 50 µl toluene added. The sample was incubated on ice for 5 min. The assay mixture was as follows:

	Assay concentration	Stock concentration
sodium phosphate buffer (pH 7.0, H_2 saturated)	93.75 mM	100 mM
benzyl viologen	12.5 mg ml^{-1}	50 mg ml^{-1}

The assay mixture was titrated with a solution of sodium dithionite (10 mg ml^{-1} in 10 mM NaOH) until the OD at 600 nm was stable. The reaction was started by the addition of 20 µl of the permeabilised cell suspension to the assay mixture and the increase in absorbance at 600 nm measured as a function of time using water as a blank. An extinction co-efficient value of 7400 $M^{-1}cm^{-1}$ at 600 nm was assumed for reduced benzyl viologen (Ballantine and Boxer, 1986). The benzyl viologen reduction rate and estimated protein concentration were used to calculate a specific hydrogenase activity (µmol benzyl viologen reduced min^{-1} (mg $protein^{-1}$)).

2.8.5.5 Gas evolution assays

Bacterial strains were grown anaerobically in 6″ test tubes stoppered with metal caps. Any gas produced was trapped in inverted Durham tubes placed in the 6″ test tubes before sterilization.

2.8.5.6 Hydrogen evolution assays

Hydrogen evolution activity was assayed using a Clarke-type oxygen electrode (Hansatech Ltd) adopted for the measurement of hydrogen (Ballantine and Boxer, 1986).

The bacterial strain to be assayed was grown anaerobically overnight and 500 ml harvested by centrifugation (3200 x g, 15 min, room temperature). The cells were washed twice in 10 ml sodium phosphate buffer (100 mM) and resuspended in the same sodium phosphate buffer to a concentration of 50 mg ml^{-1}. The pH of the phosphate buffer used was 6.8 or 7.5 depending on the growth conditions of the bacterial culture. The reaction chamber (2 ml) of the electrode was filled with the 0.05 g ml^{-1} cell suspension. Oxygen in the reaction chamber was removed by adding 100 µl glucose (0.5 M) and 10 µl glucose catalase/ glucose oxidase mixture. Once the electrode trace was steady the reaction was initiated by the addition of 100 µl sodium formate (0.4 M). The 'hydrogen electrode' was calibrated with H_2 saturated distilled water. At 30 °C, 774 nmoles of hydrogen are dissolved in 1 ml H_2 saturated distilled water. A H_2 evolution rate was calculated using an estimated protein concentration (mmol hydrogen $evolved^{-1}$ min^{-1} mg $protein^{-1}$).

2.9 Analyses

Culture samples from the bioreactor were centrifuged (12000 x g, 2 min, room temperature) and the supernatant analysed for glucose or fermentation product concentration.

2.9.1 Glucose determination

Glucose concentrations in culture supernatant were measured enzymatically by the method of Bergmeyer and Bernt (1965):

$$\text{D-Glucose} + H_2O + O_2 \xrightarrow{\text{GOD}} \text{Gluconic acid} + H_2O_2$$

$$H_2O_2 + \text{o-Dianisidine}^{RED} \xrightarrow{\text{POD}} H_2O + \text{o-Dianisidine}^{OX}$$

Glucose oxidase (GOD) was used to convert glucose to gluconic acid. The hydrogen peroxide produced was estimated by peroxidase (POD)-catalysed oxidation of the dye, o-dianisidine. The coloured derivative formed was measured at 436 nm.

Reagent I

	$g\ l^{-1}$
$Na_2HPO_4.2H_2O$ (or Na_2HPO_4)	13.8 (or 11.0)
$NaH_2PO_4.2H_2O$	7.26
(or NaH_2PO_4)	(or 5.58)
peroxidase	0.02
glucose oxidase (300 U mg^{-1})	0.013

Made on day of use.

Reagent II

o-Dianisidine-HCl	5 mg ml^{-1}

Made on day of use.

Reagent III

Reagent I (10 ml) mixed with 0.1 ml reagent II. Made on day of use.

Samples were diluted to give expected glucose concentrations of less than 0.5 mM. Reagent III (1.25 ml) was added to 100 µl of each sample and incubated at room temperature for 35 min. The absorbance at 436 nm was measured using distilled water as a blank. A calibration curve was generated using glucose standards at 0, 0.1, 0.2, 0.3, 0.4, 0.5 and 1.0 mM.

2.9.2 Ethanol determination

Ethanol concentration in culture supernatants was measured using an enzymic bioanalysis kit specific for ethanol (Boehringer Mannheim), based on the method of Beutler (1984).

$$\text{Ethanol} + NAD^+ \xrightarrow{\text{ADH}} \text{acetaldehyde} + NADH + H^+$$

$$\text{Acetaldehyde} + NAD^+ + H_2O \xrightarrow{\text{Al-DH}} \text{acetic acid} + NADH + H^+$$

Ethanol is oxidised to acetaldehyde in the presence of the enzyme alcohol dehydrogenase (ADH). Concomitantly, NAD^+ is reduced to NADH. Subsequently, acetaldehyde is oxidised by aldehyde dehydrogenase (Al-DH) to acetic acid whilst a further NAD is reduced to NADH. NADH formation is determined by means of absorbance at 340 nm.

Bottle 1 (Boehringer Mannheim)

potassium diphosphate buffer (pH approximately 9)
stabilizers

Bottle 2 (Boehringer Mannheim)

NAD	4 mg/tablet
aldehyde dehydrogenase (Al-DH)	0.8 U/tablet
stabilizers	

Reaction mixture 2*

Bottle 1	3 ml
Bottle 2	1 tablet

Samples were diluted to give an expected ethanol concentration of less than 2 mM. Reaction mixture 2* (1 ml) was added to 33 µl of each sample, mixed, and the

absorbance (A_1) measured at 340 nm. Alcohol dehydrogenase (17 µl, approx 4 – 5 U μl^{-1}) was added, the sample mixed and the absorbance (A_2) at 340 nm measured after 10 min incubation at room temperature. A blank was prepared by adding distilled water in place of the sample to reaction mixture 2*. The change in absorbance was noted ($\Delta A = (A2 - A1)_{SAMPLE} - (A2 - A1)_{BLANK}$). An extinction coefficient value of 6.3 mM^{-1} cm^{-1} at 340 nm was used for NADH. The accuracy of the assay was checked with an ethanol standard solution provided with the kit.

2.9.3 Formate determination

Formate concentration in culture supernatants was measured using an enzymic bioanalysis kit specific for formate (Boehringer Mannheim), based on the methods of Höpner and Knappe (1974) and Schaller and Triebig (1984).

$$\text{Formate} + NAD^+ + H_2O \xrightarrow{FDH} \text{bicarbonate} + NADH + H^+$$

Formate is quantitatively oxidised in the presence of formate dehydrogenase (FDH) by nicotinamide-adenine dinucleotide (NAD) to bicarbonate. The increase in NADH is measured by means of absorbance at 340 nm.

Bottle 1 (Boehringer Mannheim)

potassium phosphate buffer (pH approximately 7.5)
stabilizers

Bottle 2 (Boehringer Mannheim)

NAD	420 mg
Li salt	

Reaction mixture 2

Dissolve contents of bottle 2 with the contents of bottle 1.

Samples were diluted to give an expected formate concentration of less than 2 mM. Reaction mixture 2 (0.333 ml) was mixed with 33 µl sample and 663 µl distilled water. The absorbance (A_1) at 340 nm measured after incubation at room temperature for 5 min. Formate dehydrogenase (17 µl, 50 - 55 U ml^{-1}) was added, the sample mixed and the absorbance (A_2) at 340 nm measured after 20 min incubation at room temperature. A blank was prepared by adding distilled water in place of the sample to reaction mixture 2. The change in absorbance was noted ($\Delta A = (A2 - A1)_{SAMPLE} - (A2 - A1)_{BLANK}$). A calibration curve was generated using formate standards at 0 mM, 0.5 mM, 1.0 mM and 1.5 mM.

2.9.4 HPLC measurement of organic acid concentration

HPLC was used to detect and quantify organic acids in culture supernatants. The supernatants were filtered through 0.2 µM filters before analysis. A pump (Waters 625), autosampler (Waters 700), Shodex Ionopak KC-811 column (8 mm x 300 mm), *uv* spectrophotometer (Waters 490E, set at 210 nm), and differential refractometer (Waters 410) were coupled to a computer equipped with control and data acquisition and analysis software (Millennium 2010). Samples were eluted using 0.1 % orthophosphoric acid at a flow rate of 0.5 or 1 ml min^{-1}. The method of detection was either the *uv* spectrophotometer or the refractometer. Standard solutions of various organic acids (0.1 - 100 mM) were used as standards to calibrate peak area with concentration.

2.9.5 Dry weight ($\bar{x}$)

Culture samples (3 x 10 ml) were removed from the chemostat vessel and the cells harvested by centrifugation (3200 x g, 20 min, room temperature). The cells were washed with 2 ml distilled water and centrifuged as before. The cells were resuspended in 1 ml distilled water and transferred to pre-dried and weighed metal caps. The cell pellets in the centrifuged tubes were resuspended in 1 ml distilled water and the resuspended cells were then added to the metal caps. The metal caps were incubated in a 100 °C oven overnight before being re-weighed and the average dry weight of cells (g l^{-1}) calculated.

The dry weight of cells (g l^{-1}) was also estimated using a calibration curve of OD at 430 nm against dry weight (section 2.6; Appendix III).

2.10 Metabolic rates

2.10.1 Maximum growth rate (μ_{max})

The OD at 430 nm of a growing culture was plotted against time as a semi-log plot. The doubling time of the culture during exponential growth was measured using the linear portion of the graph and this value used to calculate the maximum growth rate of the culture:

$$\mu_{max} = \frac{\ln 2}{\text{doubling time (h)}}.$$

2.10.2 Rate of substrate utilisation ($q_{SUBSTRATE}$)

The rate of substrate utilisation was expressed as mmol h^{-1} (g dry weight)$^{-1}$ and calculated using the following equation:

$$q_{SUBSTRATE} = \frac{(f \times S_R) - (f^i \times s)}{X \times V}$$

f = medium input flow rate (l h^{-1})

f^i = output flow rate (i.e. medium plus titrant, 1 h^{-1})

S_R = reservoir substrate input concentration (mM)

s = residual substrate concentration in the culture filtrate (mM)

X = steady-state cell density (g l^{-1})

V = working volume (l)

2.11 Steady-state analysis

2.11.1 Maintenance energy (M_e)

The maintenance energy (mmol (g dry weight)$^{-1}$ h^{-1}) is the amount of substrate required to maintain cell viability. The value is determined from the Y intercept of the plot of $q_{SUBSTRATE}$ against dilution rate (h^{-1}).

2.11.2 Maximum biomass yield (Y_{MAX})

The maximum yield (g dry weight (mol substrate)$^{-1}$) is determined as the inverse of the slope of the plot of $q_{SUBSTRATE}$ against dilution rate (h^{-1}).

2.12 Solutions

2.12.1 TE

	mM
Tris base	10
EDTA	1

pH of the buffer was adjusted to 8.0 with HCl.

2.12.2 RNase (DNase free; Sambrook *et al.*, 1989)

pancreatic RNase (RNaseA)	100 mg
Tris	12 mg
NaCl	9 mg
sterile distilled water	10 ml

Heated to 100 °C for 15 min and allowed to cool to room temperature. Stored at –20°C.

2.12.3 P1 dilution fluid

NaCl	0.3 g
peptone	1 g
$MgSO_4$ (0.5 M)	1 ml
tris (1 M, pH 7.8)	10 ml
distilled water	up to 1 l

2.12.4 Saline

	$g\ l^{-1}$
NaCl	37.2
EDTA	8.75

2.12.5 EDTA

EDTA	37.25 $g\ l^{-1}$

pH adjusted to 8.0 with NaOH. Stored at –20 °C.

3. HYDROGENASE SPECIFIC PHENOTYPIC ANALYSIS OF THE *hyf* OPERON OF *E. coli.*

3.1 Introduction

Three [Ni-Fe] hydrogenases (hydrogenase-1,-2 & -3) have been identified and characterised in *E. coli* (reviewed by Sawers 1994; Ballantine & Boxer, 1985, 1986; Sawers *et al.*, 1985; Sawers & Boxer, 1986; Böhm *et al.*, 1990; Menon *et al.*, 1990, 1994; Sauter *et al.*, 1992). Hydrogenases-1 and -2, are uptake hydrogenases, coupling hydrogen oxidation to the energy conserving reduction of menaquinone. Hydrogenase-1 is synthesised during fermentative growth, however its exact physiological role is still uncertain. Hydrogenase-2 is synthesised during anaerobic growth on non-fermentable carbon sources such as hydrogen and fumarate. It is thought that hydrogenase-2 has a respiratory function, allowing cells to gain energy from the oxidation of molecular hydrogen under these growth conditions. Hydrogenase-3, like hydrogenase-1, is expressed during fermentative growth. Together with formate dehydrogenase H (Fdh-H), it forms the formate hydrogenlyase (Fhl-1) complex, catalysing the non-energy conserving breakdown of formate to CO_2 and H_2. The existence of a fourth hydrogenase in *E. coli* was postulated after sequence analysis of the *E. coli* chromosome (Andrews *et al*, 1997). It was proposed that the hydrogenase-4 (*hyf*) operon encoded a [Ni-Fe] hydrogenase complex that together with Fdh-H formed an energy-conserving (proton translocating) formate hydrogenlyase (Fhl-2).

Initial experiments showed that a *hyfA-lacZ* translational fusion was expressed during fermentative growth indicating that the *hyf* promoter is active and that *hyf* is likely to be a functional operon (M. C. Berry, S. C. Andrews & B. C. Berks, unpublished observations; Andrews *et al,* 1997). Further experiments with a *hyfA-lacZ* transcriptional fusion strain (DS5) revealed maximum expression during fermentative growth at acidic pH in the presence of formate and the absence of

electron acceptors (sections 4.9 & 4.10.1). To investigate the phenotype of the *hyf* operon, deletions in the *hyfA-B*, *hyfB-R* and *hyfR* regions were constructed and introduced into the chromosome of *E. coli* strain MC4100 using the gene replacement method described by Hamilton *et al.* (1989) (Y. S. Chang, P. Golby & S. C. Andrews, unpublished). Initial experiments with these mutant strains revealed no significant effect of the deletions on growth rate or the reduction of nitrate (Y. S. Chang & S. C. Andrews, unpublished).

In the studies described in this chapter, early experiments were established to obtain evidence for a fourth [NiFe] hydrogenase in *E. coli* encoded by the *hyf* operon. Immunoblotting experiments were carried out with anti-HycE (large subunit of hydrogenase-3) serum to try and detect the proposed large subunit of hydrogenase-4 (HyfG) and demonstrate that *hyf* gene products are present. ^{63}Ni incorporation experiments were also carried out to see whether nickel is associated with HyfG. Sauter and co-workers (1992), detected no hydrogenase activity attributable to a fourth hydrogenase in the hydrogenase triple mutant HDJ123 (*Δhya, Δhyb, ΔhycB-H*). Therefore, hydrogenase assays, Fdh-H assays, gas production experiments and H_2 production assays were performed in order to detect activity attributable to the *hyf* operon.

3.2 Construction of *hyf* deletion strains JRG3615, JRG3618 and JRG3621

A strategy for constructing the *hyf* deletion strains JRG3615 (MC4100, *ΔhyfA-B::spc*), JRG3618 (MC4100, *ΔhyfR::spc*) and JRG3621 (MC4100, *ΔhyfB-R::spc*) was designed by Dr S. C. Andrews and executed by Dr P. Golby and Mr Y. S. Chang. DNA fragments were constructed with the appropriate *hyf* regions deleted and replaced with a spectinomycin cassette (Appendix I). These DNA fragments were cloned into vector pMAK705 (Hamilton *et al.*, 1989) to create plasmids pGS1037, pGS1038 and pGS1039 (Table 2.2). These plasmids were then used to transfer the deletions to the chromosome of the wildtype strain MC4100 using the method of Hamilton *et al.* (1989). The mutant strains were verified by PCR and Southern hybridisation and were found to carry deletions of the correct size.

3.3 Construction of *hyf* encoding multicopy plasmids pGS1020 and pGS1087

Plasmids pGS1020 and pGS1087 are multicopy plasmids encoding the entire *hyf* operon (*hyfA-focB*) and the *hyfR* gene respectively (Table 2.2). A strategy for constructing the plasmids was designed and executed by Dr P. Golby.

Plasmid pGS1020 was constructed by cloning the approximately 15 kb *Eco*RI fragment containing the whole of the *hyf* operon into *Sma*I and *Eco*RI digested pSU18.

Plasmid pGS1087 was constructed as follows. A 2.03 kb *Eco*RI-*Nsi*I fragment containing the whole of the *hyfR* gene was cloned in *Eco*RI-*Pst*I digested pUC118 to give pGS1018. The 2.03 kb *Nde*I-*Hind*III fragment of pGS1018 (containing the *hyfR* gene) was cloned into *Nde*I and *Hind*III digested pET21a to give pGS1019. The approximately 2 kb *Xba*I-*Hind*III fragment of pGS1019 (containing the *hyfR* gene) was cloned into *Xba*I and *Hind*III digested pSU18 to give pGS1087.

3.4 Construction of *hyc/hyf* deletion strain JRG3934

Strain JRG3934 (MC4100, Δ*hycB-H*::*cat*, Δ*hyfB-R*::*spc*) was created as a control strain for the immunological detection of the HyfG polypeptide using anti-HycE serum (section 3.4). The strain was created using a P1 lysate grown on strain HDJ123 (MC4100, Δ*hycB-H*::*cat*) to transfer the Δ*hycB-H*::*cat* deletion to strain JRG3618 (MC4100, Δ*hyfB-R*::*spc*) by transduction.

3.4.1 PCR amplification of the *hycB-H* region

To confirm that predicted alterations had been made to the chromosome in strain JRG3934, the *hycB-H* region was amplified using PCR with primers designed to amplify the *hycA-hycI* region (primers *hyc*-F and *hyc*-R) (Table 2.3; section 2.7.3.3). As controls, the corresponding region was amplified from MC4100, HDJ123 and JRG3621. The PCR product of JRG3934 and HDJ123 was approximately 1800 bp compared with approximately 7590 bp for MC4100 and JRG3621 (Fig 3.1A & B). Different PCR protocols were used to amplify the 1931 bp DNA fragment from

JRG3934 and HDJ123 and the 7556 bp DNA fragment from MC4100 (section 2.7.3.3). The size of the PCR products were in agreement with predicted sizes from known sequence data (Table 3.1).

3.4.2 PCR amplification of the *hyfB-R* region

To confirm that predicted alterations had been made to the chromosome in strain JRG3634, the *hyfA-R* region was amplified using PCR with primers designed to amplify the *hyfB-hyfR* region (primers *hyfBR*-F1 and *hyfBR*-R1) (Table 2.3; section 2.7.3.1). As controls, the corresponding region was amplified from MC4100, JRG3621 and HDJ123. The PCR product of JRG3634 and JRG3621 was approximately 2190 bp compared with over 10000 bp for MC4100 and JRG3615 (Fig. 3.2). The size of the PCR products were in agreement with predicted sizes from known sequence data (Table 3.2).

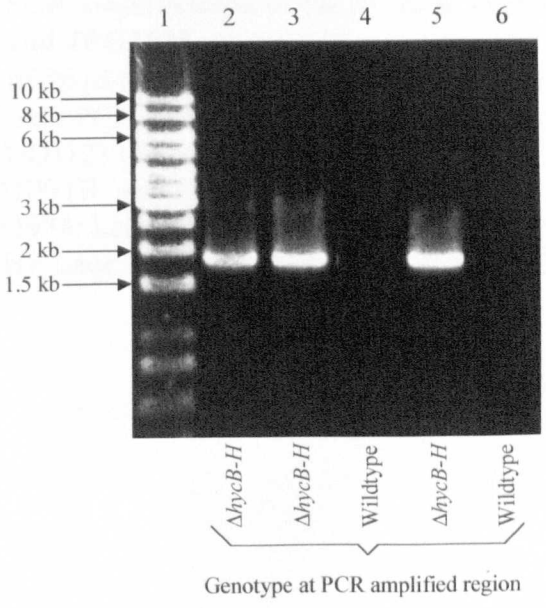
3.5 Construction of *hyc/hyf* deletion strain JRG3933

Strain JRG3933 (MC4100, Δ*hycE*, Δ*hyfB*-R::*spc*) was created as a control strain for ^{63}Ni incorporation experiments (section 3.7). The strain was created using a P1 lysate grown on strain JRG3621 (MC4100, Δ*hyfB-R*::*spc*) to transfer the Δ*hyfB*-R::*spc* deletion to strain HD705 (MC4100, Δ*hycE*) by transduction.

3.5.1 PCR amplification of the *hycE* region

To confirm that predicted alterations had been made to the chromosome in strain JRG3933, the *hycE* region was amplified using PCR with primers designed to amplify the *hycA-hycI* (primers *hyc*-F and *hyc*-R) (Table 2.3; section 2.7.3.3). As controls, the corresponding region was amplified from MC4100 and HD705. The PCR product of JRG3933 and HD705 was approximately 5750 kb compared with approximately 7590 kb for MC4100 (Fig. 3.3). The sizes of the PCR products were as predicted from the sequence data (Table 3.3).

(A)


1
2
3
4
5
6
10 kb
8 kb
6 kb
3 kb
2 kb
1.5 kb
ΔhycB-H
ΔhycB-H
Wildtype
ΔhycB-H
Wildtype
Genotype at PCR amplified region

(B)

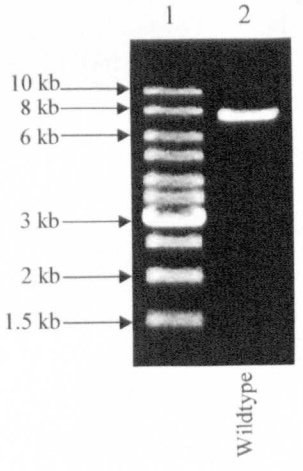

1
2
10 kb
8 kb
6 kb
3 kb
2 kb
1.5 kb
Wildtype

Fig. 3.1 (Page 96). PCR amplification of the *hycB-H* region from JRG3934, MC4100, HDJ123 and JRG3621.

PCR products (2 μl of 50 μl reaction) were electrophoresed on an agarose gel (0.7 %; section 2.7.4). Different PCR protocols were used to amplify the 1931 bp fragment from JRG3934 and HDJ123 (A; section 2.7.3.2 & 2.7.3.3) and the 7556 bp DNA fragment from MC4100 (B; section 2.7.3.1 & 2.7.3.3). (A): Lane 1, 1 kb DNA ladder; Lane 2, JRG3934; Lane 3, JRG3934; Lane 4, MC4100; Lane 5, HDJ123; Lane 6, JRG3621. (B): Lane 1, 1 kb DNA ladder; Lane 2, MC4100.

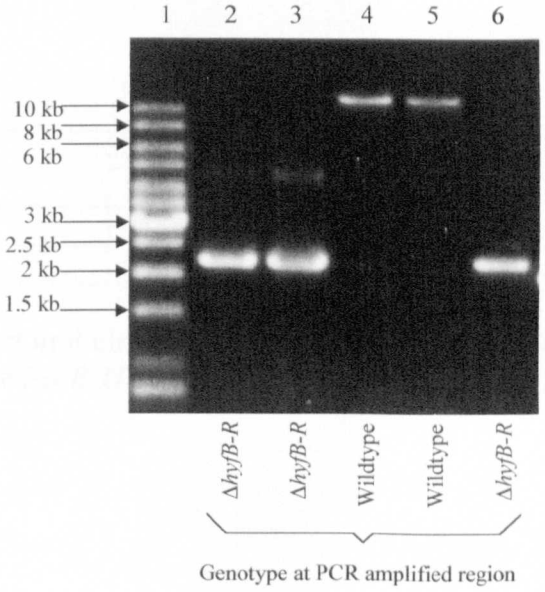


Fig. 3.2. PCR amplification of the *hyfB-R* region from JRG3634, MC4100, JRG3621 and HDJ123.

PCR products (2 µl of 50 µl reaction) were electrophoresed on an agarose gel (0.7 %; section 2.7.4). Lane 1, 1 kb DNA ladder; Lane 2, JRG3934; Lane 3, JRG3934; Lane 4, MC4100; Lane 5, HDJ123; Lane 6, JRG3621.

Strain	Size (bp)	
	Expected	Observed (estimated)
MC4100	7556	7590
JRG3934	1931	1800

Table 3.1. Expected and observed (estimated) sizes of PCR products from the amplification of the *hycB-H* region from MC4100 and JRG3634.

Strain	Size (bp)	
	Expected	Observed (estimated)
MC4100	10864	>10000
JRG3934	2229	2190

Table 3.2. Expected and observed (estimated) sizes of PCR products from the amplification of the *hyfB-R* region from MC4100 and JRG3634.

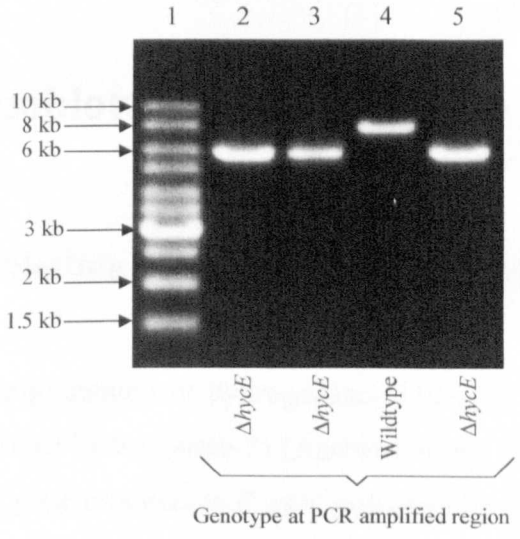


Fig. 3.3. PCR amplification of the *hycE* region from JRG3933, MC4100 and HD705.

PCR products (2 µl of 50 µl reaction) were electrophoresed on an agarose gel (0.7 %; section 2.7.4). Lane 1, 1 kb DNA ladder; Lane 2, JRG3933; Lane 3, JRG3933; Lane 4, MC4100; Lane 5, HD705.

Strain	Size (bp)	
	Expected	Observed (estimated)
MC4100	7556	7760
JRG3933	5914	5750

Table 3.3. Expected and observed (estimated) sizes of PCR products from the amplification of the *hycA-I* region from MC4100 and JRG3933.

3.5.2 PCR amplification of the *hyfB-R* region

To confirm that predicted alterations had been made to the chromosome in strain JRG3633, the *hyfA-R* region was amplified using PCR as described in section 3.4.2.

3.6 Immunoblotting with anti-HycE serum

3.6.1 Immunoblotting analysis of mutants carrying deletions in *hyf* and *hyc*

HyfG (proposed large subunit of hydrogenase-4) has 73% amino acid identity with HycE (large subunit of hydrogenase-3) (Andrews *et al.*, 1997). In order to establish the presence of *hyf* gene products in *E. coli*, polyclonal anti-HycE serum (supplied A. Böck, Universität München) was used in an attempt to detect HycE and HyfG in whole cell extracts from mutant strains carrying deletions in the *hyc* and *hyf* operons (Fig. 3.4; sections 2.8.1 & 2.8.3).

The 61 kD mature form of HycE was found to be present in the wildtype strain only (Fig. 3.4, lane 6). Extracts from HD709 (Δ*hycI*) accumulated HycE in its larger unprocessed form (65 kD) that migrated more slowly that the mature protein in the SDS gel (Fig 3.4, lane 4). This is because the *hycI* gene encodes the protease, HycI, required for the C-terminal processing of HycE (Rossmann *et* al., 1995). As expected, HycE was absent from the Δ*hycA-H* and Δ*hycE* strains (Fig. 3.4, lanes 2, 3 & 5).

A band of unknown identity (designated Band A) was detected in all strains. Band A is not a polypeptide specific to the anti-rabbit antiserum (secondary antiserum) used in the immunoblotting procedure (section 2.8.3), as immunoblotting solely with this anti-rabbit antiserum did not reveal Band A or any other polypeptides (Fig. 3.5). The anti-HycE antiserum was produced by overproducing the HycE polypeptide, purifying on SDS-polyacrylamide gels and raising antiserum in a rabbit (Sauter *et al.*, 1992). Band A could possibly be a polypeptide co-eluted from the SDS-polyacrylamide gel with the overproduced unprocessed HycE polypeptide. The detection of Band A with this anti-HycE antiserum has not been previously reported (Sauter *et al.*, 1992). In addition, a number of faint unknown bands were detected,

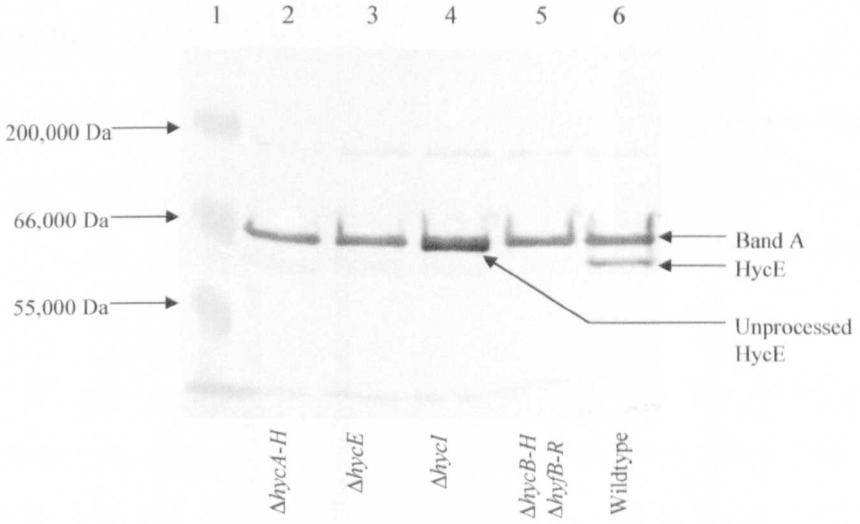


Fig. 3.4. Immunoblotting analysis of extracts from mutants carrying deletions in the *hyc* and *hyf* operons with antisera raised against HycE. Strains were grown anaerobically in TYEP (pH 6.6; section 2.5.2.2) + 0.4 % Glucose + 30 mM sodium formate + 5 μM nickel chloride. Preparations of extracts and immunoblotting procedure were carried out as described in sections 2.8.1 and 2.8.3. Lane 1, Sea Blue Marker; Lane 2, HD700 (Δ*hycA-H*); Lane 3, HD705 (Δ*hycE*); Lane 4, HD709 (Δ*hycI*); Lane 5, JRG3934 (Δ*hycB-H* Δ*hyfB-R*); Lane 6, MC4100 (wildtype).

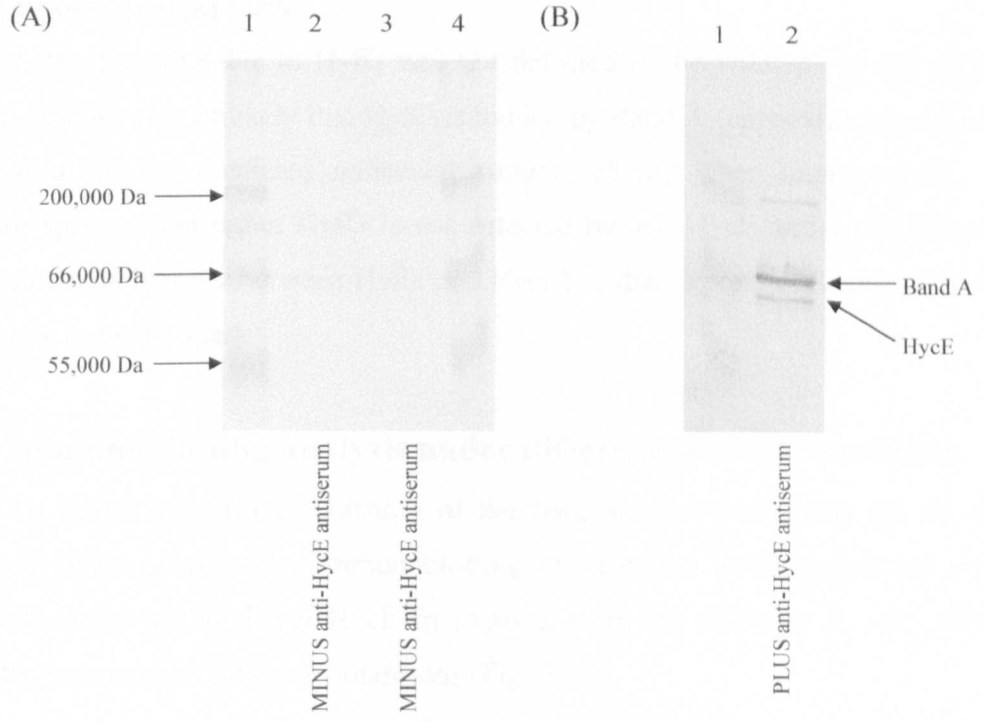


Fig. 3.5. Immunoblotting analysis of extracts from the wildtype strain MC4100, (A) solely with anti-rabbit antiserum (secondary antiserum), or (B) with both anti-HycE antiserum and anti-rabbit antiserum (normal immunoblotting procedure, sections 2.8.1 & 2.8.3). MC4100 was grown anaerobically. (A) Lane 1, Sea Blue Marker; Lane 2, MC4100 grown in LB + 0.4 % glucose + 30 mM sodium formate; Lane 3, MC4100 grown in TYEP (pH 6.6; section 2.5.2.2) + 0.4 % glucose + 30 mM sodium formate; Lane 4, Sea Blue Marker. (B) Lane 1, Sea Blue Marker; Lane 2, MC4100 grown in LB + 0.4 % glucose + 30 mM sodium formate.

however these bands were present in all strains and therefore are not *hyc* or *hyf* operon encoded polypeptides.

A band attributable to HyfG was not detected in the wildtype or any of the deletion strains. It is unlikely that HyfG is hidden by Band A (approximately 65 kD) as it has a smaller predicted molecular weight (63 kD when unprocessed). It therefore appears that either HyfG is not detected by anti-HycE antiserum (despite 73% amino acid identity between HyfG and HycE) or that expression levels of HyfG are too low to be detected.

3.6.2 Immunoblotting analysis under different growth conditions

To further confirm the identity of the band suspected of being the 61 kD processed HycE polypeptide, immunoblotting experiments were carried out with anti-HycE antiserum to detect HycE in extracts from the wildtype *E. coli* strain, MC4100, grown under different conditions (Fig. 3.6).

It has been previously shown that expression of the *hyc* operon, including *hycE*, is favoured by the absence of oxygen or other external electron acceptors (e.g. nitrate), the presence of formate and molybdate and an acidic pH (Peck & Gest, 1957; Wimpenny & Cole, 1967; Pecher *et al.*, 1983; Birkmann *et al.*, 1987b; Schlensog *et al.*, 1989; Schlensog & Böck, 1990; Rossmann *et al.*, 1991). Therefore, as expected, immunoblotting revealed an increase in the amount of HycE polypeptide in extracts from MC4100 grown anaerobically, in the absence of the external electron acceptor nitrate and in the presence of formate (Fig. 3.6, lanes 2, 4, 7 and 9). Although the presence of 30 mM sodium nitrate reduced the amount of immunologically detectable HycE, small amounts of the polypeptide were still detected (Fig. 3.6, Lane 5). The repression of *hyf* expression by nitrate is an indirect effect of depletion of the formate pool by the activity of Fdh-N (Rossmann *et al.*, 1991). It appears that supplementation of the medium with 30 mM sodium formate may have been enough to at least partially overcome this nitrate drainage effect.

The quantity of the 61 kD processed form of HycE increased, when the growth medium was supplemented with 5 μM nickel chloride (Fig. 3.6, lanes 7 and 9). Transcription of the *hycB* gene has been shown to be independent of nickel by studying *hycB-lacZ* expression in nickel-depleted medium (Zinoni *et al.*, 1984). C-terminal processing of HycE requires prior insertion and incorporation of nickel into

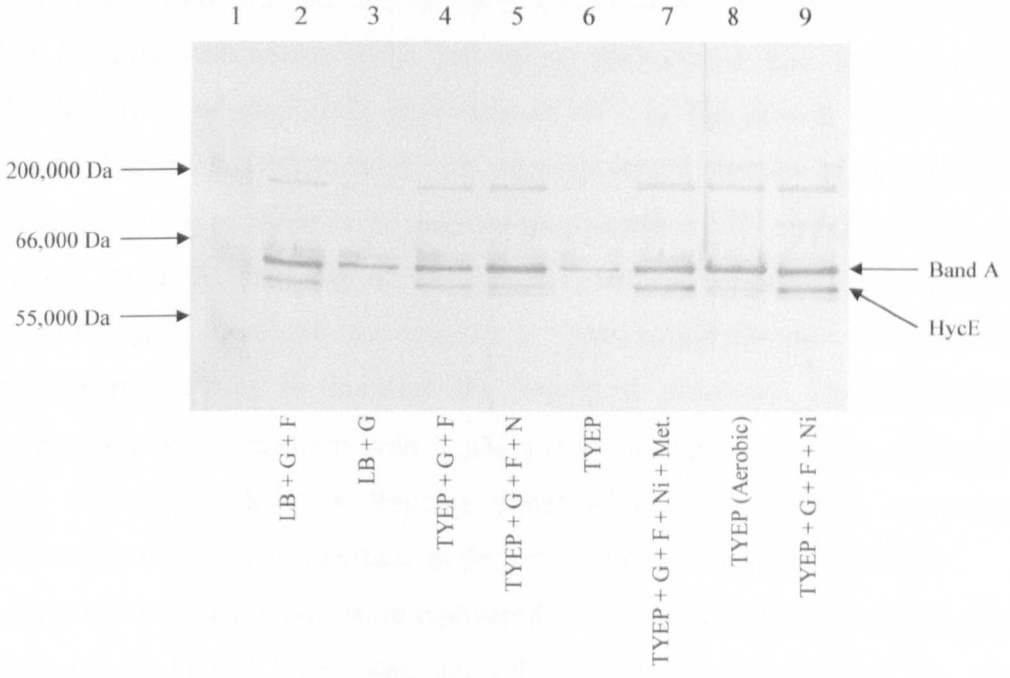


Fig. 3.6. Effect of different growth conditions on the amount of immunologically detectable HycE polypeptide. Preparations of extracts and immunoblotting procedure carried out as described in sections 2.8.1 and 2.8.3. MC4100 (wildtype) was grown anaerobically unless otherwise stated. Lane 1, Sea Blue Marker; Lane 2, LB + 0.4 % glucose + 30 mM sodium formate ; Lane 3, LB + 0.4 % glucose; Lane 4, TYEP (pH 6.6) + 0.4 % glucose + 30 mM sodium formate; Lane 5, TYEP (pH 6.6) + 0.4 % glucose + 30 mM sodium formate + 30 mM sodium nitrate; Lane 6, TYEP (pH 6.6); Lane 7, TYEP (pH 6.6) + 0.4 % glucose + 30 mM sodium formate + 5 μM nickel chloride + 1 μM sodium molybdate + 1 μM sodium selenite; Lane 8, TYEP (pH 6.6) (aerobic growth); Lane 9, TYEP (pH 6.6) + 0.4 % glucose + 30 mM sodium formate + 5 μM nickel chloride. Key to supplements: G, 0.4 % glucose; F, 30mM sodium formate; N, 30 mM sodium nitrate; Ni, 5 μM nickel chloride; Met., 1 μM sodium molybdate + 1 μM sodium selenite.

the HycE precursor (Rossmann *et al.*, 1994). It is possible that in the medium used the quantity of Ni^{2+} available for incorporation into HycE limits the amount of HycE that can be processed. Considering that only trace quantities of Ni^{2+} are present in the rich medium used, and that rich media is known to be a chelator of metal ions including Ni^{2+}, then this seems likely. By varying the concentration of added nickel chloride the effect of increasing quantities of Ni^{2+} in the growth medium was investigated (Fig. 3.7). Unfortunately any decrease in the quantity of the larger 65 kD unprocessed form of HycE in the medium supplemented with nickel chloride was masked by Band A, because of its almost identical migration through the polyacrylamide gel. Increasing the quantity of added nickel chloride to the growth medium did not appear to increase the levels of processed HycE detected. Supplementation of the medium with 1 μM nickel chloride is possibly enough to overcome the nickel chelating limiting effect of rich media, and so further supplementation would have no effect on the levels of processed HycE detected.

Two types of rich media were compared in the immunoblotting experiments, LB (section 2.5.2.1) and TYEP (section 2.5.2.2; Fig. 3.6, lanes 2 and 4). No distinguishable difference was observed in the quantity of HycE detected between the two media.

3.7 Immunoblotting with anti-HyfG, anti-HyfI, anti-HyfR, anti-HyfH and anti-MalE-HyfH sera

A strategy for the preparation of anti-Hyf sera was designed by Dr S. C. Andrews and executed by Dr P. Golby. Dr P. Golby carried out all immunoblotting experiments with anti-Hyf antisera.

The proposed large and small subunits of hydrogenase-4 (HyfG and HyfI respectively), the TYKY homologue (HyfH) and the proposed *hyf* regulator (HyfR) were overexpressed to approximately 20% of total cell protein as MalE fusions using the pMal-c2 vector. The fusion proteins were purified on a maltose column, MalE cleaved using factor Xa and the proteins purified by SDS-PAGE and electroelution. The purified HyfG, HyfI and HyfR polypeptides were used to raise antibodies in rabbits. The HyfH-MalE fusion protein could not be cleaved with factor Xa so the entire hybrid protein was used to raise antibodies.

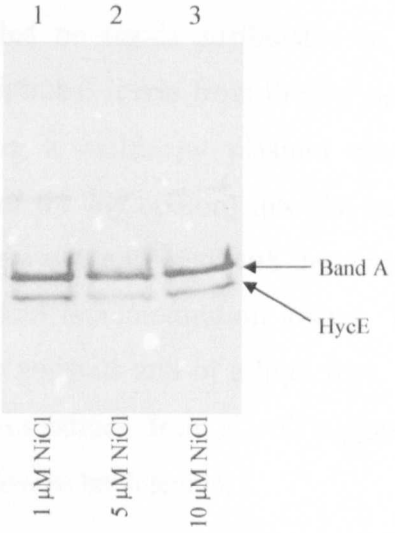


Fig. 3.7. Effect of nickel chloride concentration on the amount of immunologically detectable HycE polypeptide. Preparations of extracts and immunoblotting procedure carried out as described in sections 2.8.1 and 2.8.3. MC4100 (wildtype) was grown anaerobically in TYEP (pH 6.6; section 2.5.2.2) + 0.4 % glucose + 30 mM sodium formate. Lane 1, 1 μM nickel chloride; Lane 2, 5 μM nickel chloride; Lane 3, 10 μM nickel chloride.

HyfH was also overexpressed to approximately 20% of total cell protein as a His-tag fusion using the pET21a vector. The resulting fusion protein was insoluble, but was purified under denaturing conditions. The purified insoluble polypeptides were used to raise antibodies in rabbits.

Immunoblotting with the anti-Hyf antisera was carried on whole cell extracts from various *E. coli* strains (wildtype and mutant strains carrying deletions in the *hyc* and *hyf* operons) grown fermentatively in the presence of formate. These immunoblotting experiments revealed no bands attributable to the relevant Hyf polypeptides even when the transcriptional levels from the *hyf* operon were further increased 1000 fold by introducing a multicopy plasmid encoding *hyfR* (σ^{54}-dependent transcriptional activator of the *hyf* operon) into the cells (section 4.11). The specificity of the anti-Hyf antisera was analysed with the purified antigens both by immunoblotting and enzyme linked immunotitration assay. The antisera were found to be both immunoreactive to antigens and of a high titre. Failure to detect Hyf polypeptides with these antisera in extracts from *E. coli* suggests that expression levels of HyfG polypeptides are too low to be detected.

3.8 ^{63}Ni incorporation experiments

HyfG is homologous to the large subunits of the [Ni-Fe] (and [Ni-Fe-Se]) hydrogenases and residues acting as ligands for the [Ni-Fe] centre at the active site are conserved in HyfG (Cys 243, 246, 517 and 520; Andrews *et al.*, 1997). ^{63}Ni incorporation experiments were performed in an attempt to identify HyfG polypeptides associated with ^{63}Ni. Wildtype and hydrogenase-1, -2, -3 and -4 mutant strains were grown anaerobically in the presence of approximately 5 μM $^{63}NiCl_2$ (approximately 50 $\mu Ci.ml^{-1}$). Cell free protein extracts were prepared, separated electrophoretically and the dried gels autoradiographed (section 2.8.4).

3.8.1 Optimisation of ^{63}Ni incorporation protocol

The experimental protocol (section 2.8.4) required considerable optimisation before ^{63}Ni associated polypeptides were visualised.

3.8.1.1 Autoradiography

The most effective autoradiography method for detecting radioactivity emitted from ^{63}Ni was determined by exposure of the imaging system to ^{63}Ni standards prepared from the $^{63}NiCl_2$ stock solution (Fig. 3.8).

Amersham Pharmacia Biotech recommended detecting radioactivity emitted from ^{63}Ni by autoradiography using Hyperfilm MP film (Amersham Pharmacia Biotech) preflashed with Sensitise™ Preflash Unit (Amersham Pharmacia Biotech) and an exposure temperature of –70 °C (Fig. 3.8A). Preflashing film with the Sensitise™ Preflash Unit did not improve sensitivity of the film to radioactivity emitted by ^{63}Ni (Fig. 3.8A and B). However, Kodak BioMax MS film was found to be more sensitive to radioactivity emitted from ^{63}Ni (Fig. 3.8C). This increased sensitivity was further increased with a Kodak BioMax Transcreen HE Intensifying screen (Fig. 3.8D). Kodak BioMax MS film with a Kodak BioMax Transcreen HE Intensifying screen and an exposure temperature of –70°C were used in all experiments to detect electrophoretically separated ^{63}Ni associated proteins in dried polyacrylamide gels (section 2.8.4). It was noted that the Kodak BioMax Transcreen HE Intensifying screen reduced the resolution of radioactivity detected, however this was overshadowed by the improved sensitivity of autoradiography.

A storage phosphor imaging system with a storage phosphor screen for P^{32} was compared with film autoradiography but was not as sensitive at detecting radioactivity emitted from ^{63}Ni (Fig. 3.9). Although up to 100 times more sensitive than film, this system is dependent on isotope and sample type and a tritium storage phosphor screen better suited to detecting ^{63}Ni was not available for experimentation.

3.8.1.2 Preparation of cell free extracts

Having optimised autoradiography for the detection of ^{63}Ni, experiments were carried out to ensure that the optimised autoradiography method was sensitive enough to detect ^{63}Ni in cell cultures, harvested cells and most importantly Triton X-100 solubilised cell free protein extracts (Fig. 3.10). Radioactivity emitted from ^{63}Ni was detected in Triton X-100 solubilised cell free protein extracts (Fig. 3.10D).

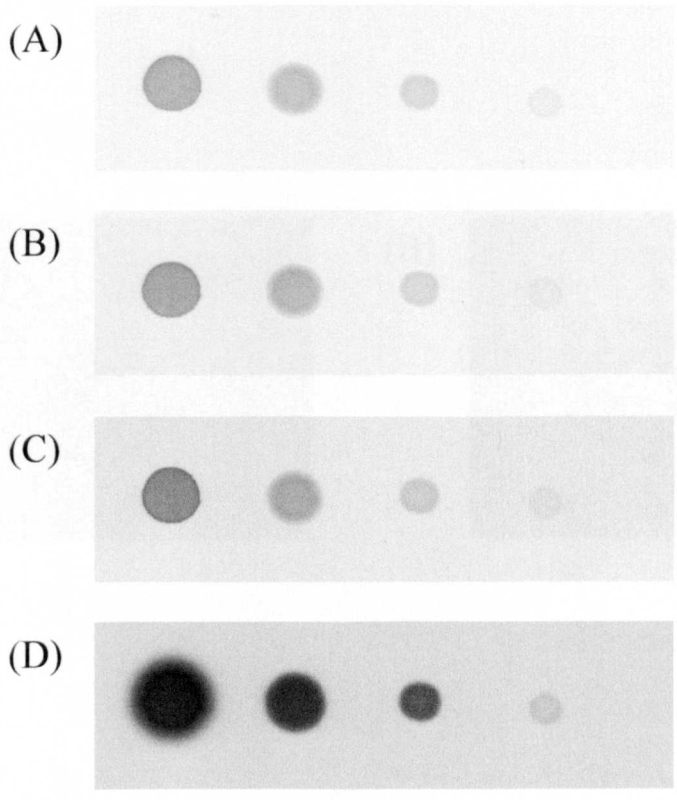


Fig. 3.8. Sensitivity of autoradiography films to radioactivity emitted by ^{63}Ni, and the effect of intensifying screens and preflashing to film sensitivity. The $^{63}NiCl_2$ stock solution (specific activity 27.92 $mCi.ml^{-1}$) was diluted 10, 100, 1000 and 10000 fold, and 5 µl from each dilution was added to a nitrocellulose filter and dried. These ^{63}Ni standards were exposed to autoradiography film for 2 hours at –70 °C. The same standard Kodak developer and developing time was used to develop all autorads.
(A) Hyperfilm MP (Amersham Pharmacia Biotech).
(B) Hyperfilm MP (Amersham Pharmacia Biotech) preflashed with Sensitize™ Pre-flash unit (Amersham Pharmacia Biotech).
(C) Kodak BioMax MS film.
(D) Kodak BioMax MS film with Kodak BioMax TranScreen HE Intensifying Screen.

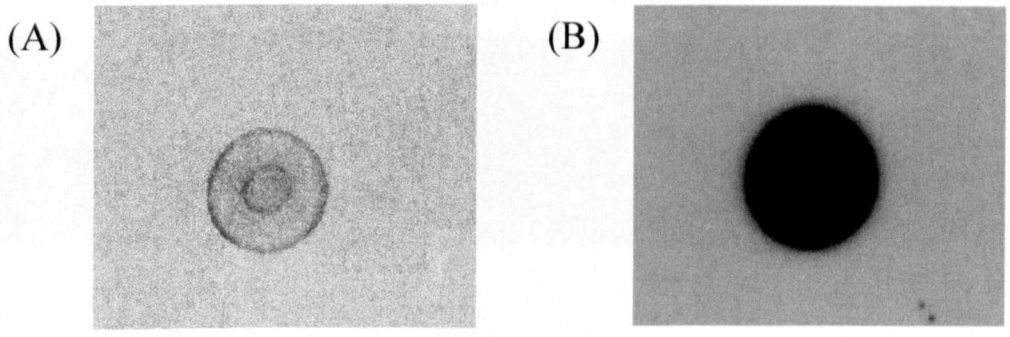


Fig. 3.9. Sensitivity of a storage phosphor imaging system (with a storage phosphor screen specific for P^{32}) to radioactivity emitted by ^{63}Ni. The $^{63}NiCl_2$ stock solution (specific activity 27.92 $mCi.ml^{-1}$) was diluted approximately 6000 fold and 10 µl added to a nitrocellulose filter and dried. This ^{63}Ni standard was used to test the sensitivity of the following imaging systems to radioactivity emitted from ^{63}Ni:
(A) Phosphor imaging system. Exposure was at room temperature for 1 week.
(B) Kodak BioMax MS film with Kodak BioMax Transcreen HE Intensifying Screen. Exposure was at –70°C for 1 week.

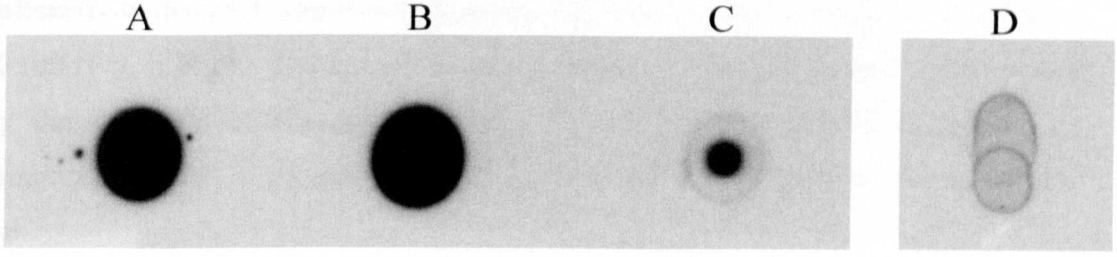


Fig. 3.10. Detection of ^{63}Ni in Triton X-100 solubilised cell free protein extracts. Autoradiograph of ^{63}Ni. MC4100 (wildtype) was grown to late exponential phase on TYEP (pH 6.6; section 2.5.2.2) supplemented with 0.4 % w/v glucose, 50 mM sodium formate and 5 μM $^{63}NiCl_2$ (approximately 50 $\mu Ci.ml^{-1}$). Cells were harvested by centrifugation and Triton X-100 solubilised cell free protein extracts were prepared as described in section 2.8.4. Autoradiography as described in section 2.8.4, except exposure was at –70°C for 1 week.

A Cell culture at late exponential phase (10 μl added to nitrocellulose filter).

B Cell culture supernatant (10 μl).

C Harvested cells resuspended in 20 μl 1 % w/v SDS (10 μl).

D Triton X-100 solubilised cell free protein extracts (4 x 10 μl).

3.8.1.3 Separation of ^{63}Ni associated polypeptides in cell free extracts by polyacrylamide gel electrophoresis (PAGE)

Initial experiments to separate ^{63}Ni associated polypeptides in Triton X-100 solubilised cell free extracts were carried out using a native Tris-HCl PAGE system with 7.5% polyacrylamide gels (Fig. 3.11A). The resolution of bands of radioactivity detected with this system was improved by the addition of 0.1% Triton X-100 (Fig. 3.11B). The sensitivity of the optimised autoradiography is illustrated by the detection of the gel front (Fig. 3.11B). The gel front is stained by bromophenol blue in the gel loading buffer and emits less light than the rest of the gel.

Amersham Pharmacia Biotech recommended soaking gels in Amplify Fluorographic Reagent (Amersham Pharmacia Biotech) for 15 minutes prior to drying and autoradiography to increase the sensitivity of detection of ^{63}Ni by autoradiography. However this was found to have no effect on the clarity and intensity of bands of radioactivity detected and Amplify was not used in further experiments (Fig. 3.12).

To improve electrophoretic separation of ^{63}Ni associated polypeptides the percentage of acrylamide in the separating gels was reduced from 7.5% to 5%. However no bands of radioactivity were visualised from Triton X-100 solubilised cell free extracts separated on 5% acrylamide gels (Fig. 3.13A). Two further PAGE systems were compared for their ability to separate and resolve ^{63}Ni associated polypeptides (Fig. 3.13B & C). Of the three PAGE systems a modified Laemmli Tris-HCl PAGE system with a 5% polyacrylamide gels containing 0.1% Triton X-100 was found to be most effective at resolving ^{63}Ni associated polypeptides (Laemmli, 1970; section 2.8.4; Fig. 3.13B). This PAGE system was used in all further ^{63}Ni incorporation experiments (Fig. 3.14 & Fig. 3.15). ^{63}Ni associated polypeptides were also detected using a Tricine PAGE system with 5% polyacrylamide gels containing 0.1% Triton X-100 (Fig. 3.13C; M. Quail, unpublished).

Gel loading buffers with and without β-mercaptoethanol were compared (Fig. 3.14). The presence of β-mercaptoethanol in the gel loading buffer was found to reduce the intensity and resolution of ^{63}Ni associated polypeptides separated by PAGE and was not used in ^{63}Ni incorporation experiments.

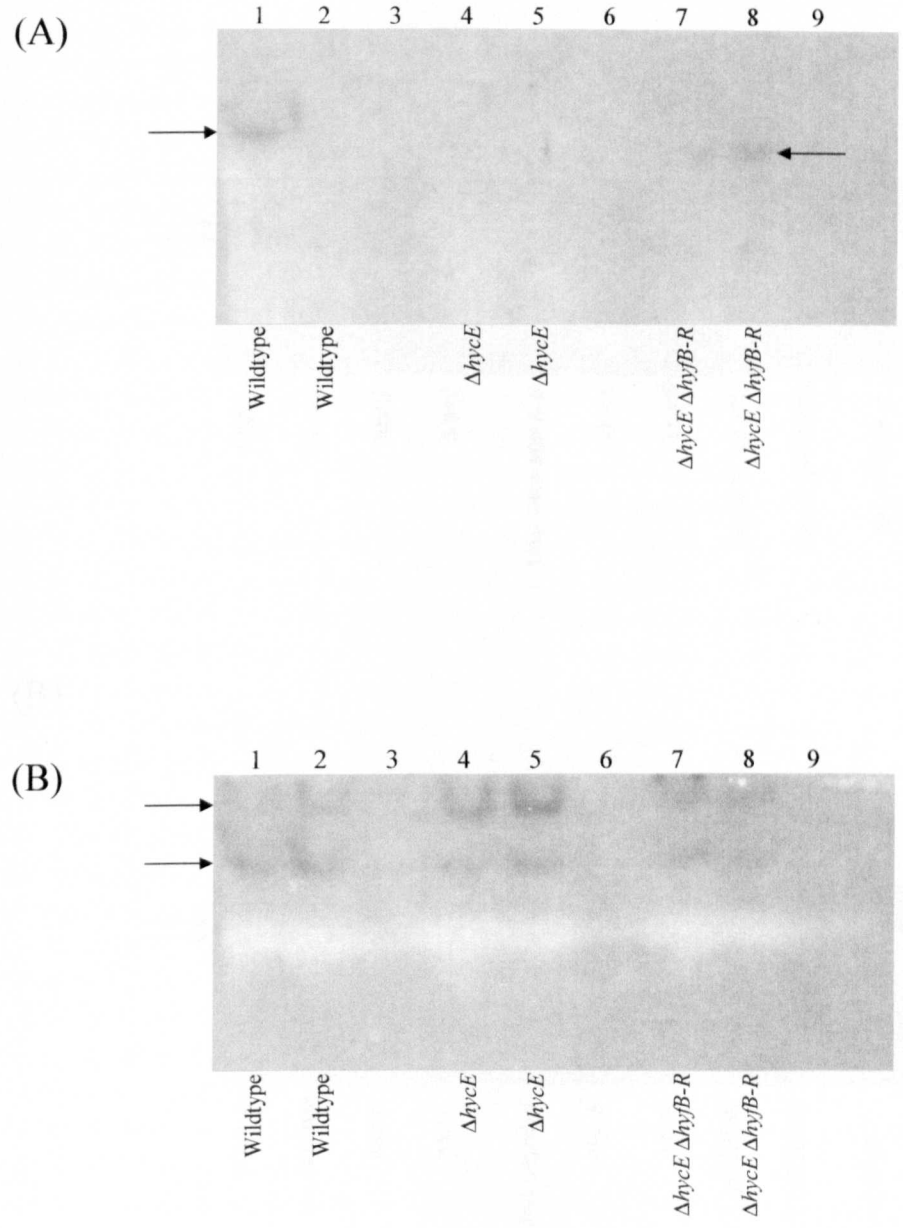


Fig. 3.11. Autoradiograph of ^{63}Ni incorporated proteins separated by native PAGE (A) and 0.1 % Triton X-100 PAGE (B) systems. Strains grown, Triton X-100 solublised cell free protein extracts prepared and autoradiography as described in section 2.8.4. Lanes 1 & 2, MC4100 (wildtype); lanes 4 & 5, HD705 (Δ*hycE*) ; lanes 7 & 8, JRG3933 (Δ*hycE* Δ*hyfB-R*); lanes 3, 6 & 9, empty. Arrows highlight the position of detected bands. (A) Proteins were separated by PAGE in 3.75 % (stacking gel) and 7.5 % (separating gel) polyacrylamide gels containing 100 mM Tris-phosphate (pH 5.5; stacking gel) and 70 mM Tris-HCl (pH 7.5; separating gel). Gels were electrophoresed in 80 mM Tris and 20 mM sodium barbitone, pH 7.5 at 4 ^{o}C for 1 h at 100 V. Native PAGE loading buffer containing 0.1 % w/v Triton X-100 was used (section 2.8.4). (B) PAGE system as described for (A), but 0.1 % Triton X-100 added to gels and buffers.

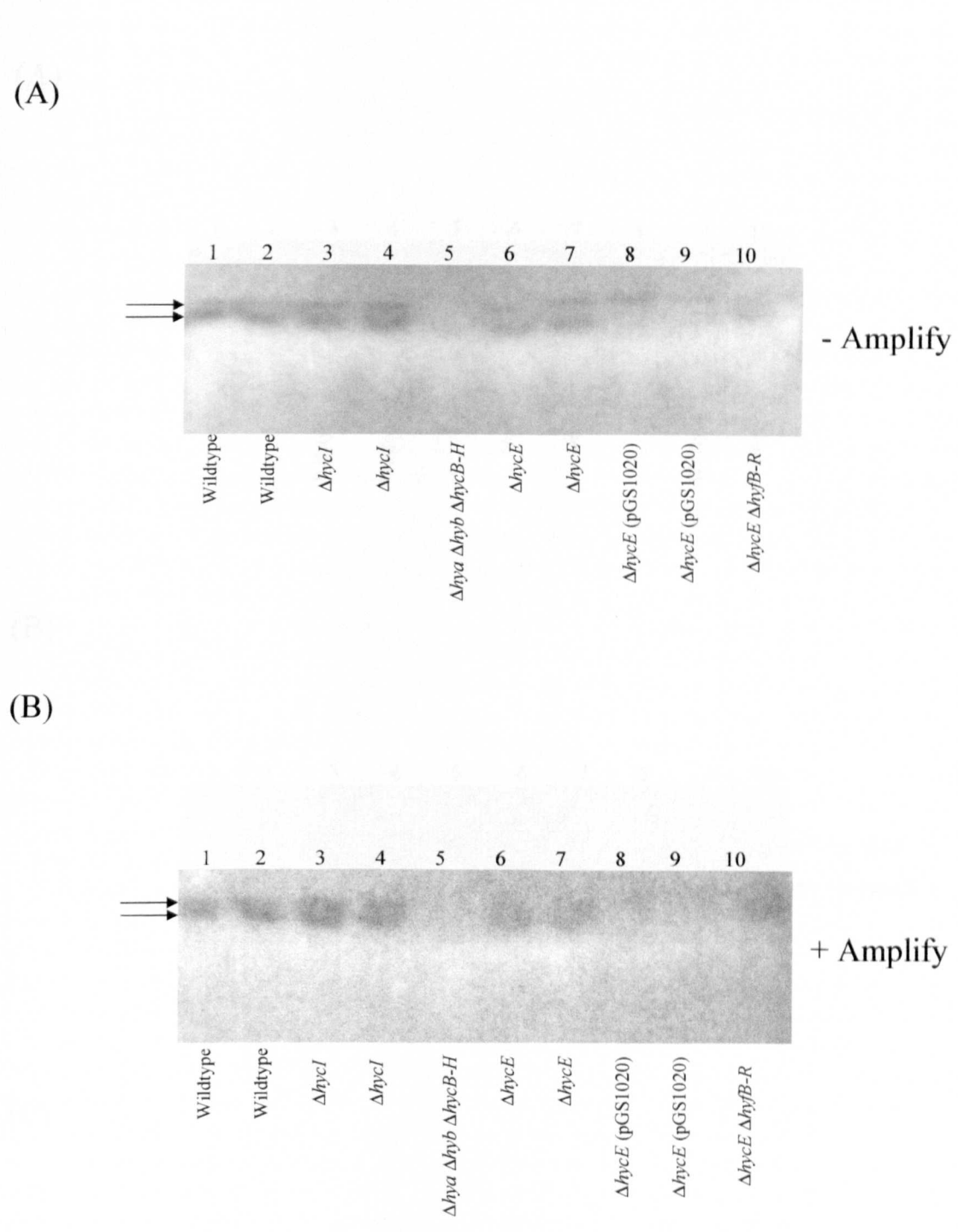
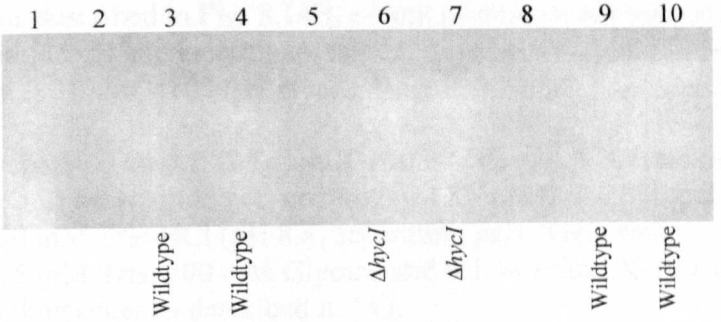
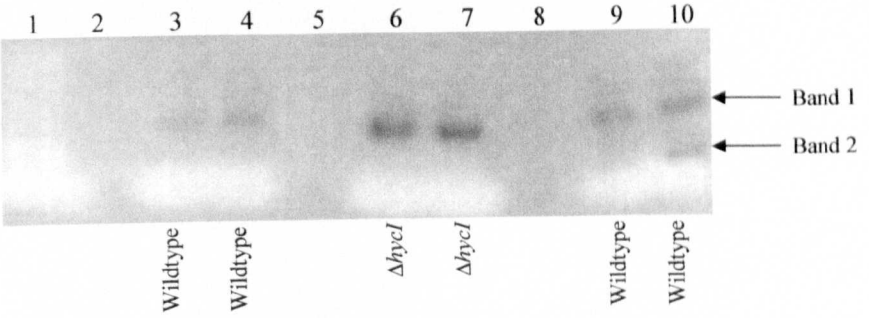


Fig. 3.12. Autoradiograph showing the detection of ^{63}Ni incorporated proteins in acrylamide gels with (B) and without (A) Amplify Fluorographic Reagent (Amersham Pharmacia Biotech). Strains grown, Triton X-100 solublised cell free protein extracts prepared and autoradiography as described in section 2.8.4. PAGE system as described in fig. 3.14B, except electrophoresis for 2 h 40 min at 50 V. Lane 1 & 2, MC4100 (wildtype); lanes 3 & 4, HD709 (Δ*hycI*); lane 5, HDJ123 (Δ*hya* Δ*hyb* Δ*hycB-H*); lanes 6 & 7, HD705 (Δ*hycE*); lanes 8 & 9, HD705 (Δ*hycE*) transformed with pGS1020 (multicopy plasmid encoding *hyfB-R*); JRG3933 (Δ*hycE* Δ*hyfB-R*). Arrows highlight the position of detected bands.

(A)


1 2 3 4 5 6 7 8 9 10
Wildtype
Wildtype
Δhyc1
Δhyc1
Wildtype
Wildtype

(B)


1 2 3 4 5 6 7 8 9 10
Band 1
Band 2
Wildtype
Wildtype
Δhyc1
Δhyc1
Wildtype
Wildtype

(C)

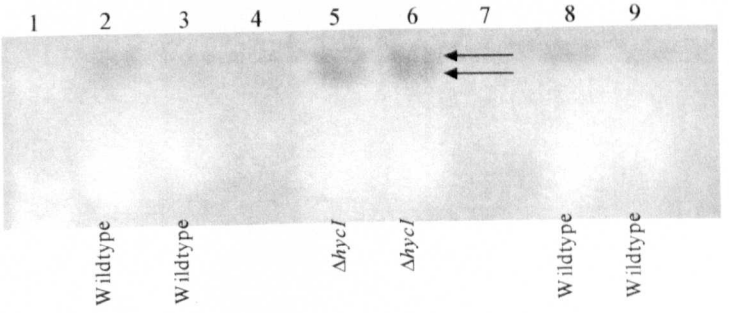

1 2 3 4 5 6 7 8 9
Wildtype
Wildtype
Δhyc1
Δhyc1
Wildtype
Wildtype

Fig. 3.13 (Page 116). Autoradiograph of ^{63}Ni incorporated proteins I. Strains grown, Triton X-100 solubilised cell free protein extracts prepared and autoradiography as described in section 2.8.4. Arrows highlight the position of detected bands.
(A) PAGE system as described in Fig. 3.14B, except proteins were separated by a 5 % separating acrylamide gel and electrophoresis for 1 h at 60 V. Lane 1, Sea blue marker; lanes 3, 4, 9 & 10, MC4100 (wildtype); lanes 6 & 7, HD709 (Δ*hycI*); lanes 5 & 8, empty.
(B) Proteins were separated by 0.1 % Triton X-100 PAGE in 4.5 % (stacking gel) and 5 % (separating gel) acrylamide gels containing 125 mM Tris-HCl (pH 6.8; stacking gel) and 150 mM Tris-HCl (pH 8.8; separating gel). Gels were electrophoresed in 25 mM Tris, 200 mM Glycine and 0.1 % Triton X-100 at 4 °C for 1 h at 60 V. Gel loading order as described in (A).
(C) Proteins were separated by 0.1 % triton X-100 PAGE in 4 % (stacking gel) and 5 % (separating gel) acrylamide gels containing 100 mM Tricine (pH 8; stacking gel) and 200 mM Tricine (pH 8; separating gel). Gels were electrophoresed in 200 mM Tricine (pH 8.0) and 0.1 % Triton X-100 at 4 °C for 1 h at 60 V. Lane 1, Sea blue marker; lanes 2, 3, 8 & 9, MC4100 (wildtype); lanes 5 & 6, HD709 (Δ*hycI*); lanes 4 & 7, empty.

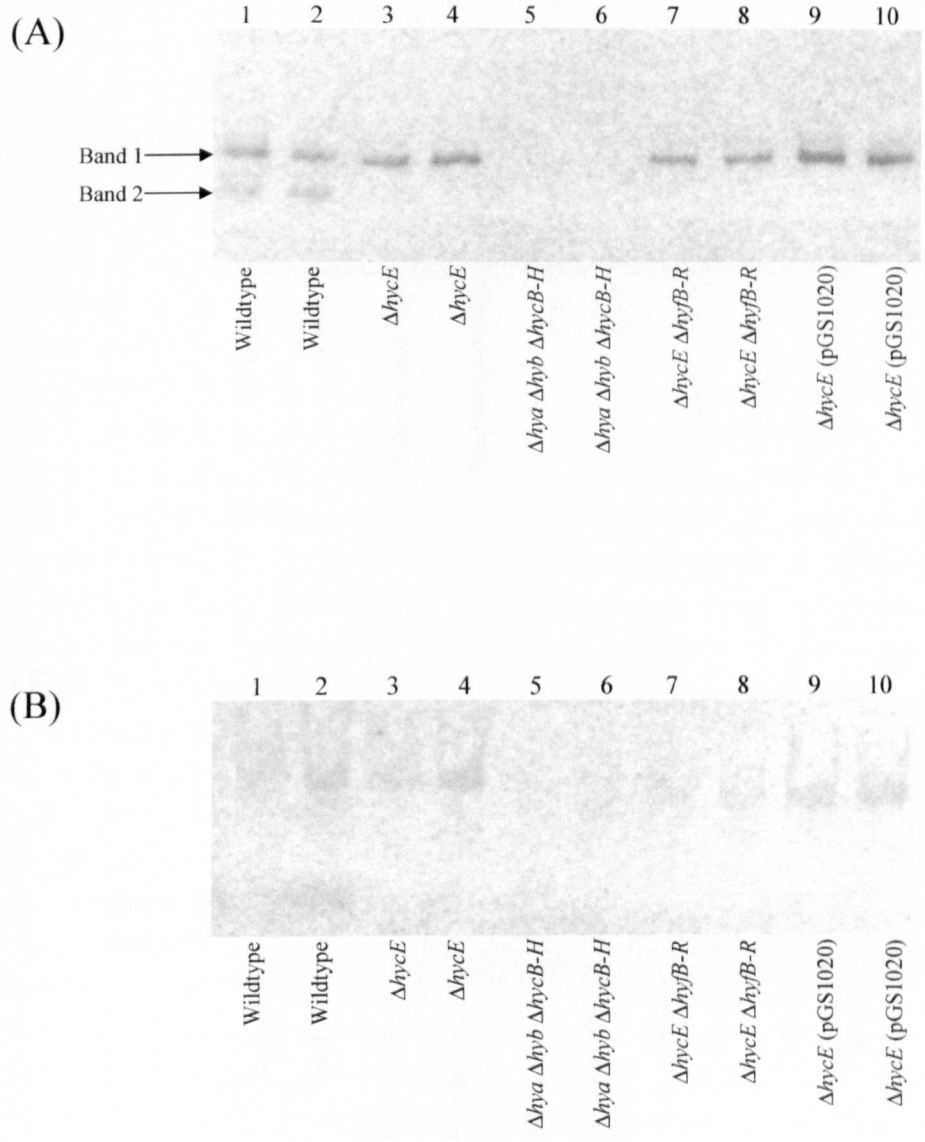
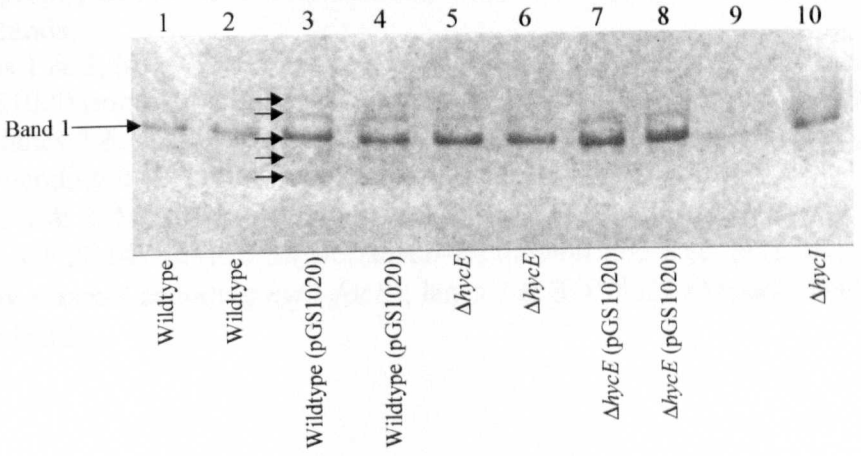


Fig. 3.14. Autoradiograph of ^{63}Ni incorporated proteins II. Strains grown, Triton X-100 solubilised cell free protein extracts prepared and autoradiography as described in section 2.8.4. Lane 1 & 2, MC4100 (wildtype); lanes 3 & 4, HD705 (Δ*hycE*); lanes 5 & 6, HDJ123 (Δ*hya* Δ*hyb* Δ*hycB-H*); lanes 7 & 8, JRG3933 (Δ*hycE* Δ*hyfB-R*); lanes 9 & 10, HD705 (Δ*hycE*) transformed with pGS1020 (multicopy plasmid encoding *hyfA-focB*). Arrows highlight the position of detected bands.
(A) PAGE system as described in section 2.8.4.
(B) As (A) but Native PAGE loading buffer containing 0.1 % w/v Triton X-100 also containing 10 % β-mercaptoethanol.

(A)


1
2
3
4
5
6
7
8
9
10
Band 1
Wildtype
Wildtype
Wildtype (pGS1020)
Wildtype (pGS1020)
ΔhycE
ΔhycE
ΔhycE (pGS1020)
ΔhycE (pGS1020)
ΔhycI

(B)

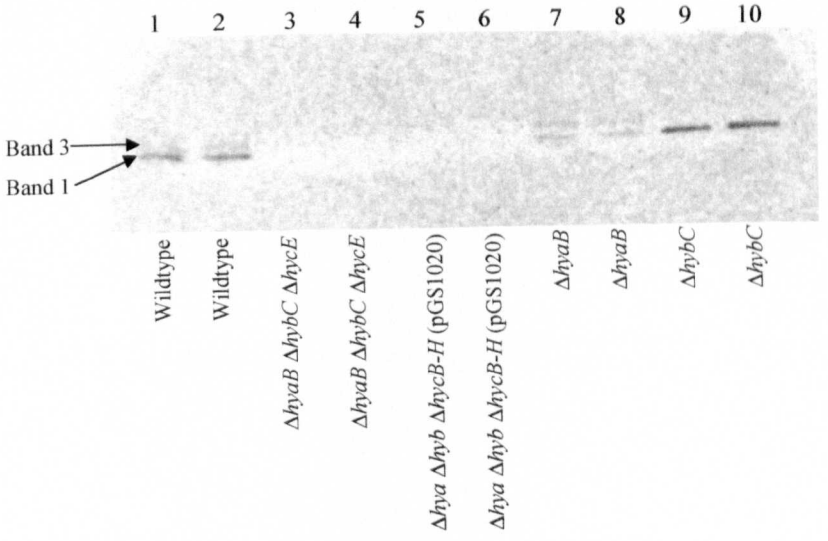

1
2
3
4
5
6
7
8
9
10
Band 3
Band 1
Wildtype
Wildtype
ΔhyaB ΔhybC ΔhycE
ΔhyaB ΔhybC ΔhycE
Δhya Δhyb ΔhycB-H (pGS1020)
Δhya Δhyb ΔhycB-H (pGS1020)
ΔhyaB
ΔhyaB
ΔhybC
ΔhybC

Fig. 3.15 (Page 119). Autoradiograph of ^{63}Ni incorporated proteins III. Strains grown, Triton X-100 solubilised cell free protein extracts prepared, PAGE and autoradiography as described in section 2.8.4. Arrows highlight the position of detected bands.
(A) Lanes 1 & 2, MC4100 (wildtype); lanes 3 & 4, MC4100 (wildtype) transformed with pGS1020 (multicopy plasmid encoding *hyfA-focB*); lanes 5 & 6, HD705 (Δ*hycE*); lanes 7 & 8, HD705 (Δ*hycE*) transformed with pGS1020 (multicopy plasmid encoding *hyfA-focB*); lane 9, empty; lane 10, HD709 (Δ*hycI*).
(B) Lanes 1 & 2, MC4100 (wildtype); lanes 3 & 4, FTD147 (Δ*hyaB* Δ*hybC* Δ*hycE*); lanes 5 & 6, FTD147 (Δ*hyaB* Δ*hybC* Δ*hycB-H*) transformed with pGS1020 (multicopy plasmid encoding *hyfA-focB*); lanes 7 & 8, FTD22 (Δ*hyaB*); lanes 9 & 10, FTD67 (Δ*hybC*).

Dried gels were subject to autoradiography for at least two weeks, as shorter exposures reduced the intensity of radioactivity detected. However exposure for longer than 2 weeks did not improve the intensity of bands on autoradiographs (observations not shown).

3.8.2 Detection of ^{63}Ni associated polypeptides

Despite the detection of ^{63}Ni associated polypeptides and significant optimisation of the PAGE system, separation of these polypeptides was minimal. The ^{63}Ni associated polypeptides did not enter separating gels very far and were observed just below the stacking/separating gel interface. However using the optimised protocol for PAGE and autoradiography the following observations could be made;

(i) No ^{63}Ni associated polypeptides, attributable to the large subunit of hydrogenase-4 (HyfG), were detected in the hydrogenase-1, -2 and -3 triple mutants HDJ123 (Δ*hya*, Δ*hyb*, Δ*hycB-H*; Fig.3.12A & B, lane 5; Fig. 3.14A & B, lanes 5 & 6) and FTD147 (Δ*hyaB*, Δ*hybC*, Δ*hycE*; Fig. 3.15B, Lanes 3, 4, 5 & 6). It seems that either HyfG is not a nickel enzyme (despite residues acting as ligands for the [Ni-Fe] centre being conserved in HyfG) or that expression levels of HyfG are too low to be detected. Alternatively, ^{63}Ni may have dissociated from HyfG during the preparation of Triton X-100 solubilised cell free protein extracts or PAGE.

(ii) Up to five ^{63}Ni associated polypeptides were detected in extracts from the wildtype strain (MC4100; Fig. 3.15A, lane 3). These polypeptides are likely to correspond to the large subunits of hydrogenases-1, -2 and -3 (HyaB, HybC & HycE respectively) and corresponding processing intermediates.

The ^{63}Ni incorporation experiments of Theodoratou and co-workers (2000) detected at least two nickel-containing precursors of HycE, which were suggested to represent maturation intermediates. However these pre-HycE forms were only detected in a Δ*hycI* mutant strain (HD709) lacking the specific maturation endopeptidase.

(iii) One of the bands of radioactivity detected (Band 1) was more intense than the other bands detected in extracts from the wildtype (Fig. 3.13B; Fig. 3.14A; Fig.

3.15A & B). This intense band was absent in the Δ*hyaB* mutant (FTD22) and is possibly the hydrogenase-1 large subunit, HyaB (Fig. 3.15B, Lanes 7 & 8).

This is consistent with observations that in the cell under fermentative growth conditions hydrogenase-1 is more abundant than hydrogenase-2 (Ballantine & Boxer, 1985; Richard *et al.*, 1999).

(iv) A faster migrating band (Band 2), absent in extracts from HD705 (Δ*hycE*), was detected in extracts from MC4100 on a number of autoradiographs (Fig. 3.13B; Fig. 3.14A; Fig. 3.15A).

The ^{63}Ni incorporation experiments of Theodoratou and co-workers (2000) resolved two bands of radioactivity in extracts from MC4100. They found that the faster migrating of the two bands was absent in extracts from Δ*hycE* and comigrated with material reacting with anti-HycE sera. They attributed the identity of this faster migrating band to HycE. The slower migrating band was present in extracts from Δ*hycE* and was attributed to the large subunits of either hydrogenase-1 and -2.

However this faster migrating band is absent from some cell free protein extracts prepared from the wildtype strain (MC4100; Fig. 3.13B, lanes 3 & 4). The reason for this is unknown but it is possible that the polypeptide is labile or that its expression is low making detection difficult.

(v) One band of radioactivity (Band 3) detected in extracts from the wildtype (MC4100) and Δ*hyaB* mutant (FTD22), was absent in the Δ*hybC* mutant (FTD67) and is possibly the hydrogenase-2 large subunit, HybC (Fig. 3.15B, lanes 9 & 10).

3.9 Analysis of hydrogenase-1, -2, -3 and -4 mutant strains for enzymic activity

The hydrogen dependent and formate dependent reduction of benzyl viologen are measures of total hydrogenase and Fdh-H activities, respectively. In an attempt to detect hydrogenase or formate dehydrogenase activities attributable to hydrogenase-4, enzyme assays were performed on hydrogenase-1, -2, -3 and -4 mutant strains grown to late exponential phase under fermentative conditions.

3.9.1 Total hydrogenase activity

Sauter and co-workers (1992) detected no hydrogenase activity at pH 7.0 in the hydrogenase-1, -2 and -3 triple mutant HDJ123 (*Δhya, Δhyb, ΔhycB-H*) grown fermentatively to late exponential phase on TYEP medium (pH 6.5; section 2.5.2.2) supplemented with glucose (0.4% w/v) and sodium formate (30 mM). They concluded that under these growth conditions, hydrogenases-1, -2 and –3 are the only ones formed which can catalyse the hydrogen dependent reduction of benzyl viologen.

This hydrogenase assay has been repeated here with the same strain grown under the same conditions (except medium supplemented with 50 mM sodium formate) (section 2.8.5.4). Again no hydrogenase activity was detected at pH 7.0 (data not shown). Previously it was suggested hydrogenase activity attributable to hydrogenase-4 was not detected in this triple mutant because the Hyf system may not be expressed under these growth conditions (Andrews *et al.*, 1997). However, initial expression studies with a *hyfA-lacZ* transcription fusion (DS5) found that optimal *hyf* expression was observed under these growth conditions i.e. fermentatively at acidic pH in the presence of formate and the absence of electron acceptors (section 4.10.1). Andrews and co-workers (1997) also suggested that because the *hycB-H* deletion of strain HDJ123 contains a chloramphenicol cassette that could exert a polar effect on *hycI* expression, the failure to detect hydrogenase-4 activity might be due to the lack of the protease HycI proposed to be required for HyfG processing. Another hydrogenase-1, -2 and -3 triple mutant (FTD147; *ΔhyaB ΔhybC ΔhycE*) has been constructed carrying chromosomal in frame deletions solely in the genes encoding the large subunits of hydrogenase-1, -2 and -3 (F. Sargent, unpublished). This strain was constructed to preserve identifiable regulatory elements, coding sequences, stop codons, and Shine-Dalgarno sequences of genes flanking the deletions. The strain (FTD147) was grown fermentatively to late exponential phase on TYEP medium (pH 6.6; section 2.5.2.2) supplemented with glucose (0.4% w/v) and sodium formate (50 mM) (i.e. optimal *hyf* expression conditions) and assayed for total hydrogenase activity at pH 7.0 (Fig. 3.16) (section 2.8.5.4). Hydrogenase activity was abolished by the deletion of the *hyaB*, *hybC* and *hycE* genes. The level of hydrogenase activity (approximately 1.75 μmol benzyl viologen min^{-1} (mg protein)$^{-1}$ for MC4100) obtained was approximately 2.4 fold lower than that (4.19 μmol benzyl viologen

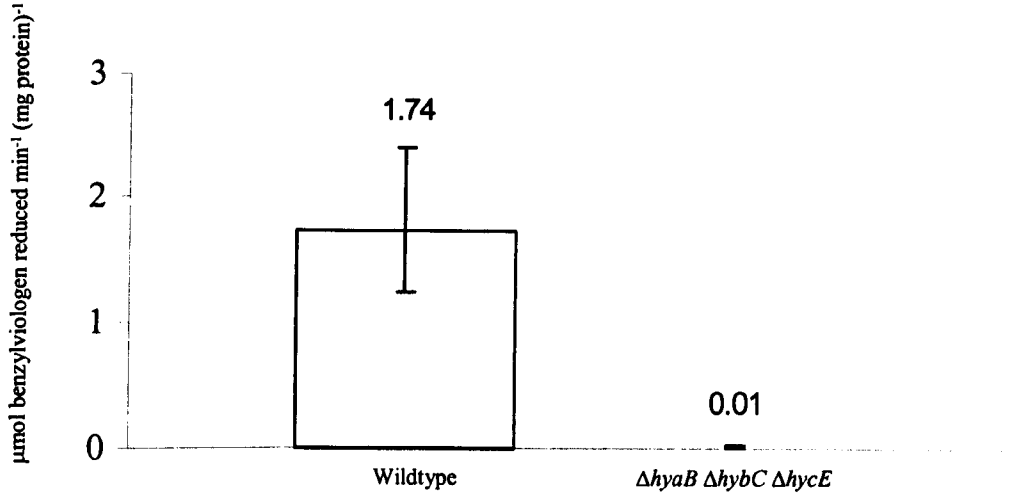


Fig. 3.16. Total hydrogenase activity (H_2 dependent reduction of benzyl viologen) detected in wildtype (MC4100) and hydrogenase-1, -2 and –3 triple mutant (FTD147; *ΔhyaB ΔhybC ΔhycE*) strains. Cells were grown fermentatively to late exponential phase in TYEP medium (pH 6.6; section 2.5.2.2) supplemented with 0.4 % glucose, 50 mM sodium formate and 5 µM nickel chloride. The H_2 dependent reduction of benzyl viologen was used to measure total hydrogenase activity at pH 7.0 in whole cells made permeable by toluene. Values are the average of four to six determinations from independent cultures and error bars give the maximal and minimal values.

min^{-1} (mg protein)$^{-1}$) previously reported by Sauter *et al.* (1992). These differences are likely to be largely due to differences in sample preparation. Sauter and co-workers (1992) measured hydrogenase activity in crude extracts (Triton X-100-dispersed membranes), whereas assays reported here measured hydrogenase activity in toluene permebilised whole cells (Ballantine *et al.*, 1986). Other factors such as assay temperature may also have contributed to differences observed. The levels of hydrogenase activity measured vary up to 100 fold between different reported studies (Sawers *et al.*, 1985; Stoker *et al.*, 1989; Sauter *et al.*, 1992; Sawers, 1994).

Further hydrogenase assays were performed with wildtype (MC4100) and Δ*hyfB-R* (JRG3621) strains, to look for differences in activity between the strains that could be attributed to the *hyf* operon. As expected, hydrogenase activity was detected in both these strains, however difficulties were encountered obtaining consistent data and differences observed were not reproducible (data not shown). Previous studies have shown that hydrogenase-3, the most active hydrogenase enzyme under these growth conditions, is extremely labile (Sauter *et al.*, 1992; Sawers *et al.*, 1985). This liability of hydrogenase-3 during the harvesting of cells and the preparation of samples (no anaerobic preparations were taken) may have contributed to the inconsistency of results obtained. It is also possible that hydrogenase-4 enzyme activity is extremely labile, a factor which may explain why hydrogenase-4 has not been previously discovered and characterised.

3.9.2 Fdh-H activity

Fdh-H assays were performed with wildtype (MC4100), Δ*hycE* (HD705) and Δ*hyfBR* (JRG3621) strains to detect any differences in activity attributed to the *hyf* operon. For Fdh-H assays, improvements were made to the anaerobic cuvettes used, to reduce inconsistencies that may occur from the liability of the hydrogenase-3 enzyme. Rubber bungs previously used to seal anaerobic cuvettes were replaced with plastic stoppers, which maintained an anaerobic environment inside the cuvette more effectively. Holes (0.7 mm diameter) were created in the plastic stoppers to allow the addition of sodium dithionite and toluene permeabilised whole cell samples to the cuvette with a Hamilton syringe. The wildtype (MC4100), Δ*hycE* (HD705) and Δ*hyfB-R* (JRG3621) strains were grown to late exponential phase on TYEP (pH 6.6; section 2.5.2.2) supplemented with glucose (0.4% w/v) and sodium formate (50

mM) (i.e. optimal *hyf* expression conditions) and assayed for Fdh-H activity at pH 6.6 (Fig. 3.17) (section 2.8.5.3). Deletion of *hycE* reduced the Fdh-H activity to approximately 30% of wildtype enzyme activity. Sauter and co-workers (1992) also reported a reduction in Fdh-H activity by deletion of *hycE*, but activity was reduced to approximately 10% of that of the wildtype. However, another more recent study has shown this reduction in Fdh-H activity in a *hycE* mutant to be approximately 40% of that of the wildtype (Bagramyan *et al.*, unpublished). Fdh-H activity was not affected by the deletion of *hyfB-R*. The level of Fdh-H activity (approximately 0.206 μmol benzyl viologen min^{-1} (mg protein)$^{-1}$ for MC4100) obtained was approximately 14 fold lower than that (2.83 μmol benzyl viologen min^{-1} (mg protein)$^{-1}$) previously reported by Sauter *et al.* (1992). These differences are likely to be due to sample preparation as described in section 3.9.1.

3.9.3 Total hydrogenase and Fdh-H activities at a slightly alkaline pH

Bagramyan and co-workers (2000) detected hydrogen production in *E. coli* grown fermentatively at pH 7.5, which was not observed in mutants carrying deletions in the genes of the *hyf* operon. This *hyf* dependent hydrogen production was absent in medium supplemented with formate and/or the F_0F_1-ATPase inhibitor DCCD, or upon osmotic stress. Proton-potassium exchange at pH 7.5 was also lost in these *hyf* mutants. It was proposed that hydrogenase-4 participates in the production of H_2 and proton-potassium exchange in *E. coli* grown fermentatively at pH 7.5, and that hydrogenase-3 is responsible for the production of hydrogen during growth at pH 6.5 or in medium containing formate.

The formate dependent reduction of benzyl viologen is a measure of formate dehydrogenase-H activity in *E. coli*. Further work by Bagramyan and co-workers (unpublished) detected weak Fdh-H activity in *E. coli* grown fermentatively at pH 7.5, which was increased by the addition of ATP to the assay reaction mixture. This ATP dependent Fdh-H activity was inhibited by DCCD, reduced in hydrogenase-4 mutants and lost in Fdh-H, F_0F_1-ATPase and HycB mutants. Bagramyan and co-workers (unpublished) proposed that Fdh-H and hydrogenase-4 combine to form a second Fhl system (Fhl-2) in *E. coli* that is driven by the proton gradient established by F_OF_1-ATPase and that HycB would serve to transfer electrons from Fdh-H to

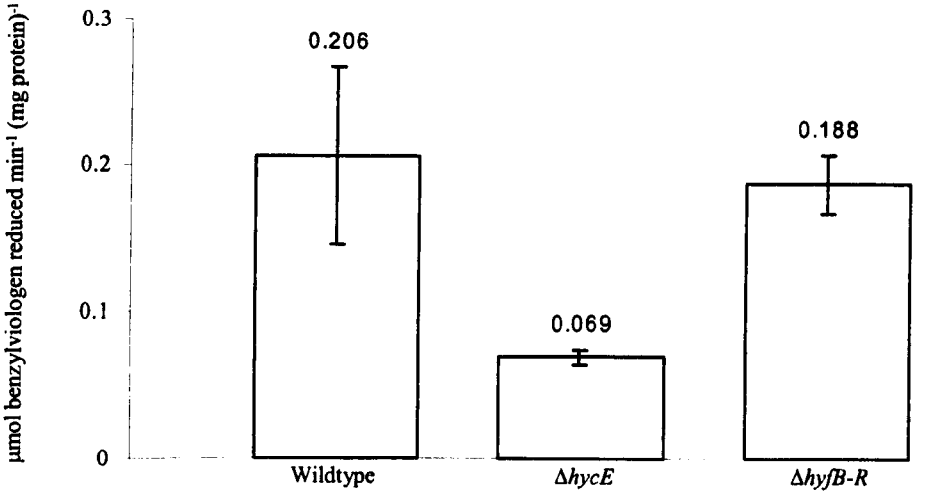


Fig. 3.17. Fdh-H activity (formate dependent reduction of benzyl viologen) detected in wildtype (MC4100), Δ*hycE* (HD705) and Δ*hyfB-R* (JRG3621) strains. Cells were grown fermentatively to late exponential phase in TYEP medium (pH 6.6; section 2.5.2.2) supplemented with 0.4 % glucose and 50 mM sodium formate. The formate dependent reduction of benzyl viologen was used to measure Fdh-H activity at pH 6.6 in whole cells made permeable by toluene. Values are the average of two determinations from independent cultures and error bars give the maximal and minimal values.

hydrogenase-4 and thus couple formate dehydrogenation with H_2 production in Fhl-2. The physiological purpose of this Fhl-2 pathway is uncertain, but it may be required for the generation of CO_2 during fermentation at high pH (above pH 7) for use in the generation of oxalate by phophoenolpyruvate carboxylate. This in turn could be used for biosynthesis or for the consumption of reducing equivalents and for H_2-dependent fumarate respiration (Table 1.1). It is important to note that in these studies growth at pH 7.5 refers to growth in medium buffered to pH 7.5 and the pH of this medium will drop to as low as pH 6.5 during fermentative growth. Bagramyan and co-workers (unpublished) reported that the *hyf* mutations had no major effect on growth under fermentative growth conditions.

Hydrogenase assays were performed at pH 7.5 to try and detect hydrogenase activity attributable to hydrogenase-4 in the hydrogenase-1, -2 and -3 triple mutant FTD147 (*ΔhyaB, ΔhybC, ΔhycE*) transformed with a multicopy plasmid encoding the *hyfR* gene (pGS1087). Introduction of pGS1087 into the *hyfA-lacZ* transcription fusion strain DS5 enhanced *hyf* expression >1000 fold (section 4.11). The wildtype (MC4100) and hydrogenase-1, -2 and -3 mutant (FTD147) strains (both transformed with pGS1087) were grown to late exponential phase on LB (section 2.5.2.1) buffered to pH 7.5 with 0.1 M Tris phosphate and supplemented with glucose (0.4% w/v). The harvested cells were assayed for total hydrogenase activity at pH 7.5. Hydrogenase activity was detected in the wildtype (MC4100) but abolished by the deletion of the *hyaB, hybC and hycE* genes (FTD147) (data not shown). These assays were performed in the absence of added ATP, which Bagramyan and co-workers (unpublished) suggested would be required to drive Fhl-2 at slightly alkaline pH. Bagramyan and co-workers (unpublished) did not assay for total hydrogenase activity, however they reported trace *hyf* dependent Fdh-H activity even in the absence of added ATP.

The Fdh-H assays of Bagramyan and co-workers (unpublished) were repeated in wildtype (MC4100) and *hyfB-R* deletion (JRG3621) strains. The strains were grown fermentatively to late exponential phase on TYEP (pH 7.6; section 2.5.2.2) supplemented with glucose (0.4% w/v) and assayed for Fdh-H activity at pH 7.6 in the presence of 12.5 mM ATP (Fig. 3.18) (section 2.8.5.3). Deletion of *hyfB-R* had no effect on formate dehydrogenase-H activity detected. Bagramyan and co-workers

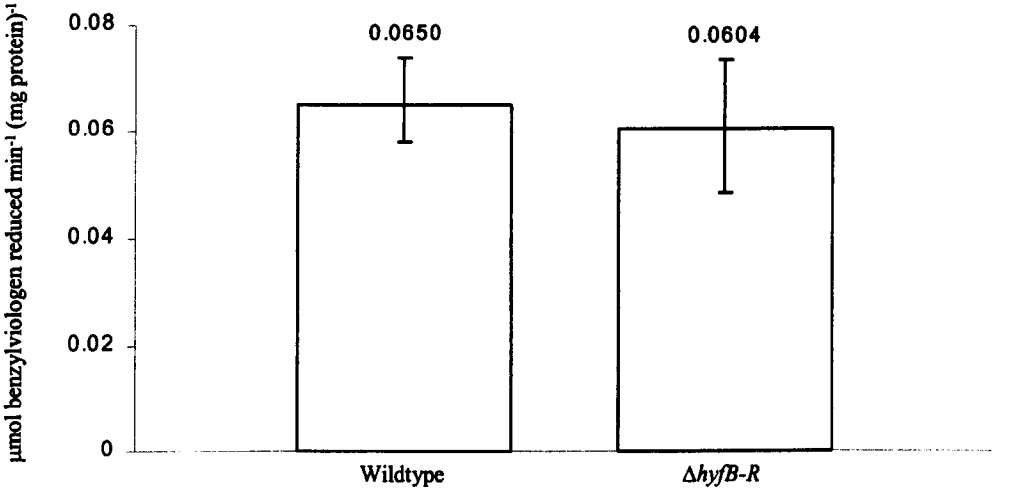


Fig. 3.18. Fdh-H activity (formate dependent reduction of benzyl viologen) detected at pH 7.6 in wildtype (MC4100) and Δ*hycE* (HD705) strains. Cells were grown fermentatively to late exponential phase in TYEP medium (pH 7.6; section 2.5.2.2) supplemented with 0.4 % glucose. The formate dependent reduction of benzyl viologen was used to measure Fdh-H activity at pH 7.6 in whole cells made permeable by toluene (section 2.8.5.3). Values are the average of four determinations from independent cultures and error bars give the maximal and minimal values.

(unpublished) reported that deletion of *hyfB-R* reduced Fdh-H activity to approximately 20% of that of the wildtype.

3.10 Gas production experiments

The Fhl-1 complex, encoded by *fdhF* and the *hyc* operon, catalyses the non-energy conserving breakdown of formate to H_2 and CO_2 gas (Zinoni *et al.*, 1986; Bohm *et al.*, 1990). Andrews and co-workers (1997) proposed that the *hyf* operon (together with *fdhF*) encoded a second Fhl complex (Fhl-2) in *E. coli* that catalysed the energy conserving breakdown of formate to H_2 and CO_2 gas. Gas production assays were performed to try and detect gas evolution attributable to this Fhl-2 complex (section 2.8.5.5). Strains carrying deletions in the *hyc* and *hyf* operons were grown fermentatively to late exponential phase and any gas produced was collected in inverted Durham tubes (Table 2.4; Schlensog *et al.*, 1989) (section 2.8.5.5). Gas production was not measured quantitatively and only the presence or absence of gas production was recorded. Gas production was observed in wildtype (MC4100) and *hyfB-R* deletion strains (JRG3621). This gas production was attributed to Fhl-1 as deletion of the *hycE* gene abolished gas production. No gas production attributable to the *hyf* operon and therefore Fhl-2 was detected, even in the presence of a multicopy plasmid encoding *hyfR* (pGS1087) which was found to increase expression from the *hyf* operon > 1000 fold (section 4.11). Andrews and coworkers (1997) proposed that H_2 evolution from Fhl-2 was only favourable at low PH_2 (i.e. when H_2 is being removed from the environment as it is being produced). These conditions do not exist in the gas production experiments performed where any H_2 produced was allowed to build up. In a more recent study by Bagramyan and coworkers (2000, unpublished), H_2 production attributable to the *hyf* operon was detected at slightly alkaline pH using a hydrogen detecting electrode. This *hyf* dependent H_2 production was almost abolished at pH 6.5. The pH of the two growth media used in the gas production assays described here (Table 2.4) would have decreased to below pH 6.5 during the course of fermentative growth. Gas production assays were used throughout this study to check the phenotype of wildtype and mutant strains.

Strain	Relevant Genotype	Plasmid	Gas production	
			TYEP (pH 6.5) + 0.4% Glucose + 50 mM Formate	Standard minimal medium for bioreactor work (pH 7.0) + 20 mM Glucose
MC4100	Wildtype	-	+	+
HD705	MC4100, Δ*hycE*	-	-	-
HD705	MC4100, Δ*hycE*	pGS1087	-	ND
JRG3621	MC4100, Δ*hyfB-R*	-	+	+
JRG3933	HD705, Δ*hyfB-R*	-	-	-

Table 3.4. Gas production in *hyc* and *hyf* deletion strains. Cells were grown to late exponential phase and any gas produced was trapped in inverted Durham tubes. Observations were taken from experiments carried out in duplicate. Key to observations: +, gas production observed; -, no gas production observed; ND, not done.

3.10 H_2 production assays

A Clarke-type oxygen electrode (Hansatech Ltd) adopted for the measurement of H_2 was used to try and detect H_2 production attributable to the *hyf* operon (section 2.8.5.6; Ballantine and Boxer, 1986). Bagramyan and co-workers (2000, unpublished) detected H_2 production attributable to the *hyf* operon in cells grown and assayed at pH 7.5. They found that deletion of *hycE* had no effect on hydrogen evolution but deletion of *hyfB-R* virtually abolished (up to 17 fold reduction) hydrogen production at pH 7.5. This *hyf* attributed activity was partly dependent on the presence of 3 mM ATP and the 16Fe ferredoxin HycB. Bagramyan and co-workers (unpublished) proposed that ATP was required to drive the *hyf* encoded Fhl complex (Fhl-2) via a proton gradient established by F_0F_1-ATPase, and that HycB would serve to transfer electrons from Fdh-H to hydrogenase-4 and thus couple formate dehydrogenation with H_2 production in Fhl-2.

H_2 production assays were performed at pH 6.8 or pH 7.5 in the hydrogenase-1, -2 and –3 triple mutant FTD147 (*ΔhyaB, ΔhybC, ΔhycE*) transformed with a multicopy plasmid encoding the *hyfR* gene (pGS1087; enhanced *hyf* expression > 1000 fold; section 4.11). The wildtype (MC4100) and hydrogenase-1, -2 and –3 mutant (FTD147) strains (both transformed with pGS1087) were grown to late exponential phase on LB (section 2.5.2.1) buffered to either pH 6.5 and pH 7.5 with 0.1 M Tris phosphate buffer and supplemented with 0.4% w/v glucose. Harvested whole cells were assayed for hydrogen production at pH 6.8 or pH 7.5 (Fig. 3.19) (section 2.8.5.6). Hydrogen production at both slightly acidic and alkaline pH was abolished by the deletion of *hyaB*, *hybC* and *hycE* genes (FTD147) and no H_2 production attributable to the *hyf* operon was detected. It is possible that the large subunits of hydrogenase-1 and –2 (HyaB and HybC), respectively, absent in strain FTD147 are also required, as HycB is proposed to be, for Fhl-2 to be functional. H_2 production in the wildtype was reduced approximately 14 fold by an increase in growth (and assay) pH from 6.5 (6.8) to 7.5 (7.5). A similarly large reduction in *hyc* expression with an increase in extracellular pH from 6.5 to 7.5 has been reported (Rossmann *et al.*, 1991). However, Bagramyan and co-workers (unpublished) found that hydrogen production in the wildtype was only slightly affected by pH (production was approximately 1.2 fold higher at pH 6.5 than at 7.5). It should be noted however that hydrogen production assays reported here were carried out

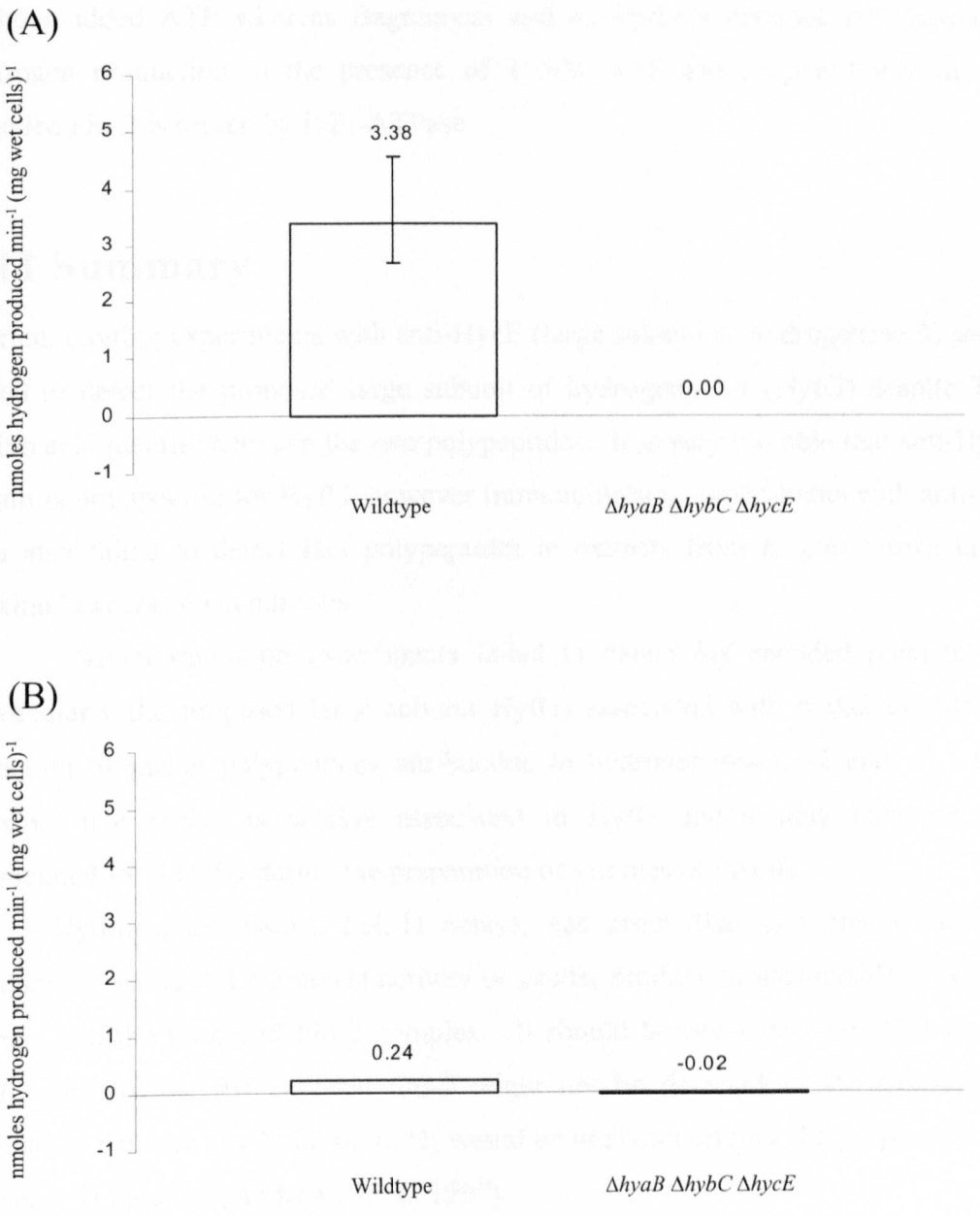


Fig. 3.19. Hydrogen production detected at pH 6.8 and pH 7.5 in wildtype (MC4100) and hydrogenase-1, -2 and –3 triple mutant (FTD147; Δ*hyaB* Δ*hybC* Δ*hycE*) strains transformed with pGS1087 (multicopy plasmid encoding the *hyfR* gene). Cells were grown anaerobically to late exponential phase in L broth medium (section 2.5.2.1) supplemented with 0.4 % glucose and 25 μg/ml chloramphenicol. The medium was buffered to pH 6.6 (A) and pH 7.5 (B) with 0.1 M Tris phosphate buffer. A Clarke-type electrode was modified to measure hydrogen production and uptake at pH 6.8 (A) and pH 7.5 (B) in whole cell suspensions. The value for MC4100 at pH 6.8 is the average of three determinations from two independent cultures and error bars give the maximal and minimal values. Other values are from single determinations.

without added ATP whereas Bagramyan and co-workers detected *hyf* dependent hydrogen production in the presence of 3 mM ATP and proposed that the *hyf* encoded Fhl-2 is driven by F_0F_1-ATPase.

3.11 Summary

Immunoblotting experiments with anti-HycE (large subunit of hydrogenase-3) serum failed to detect the proposed large subunit of hydrogenase-4 (HyfG) despite 73% amino acid identity between the two polypeptides. It is very possible that anti-HycE serum is not specific for HyfG, however immunoblotting experiments with anti-Hyf sera also failed to detect Hyf polypeptides in extracts from *E. coli* grown under maximal expression conditions.

^{63}Ni incorporation experiments failed to detect *hyf* encoded polypeptides (particularly the proposed large subunit HyfG) associated with nickel despite the detection of nickel polypeptides attributable to hydrogenases-1, -2 and –3. It is possible that nickel is weakly associated to HyfG and/or may have become dissociated from HyfG during the preparation of samples or PAGE.

Hydrogenase assays, Fdh-H assays, gas production experiments and H_2 production assays did not detect activity or gas/H_2 production attributable to the *hyf* operon and the proposed Fhl-2 complex. It should however be noted that gas/H_2 production by the *hyf* encoded Fhl-2 might not be detected in the experiments conducted because any build up in H_2 would be unfavourable for the proposed Fhl-2 complex to function (Andrews *et al.*, 1997).

Taken together all these experiments failed to detect *hyf* gene products and activity attributable to the *hyf* operon. It is possible that expression levels of Hyf polypeptides are too low to be detected in these experiments.

4. REGULATION OF *hyf* OPERON EXPRESSION AND FURTHER EVIDENCE FOR *hyfR* ENCODING A σ^{54}-DEPENDENT TRANSCRIPTIONAL REGULATOR OF THE *hyf* OPERON.

4.1 Introduction

Expression of the hydrogenase-1 (*hya*) and hydrogenase-2 (*hyb*) operons is induced by anaerobiosis and repressed by nitrate (Bronsted & Atlung, 1994; Richard *et al.*, 1999). The *hyb* operon has also been found to be catabolite repressed (Richard *et al.*, 1999). Expression of the hydrogenase-3 (*hyc*) operon, and other genes of the formate regulon, is induced by the absence of oxygen and other external electron acceptors (e.g nitrate), the presence of formate and molybdate, and an acidic pH (Schlensog *et al.*, 1989; Schlensog & Böck, 1990; Rossmann *et al.*, 1991).

To study the regulation of hydrogenase-4 (*hyf*) operon expression, a λ*hyfA-lacZ* transcriptional fusion phage containing the *hyfA* promoter region and part of the *hyfA* coding region linked 'in phase' to the β-galactosidase reporter gene, was constructed and established as a single copy prophage in MC4100 (Δ*lac*). Initial expression studies with this strain (DS5) found that optimal *hyf* expression, as for *hyc*, is observed when *E. coli* is grown anaerobically in the presence of formate and in the absence of exogenous electron acceptors (P.Golby, unpublished data).

In the studies described in this chapter, regulation of the *hyf* operon was examined by studying *hyf* expression under different growth conditions and in different mutant backgrounds. Also the effect of HyfR (proposed σ^{54}-dependent transcriptional regulator of the *hyf* operon) on expression of the *hyf* operon, *hyc*

operon and *fdhF* gene was examined. β-galactosidase activities reported in this chapter were assayed as described in section 2.8.5.1. All strains were grown anaerobically in 8 ml Bijoux as described in section 2.6.1.3 and all media used are specified in the figure legends.

4.2 Construction of the *hyfA-lacZ* transcriptional fusion strain, DS5

A strategy for constructing the *hyfA-lacZ* transcriptional fusion strain, DS5 (MC4100, λ*hyfA-lacZ bla*) was designed and executed by Dr P. Golby.

The *hyfA-lacZ* transcriptional fusion was created by ligating the 0.957 kb *Eco*RV-*Hin*dIII fragment containing the entire *bcp-hyf* intergenic region upstream of *hyfA* and part of the *hyfA*-coding region into the *Sma*I site of pRS415 (Simons *et al.*, 1987) to generate pGS935. The fusion was then transferred to λRZ5 by *in vivo* recombination in strain RK4353 (pGS935) according to Spiro and Guest (1987). The corresponding λ*hyfA-lacZ* fusion phage (λ*hyfA-lacZ bla*) was established as a single-copy prophage in MC4100 (Δ*lac*). Lysogens were selected initially with *λh80del9c* and monolysogens were identified by their sensitivity to *λc190c17*.

4.3 Construction of *hyfR* deletion strain DS6

Strain DS6 (MC4100, λ*hyfA-lacZ bla*, Δ*hyfR*::*spc*) was created to study the effect of HyfR (proposed σ^{54}-dependent transcriptional regulator of the *hyf* operon) on *hyf* operon expression. The strain was produced using a P1 lysate grown on strain JRG3618 (MC4100, Δ*hyfR*::*spc*) to transfer the Δ*hyfR*::*spc* deletion to strain DS5 (MC4100, λ*hyfA-lacZ bla*) by transduction.

4.3.1 PCR amplification of the *hyfR* region

To confirm that the predicted chromosomal alterations had been made to the chromosome in strain DS6, the *hyfR* region was amplified using PCR with primers designed to amplify the *hyfR* region (primers *hyfR*-F1 & *hyfR*-R1) (Table 2.3; section 2.7.3.2). As controls, the corresponding region was amplified from MC4100, JRG3618 and DS5. The PCR product of DS6 and JRG3618 was approximately 2140

bp compared with approximately 1700 bp for MC4100 and DS5 (Fig. 4.1). The size of the PCR products were in agreement with predicted sizes from known sequence data (Table 4.1).

4.3.2 PCR amplification of the *hyfA-lacZ* region

To confirm that predicted alterations had been made to the chromosome in strain DS6, the *hyfA-lacZ* region was amplified using PCR with primers designed to amplify the *hyfA-lacZ* region (primers *hyfRB*-1F & *lacZ*-R) (Table 2.3; section 2.7.3.2). As controls, PCR with these primers was also carried out with genomic DNA purified from MC4100, DS5 and JRG3618. The PCR products of DS6 and DS5 were approximately 740 bp, which is in agreement with the predicted size from known sequence data (Fig. 4.2). An amplification product was not detected for MC4100 and JRG3618 because the *hyfA-lacZ* fusion was not present in these strains.

4.4 Construction of *fhlA* mutant strain DS7

Strain DS7 (MC4100, λ*hyfA-lacZ bla*, *fhlA*::λp*lac*Mu53 *kan*) was created to study the effect of FhlA (σ^{54}-dependent transcriptional activator of the 'formate regulon') on *hyf* operon expression. The strain was produced using a P1 lysate grown on strain SV83 (MC4100, *fhlA*::λp*lac*Mu53 *kan*) to transfer the *fhlA*::λp*lac*Mu53 *kan* mutation to strain DS5 (MC4100, λ*hyfA-lacZ bla*) by transduction.

4.4.1 PCR amplification of the *hyfA-lacZ* region

The presence of the *hyfA-lacZ* fusion in strain DS7 was successfully verified using PCR as described in section 4.3.2.

4.4.2 Phenotypic confirmation of the *fhlA* mutation

To confirm the presence of the *fhlA* mutation (*fhlA*::λp*lac*Mu53), strain DS7 was analysed for gas production (section 2.8.5.5). As controls, strains MC4100, DS5 and SV83 were also analysed for gas (H_2) forming capacity. Schlensog and co-workers (1989) reported that gas production (as an overall measure of formate hydrogenlyase

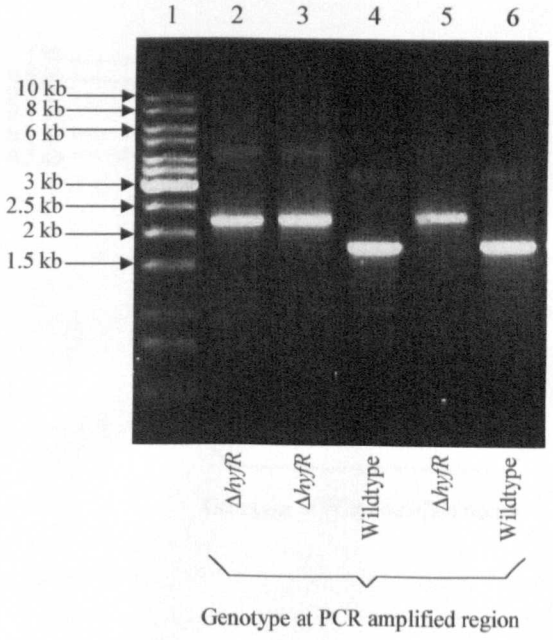


Fig. 4.1. PCR amplification of the *hyfR* region from DS6, MC4100, JRG3618 and DS5.
PCR with primers designed to amplify the *hyfR* region (primers *hyfR*-R1 & *hyfR*-R1) (Table 2.3; section 2.7.3.2). PCR products (2 µl of 50 µl reaction) were electrophoresed on an agarose gel (0.7 %; section 2.7.4). Lane 1, 1 kb DNA ladder; Lane 2, DS6; Lane 3, DS6; Lane 4, MC4100; Lane 5, JRG3618; Lane 6, DS5.

Strain	Size (bp)	
	Expected	Observed (estimated)
MC4100	1751	1700
DS6	2190	2140

Table 4.1. Expected and observed (estimated) sizes of PCR products from the amplification of the *hyfR* region from MC4100 and DS6.

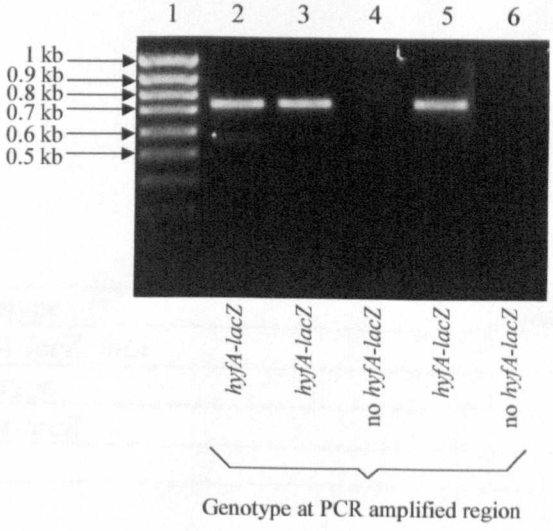


Fig. 4.2. PCR amplification of the *hyfA-lacZ* region from DS6, MC4100, DS5 and JRG3618.

PCR with primers designed to amplify the *hyfA-lacZ* region (primers *hyfRB*-1F & *lacZ*-R) (Table 2.3; section 2.7.3.2). PCR products (2 µl of 50 µl reaction) were electrophoresed on an agarose gel (0.7 %; section 2.7.4). Lane 1, 100 bp DNA ladder; Lane 2, DS6; Lane 3, DS6; Lane 4, MC4100; Lane 5, DS5; Lane 6, JRG3618.

Strain	Genotype	Gas production
DS7	λ*hyfA-lacZ*, *fhlA*$^-$	-
MC4100	Wildtype	+
DS5	λ*hyfA-lacZ*	+
SV83	*fhlA*$^-$	-

Table 4.2. Gas production in DS7, MC4100, DS5 and SV83.
Cells were grown to late exponential phase in TYEP (pH 6.6) supplemented with 0.4% glucose and 30 mM sodium formate, and any gas production was trapped in inverted Durham tubes. Observations were taken from experiments carried out in duplicate. Key to observations: +, gas production observed; -, no gas production observed.

activity) was lost in strain SV83. Gas production was detected in strains MC4100 and DS5 but gas forming capacity was absent in strains DS7 and SV83 (Table 4.2).

4.5 Construction of *fhlA* mutant strain DS8

Strain DS8 (MC4100, *fhlA*::λp*lac*Mu53 *kan*) was created as a control strain to measure what proportion of β-galactosidase activity detected in strain DS7 was attributable to the *fhlA*::λp*lac*Mu53 mutation (generated by the integration of λp*lac*Mu53 into the *fhlA* gene). The strain was produced using a P1 lysate grown on strain SV83 (MC4100, *fhlA*::λp*lac*Mu53 *kan*) to transfer the *fhlA*::λp*lac*Mu53 *kan* mutation to strain MC4100 (Δ*lac*).

4.5.1 Phenotypic confirmation of the *fhlA* mutation

To confirm the presence of the *fhlA* mutation (*fhlA*::λp*lac*Mu53 *kan*), strain DS8 was analysed for gas production as described in section 4.4.2.

4.6 Construction of *hycA* deletion strain DS9

Strain DS9 (MC4100, λ*hyfA-lacZ bla*, Δ*hycA*) was created to study the effect of HycA (anti-activator of the 'formate regulon') on *hyf* operon expression. The strain was produced using a P1 lysate grown on strain DS5 (MC4100, λ*hyfA-lacZ bla*) to transfer the λ*hyfA-lacZ* fusion to strain HD701 (MC4100, Δ*hycA*) by transduction. Transductants were selected on ampicillin plates.

4.6.1 PCR amplification of the *hycA* region

To confirm that predicted alterations had been made to the chromosome in strain DS9, the *hycA* region was amplified using PCR with primers designed to amplify the *hycA* region (primers *hycA*-L & *hycA*-R) (Table 2.3; section 2.7.3.4). As controls, the corresponding region was amplified from HD701, MC4100 and DS5. The PCR products of DS9 and HD701 were estimated to be between 200 and 300 bp compared with approximately 500 bp for MC4100 and DS5 (Fig. 4.3). The size of the PCR products was in agreement with predicted sizes from known sequence data (Table 4.3).

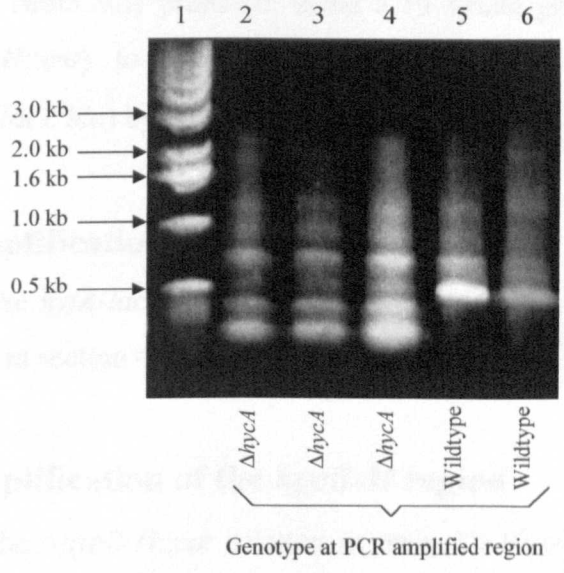


Fig. 4.3. PCR amplification of the *hycA* region from DS9, HD701, MC4100 and DS5.
PCR with primers designed to amplify the *hyfA* region (primers *hycA*-L & *hycA*-R) (Table 2.3; section 2.7.3.4). PCR products (2 µl of 50 µl reaction) were electrophoresed on an agarose gel (0.7 %; section 2.7.4). Lane 1; 1 kb DNA ladder (Gibco Brl); Lane 2, DS9; Lane 3, DS9; Lane 5, HD701; Lane 7, MC4100; Lane 8, DS5; Lanes 4 and 6, empty. For all strains in addition to amplification of the band of expected size a number of fainter larger bands were also amplified (Lanes 2 to 6). This is likely to be a consequence of the long extension time used for this PCR protocol.

Strain	Size (bp)	
	Expected	Observed (estimated)
MC4100	521	Approximately 500
DS9	243	<500

Table 4.3. Expected and observed (estimated) sizes of PCR products from the amplification of the *hycA* region from MC4100 and DS9.

4.7 Construction of *hycB-H* deletion strain DS10

Strain DS10 (MC4100, λ*hyfA-lacZ bla*, Δ*hycB-H*::*cat*) was created to study the effect of the Fhl-1 complex (encoded by the *hyc* operon and the *fdhF* gene) on *hyf* operon expression. The strain was produced using a P1 lysate grown on strain HDJ123 (JW136, Δ*hycB-H*::*cat*) to transfer the Δ*hycB-H*::*cat* deletion to strain DS5 (MC4100, λ*hyfA-lacZ bla*) by transduction.

4.7.1 PCR amplification of the *hyfA-lacZ* region

The presence of the *hyfA-lacZ* fusion in strain DS10 was successfully verified using PCR as described in section 4.3.2.

4.7.2 PCR amplification of the *hycB-H* region

The presence of the Δ*hycB-H*::*cat* deletion in strain DS10 was successfully verified using PCR as described in section 3.4.1.

4.8 Construction of the *ntrA* deletion strain DS11

Strain DS11 (MC4100, λ*hyfA-lacZ bla*, Δ(*ntrA208*::Tn*10*)) was created to study the effect of NtrA (encodes the sigma factor, σ^{54}) on *hyf* operon expression. The strain was produced using a P1 lysate grown on strain BN450 (Δ(*ntrA208*::Tn*10*) Δ(*srl-recA*)306::Tn*10*) to transfer the Δ(*ntrA208*::Tn*10*) mutation to strain DS5 (MC4100, λ*hyfA-lacZ bla*) by tranduction and selection for tetracycline resistance and glutamine auxotrophy. An *ntrA* mutant is a glutamine auxotroph when grown in glucose containing medium.

4.8.1 PCR amplification of the *hyfA-lacZ* region

The presence of the *hyfA-lacZ* fusion in strain DS11 was successfully verified using PCR as described in section 4.3.2.

4.8.2 Phenotypic confirmation of the *ntrA* mutation

To confirm the presence of the *ntrA* mutation (Δ(*ntrA208*::Tn*10*)), strain DS11 was analysed for gas production (section 2.8.5.5). As controls, strain MC4100, DS5 and BN450 were also analysed for gas forming capacity. Birkmann and co-workers (1987b) reported that gas production (as an overall measure of formate hydrogenlyase activity) was lost in *ntrA* mutant strains. Gas production was detected in strains MC4100 and DS5, but gas forming capacity was lost in strains DS11 and BN450 (Table 4.4).

4.9 Expression and transcriptional organisation of the *hyf* operon

To demonstrate that the *hyf* operon is expressed and to investigate its regulation, a *hyfA-lacZ* fusion containing the entire *bcp-hyf* intergenic region upstream of *hyfA* together with part of the *hyfA* coding region was constructed (section 4.2). Expression activity attributable to the *hyfA*-lacZ fusion was detected when the fusion was present both in multicopy (in strains transformed with pGS935) and in single copy (strain DS5) indicating that *hyf* is likely to be a functional operon. Initial experiments executed by Dr P. Golby with the single copy *hyfA-lacZ* fusion strain DS5 (MC4100, λ*hyfA-lacZ bla*) found that optimal expression of the *hyf* operon, like the *hyc* operon, was observed when *E. coli* was grown anaerobically, in the presence of formate and the absence of exogenous electron acceptors (P. Golby, unpublished data).

The *hyfR* and *focB* genes are preceded by 29 bp and 20 bp, respectively, of non-coding DNA, suggesting that unlike the other genes in the *hyf* operon, they are not subject to translational coupling and may thus be under control of independent promoters. To investigate whether these two genes possess independent promoters, *hyfR-lacZ* and *focB-lacZ* transcriptional fusions were constructed by Dr P. Golby. No expression activity was observed for the *hyfR-lacZ* and *focB-lacZ* fusions in multicopy. Therefore the *hyfR* and *focB* genes are likely to be transcribed as a single transcript with the rest of the *hyf* operon (P. Golby, unpublished data).

Reverse-transcriptase mediated primer-extension analysis was carried out by Dr P. Golby to investigate the location of the transcriptional start site of the *hyf*

Strain	Genotype	Gas production
DS11	λ*hyfA-lacZ*, Δ(*ntrA208*::Tn*10*)	-
MC4100	Wildtype	+
DS5	λ*hyfA-lacZ*	+
BN450	Δ(*ntrA208*::Tn*10*)	-

Table 4.4. Gas production in DS11, MC4100, DS5 and BN450.
Cells were grown to late exponential phase in TYEP (pH 6.6) supplemented with 0.4% glucose and 30 mM sodium formate, and any gas production was trapped in inverted Durham tubes. Observations were taken from experiments carried out in duplicate. Key to observations: +, gas production observed; -, no gas production observed.

operon. RNA was isolated from wildtype cells grown anaerobically in L broth supplemented with 0.4% glucose and 50 mM sodium formate (optimal *hyf* expression conditions). Using two different priming sites a transcriptional start site was detected 30 bp upstream of the *hyfA* gene. This start site is 12 bp downstream from the strongly predicted σ^{54}-dependent promoter indicating that transcription from the *hyf* operon is σ^{54} dependent. No primer extension products were detected for *hyfR* and *focB* supporting the conclusion that these genes do not possess independent promoters and are therefore likely to be co-transcribed with the rest of the *hyf* operon (P. Golby, unpublished).

Northern blotting experiments were carried out by Dr P. Golby using RNA isolated from MC4100 (wt) and HD700 (MC4100, Δ*hycA-H*) grown anaerobically on L broth supplemented with 0.4% glucose and 50 mM sodium formate (optimal *hyf* expression conditions). No *hyf* hybridising transcripts were detected using *hyfB-R*, *hyfR* and *focB* probes. Failure to detect *hyf* hybridising transcripts suggests that expression levels of the *hyf* operon are too low to be detected. Also because it appears that the *hyf* operon is transcribed as a single transcript, the size of the transcript (approximately 14000 bp) would make it prone to fragmentation *in vivo* and during RNA purification. No Northern blotting data has been reported for the 8 kb *hyc* operon, which is also transcribed as a single transcript.

4.10 The effect of growth conditions on *hyfA-lacZ* expression

4.10.1 The effect of pH on *hyfA-lacZ* expression

Expression of the *hyc* operon was found to be strongly pH-dependent, with increasing extracellular pH resulting in a reduction in expression (Rossmann *et al.*, 1991). Supplementation with sodium formate greatly relieved this lack of induction at alkaline pH with expression of the *hyc* operon becoming less pH dependent.

Strain DS5 was used to assess the effect of external pH on *hyf* operon expression in the presence and absence of added sodium formate (Fig. 4.4 & 4.5). It is important to note that the pH values indicated on Fig. 4.5 are the initial pH values of the media. Media having an initial pH of 6.1, 7.0 or 7.8, although buffered,

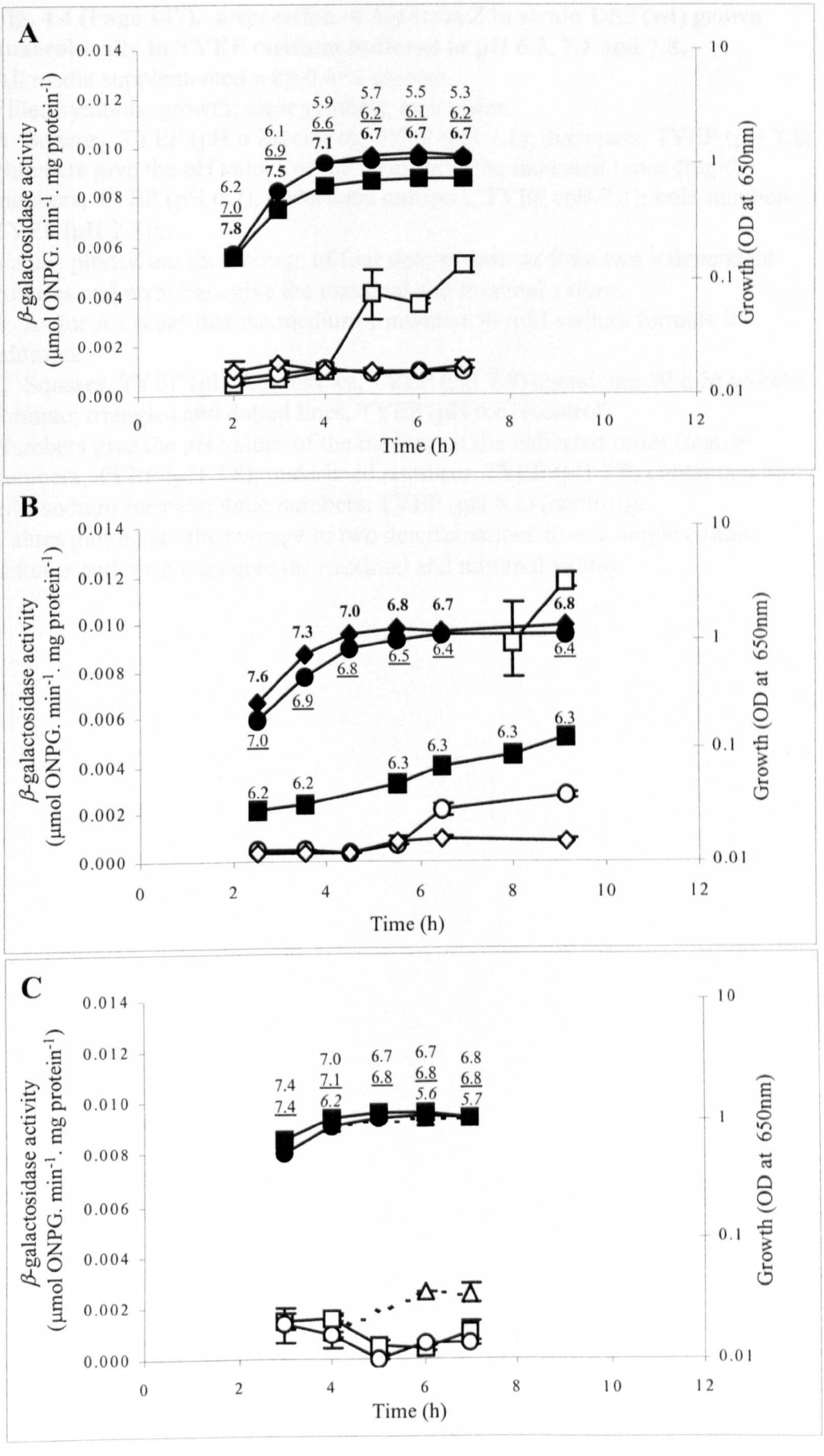

A
B
C
β-galactosidase activity (μmol ONPG. min-1. mg protein-1)
Growth (OD at 650nm)
Time (h)
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014
0.01
0.1
1
10
0
2
4
6
8
10
12

Fig. 4.4 (Page 147). Expression of *hyfA*::*lacZ* in strain DS5 (wt) grown anaerobically in TYEP medium buffered to pH 6.2, 7.1 and 7.8.
All media supplemented with 0.4 % glucose.
Filled symbols, growth; clear symbols, expression.
A Squares, TYEP (pH 6.2); circles, TYEP (pH 7.1); diamonds, TYEP (pH 7.8).
Numbers give the pH values of the medium at the indicated times (regular numbers, TYEP (pH 6.2); underlined numbers, TYEP (pH 7.1); bold numbers, TYEP (pH 7.8)).
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.
B As for **A** except that the medium contained 50 mM sodium formate in addition.
C Squares, TYEP (pH 7.8); circles, TYEP (pH 7.8) containing 50 mM sodium formate; triangles and dotted lines, TYEP (pH 6.6) (control).
Numbers give the pH values of the medium at the indicated times (regular numbers, TYEP (pH 7.8); underlined numbers, TYEP (pH 7.8) containing 50 mM sodium formate; italic numbers, TYEP (pH 6.6) (control)).
Values plotted are the average of two determinations from a single culture cultures and error bars give the maximal and minimal values.

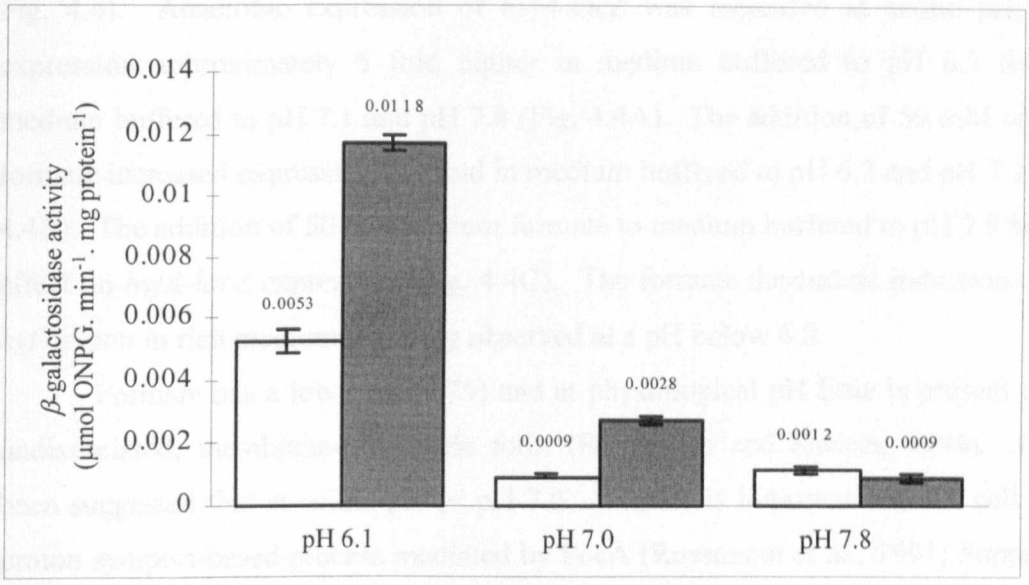


Fig. 4.5. Summary of Figs 4.4A and B.
Expression of *hyfA*::*lacZ* in strains DS5 (wt) growing anaerobically in TYEP medium buffered to pH 6.2, 7.1 and 7.8. TYEP supplemented with 0.4 % glucose.
Clear bars, no formate; filled bars, in the presence of 50 mM sodium formate. Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.
Summary of **Fig. 4.4A and B**. Values plotted are taken from stationary phase: **Fig. 4.4A**, 7 h; **Fig. 4.4B**, 9.2 h.

had a final pH (after fermentative growth) as low as 5.3, 6.2 or 6.7 respectively (see Fig. 4.4). Anaerobic expression of *hyfA-lacZ* was increased at acidic pH, with expression approximately 5 fold higher in medium buffered to pH 6.2 than in medium buffered to pH 7.1 and pH 7.8 (Fig. 4.4A). The addition of 50 mM sodium formate increased expression 2-3 fold in medium buffered to pH 6.2 and pH 7.1 (Fig. 4.4B). The addition of 50 mM sodium formate to medium buffered to pH 7.8 had no effect on *hyfA-lacZ* expression (Fig. 4.4C). The formate dependent induction of the *hyf* operon in rich medium was only observed at a pH below 6.8.

Formate has a low pKa (3.75) and at physiological pH little is present in the undissociated, membrane-permeable form (Suppmann and Sawers, 1994). It has been suggested that at acidic pH (< pH 7.0), formate is imported into the cell via a proton symport-based process mediated by FocA (Rossmann *et al.*, 1991; Suppmann & Sawers, 1994). The results presented above are consistent with the theory that dependence of *hyf* operon expression at low pH is channelled via the intracellular concentration of formate, with formate only being able to accumulate to sufficient levels in the cell when the external medium is at pH 6.8 or below.

4.10.2 Inhibition of *hyf* expression by complex medium

Colonies of the *hyfA-lacZ* strain DS5 (wt) on aerobic agar plates containing 40 μM ml^{-1} X-Gal were more strongly Lac^+ on M9 minimal medium than on rich medium. This suggested that *hyfA-lacZ* expression is greater on M9 minimal medium (section 2.5.3.1) than on L broth complex medium (section 2.5.2.1). When grown in liquid M9 minimal medium under anaerobic glucose fermentative conditions, β-galactosidase activity was approximately two-fold greater than in TYEP complex medium (section 2.5.2.2) (Figs 4.6 & 4.7). Addition of 1% tryptone to the M9 minimal medium did not reverse this increase but further enhanced expression approximately 2 fold. It should be noted that the initial pH of the TYEP complex and M9 minimal medium used was different, with TYEP buffered to a pH of 6.6 and M9 minimal medium having an initial pH of approximately 7.0. However, this difference in pH between the two media is unlikely to account for the elevated expression observed in M9 minimal medium, as the higher pH of M9 minimal medium would have a negative effect on *hyf* expression (see effect of pH, section 4.10.1). Also the pH of TYEP and of M9 minimal medium supplemented with 1%

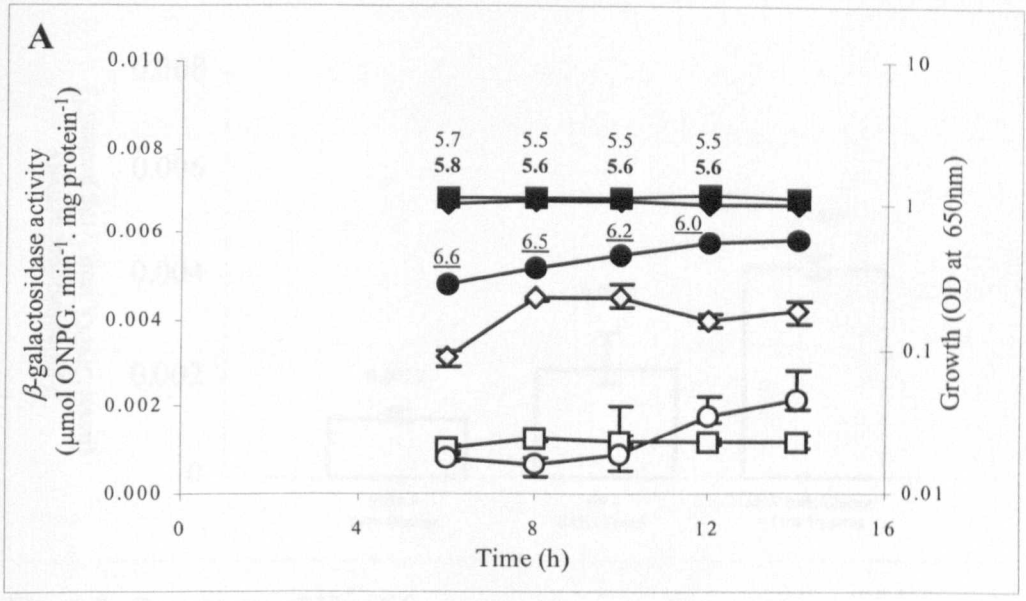


Fig. 4.6. Expression of *hyfA*::*lacZ* in strain DS5 (wt) grown anaerobically in rich and minimal media.

Squares, TYEP (pH 6.6); circles, M9 minimal medium; diamonds, M9 minimal medium containing 1.0 % w/v tryptone.

All media supplemented with 0.4 % glucose.

Filled symbols, growth; clear symbols, expression.

Numbers give the pH values of the medium at the indicated times (regular numbers, TYEP (pH 6.6); underlined numbers, M9 minimal medium; bold numbers, M9 minimal medium containing 1.0 % w/v tryptone).

Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

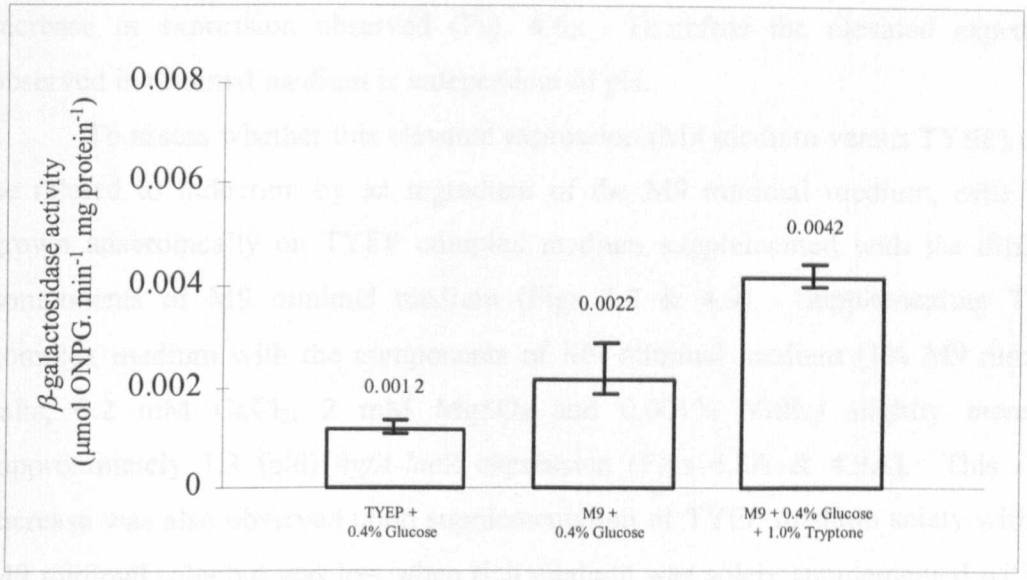


Fig. 4.7. Summary of Fig. 4.6.
Expression of *hyfA*::*lacZ* in strains DS5 (wt) growing anaerobically in rich and minimal media.
Values plotted are taken from stationary phase (14 h).
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

tryptone was almost identical during growth and so cannot account for the 3.5 fold increase in expression observed (Fig. 4.6). Therefore the elevated expression observed in minimal medium is independent of pH.

To assess whether this elevated expression (M9 medium versus TYEP) could be related to induction by an ingredient of the M9 minimal medium, cells were grown anaerobically on TYEP complex medium supplemented with the different components of M9 minimal medium (Figs 4.8 & 4.9). Supplementing TYEP complex medium with the components of M9 minimal medium (1% M9 minimal salts, 0.2 mM $CaCl_2$, 2 mM $MgSO_4$ and 0.001% $VitB_1$) slightly increased (approximately 1.3 fold) *hyfA-lacZ* expression (Figs 4.8A & 4.9A). This slight increase was also observed upon supplementation of TYEP medium solely with 1% M9 minimal salts but was lost when rich medium was solely supplemented with the other components of M9 minimal medium (0.2 mM $CaCl_2$, 2 mM $MgSO_4$ and 0.001% $VitB_1$). To assess which ingredient of M9 minimal salts was responsible for this 1.3 fold increase, TYEP complex medium was supplemented with the different components of M9 minimal salts (42 mM Na_2HPO_4, 22 mM KH_2PO_4, 9 mM NaOH and 19 mM NH_4Cl) (Figs 4.8B & 4.9B). Addition of 42 mM Na_2HPO_4 and 22 mM KH_2PO_4 to TYEP rich medium elevated *hyfA-lacZ* expression 1.7 fold, accounting for the slight increase in expression observed in TYEP complex medium supplemented with the components on M9 minimal medium. Expression was not elevated in TYEP complex medium supplemented with 9 mM NaCl or 19 mM NH_4Cl. It is not known why Na_2HPO_4 and KH_2PO_4 elevated *hyfA-lacZ* expression. Addition of these salts to TYEP rich medium increased the buffering capacity of the medium reducing the amount that the pH dropped during growth (Fig. 4.8B). However this reduced drop in pH would be expected to have a negative effect on expression (see effect of pH, section 4.10.1). This 1.7 fold increase in *hyfA-lacZ* expression observed in TYEP complex medium supplemented with Na_2HPO_4 and KH_2PO_4 is not enough to account for the approximately 3.5 fold increase in *hyfA-lacZ* expression observed when M9 minimal medium is supplemented with 1% tryptone (Figs 4.6 & 4.7). Therefore elevation of *hyf* expression in minimal medium (with respect to TYEP) is not solely due to induction of expression by Na_2HPO_4 and KH_2PO_4 (ingredients of the M9 minimal medium used), but also a possible inhibition of expression by a component of TYEP complex medium other than tryptone i.e. yeast extract.

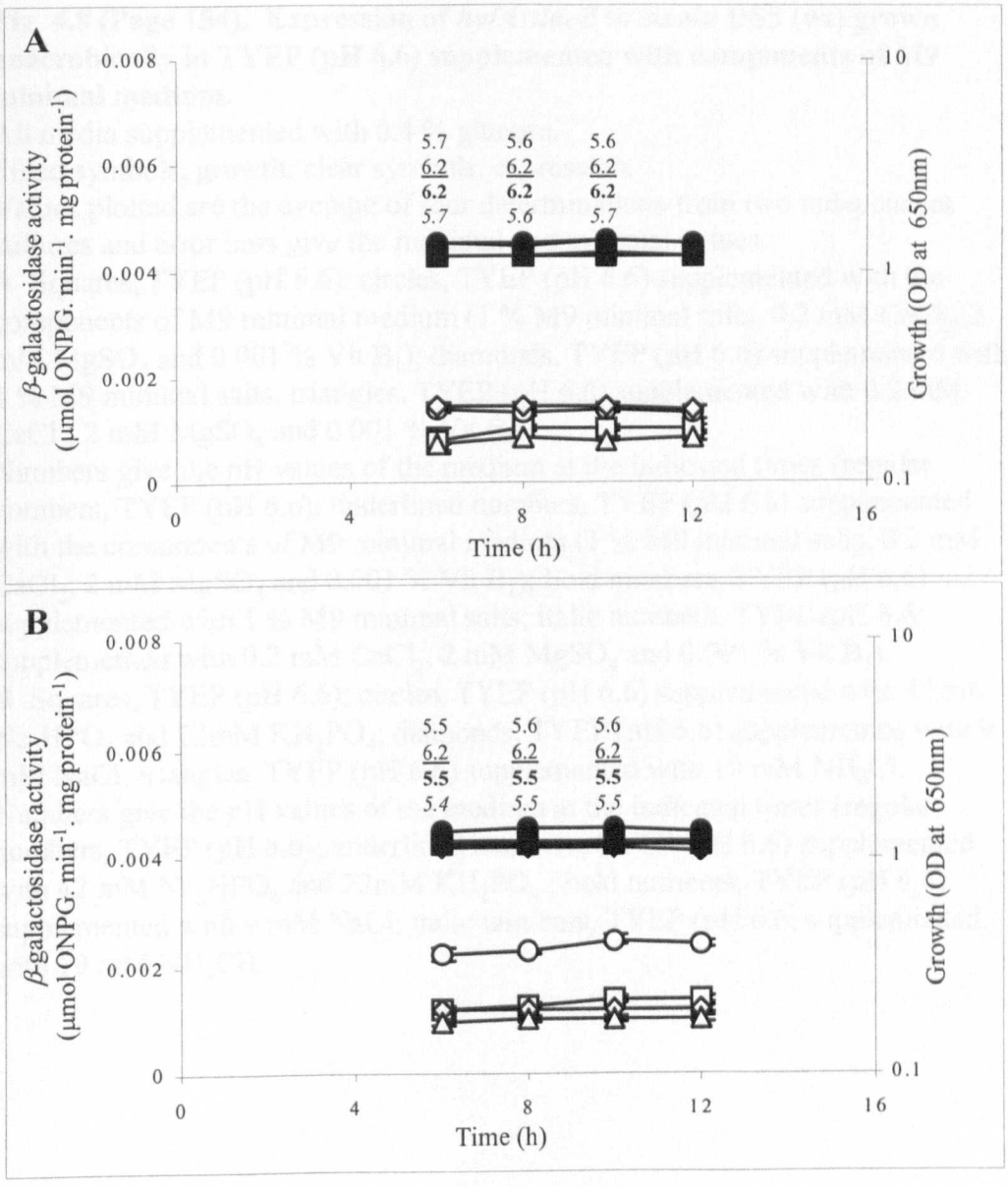

A
0.008
0.006
0.004
0.002
0
β-galactosidase activity
(µmol ONPG. min-1. mg protein-1)
10
1
0.1
Growth (OD at 650nm)
5.7 5.6 5.6
6.2 6.2 6.2
6.2 6.2 6.2
5.7 5.6 5.7
0
4
8
12
16
Time (h)
B
0.008
0.006
0.004
0.002
0
β-galactosidase activity
(µmol ONPG. min-1. mg protein-1)
10
1
0.1
Growth (OD at 650nm)
5.5 5.6 5.6
6.2 6.2 6.2
5.5 5.5 5.5
5.4 5.5 5.5
0
4
8
12
16
Time (h)

Fig. 4.8 (Page 154). Expression of *hyfA::lacZ* in strain DS5 (wt) grown anaerobically in TYEP (pH 6.6) supplemented with components of M9 minimal medium.
All media supplemented with 0.4 % glucose.
Filled symbols, growth; clear symbols, expression.
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.
A Squares, TYEP (pH 6.6); circles, TYEP (pH 6.6) supplemented with the components of M9 minimal medium (1 % M9 minimal salts, 0.2 mM $CaCl_2$, 2 mM $MgSO_4$ and 0.001 % Vit B_1); diamonds, TYEP (pH 6.6) supplemented with 1 % M9 minimal salts; triangles, TYEP (pH 6.6) supplemented with 0.2 mM $CaCl_2$, 2 mM $MgSO_4$ and 0.001 % Vit B_1.
Numbers give the pH values of the medium at the indicated times (regular numbers, TYEP (pH 6.6); underlined numbers, TYEP (pH 6.6) supplemented with the components of M9 minimal medium (1 % M9 minimal salts, 0.2 mM $CaCl_2$, 2 mM $MgSO_4$ and 0.001 % Vit B_1); bold numbers, TYEP (pH 6.6) supplemented with 1 % M9 minimal salts; italic numbers, TYEP (pH 6.6; supplemented with 0.2 mM $CaCl_2$, 2 mM $MgSO_4$ and 0.001 % Vit B_1).
B Squares, TYEP (pH 6.6); circles, TYEP (pH 6.6) supplemented with 42 mM Na_2HPO_4 and 22mM KH_2PO_4; diamonds, TYEP (pH 6.6) supplemented with 9 mM NaCl; triangles, TYEP (pH 6.6) supplemented with 19 mM NH_4Cl.
Numbers give the pH values of the medium at the indicated times (regular numbers, TYEP (pH 6.6); underlined numbers, TYEP (pH 6.6) supplemented with 42 mM Na_2HPO_4 and 22mM KH_2PO_4 ; bold numbers, TYEP (pH 6.6) supplemented with 9 mM NaCl; italic numbers, TYEP (pH 6.6; supplemented with 19 mM NH_4Cl).

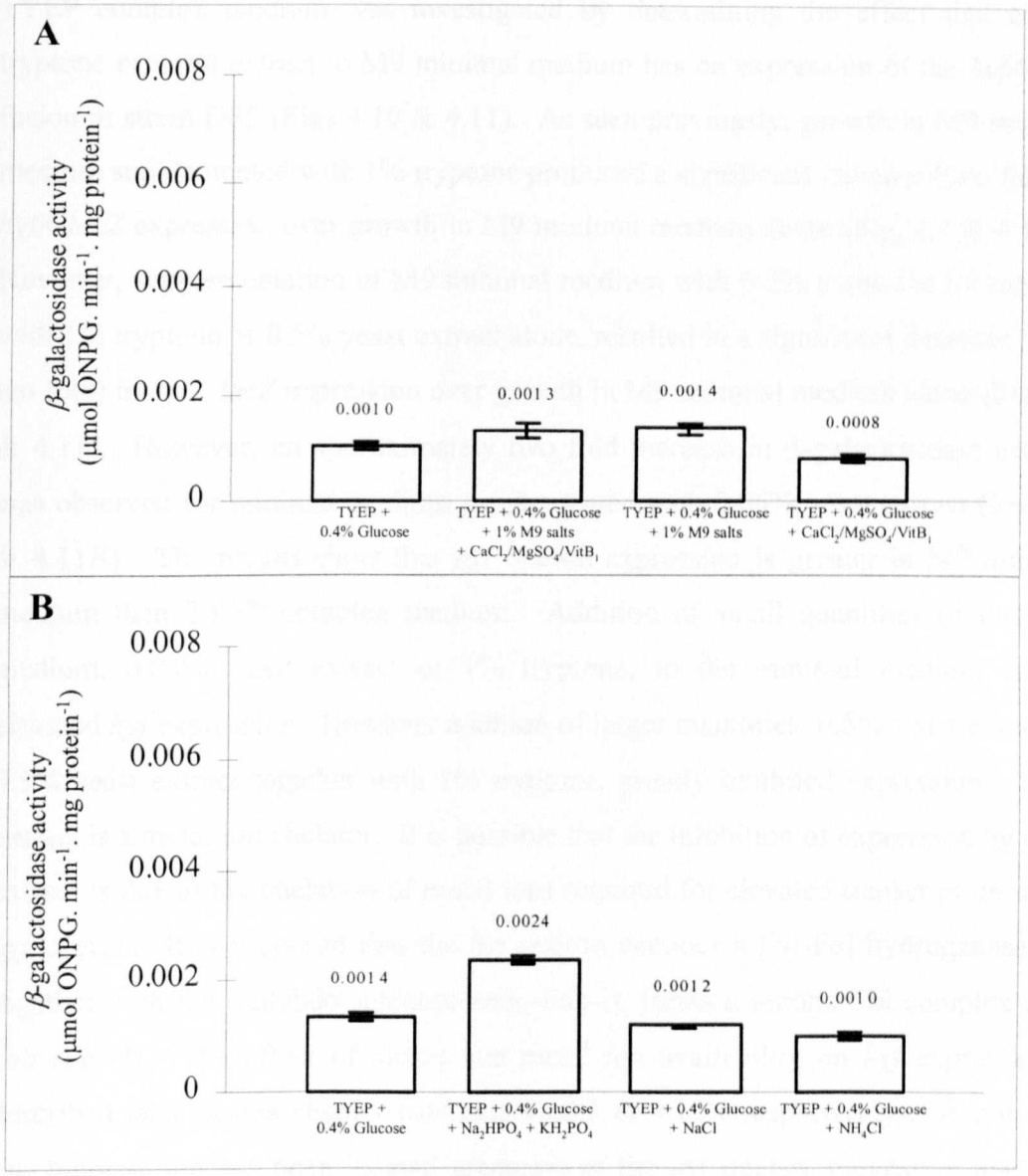


Fig. 4.9. Summary of Fig. 4.8.
Expression of *hyfA*::*lacZ* in strains DS5 (wt) growing anaerobically in TYEP (pH 6.6) supplemented with components of M9 minimal medium.
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.
A Summary of **Fig. 4.8A**. Values plotted are taken from stationary phase (12 h).
B Summary of **Fig. 4.8B**. Values plotted are taken from stationary phase (12 h).

The possibility of an inhibitory effect on *hyf* expression by a component of TYEP complex medium was investigated by determining the effect that adding tryptone or yeast extract to M9 minimal medium has on expression of the *hyfA-lacZ* fusion in strain DS5 (Figs 4.10 & 4.11). As seen previously, growth in M9 minimal medium supplemented with 1% tryptone produced a significant increase (two fold) in *hyfA-lacZ* expression over growth in M9 minimal medium alone (Fig. 4.7 & 4.11A). However, supplementation of M9 minimal medium with 0.5% yeast extract together with 1% tryptone or 0.5% yeast extract alone, resulted in a significant decrease (five-ten fold) in *hyfA-lacZ* expression over growth in M9 minimal medium alone (Fig. 4.7 & 4.11). However, an approximately two fold increase in β-galactosidase activity was observed for minimal medium supplemented with 0.05% yeast extract (Fig. 4.7 & 4.11B). The results show that *hyf* operon expression is greater in M9 minimal medium than TYEP complex medium. Addition of small quantities of complex medium, 0.05% yeast extract or 1% tryptone, to the minimal medium further elevated *hyf* expression. However addition of larger quantities, 0.5% yeast extract or 0.5% yeast extract together with 1% tryptone, greatly inhibited expression. Yeast extract is a metal ion chelator. It is possible that the inhibition of expression by yeast extract is due to the chelation of metal ions required for elevated transcription of the *hyf* operon. It is proposed that the *hyf* operon encodes a [Ni-Fe] hydrogenase that together with the molybdo-selenoprotein, Fdh-H, forms a second Fhl complex in *E. coli* (Fhl-2). The effect of nickel and metal ion availability on *hyf* expression is described later in this chapter (sections 4.10.3 & 4.10.4 respectively). A potential Fur binding site has been located upstream of the *hyf* operon suggesting that iron may play a role on the transcriptional regulation of the *hyf* operon (Andrews *et al*, 1991). Also, molybdate is required for maximal *fdhF* and *hyc* transcription (Schlensog *et al.*, 1989). However it is yet to be conclusively investigated whether iron or molybdate have an effect on *hyf* operon transcription. Also, the effect of selenium on the expression of the *hyf* operon has yet to be studied specifically.

4.10.3 Effect of nickel on *hyfA-lacZ* expression

Transcription of the *hycB* gene has been shown to be independent of nickel by studying *hycB-lacZ* expression in nickel depleted medium (Zinoni *et al.*, 1984). To

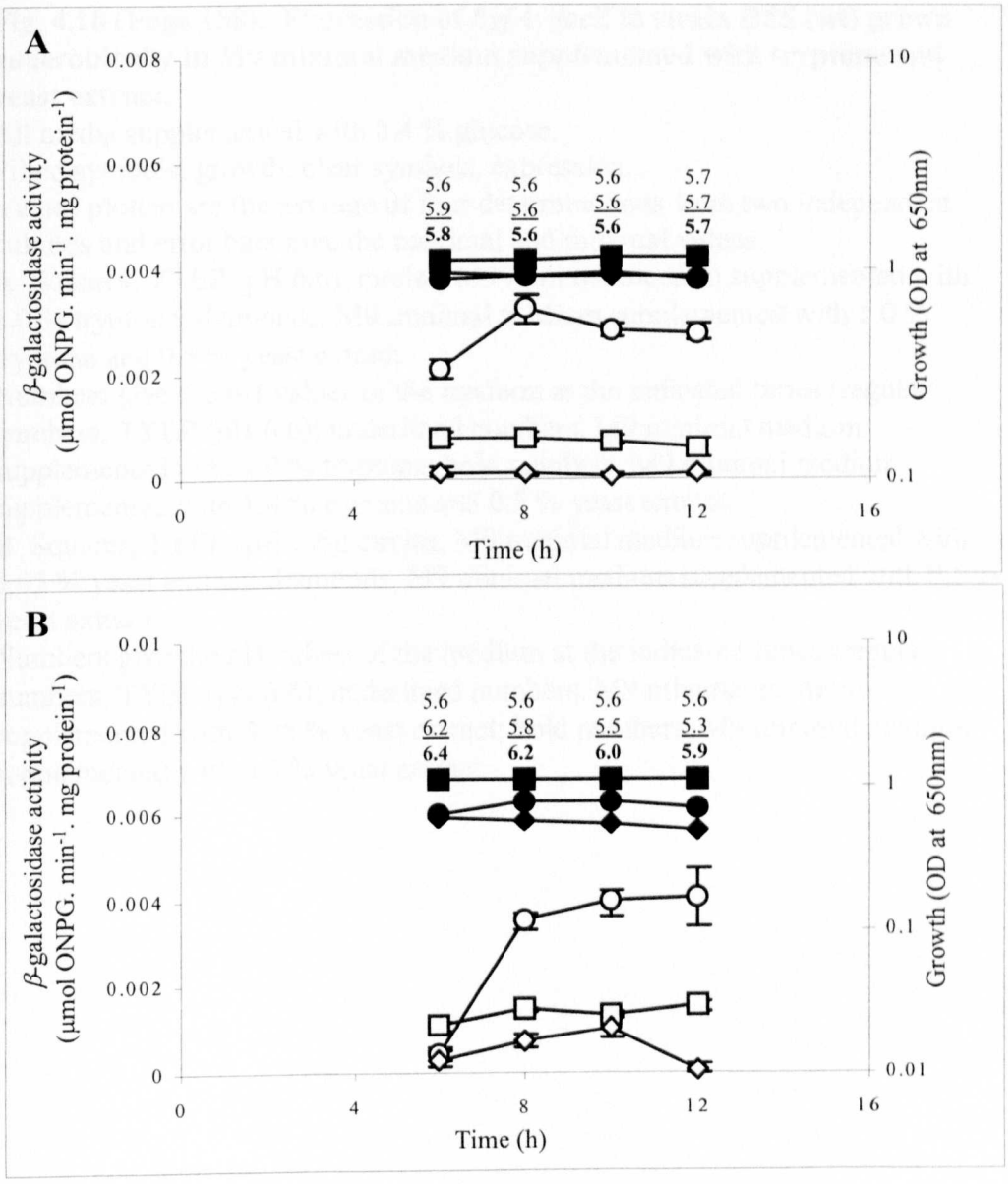

A
0.008
0.006
0.004
0.002
0
β-galactosidase activity
(µmol ONPG. min-1. mg protein-1)
10
1
0.1
Growth (OD at 650nm)
5.6 5.6 5.6 5.7
5.9 5.6 5.6 5.7
5.8 5.6 5.6 5.7
0
4
8
12
16
Time (h)
B
0.01
0.008
0.006
0.004
0.002
0
β-galactosidase activity
(µmol ONPG. min-1. mg protein-1)
10
1
0.1
0.01
Growth (OD at 650nm)
5.6 5.6 5.6 5.6
6.2 5.8 5.5 5.3
6.4 6.2 6.0 5.9
0
4
8
12
16
Time (h)

Fig. 4.10 (Page 158). Expression of *hyfA*::*lacZ* in strain DS5 (wt) grown anaerobically in M9 minimal medium supplemented with tryptone and yeast extract.
All media supplemented with 0.4 % glucose.
Filled symbols, growth; clear symbols, expression.
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.
A Squares, TYEP (pH 6.6); circles, M9 minimal medium supplemented with 1.0 % tryptone; diamonds, M9 minimal medium supplemented with 1.0 % tryptone and 0.5 % yeast extract.
Numbers give the pH values of the medium at the indicated times (regular numbers, TYEP (pH 6.6); underlined numbers, M9 minimal medium supplemented with 1.0 % tryptone; bold numbers, M9 minimal medium supplemented with 1.0 % tryptone and 0.5 % yeast extract.
B Squares, TYEP (pH 6.6); circles, M9 minimal medium supplemented with 0.05 % yeast extract; diamonds, M9 minimal medium supplemented with 0.5 % yeast extract.
Numbers give the pH values of the medium at the indicated times (regular numbers, TYEP (pH 6.6); underlined numbers, M9 minimal medium supplemented with 0.05 % yeast extract; bold numbers, M9 minimal medium supplemented with 0.5 % yeast extract.

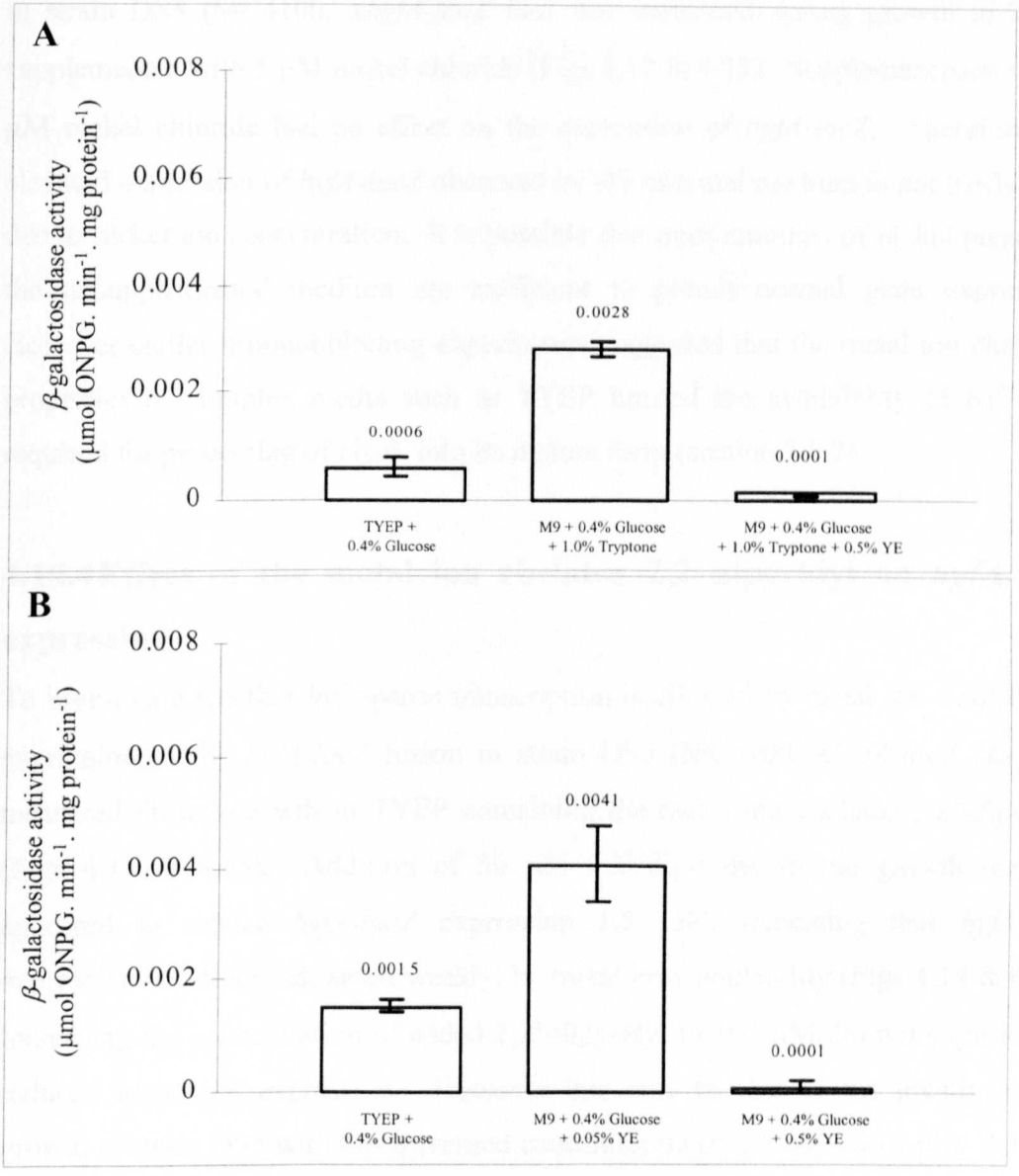


Fig. 4.11. Summary of Fig. 4.10.
Expression of *hyfA*::*lacZ* in strains DS5 (wt) growing anaerobically in M9 minimal medium supplemented with tryptone and yeast extract.
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.
A Summary of **Fig. 4.10A**. Values plotted are taken from stationary phase (12 h).
B Summary of **Fig. 4.10B**. Values plotted are taken from stationary phase (12 h).

investigate the effect of nickel on *hyf* expression, expression of the *hyfA-lacZ* fusion in strain DS5 (MC4100, λ*hyfA-lacZ bla*) was measured during growth in TYEP supplemented with 5 μM nickel chloride (Figs 4.12 & 4.13). Supplementation with 5 μM nickel chloride had no effect on the expression of *hyfA-lacZ*. Therefore, the elevated expression of *hyfA-lacZ* observed in M9 minimal medium is not likely to be due to nickel ion concentration. It is possible that trace amounts of nickel present in the unsupplemented medium are sufficient to permit normal gene expression. However earlier immunoblotting experiments suggested that the metal ion chelating properties of complex media such as TYEP limited the availability of Ni^{2+} ions required for processing of HycE into its mature form (section 3.6.2).

4.10.4Effect of the metal ion chelator 2,2 dipyridyl on *hyfA-lacZ* expression

To investigate whether *hyf* operon transcription is affected by metal ion availability, expression of the *hyfA-lacZ* fusion in strain DS5 (MC4100, λ*hyfA-lacZ bla*) was measured during growth in TYEP containing the metal ion chelator 2,2'-dipyridyl (Figs 4.14 & 4.15). Addition of 50 μM 2,2'-dipyridyl to the growth medium appeared to reduce *hyfA-lacZ* expression 1.5 fold, indicating that *hyfA-lacZ* expression is influenced, albeit weakly, by metal iron availability (Figs 4.14 & 4.15). Increasing the concentration of added 2,2'-dipyridyl to 100 μM did not significantly reduced *hyfA-lacZ* expression. However this may be due to the greatly altered growth of strain DS5 with this increased concentration of 2,2'-dipyridyl (Fig. 4.14).

4.10.5Study of *hyf* expression on agar plates

The effect of 0.4% w/v glucose, 50 mM sorbitol, 0.4% w/v ethanol, 50 mM formate, 50mM gluconate, 50 mM glucuronate and 50 mM acetate on *hyfA-lacZ* expression in strains DS5 (wt), DS6 (Δ*hyfR*) and DS7 (*fhlA*) was studied from the appearance of colonies after overnight aerobic growth on M9 minimal agar plates (section 2.5.2.2) containing one of the above compounds and 40 $\mu M.ml^{-1}$ X-Gal. It should be noted that strains were grown aerobically. These compounds were chosen randomly to screen for the effector(s) used by HyfR (proposed σ^{54}-dependent transcriptional

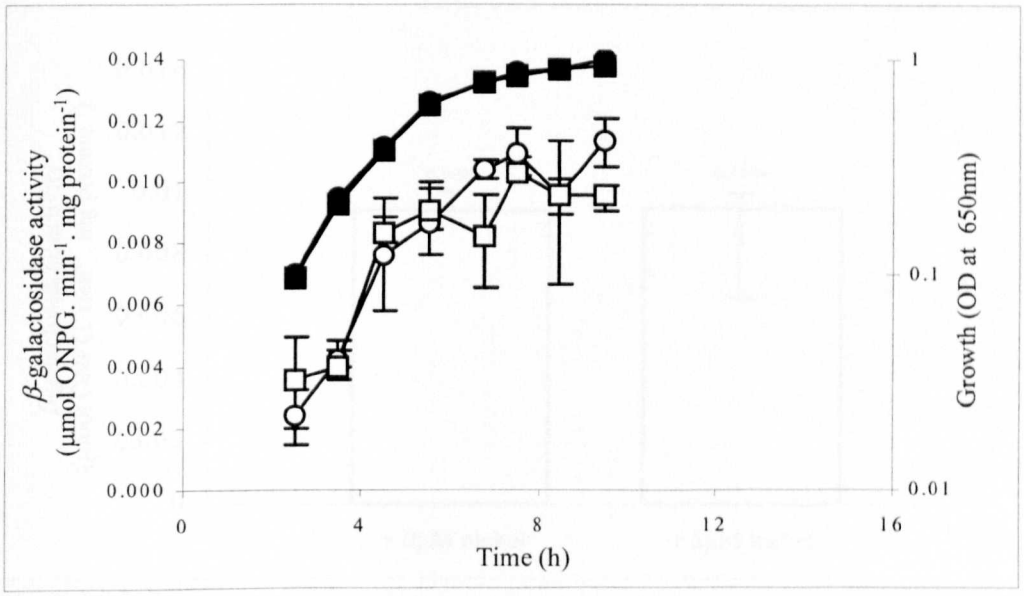


Fig. 4.12. Expression of *hyfA*::*lacZ* in strain DS5 (wt) grown anaerobically in TYEP (pH 6.6) supplemented with nickel chloride.
All media supplemented with 0.4 % glucose.
Squares, 0 μM nickel chloride; circles, 5 μM nickel chloride.
Filled symbols, growth; clear symbols, expression.
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

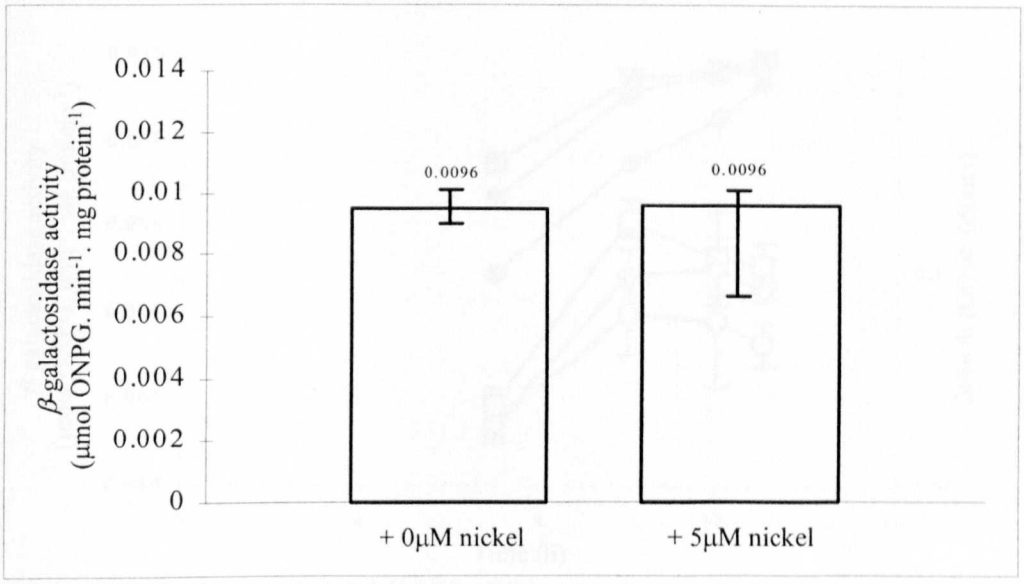


Fig. 4.13. Summary of Fig. 4.12.
Expression of *hyfA*::*lacZ* in strains DS5 (wt) growing anaerobically in TYEP (pH 6.6) supplemented with nickel chloride.
Values plotted are taken from stationary phase (8.5 h).
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

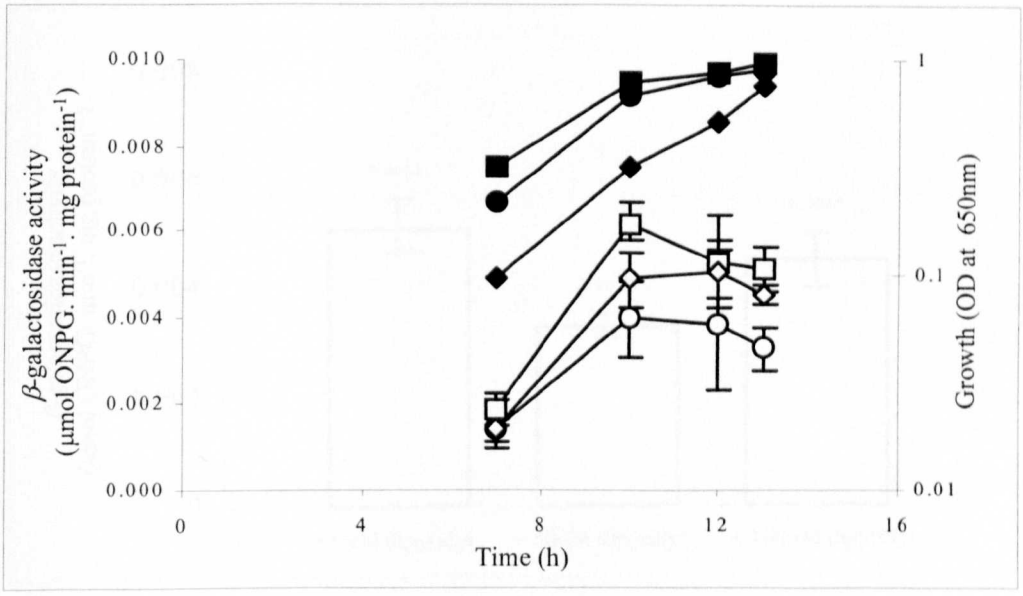


Fig. 4.14. Expression of *hyfA*::*lacZ* in strain DS5 (wt) grown anaerobically in TYEP (pH 6.6) supplemented with the iron chelator 2,2 dipyridyl.
All media supplemented with 0.4 % glucose.
Squares, 0 μM 2,2 dipyridyl; circles, 50 μM 2,2 dipyridyl; diamonds, 100 μM 2,2 dipyridyl.
Filled symbols, growth; clear symbols, expression.
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

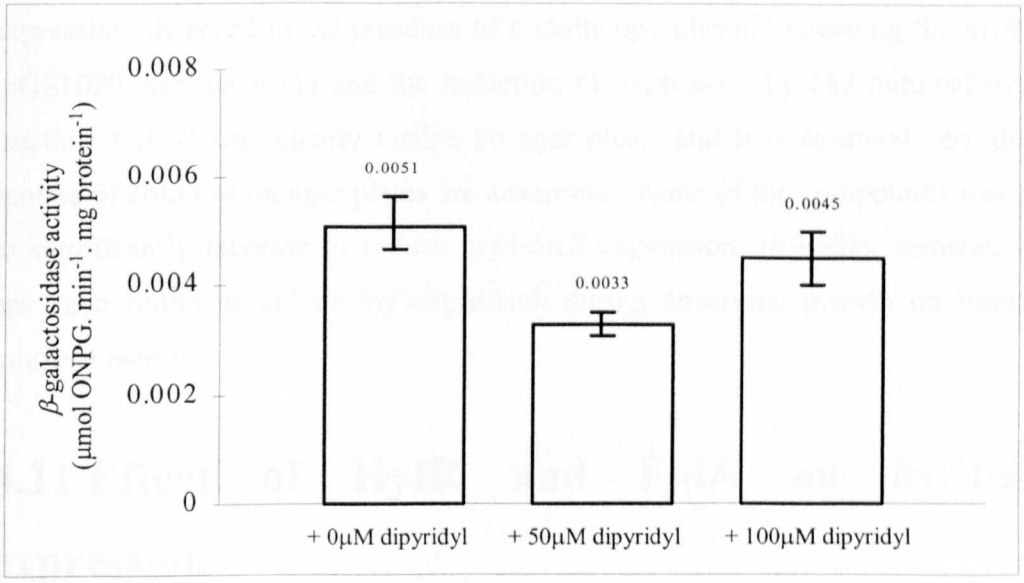


Fig. 4.15. Summary of Fig. 4.14.
Expression of *hyfA*::*lacZ* in strains DS5 (wt) growing anaerobically in TYEP (pH 6.6) supplemented with the iron chelator 2,2 dipyridyl.
Values plotted are taken from stationary phase (13 h).
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

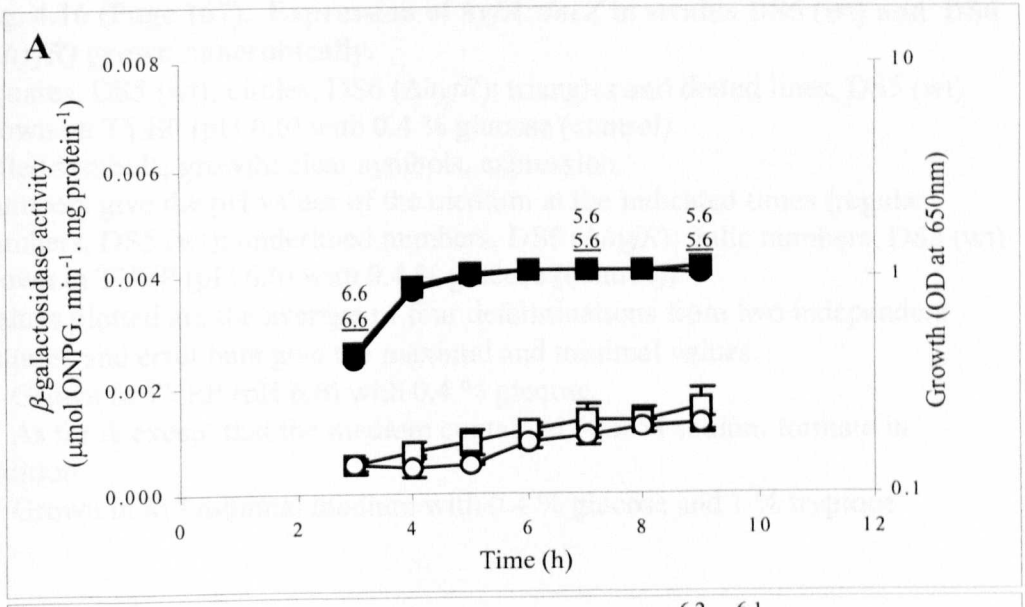
activator of the formate regulon). However both the >1000 fold increase in *hyf* expression observed in the presence of a multicopy plasmid encoding the *hyfR* gene (pGS1020; section 4.11) and the induction of expression by M9 minimal medium (section 4.10.2) was clearly visible on agar plates and it is assumed here that the centres of colonies on agar plates are anaerobic. None of the compounds was found to significantly increase or reduce *hyfA-lacZ* expression, including formate, which has been found to induce *hyf* expression during anaerobic growth on liquid M9 minimal medium.

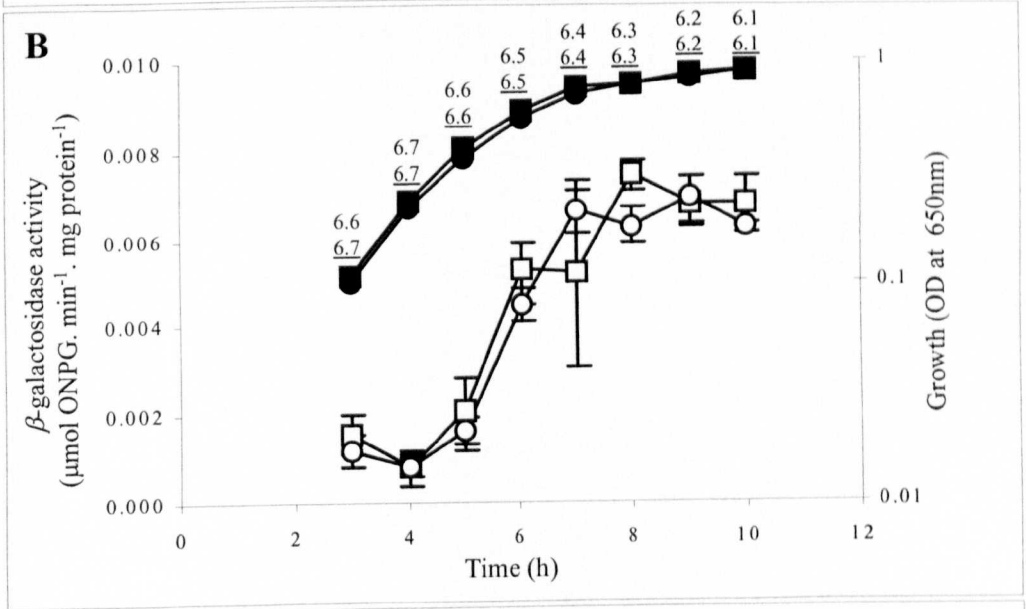
4.11 Effect of HyfR and FhlA on *hyfA-lacZ* expression

The HyfR protein is closely related to the FhlA protein of *E. coli*, a member of a group of σ^{54}-dependent transcriptional regulators (Andrews *et al.*, 1997). FhlA induces expression of the *hyc, fdhF, hydN-hypF* and *hyp* genes (the 'formate regulon') in response to elevated intracellular formate levels. A potential σ^{54}-dependent promoter and a potential FhlA and/or HyfR binding site have been detected, 12 bp and 124 bp upstream of the *hyf* coding regions respectively (Andrews *et al.*, 1997). Andrews and co-workers (1997) suggested that HyfR regulates transcription of the *hyf* operon (and other genes) in response to formate.

Mutations in *hyfR* and *fhlA* were introduced into strain DS5 (MC4100, λ*hyfA-lacZ bla*) to create strains DS6 (MC4100, λ*hyfA-lacZ bla*, Δ*hyfR*::*spc*) and DS7 (MC4100, λ*hyfA-lacZ bla*, *fhlA*::λ*plac*Mu53 *kan*) respectively (sections 4.3 and 4.4 respectively). The effect of these mutations on *hyfA-lacZ* expression was studied (Figs 4.16, 4.17, 4.18, 4.19, 4.20, 4.21, 4.22 & 4.23).

The absence of the *hyfR* gene in strain DS6 appeared to have no significant effect on *hyfA-lacZ* expression (Figs 4.16 & 4.19A). However, introduction of pGS1087, a multicopy plasmid encoding the *hyfR* gene, enhanced *hyfA-lacZ* expression >1000 fold (Figs 4.17A & 4.19B). Addition of sodium formate to the growth medium was found to have little effect on induction of *hydA-lacZ* expression by multicopy HyfR (Figs 4.17B, 4.17C & 4.19B). Note that control experiments showed that the vector, pSU18, does not significantly affect *hyfA-lacZ* expression (Fig. 4.17B). Further work is required to identify any effector molecule specific to


A
β-galactosidase activity (µmol ONPG. min-1. mg protein-1)
Growth (OD at 650nm)
Time (h)
6.6
6.6
5.6
5.6
5.6
5.6


B
β-galactosidase activity (µmol ONPG. min-1. mg protein-1)
Growth (OD at 650nm)
Time (h)
6.6
6.7
6.7
6.7
6.6
6.6
6.5
6.5
6.4
6.4
6.3
6.3
6.2
6.2
6.1
6.1

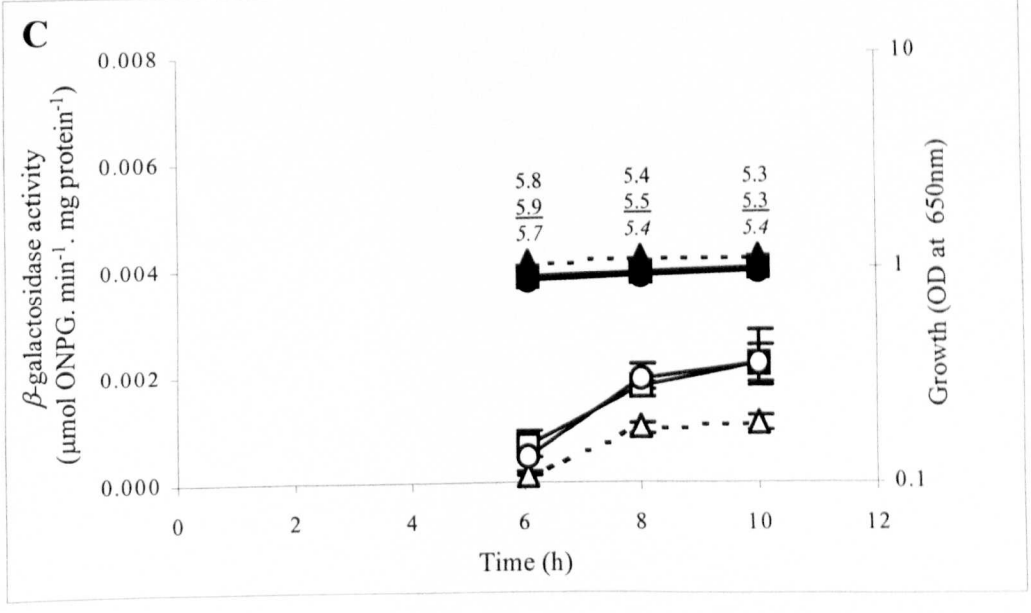

C
β-galactosidase activity (µmol ONPG. min-1. mg protein-1)
Growth (OD at 650nm)
Time (h)
5.8
5.9
5.7
5.4
5.5
5.4
5.3
5.3
5.4

Fig. 4.16 (Page 167). Expression of *hyfA*::*lacZ* in strains DS5 (wt) and DS6 (Δ*hyfR*) grown anaerobically.
Squares, DS5 (wt); circles, DS6 (Δ*hyfR*); triangles and dotted lines, DS5 (wt) grown on TYEP (pH 6.6) with 0.4 % glucose (control).
Filled symbols, growth; clear symbols, expression.
Numbers give the pH values of the medium at the indicated times (regular numbers, DS5 (wt); underlined numbers, DS6 (Δ*hyfR*); italic numbers, DS5 (wt) grown in TYEP (pH 6.6) with 0.4 % glucose (control)).
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.
A Grown in TYEP (pH 6.6) with 0.4 % glucose.
B As for **A** except that the medium contained 50 mM sodium formate in addition.
C Grown in M9 minimal medium with 0.4 % glucose and 1 % tryptone.

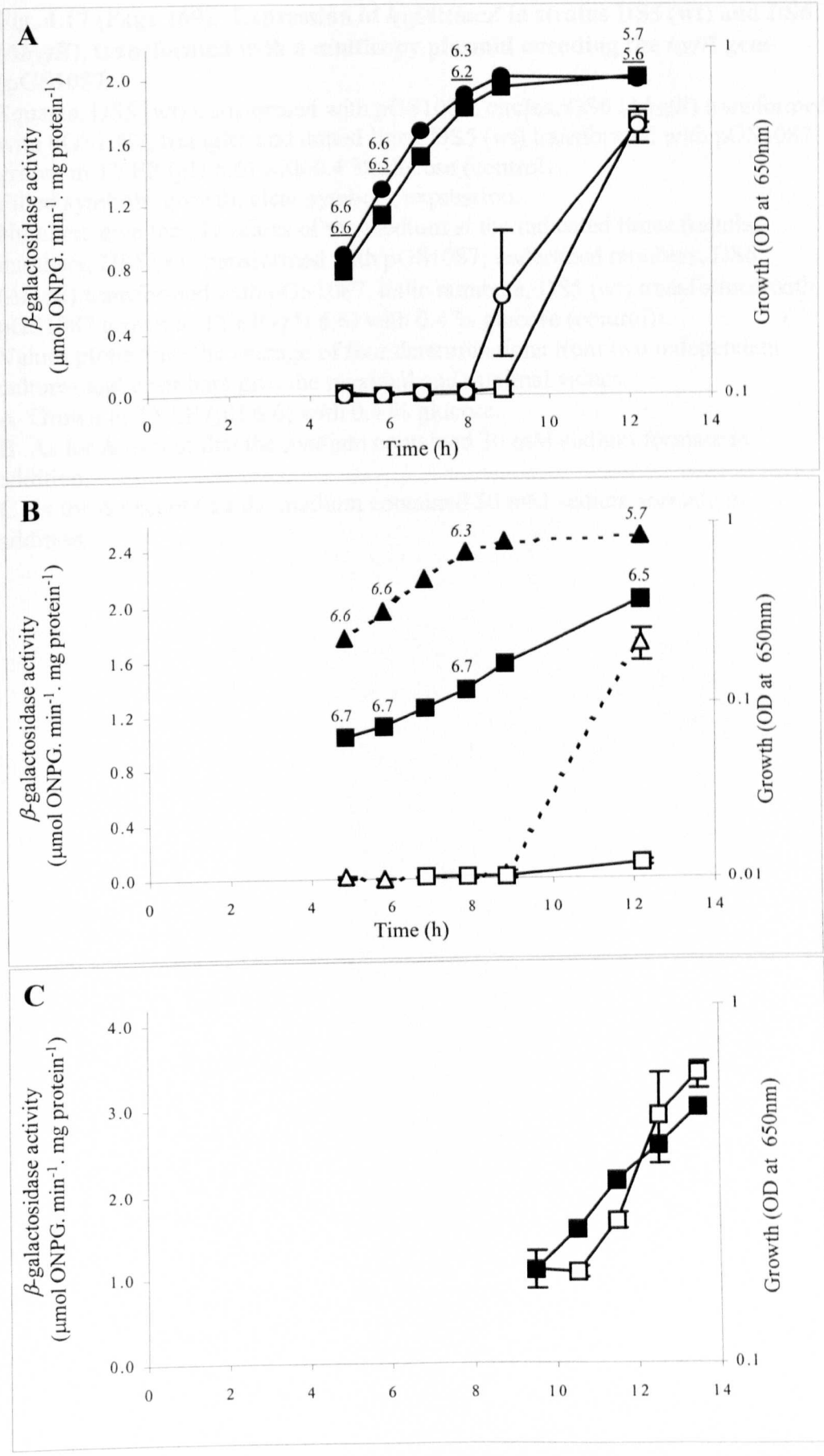

A
β-galactosidase activity (µmol ONPG. min-1. mg protein-1)
Growth (OD at 650nm)
Time (h)
6.6
6.6
6.6
6.5
6.3
6.2
5.7
5.6
B
β-galactosidase activity (µmol ONPG. min-1. mg protein-1)
Growth (OD at 650nm)
Time (h)
6.6
6.6
6.3
5.7
6.7
6.7
6.5
C
β-galactosidase activity (µmol ONPG. min-1. mg protein-1)
Growth (OD at 650nm)

Fig. 4.17 (Page 169). Expression of *hyfA*::*lacZ* in strains DS5 (wt) and DS6 (Δ*hyfR*), transformed with a multicopy plasmid encoding the *hyfR* gene (pGS1087).

Squares, DS5 (wt) transformed with pGS1087; circles, DS6 (Δ*hyfR*) transformed with pGS1087; triangles and dotted lines, DS5 (wt) transformed with pGS1087 grown in TYEP (pH 6.6) with 0.4 % glucose (control).

Filled symbols, growth; clear symbols, expression.

Numbers give the pH values of the medium at the indicated times (regular numbers, DS5 (wt) transformed with pGS1087; underlined numbers, DS6 (Δ*hyfR*) transformed with pGS1087; italic numbers, DS5 (wt) transformed with pGS1087 grown in TYEP (pH 6.6) with 0.4 % glucose (control)).

Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

A Grown in TYEP (pH 6.6) with 0.4 % glucose.

B As for **A** except that the medium contained 30 mM sodium formate in addition.

C As for **A** except that the medium contained 50 mM sodium formate in addition.

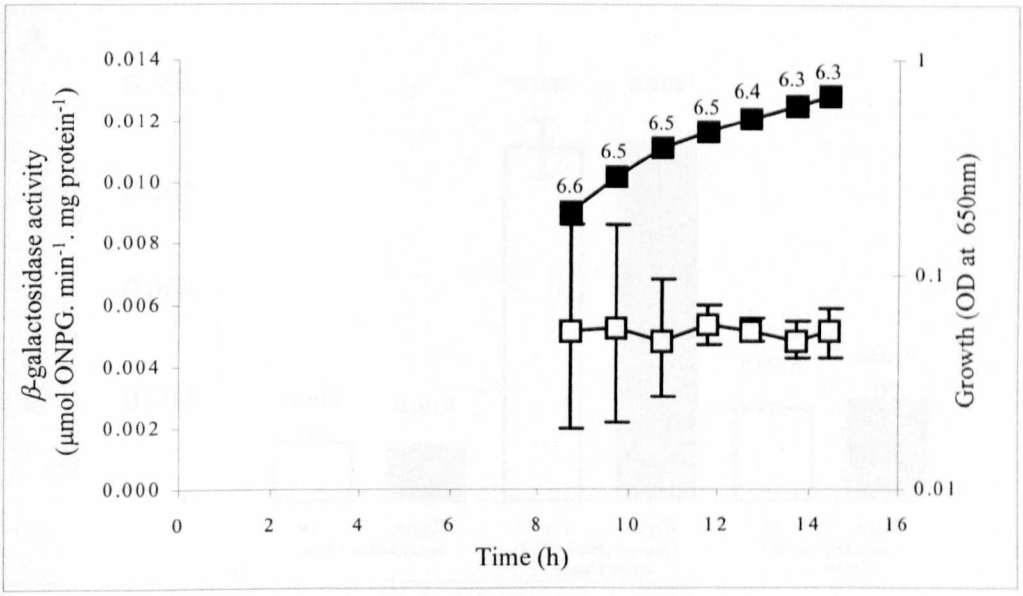


Fig. 4.18. Expression of *hyfA*::*lacZ* in strain DS5 (wt), transformed with a multicopy plasmid encoding the *hyfA-focB* genes (pGS1020).

Strains were grown anaerobically in TYEP (pH 6.6) with 0.4 % glucose and 50 mM sodium formate.

Squares, DS5 (wt) transformed with pGS1020.

Filled symbols, growth; clear symbols, expression.

Numbers give the pH values of the medium at the indicated times.

Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

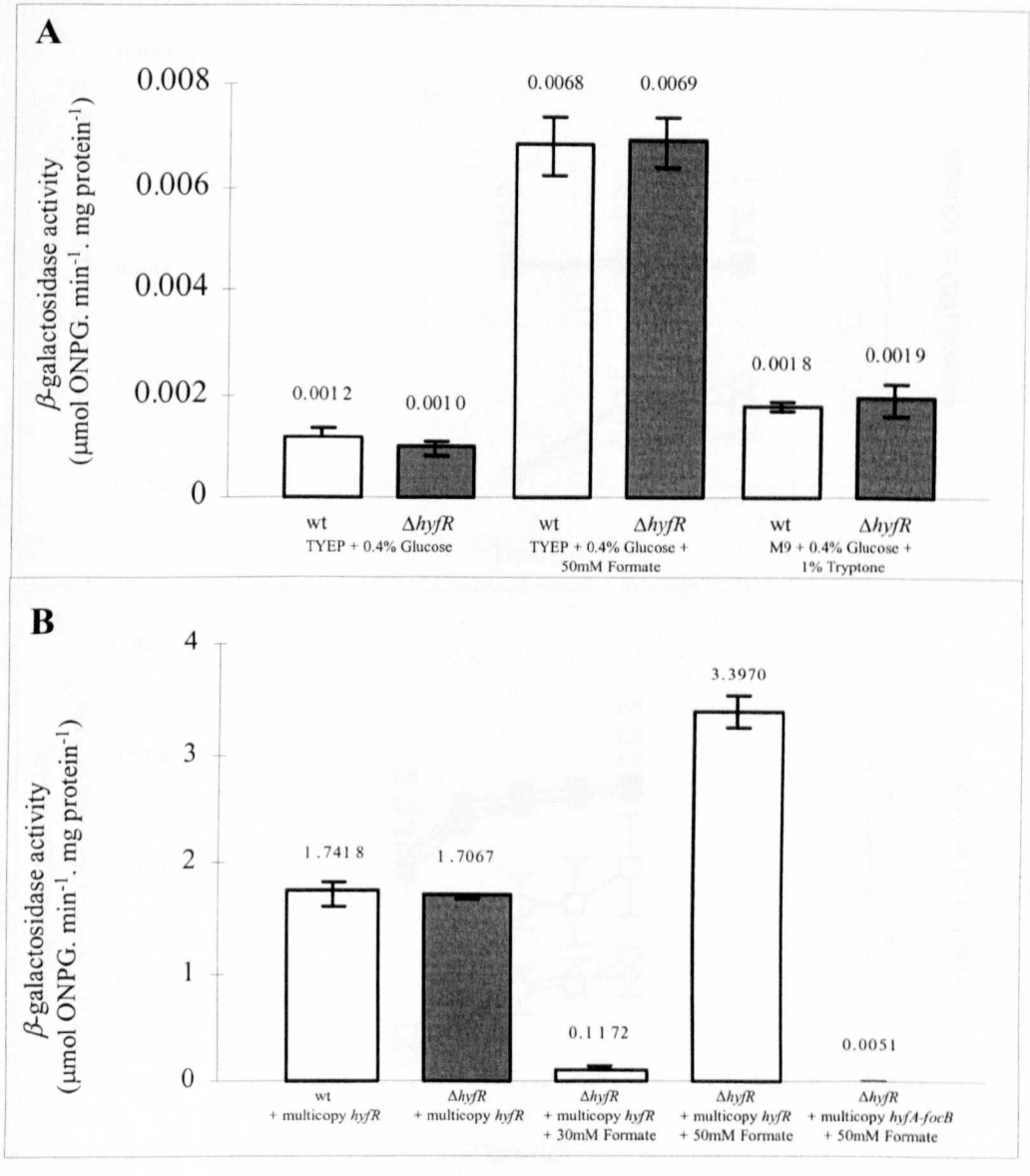


Fig. 4.19. Summary of Figs 4.16, 4.17 and 4.18.

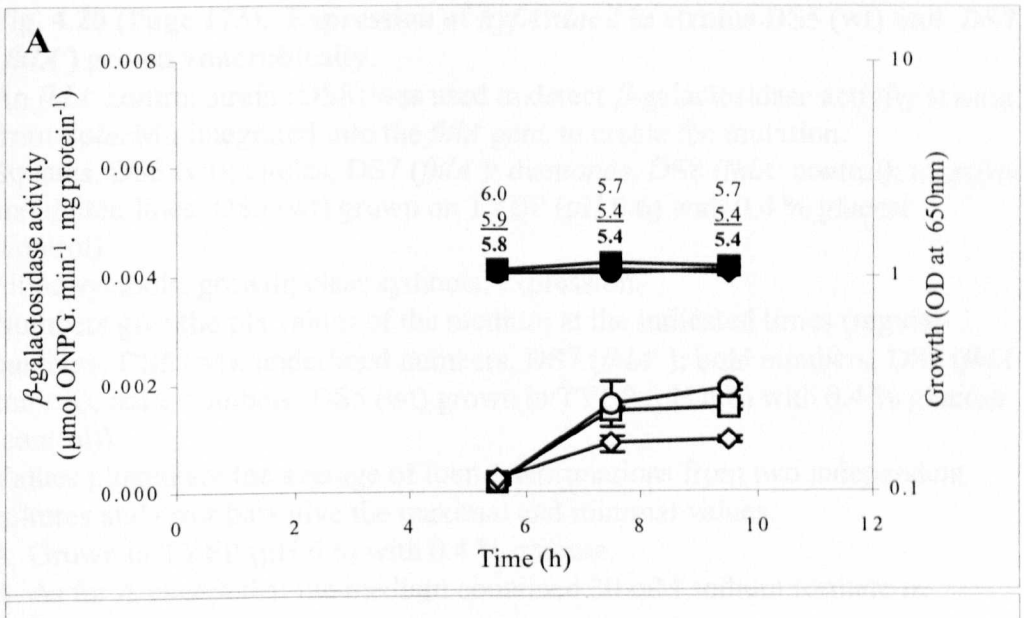
Expression of *hyfA::lacZ* in strains DS5 and DS6 (Δ*hyfR*) growing anaerobically at 37 °C.

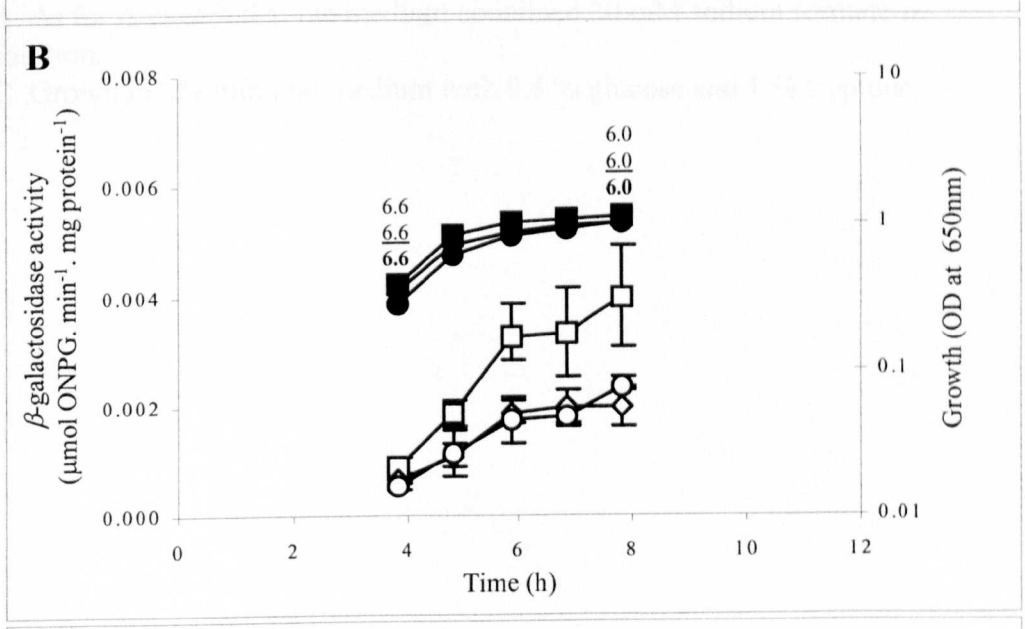
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

A Summary of **Fig. 4. 16**. Values plotted are taken from stationary phase: **Fig. 4.16A**, 6 h; **Fig. 4.16B**, 9 h; **Fig. 4.16C**, 8 h.

B Summary of **Figs 4.17 and 4.18**. Values plotted are taken from stationary phase (where possible): **Fig. 4.17A**, 12.2 h; **Fig. 4.17B**, 12.2 h; **Fig. 4.17C**, 13.5 h; **Fig. 4.18**, 14.5 h.

Grown in TYEP (pH6.6) with 0.4% glucose.


A
β-galactosidase activity (µmol ONPG. min-1. mg protein-1)
Growth (OD at 650nm)
Time (h)
6.0
5.9
5.8
5.7
5.4
5.4
5.7
5.4
5.4


B
β-galactosidase activity (µmol ONPG. min-1. mg protein-1)
Growth (OD at 650nm)
Time (h)
6.6
6.6
6.6
6.0
6.0
6.0

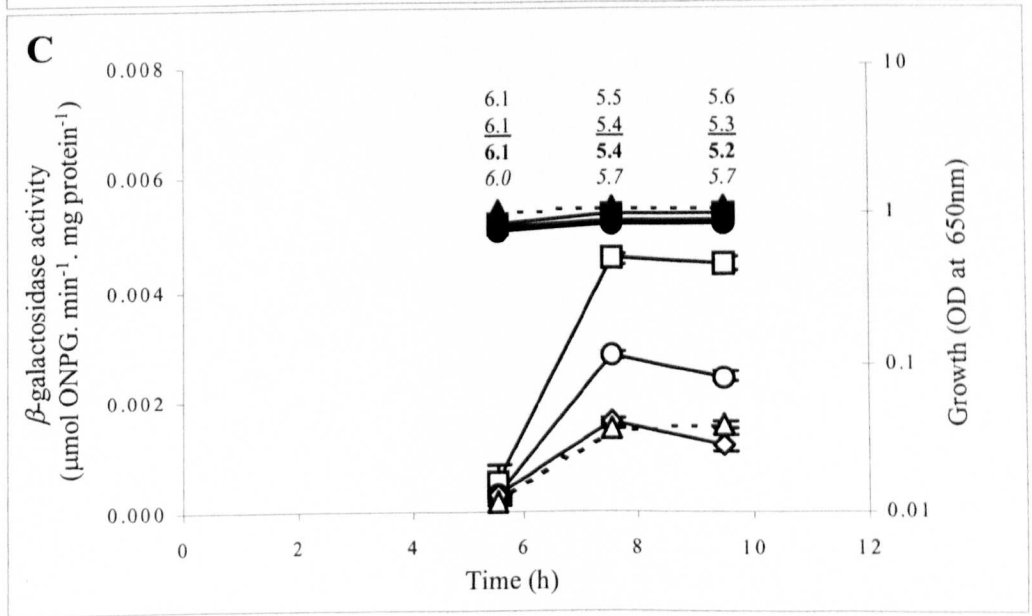

C
β-galactosidase activity (µmol ONPG. min-1. mg protein-1)
Growth (OD at 650nm)
Time (h)
6.1
6.1
6.1
6.0
5.5
5.4
5.4
5.7
5.6
5.3
5.2
5.7

Fig. 4.20 (Page 173). Expression of *hyfA*::*lacZ* in strains DS5 (wt) and DS7 (*fhlA*$^-$) grown anaerobically.

An *fhlA*$^-$ control strain (DS8) was used to detect β-galactosidase activity arising from λp*lac*Mu integrated into the *fhlA* gene to create the mutation.

Squares, DS5 (wt); circles, DS7 (*fhlA*$^-$); diamonds, DS8 (*fhlA*$^-$ control); triangles and dotted lines, DS5 (wt) grown on TYEP (pH 6.6) with 0.4 % glucose (control).

Filled symbols, growth; clear symbols, expression.

Numbers give the pH values of the medium at the indicated times (regular numbers, DS5 (wt); underlined numbers, DS7 (*fhlA*$^-$); bold numbers, DS8 (*fhlA*$^-$ control); italic numbers, DS5 (wt) grown in TYEP (pH 6.6) with 0.4 % glucose (control)).

Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

A Grown in TYEP (pH 6.6) with 0.4 % glucose.

B As for **A** except that the medium contained 30 mM sodium formate in addition.

C Grown in M9 minimal medium with 0.4 % glucose and 1 % tryptone.

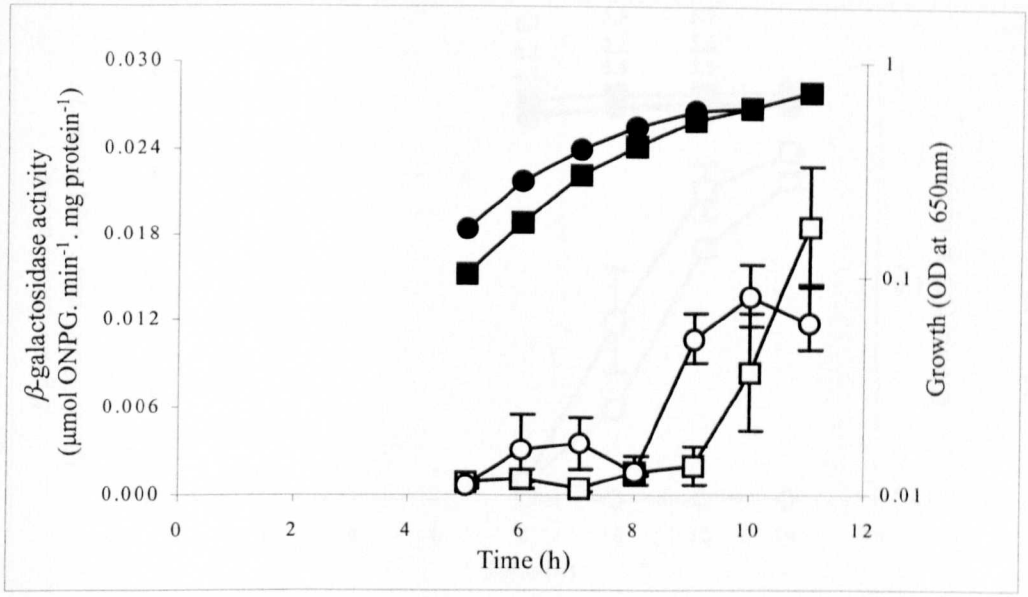


Fig. 4.21. Expression of *hyfA*::*lacZ* in strains DS5 (wt) and DS7 (*fhlA*$^-$), both transformed with a multicopy plasmid encoding the *fhlA* gene (pSH9).
Strains were grown anaerobically in TYEP (pH 6.6) with 0.4 % glucose and 50 mM formate.
Squares, DS5 (wt) transformed with pSH9; circles, DS7 (*fhlA*$^-$) transformed with pSH9.
Filled symbols, growth; clear symbols, expression.
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

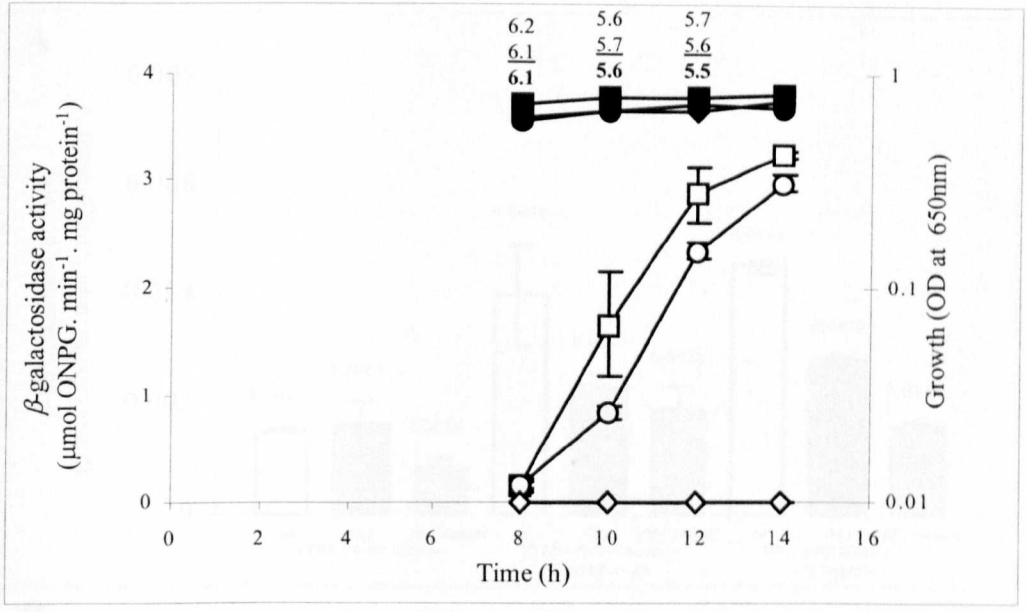


Fig. 4.22. Expression of *hyfA*::*lacZ* in strains DS5 (wt) and DS7 (*fhlA*$^-$), both transformed with a multicopy plasmid encoding the *hyfR* gene (pGS1087).

Strains were grown anaerobically in TYEP (pH 6.6) with 0.4 % glucose. An *fhlA*$^-$ control strain (DS8) was used to detect β-galactosidase activity arising from λp*lac*Mu integrated into the *fhlA* gene to create the mutation.
Squares, DS5 (wt) transformed with pGS1087; circles, DS7 (*fhlA*$^-$) transformed with pGS1087; diamonds, DS8 (*fhlA*$^-$ control) transformed with pGS1087.
Filled symbols, growth; clear symbols, expression.
Numbers give the pH values of the medium at the indicated times (regular numbers, DS5 (wt); underlined numbers, DS7 (*fhlA*$^-$); bold numbers, DS8 (*fhlA*$^-$ control)).
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

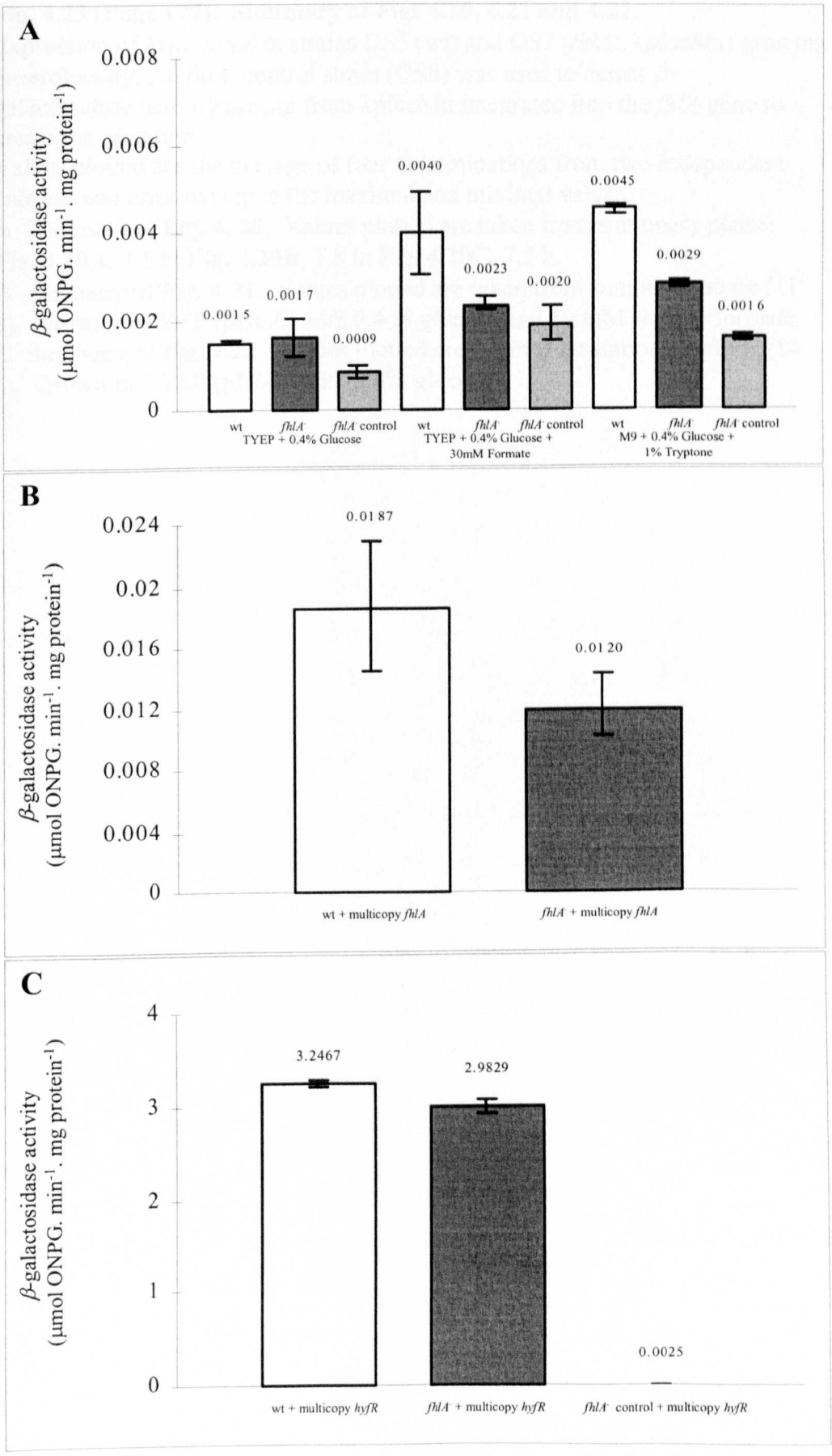

A
β-galactosidase activity
(µmol ONPG. min-1. mg protein-1)
0.008
0.006
0.004
0.002
0
0.0015
0.0017
0.0009
0.0040
0.0023
0.0020
0.0045
0.0029
0.0016
wt
fhlA-
fhlA- control
TYEP + 0.4% Glucose
TYEP + 0.4% Glucose + 30mM Formate
M9 + 0.4% Glucose + 1% Tryptone
B
β-galactosidase activity
(µmol ONPG. min-1. mg protein-1)
0.024
0.02
0.016
0.012
0.008
0.004
0
0.0187
0.0120
wt + multicopy fhlA
fhlA- + multicopy fhlA
C
β-galactosidase activity
(µmol ONPG. min-1. mg protein-1)
4
3
2
1
0
3.2467
2.9829
0.0025
wt + multicopy hyfR
fhlA- + multicopy hyfR
fhlA- control + multicopy hyfR

Fig. 4.23 (Page 177). Summary of Figs 4.20, 4.21 and 4.22.
Expression of *hyfA::lacZ* in strains DS5 (wt) and DS7 (*fhlA::λplacMu*) growing anaerobically. An *fhlA*$^-$ control strain (DS8) was used to detect *β*-galactosidase activity arising from λp*lac*Mu integrated into the *fhlA* gene to create the mutation.
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.
A Summary of **Fig. 4. 20**. Values plotted are taken from stationary phase: **Fig. 4.20A**, 7.5 h; **Fig. 4.20B**, 7.8 h; **Fig. 4.20C**, 7.5 h.
B Summary of **Fig. 4.21**. Values plotted are taken from stationary phase (11 h). Grown in TYEP (pH6.6) with 0.4 % glucose and 50 mM sodium formate.
C Summary of **Fig. 4.22**. Values plotted are taken from stationary phase (14 h). Grown in TYEP (pH6.6) with 0.4% glucose.

HyfR and to investigate whether HyfR, like FhlA, is activated by formate. Introduction of pGS1020, a multicopy plasmid encoding the entire *hyf* operon (*hyfA-focB*) surprisingly had no effect on *hyfA-lacZ* expression despite also encoding the *hyfR* gene (Figs 4.18 & 4.19B). Medium scale plasmid preparations were carried out to purify the pGS1020 plasmid from the DS5 transformants used in this experiment. The purified plasmid still appeared to carry and had not lost the large 15 kb *Eco*RI fragment containing the whole of the *hyf* operon (section 3.3). It therefore appears that the levels of HyfR in cells transformed with pGS1020 are not increased sufficient to affect expression of *hyfA-lacZ*. The above results suggest that the *hyf* operon is strongly induced by HyfR when HyfR levels are increased by supplying a multicopy plasmid expressing *hyfR*. However failure of the *hyfR* chromosomal mutation to affect *hyfA-lacZ* expression suggests that HyfR levels are too low to influence *hyf* expression when the *hyfR* gene is present as a single-copy gene within the *hyf* operon, at least when expression is measured under the fermentative growth conditions employed in this study. It should be noted that *hyfR* and *hyfA-focB* in plasmids pGS1087 and pGS1020 respectively, are under the control of the *lac* promoter.

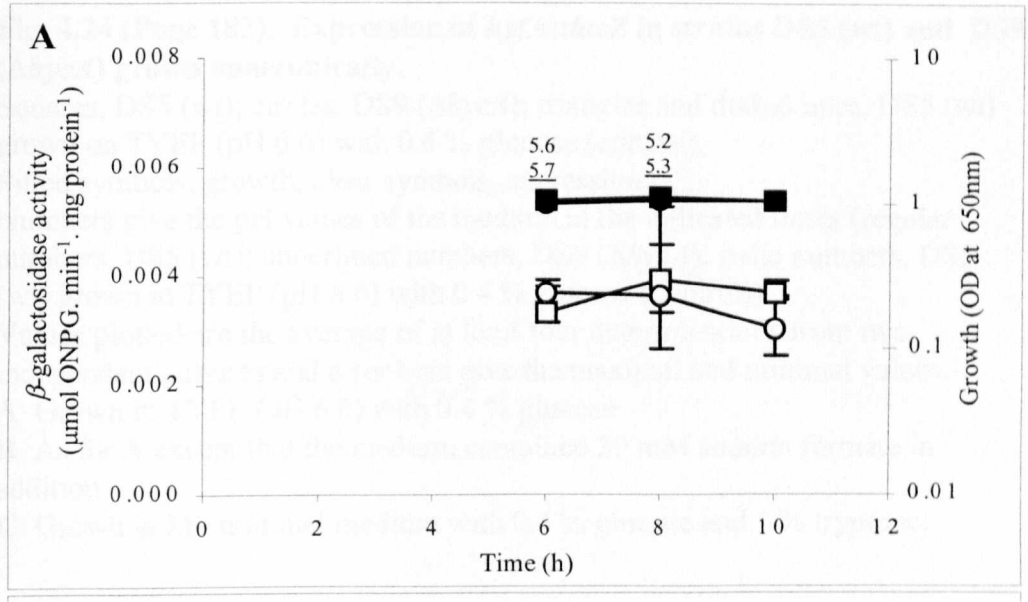
The presence of the *fhlA* mutation in strain DS7 reduced *hyfA-lacZ* expression during growth in TYEP (pH 6.6) with glucose, despite the detection of increased levels of β-galactosidase activity in this strain over those found in strain DS5 (Figs 4.20 & 4.23). This is because part of the activity detected in strain DS7 is contributed by the *fhlA-lacZ* fusion (generated by the integration of λp*lac*Mu into the *fhlA* gene in order create the *fhlA* mutation; Schlensog *et al.*, 1989). The β-galactosidase activity produced by the *fhlA-lacZ* fusion was determined by analysis of the strain DS8 (MC4100, *fhlA*::λp*lac*Mu53 *kan*). This amount can then be subtracted from the activity measured in DS7 to give the amount of β-galactosidase activity contributed by the *hyfA-lacZ* fusion.

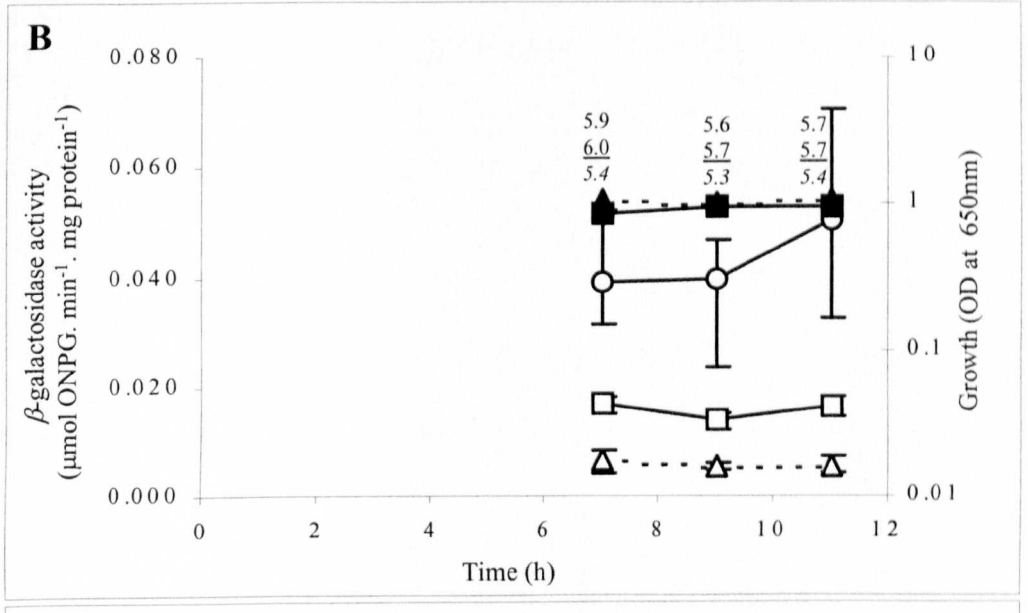
The results show that induction of expression by formate is lost in the *fhlA* mutant (Figs 4.20B & 4.23A). Also, the inductive effect of minimal medium plus 1% tryptone was greatly reduced in the *fhlA* mutant suggesting that the increased expression normally observed in this medium is due to activation of FhlA by formate generated during fermentation (Figs 4.20C & 4.23A). However a recent study by

Self and co-workers (2000) found evidence for the activation of *fdhF* and *hyc* transcription by molybdate though the action FhlA. Therefore the inductive effect of M9 minimal medium on *hyfA*-lacZ expression could also be due to increased molybdate availability. The presence of the *fhlA* gene on a multicopy plasmid increased *hyfA-lacZ* expression in strain DS7 approximately 5 fold above wildtype levels (Figs 4.21 & 4.23B). The summary figure, Fig. 4.23B, suggests that *hyfA-lacZ* expression is not fully restored however Fig. 4.21 shows that expression levels between the wildtype (DS5) and *fhlA* mutant (DS7) transformed with multicopy *fhlA* are of similar levels. The presence of the multicopy plasmid encoding the *hyfR* gene enhanced the levels of β-galactosidase activity detected >1000 fold in DS7 (MC4100, λ*hyfA-lacZ bla*, *fhlA*::λ*plac*Mu53 *kan*) (Figs 4.22 & 4.23C). Introduction of this plasmid into strain DS8 (MC4100, *fhlA*::λ*plac*Mu53 *kan*) did not have this effect on β-galactosidase activity levels, indicating that the increase in activity detected in strain DS7 transformed with the multicopy plasmid encoding *hyfR* is solely due to increased expression of the *hyfA-lacZ* fusion.

4.12 Effect of HycA on *hyfA-lacZ* expression

The product of the first gene of the *hyc* operon, *hycA*, is an anti-activator of formate regulon transcription, inhibiting FhlA by direct protein-protein interaction (S. C. Andrews, personal communication). Sauter and co-workers (1992) found that mutations in *hycA* led to an increase in Fdh-H and total hydrogenase activity, whereas overexpression of *hycA* from a multicopy plasmid nearly abolished Fdh-H and total hydrogenase activity. In order to study the effect of HycA on *hyf* operon expression, the *hyfA-lacZ* fusion, was introduced into strain HD701 (MC4100, Δ*hycA*) to create strain DS9 (MC4100, λ*hyfA-lacZ bla*, Δ*hycA*) (section 4.6). The absence of the *hycA* gene had no significant effect on expression of *hyfA-lacZ* during growth in TYEP with glucose (Figs 4.24A & 4.25). However, the inductive effect of formate on *hyfA-lacZ* expression was increased 3 fold in the *hycA* mutant (Figs 4.24B & 4.25). Expression in M9 minimal medium plus 1% tryptone was not significantly affected by deletion of the *hycA* gene (Figs 4.24C & 4.25). These results are consistent with FhlA being a transcriptional activator of the *hyf* operon and with HycA acting as an FhlA-antiactivator. In the presence of formate the anti-


A
0.008
0.006
0.004
0.002
0.000
β-galactosidase activity
(µmol ONPG. min-1. mg protein-1)
10
1
0.1
0.01
Growth (OD at 650nm)
5.6
5.7
5.2
5.3
0
2
4
6
8
10
12
Time (h)


B
0.080
0.060
0.040
0.020
0.000
β-galactosidase activity
(µmol ONPG. min-1. mg protein-1)
10
1
0.1
0.01
Growth (OD at 650nm)
5.9
6.0
5.4
5.6
5.7
5.3
5.7
5.7
5.4
0
2
4
6
8
10
12
Time (h)

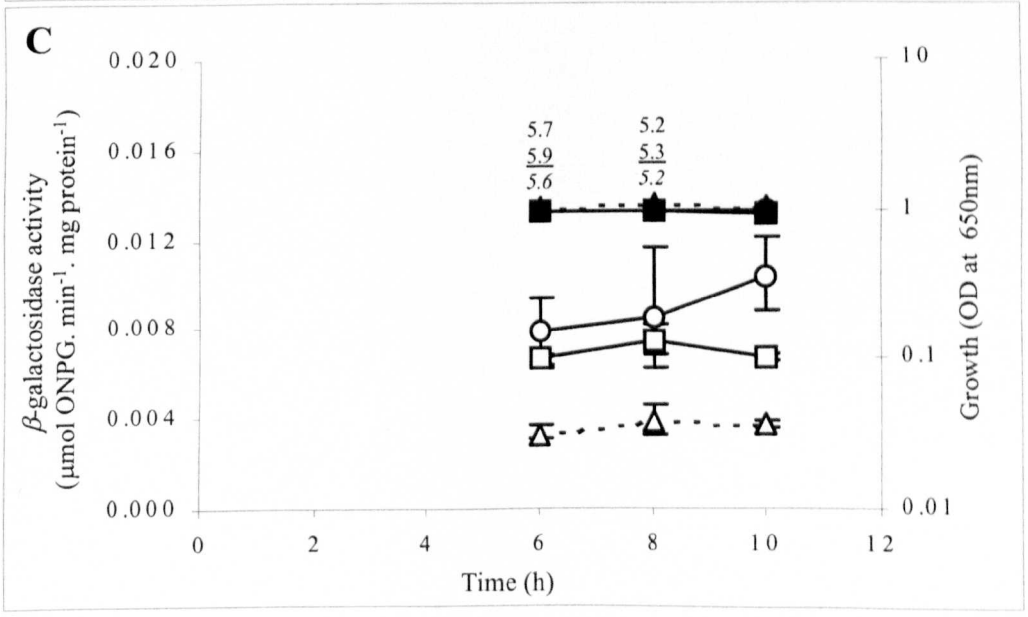

C
0.020
0.016
0.012
0.008
0.004
0.000
β-galactosidase activity
(µmol ONPG. min-1. mg protein-1)
10
1
0.1
0.01
Growth (OD at 650nm)
5.7
5.9
5.6
5.2
5.3
5.2
0
2
4
6
8
10
12
Time (h)

Fig. 4.24 (Page 182). Expression of *hyfA*::*lacZ* in strains DS5 (wt) and DS9 (Δ*hycA*) grown anaerobically.
Squares, DS5 (wt); circles, DS9 (Δ*hycA*); triangles and dotted lines, DS5 (wt) grown on TYEP (pH 6.6) with 0.4 % glucose (control).
Filled symbols, growth; clear symbols, expression.
Numbers give the pH values of the medium at the indicated times (regular numbers, DS5 (wt); underlined numbers, DS9 (Δ*hycA*); italic numbers, DS5 (wt) grown in TYEP (pH 6.6) with 0.4 % glucose (control)).
Values plotted are the average of at least four determinations from two independent cultures and error bars give the maximal and minimal values.
A Grown in TYEP (pH 6.6) with 0.4 % glucose.
B As for **A** except that the medium contained 30 mM sodium formate in addition.
C Grown in M9 minimal medium with 0.4 % glucose and 1 % tryptone.

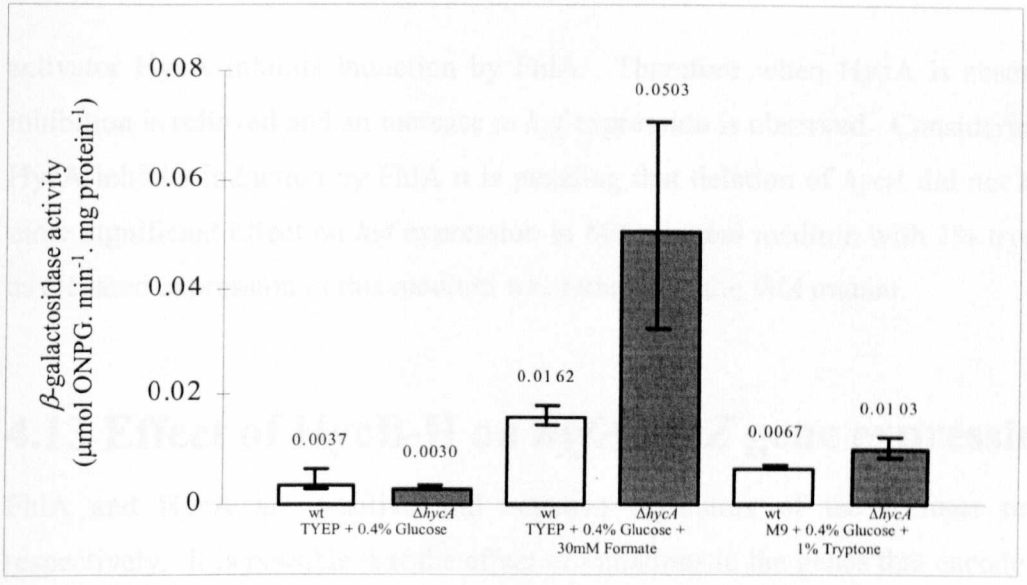


Fig. 4.25. Summary of Fig. 4.24.
Expression of *hyfA::lacZ* in strains DS5 and DS6 (Δ*hycA*) growing anaerobically at 37 °C.
Values plotted are taken from stationary phase: **Fig. 4.24A**, 10 h; **Fig. 4.24B**, 11 h; **Fig. 4.24C**, 10 h.
Values plotted are the average of at least four determinations from two independent cultures and error bars give the maximal and minimal values.

activator HycA inhibits induction by FhlA. Therefore when HycA is absent this inhibition is relieved and an increase in *hyf* expression is observed. Considering that HycA inhibits induction by FhlA it is puzzling that deletion of *hycA* did not have a more significant effect on *hyf* expression in M9 minimal medium with 1% tryptone, as elevated expression in this medium was reduced in the *fhlA* mutant.

4.13 Effect of HycB-H on *hyfA-lacZ* gene expression

FhlA and HycA are positive and negative regulators of the formate regulon respectively. It is possible that the effect of mutations in the genes that encode these proteins on *hyfA-lacZ* expression is an indirect result of their effect on production of the *hyc* encoded Fhl-1 system. A deletion in the *hycB-H* genes was introduced into strain DS5 (MC4100, λ*hyfA-lacZ bla*) to create strain DS10 (MC4100, λ*hyfA-lacZ bla*, Δ*hycB-H*::*cat*) (section 4.7). The absence of the *hycB-H* genes in strain DS10 appeared to have no significant effect on *hyfA-lacZ* expression (Figs 4.26 & 4.27).

4.14 Effect of NtrA on *hyfA-lacZ* gene expression

The *ntrA* gene product, required for the expression of genes involved in nitrogen fixation (*nif*) and regulation (*ntr*), encodes a sigma factor (σ^{54}) and has been shown to be necessary for the expression of the formate regulon (Birkmann *et al.*, 1987). Sequence analysis upstream of the *hyf* operon revealed a potential σ^{54}-dependent promoter located 110 bp downstream of the potential FhlA and HyfR binding sites (Andrews *et al.*, 1997). This promoter has been confirmed by primer extension (section 4.9).

A mutation in *ntrA* was introduced into strain DS5 (MC4100, λ*hyfA-lacZ bla*) to create strain DS11 (MC4100, λ*hyfA-lacZ bla*, Δ(*ntrA208*::Tn*10*)) (section 4.8). The effect of these mutations on *hyfA-lacZ* expression was studied (Figs 4.28, 4.29 & 4.30).

An *ntrA* mutant is a glutamine auxotroph when grown in glucose containing medium. Therefore the media used in these experiments to investigate the effect of NtrA on *hyf* expression was supplemented with 0.2% glutamine. Addition of 0.2%

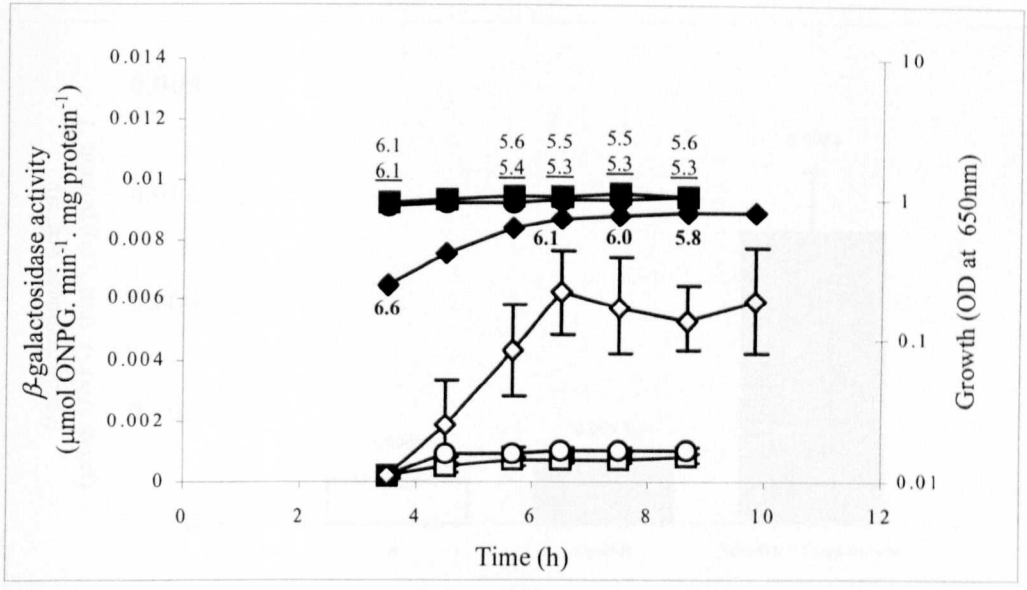


Fig. 4.26. Expression of *hyfA*::*lacZ* in strains DS5 (wt) and DS10 (Δ*hycB-H*) grown anaerobically.

Squares, DS5 (wt) grown in TYEP (pH 6.6) with 0.4 % glucose; circles, DS10 (Δ*hycB-H*) grown in TYEP (pH 6.6) with 0.4 % glucose; diamonds, DS10 (Δ*hycB-H*) grown on TYEP (pH 6.6) with 0.4 % glucose and 30 mM sodium formate.

Filled symbols, growth; clear symbols, expression.

Numbers give the pH values of the medium at the indicated times (regular numbers, DS5 (wt); underlined numbers, DS10 (Δ*hycB-H*); italic numbers, DS5 (wt) grown in TYEP (pH 6.6) with 0.4 % glucose (control)).

Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

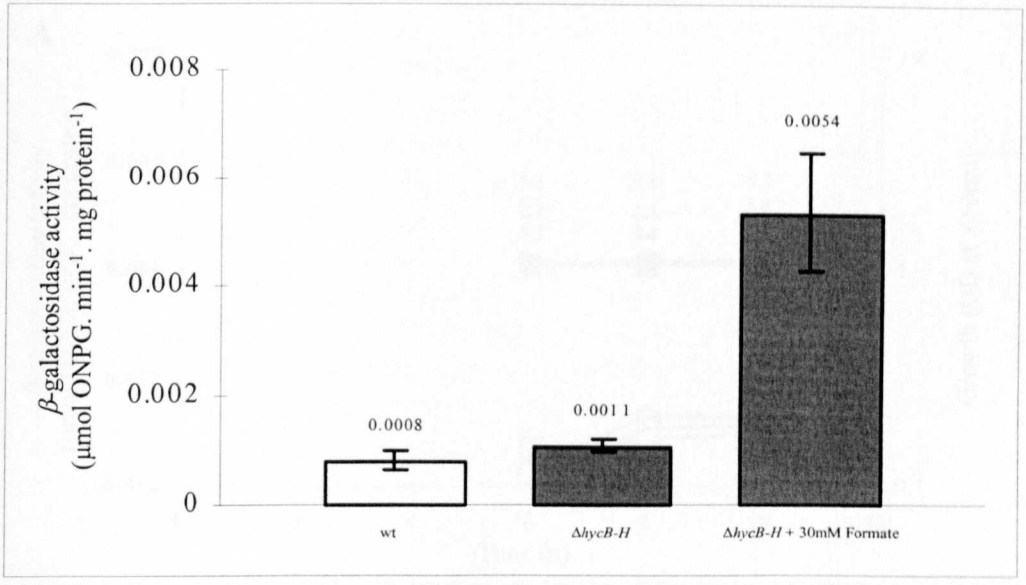


Fig. 4.27. Summary of Fig. 4.26.
Expression of *hyfA::lacZ* in strains DS5 and DS10 (Δ*hycB-H*) growing anaerobically at 37 °C. Grown in TYEP (pH 6.6) + 0.4 % glucose.
Values plotted are taken from stationary phase (8.7 h).
Values plotted are the average of at least four determinations from two independent cultures and error bars give the maximal and minimal values.

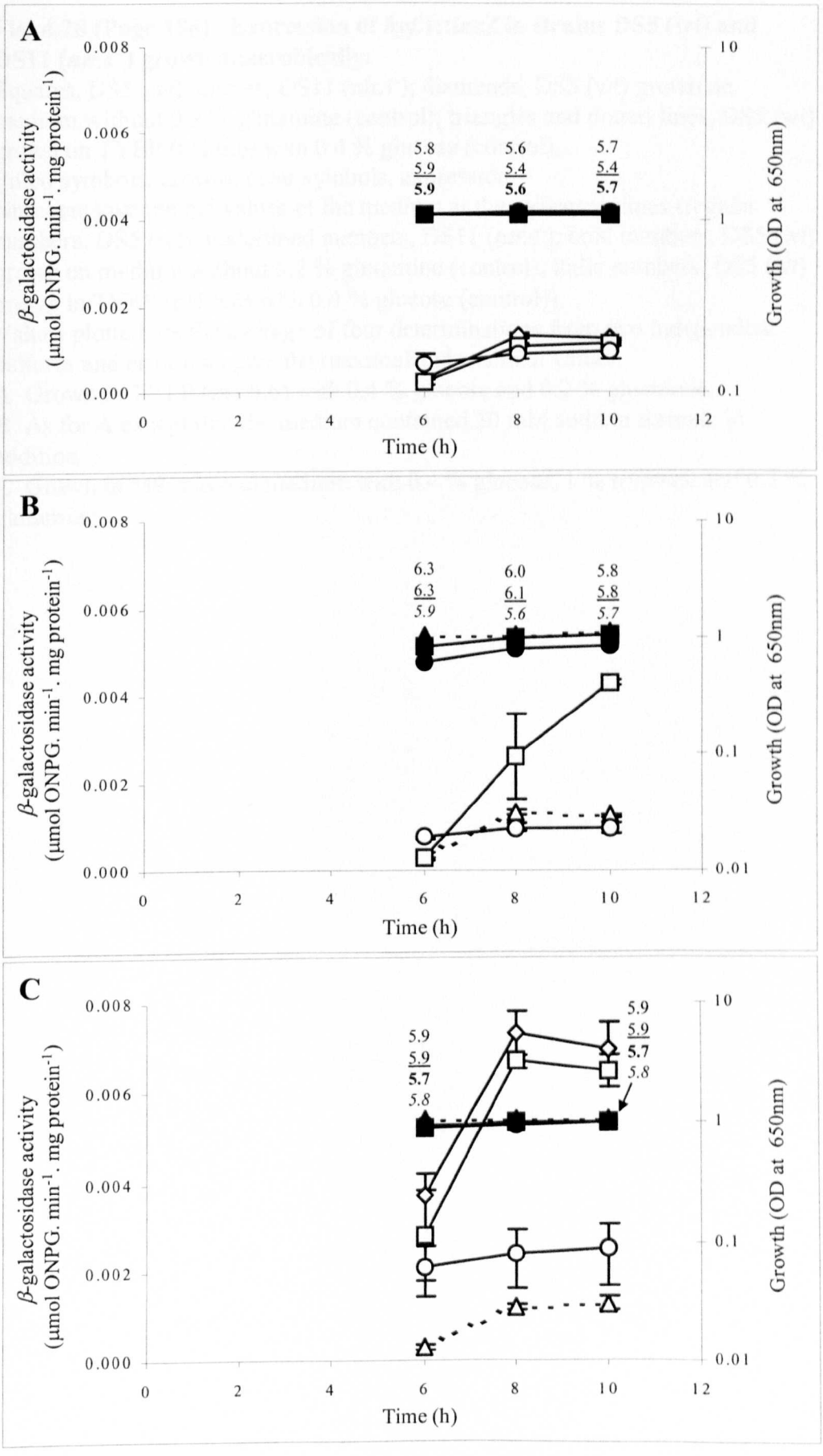

A
B
C
β-galactosidase activity (µmol ONPG. min-1. mg protein-1)
Growth (OD at 650nm)
Time (h)
0.000
0.002
0.004
0.006
0.008
0.01
0.1
1
10
0
2
4
6
8
10
12
5.8
5.9
5.9
5.6
5.4
5.6
5.7
5.4
5.7
6.3
6.3
5.9
6.0
6.1
5.6
5.8
5.8
5.7
5.9
5.9
5.7
5.8
5.9
5.9
5.7
5.8

Fig. 4.28 (Page 188). Expression of *hyfA*::*lacZ* in strains DS5 (wt) and DS11 (*ntrA*$^-$) grown anaerobically.
Squares, DS5 (wt); circles, DS11 (ntr*A*$^-$); diamonds, DS5 (wt) grown on medium without 0.2 % glutamine (control); triangles and dotted lines, DS5 (wt) grown on TYEP (pH 6.6) with 0.4 % glucose (control).
Filled symbols, growth; clear symbols, expression.
Numbers give the pH values of the medium at the indicated times (regular numbers, DS5 (wt); underlined numbers, DS11 (*ntrA*$^-$); bold numbers, DS5 (wt) grown on medium without 0.2 % glutamine (control); italic numbers, DS5 (wt) grown in TYEP (pH 6.6) with 0.4 % glucose (control)).
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.
A Grown in TYEP (pH 6.6) with 0.4 % glucose and 0.2 % glutamine.
B As for **A** except that the medium contained 30 mM sodium formate in addition.
C Grown in M9 minimal medium with 0.4 % glucose, 1 % tryptone and 0.2 % glutamine.

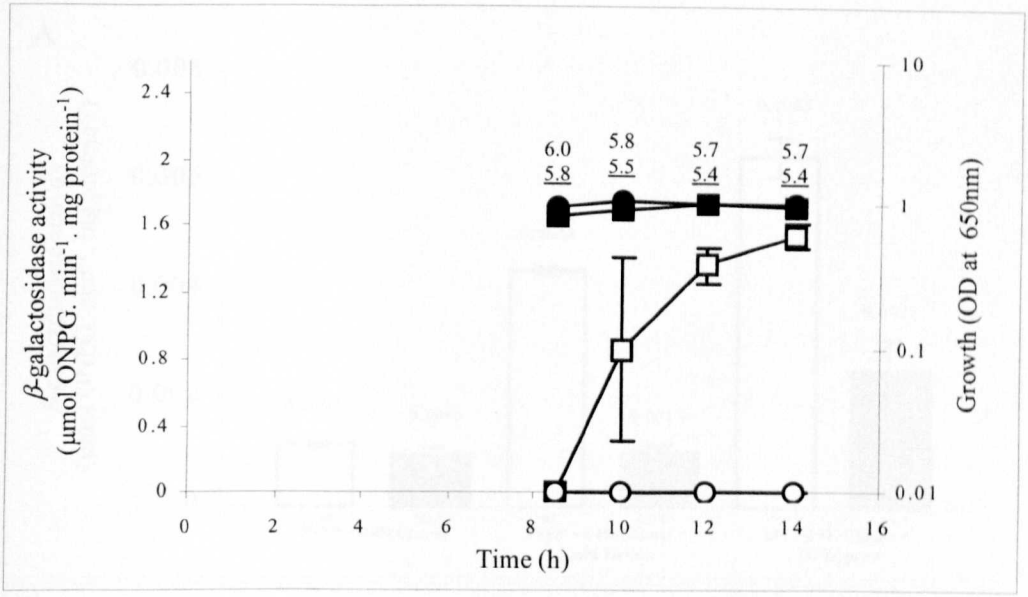


Fig. 4.29. Expression of *hyfA*::*lacZ* in strains DS5 (wt) and DS11 (*ntrA*⁻), both transformed with a multicopy plasmid encoding the *hyfR* gene (pGS1087).

Strains were grown anaerobically in TYEP (pH 6.6) with 0.4 % glucose and 0.2 % glutamine.

Squares, DS5 (wt) transformed with pGS1087; circles, DS11 (*ntrA*⁻) transformed with pGS1087.

Filled symbols, growth; clear symbols, expression.

Numbers give the pH values of the medium at the indicated times (regular numbers, DS5 (wt); underlined numbers, DS11 (*ntrA*⁻)).

Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.

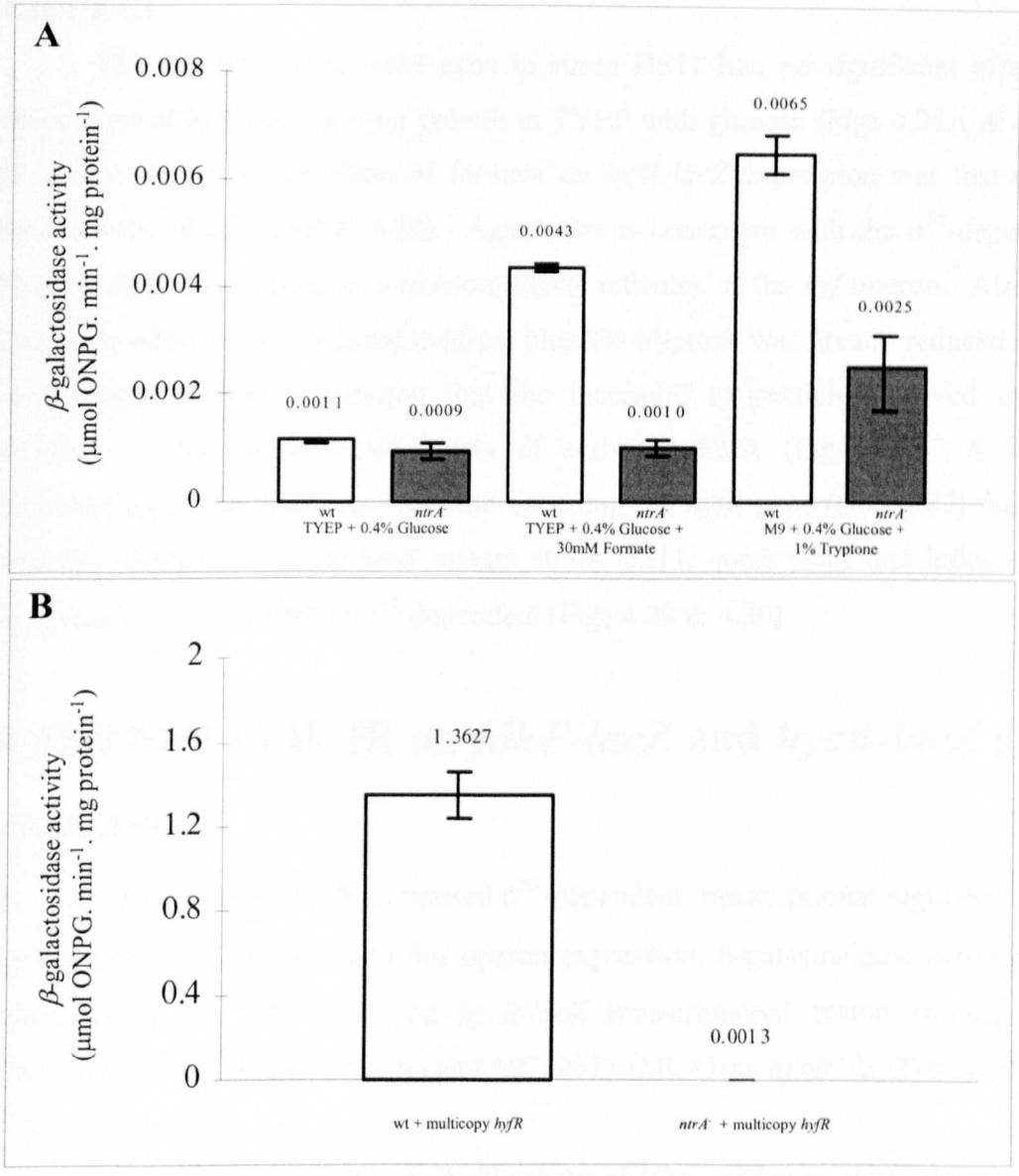


Fig. 4.30. Summary of Figs 4.28 and 4.29.
Expression of *hyfA::lacZ* in strains DS5 and DS11 *(ntrA⁻)* growing anaerobically at 37 °C. All media supplemented with 0.2 % glutamine.
Values plotted are the average of at least four determinations from two independent cultures and error bars give the maximal and minimal values.
A Summary of **Fig. 4.28.** Values plotted are taken from stationary phase: **Fig. 4.28A**, 10 h; **Fig. 4.28B**, 10 h; **Fig. 4.28C**, 10 h.
B Summary of **Fig. 4.29**. Values plotted are taken from stationary phase (12 h). Grown in TYEP (pH6.6) with 0.4% glucose.

glutamine to the growth medium had no significant effect on *hyf* expression (Figs 4.28A & C).

The absence of the *ntrA* gene in strain DS11 had no significant effect on expression of *hyfA-lacZ* during growth in TYEP with glucose (Figs 4.28A & 4.30). However, the inductive effect of formate on *hyfA-lacZ* expression was lost in the *ntrA* mutant (Figs 4.28B & 4.30). Again, this is consistent with the σ^{54}-dependent FhlA protein functioning as a transcriptional activator of the *hyf* operon. Also, the inductive effect of M9 minimal medium plus 1% tryptone was greatly reduced in the *ntrA* mutant further suggesting that the increased expression observed in this medium is due to increased levels of activated FhlA (Figs 4.28C & 4.30). Introduction of the multicopy plasmid encoding the *hyfR* gene (pGS1087) failed to enhance expression in the *ntrA* mutant strain DS11, confirming that induction of *hyfA* expression by HyfR is σ^{54}-dependent (Figs 4.29 & 4.30).

4.15 Effect of HyfR on *fdhF-lacZ* and *hycB-lacZ* gene expression

To study the effect of HyfR (proposed σ^{54}-dependent transcriptional regulator of the *hyf* operon) on *fdhF* gene and *hyc* operon expression, β-galactosidase activity was measured in the *fdhF-lacZ* and *hycB-lacZ* transcriptional fusion strains, M9S (MC4100, *fdhF*::Mu *d*(*bla lac*)ts) and MC10613 (MC4100, *hycB'-lacZ*) respectively (Figs 4.31 & 4.32).

The presence of multicopy *hyfR* enhanced *fdhF-lacZ* expression almost 2 fold (Figs 4.31A & 4.32A). This increase in expression is very small relative to the >1000 fold increase observed for *hyf* expression in the presence of multicopy *hyfR*, and could possibly be attributed to the altered growth of strain M9S when transformed with the multicopy *hyfR* plasmid. However, this increase in *fdhF* expression is consistent with HyfR being a transcriptional activator of the *hyf* encoded Fhl-2 complex requiring the *fdhF* gene product Fdh-H.

The presence of multicopy *hyfR* had no significant effect on *hycB-lacZ* expression (Figs 4.31B & 4.32B).

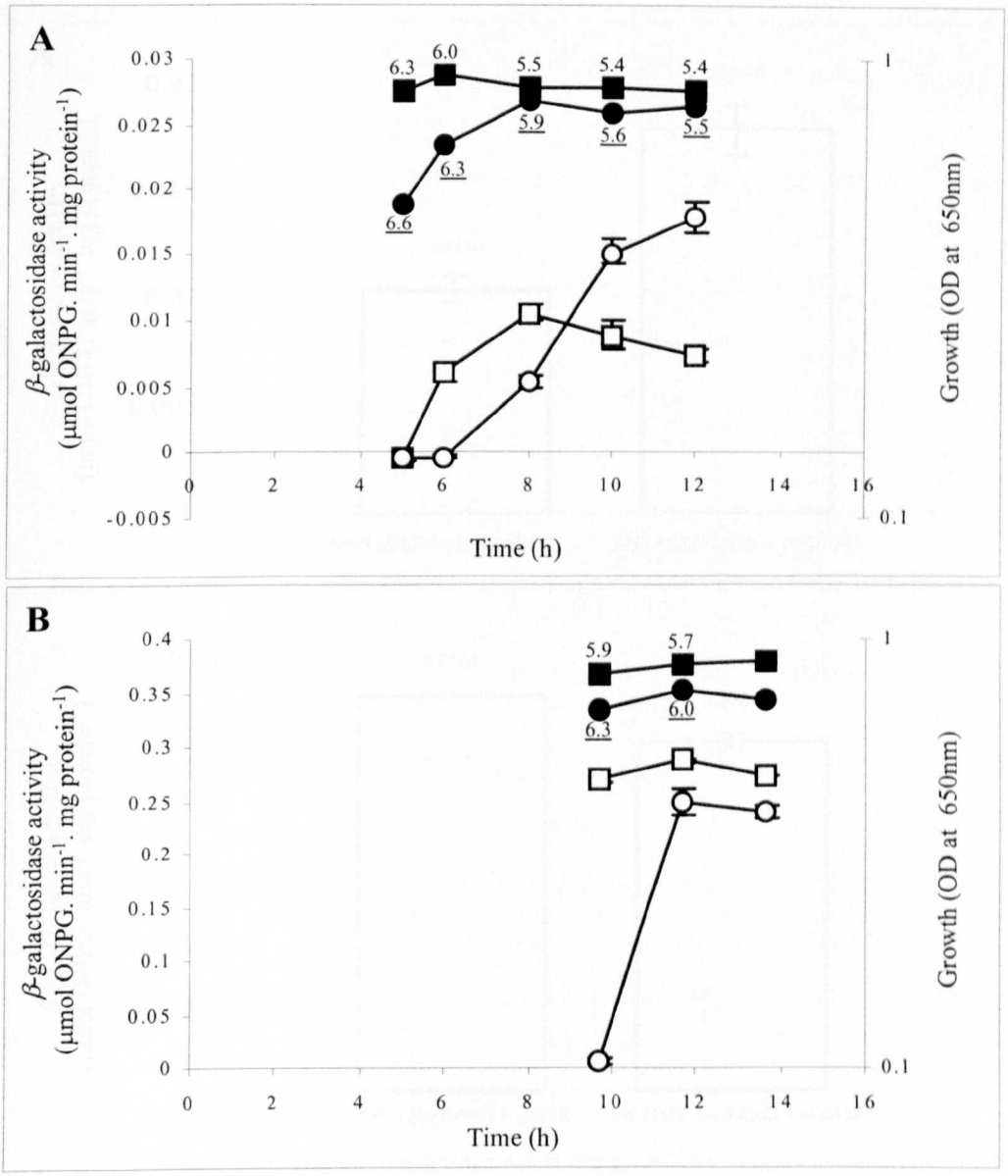


Fig. 4.31. Expression of *fdhF*::*lacZ* in strain M9S (wt) and *hycB*::*lacZ* in strain MC10613 (wt), both transformed with a multicopy plasmid encoding the *hyfR* gene (pGS1087).

Strains were grown anaerobically in TYEP (pH 6.6) with 0.4 % glucose.
Filled symbols, growth; clear symbols, expression.
Values plotted are the average of four determinations from two independent cultures and error bars give the maximal and minimal values.
A Squares, M9S transformed with pSU18; circles, M9S transformed with pGS1087.
Numbers give the pH values of the medium at the indicated times (regular numbers, M9S transformed with pSU18; underlined numbers, M9S transformed with pGS1087).
B Squares, MC10613 transformed with pSU18; circles, MC10613 transformed with pGS1087.
Numbers give the pH values of the medium at the indicated times (regular numbers, MC10613 transformed with pSU18; underlined numbers, MC10613 transformed with pGS1087).

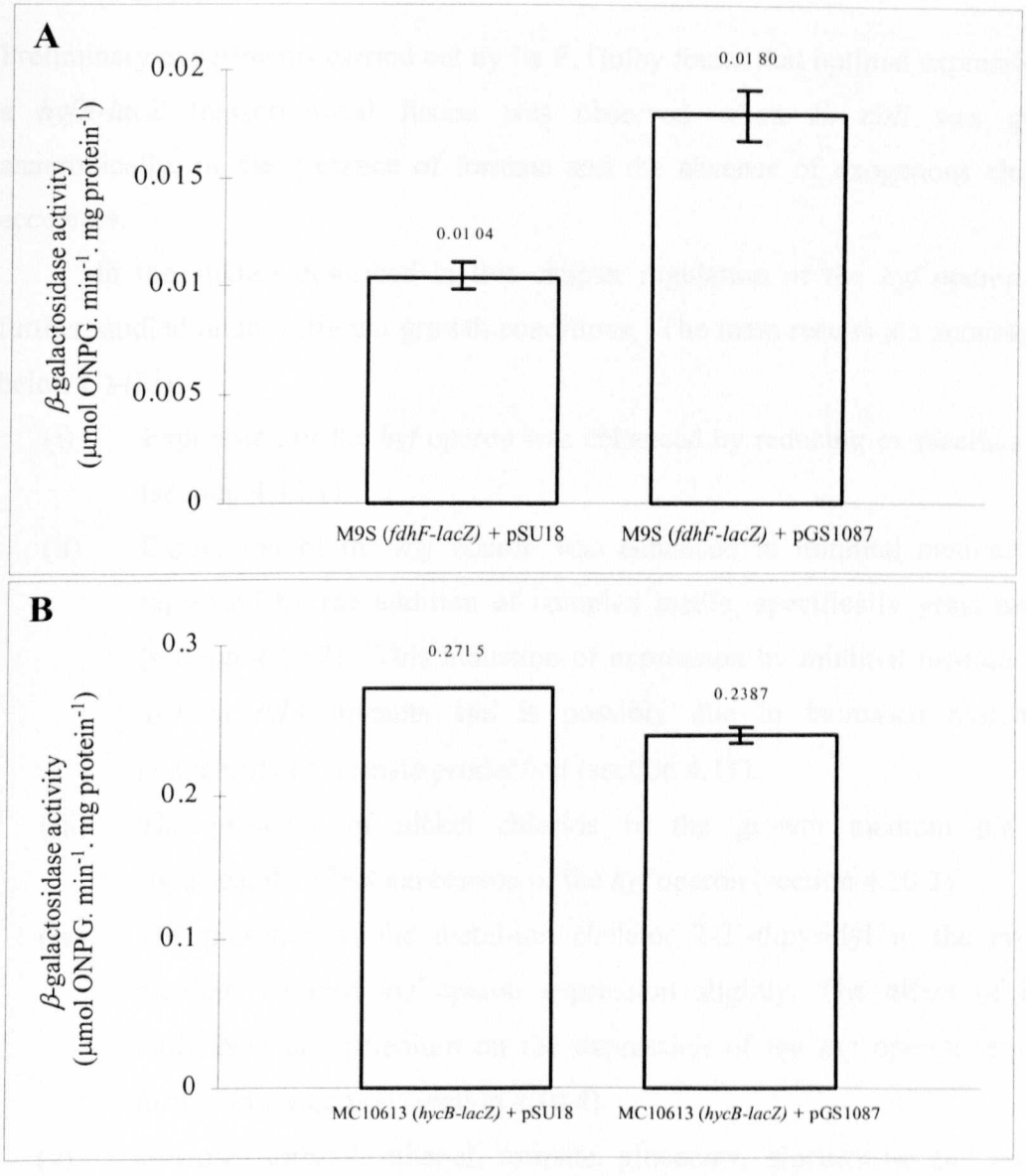


Fig. 4.32. Summary of Fig. 4.31.
Expression of *fdhF*::*lacZ* in strain M9S (wt) and *hycB*::*lacZ* in strain MC10613 (wt), both transformed with a multicopy plasmid encoding the *hyfR* gene (pGS1087). Strains were grown anaerobically in TYEP (pH 6.6) with 0.4 % glucose.
Values plotted are the average of at least four determinations from two independent cultures and error bars give the maximal and minimal values.
A Summary of **Fig. 4.31A**. Values plotted are taken from stationary phase. M9S transformed with pSU18, 8 h; M9S transformed with pGS1087, 12 h.
B Summary of **Fig. 4.31B**. Values plotted are taken from stationary phase (14 h).

4.16 Summary

Preliminary experiments carried out by Dr P. Golby found that optimal expression of a *hyfA-lacZ* transcriptional fusion was observed when *E. coli* was grown anaerobically, in the presence of formate and the absence of exogenous electron acceptors.

In the studies described in this chapter regulation of the *hyf* operon was further studied under different growth conditions. The main results are summarised below (i)-(iii):

(i) Expression of the *hyf* operon was enhanced by reducing extracellular pH (section 4.10.1).

(ii) Expression of the *hyf* operon was enhanced in minimal medium and repressed by the addition of complex media, specifically yeast extract (section 4.10.2). This induction of expression by minimal medium was lost in *fhlA* mutants and is possibly due to increased molybdate availability or formate production (section 4.11).

(iii) The presence of nickel chloride in the growth medium did not significantly affect expression of the *hyf* operon (section 4.10.3).

(iv) The presence of the metal-ion chelator 2-2'-dipyridyl in the growth medium reduced *hyf* operon expression slightly. The effect of iron, molybdate and selenium on the expression of the *hyf* operon requires further investigation (section 4.10.4).

(v) Glucose, sorbitol, ethanol, formate, gluconate, glucuronate and acetate did not significantly enhance *hyf* operon expression on M9 minimal agar plates (section 4.10.5).

A study of *hyf* expression in different mutant backgrounds revealed that the *hyf* operon, like the genes of the formate regulon, appears to be regulated by the σ^{54}-dependent transcriptional activator FhlA. This conclusion is supported by the evidence listed below (i) – (vi):

(i) Induction of *hyf* expression by formate and low pH (sections 4.9 & 4.10.1).

(ii) Loss of formate induction of *hyf* expression in an *fhlA* mutant strain (section 4.11).

(iii) Inhibition of *hyf* expression in the presence of oxygen (formate not produced), the presence of exogenous electron acceptors (e.g. nitrate; formate pool drained by activity of Fdh-N) and non-acidic pH (formate not transported into the cell and intracellular formate pool low) (sections 4.9 & 4.10.1).

(iv) Formate induction of *hyf* expression elevated in a Δ*hycA* (*hycA* encodes an FhlA anti-activator) strain (section 4.12).

(v) Formate induction of *hyf* expression unaffected by a Δ*hycB-H* mutation. FhlA appears to activate *hyf* transcription directly, not via hydrogenase-3 activity (section 4.13).

(vi) Loss of formate induction of *hyf* expression in an *ntrA* mutant strain. Induction dependent on σ^{54} (section 4.14).

Deletion of *hyfR*, encoding the proposed σ^{54}-dependent transcriptional activator of the *hyf* operon did not affect *hyf* operon expression under the growth conditions tested. However introduction of multicopy *hyfR* increased *hyf* operon expression >1000 fold (section 4.11). This suggests that the *hyf* operon is regulated by HyfR but not under the fermentative growth conditions employed in this study. Further work is required to identify any effector molecule specific to HyfR and to investigate whether HyfR, like FhlA, is activated by formate. Introduction of multicopy *hyfR* failed to increase expression of the *hyf* operon in an *ntrA* mutant, confirming that HyfR is σ^{54}-dependent (section 4.14).

Regulation of the *hyf* operon is depicted in Fig. 4.33.

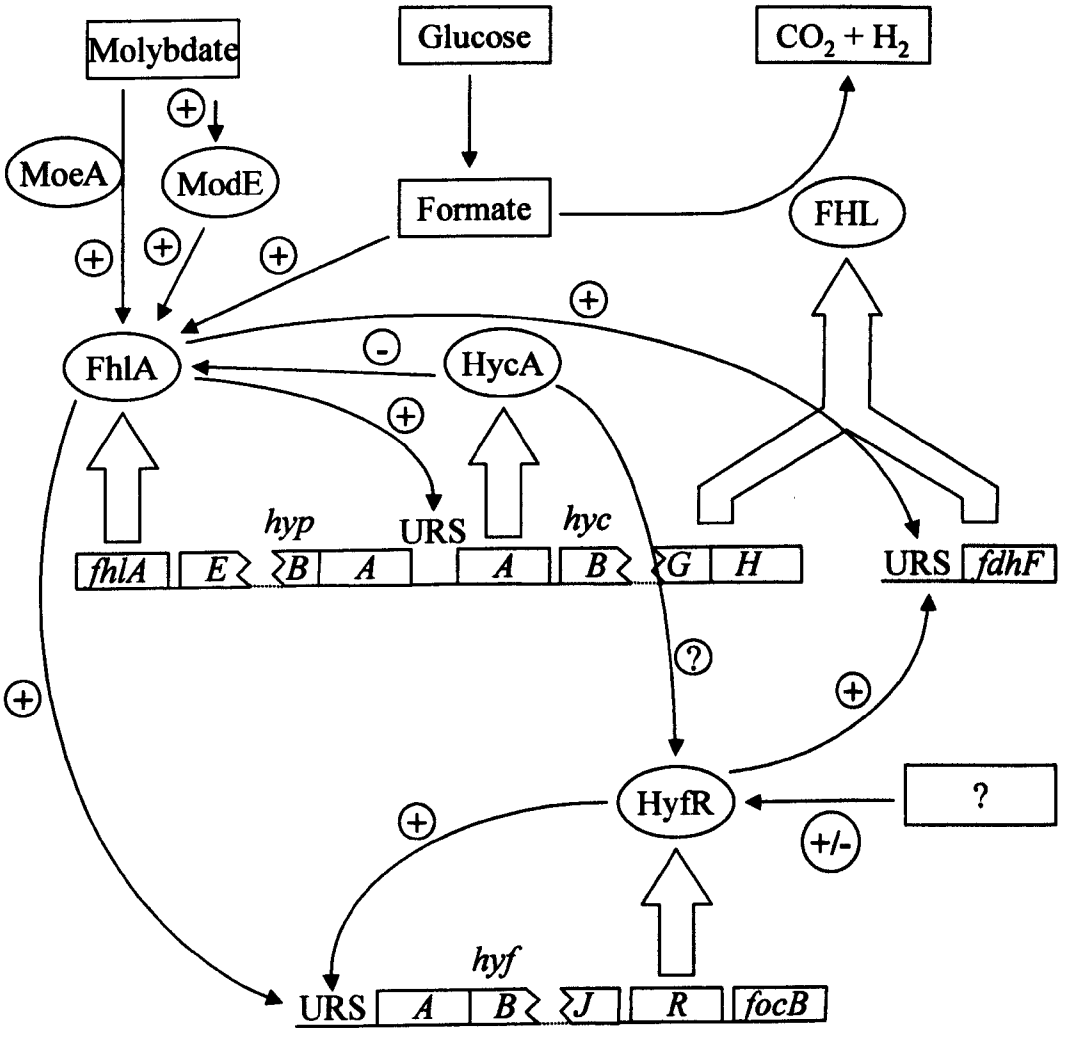


Fig. 4.33. A model depicting the transcriptional regulation of the formate regulon and the *hyf* operon.
Abbreviations: URS, upstream regulatory sequence; +, activation; -, inhibition.

5. USE OF BIOREACTORS TO ANALYSE THE GROWTH PROPERTIES OF *hyc* AND *hyf* MUTANTS.

5.1 Introduction

Andrews and co-workers (1997) proposed that the hydrogenase-4 (*hyf*) operon of *E. coli* encodes a [Ni-Fe] hydrogenase complex, which together with Fdh-H formed an energy conserving (proton translocating) formate hydrogenlyase (Fhl-2). Hydrogenase assays, Fdh-H assays, gas production experiments and H_2 production assays were established to investigate the phenotype of the *hyf* operon (sections 3.9, 3.10 & 3.11). These experiments detected no activity or gas/H_2 production attributable to the *hyf* operon and the proposed Fhl-2 complex. However a recent study by Bagramyan and co-workers (2000, unpublished) detected Fdh-H activity and H_2 production attributable to the *hyf* operon in *E. coli* grown in medium buffered at slightly alkaline pH (pH 7.5).

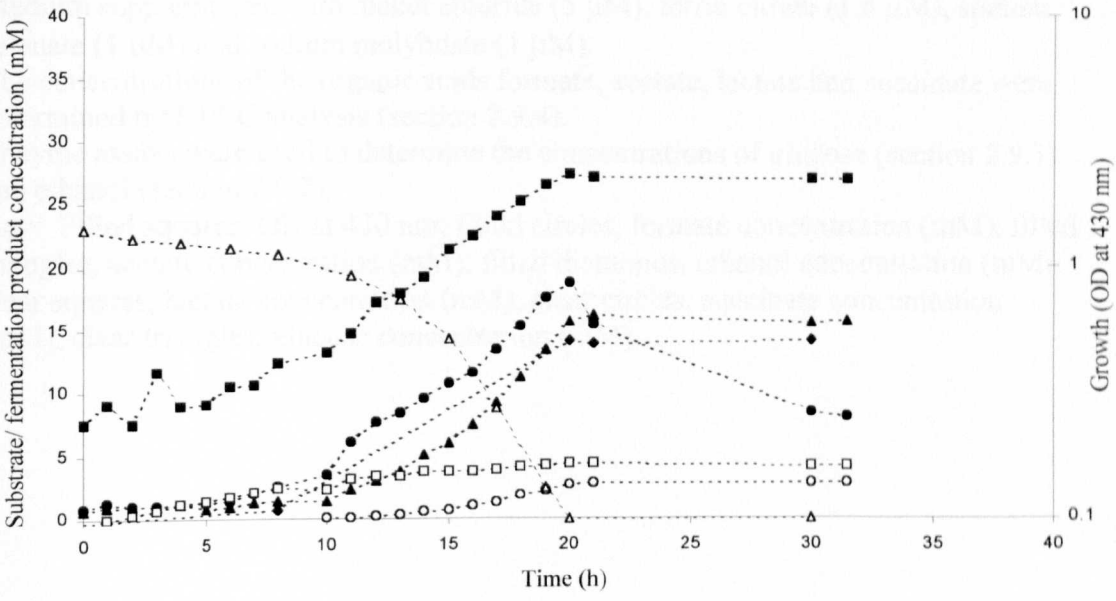
Bioreactors can be used to determine the importance of specific genes in *E. coli* metabolism. A recent study by Abdel-Hamid and co-workers (2001) used batch and continuous cultures to quantify the contribution of the pyruvate oxidase gene (*poxB*) on the growth efficiency and metabolism of *E. coli*. In this chapter anaerobic controlled batch culture experiments were conducted at pH 6.5 and pH 7.5 to further investigate the effects of deletions in the *hyf* operon upon *E. coli* growth and metabolism. Aerobic and anaerobic glucose-limited continuous culture experiments were carried out in an attempt to elucidate a phenotype for the *hyf* mutant.

5.2 Anaerobic controlled batch cultivation of the wildtype (MC4100) and a *hyfB-R* (JRG3621) mutant at pH 6.5 and 7.5

Expression of the *hyf* operon has been found to be strongly pH dependent, with decreasing extracellular pH resulting in an increase in expression. Anaerobic growth in medium buffered to pH 6.1 resulted in an approximately five fold increase in *hyf* expression over anaerobic growth in medium buffered to pH 7.0 or 7.8 (section 4.10.1). However a recent study detected Fdh-H activity and H_2 production attributable to the *hyf* operon in *E. coli* grown in rich medium buffered to slightly alkaline pH (pH 7.5). This *hyf* dependent Fdh-H activity and H_2 production was not detected in *E. coli* grown in medium buffered at pH 6.5 (Bagramyan *et al.*, 2000, unpublished; sections 3.9.2 and 3.11). In contrast, hydrogenase assays, Fdh-H assays and H_2 production experiments presented in this thesis did not detect activity/H_2 production attributable to the *hyf* operon in rich medium buffered to pH 7.5 (section 3.9.3 & section 3.11).

To investigate the effects of the *hyf* operon on anaerobic growth and metabolism at slightly acidic and alkaline pH, wildtype (MC4100) and *hyfB-R* deletion (JRG3621) strains were grown at pH 6.5 and pH 7.5 in anaerobic controlled batch culture in standard minimal medium (section 2.5.3.2) supplemented with glucose (20 mM), nickel chloride (5 μM), ferric citrate (1.6 μM), sodium selenate (1μM) and sodium molybdate (1 μM). The bioreactor was inoculated with an overnight culture pre-grown in shake flasks in the same medium (section 2.6.1.4). Any fermentation gases produced were left to accumulate in the bioreactor headspace and pressure build up was allowed to escape via a tube running from the bioreactor to an inverted water filled measuring cylinder (gas trap). The culture was sampled every hour during exponential phase and at several points during lag and stationary phase. Growth was monitored by measuring the OD at 430 nm of diluted culture samples with time. The concentration of fermentation products and unutilised substrate in the culture sample supernatants were determined by HPLC analysis (formate, acetate, lactate, succinate; section 2.9.4) and enzyme assay (formate, ethanol, glucose: sections 2.9.3, 2.9.2 & 2.9.1 respectively). The results are shown in Fig. 5.1, Fig. 5.2 & Table 5.1.

A MC4100 (*wt*)



B JRG3621 (Δ*hyfB-R*)

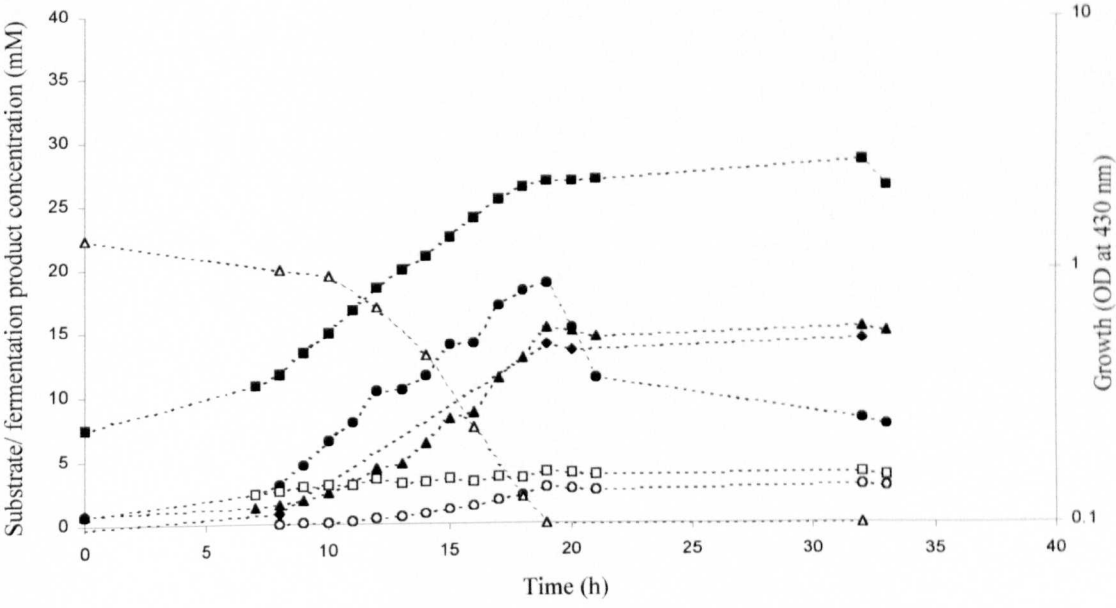
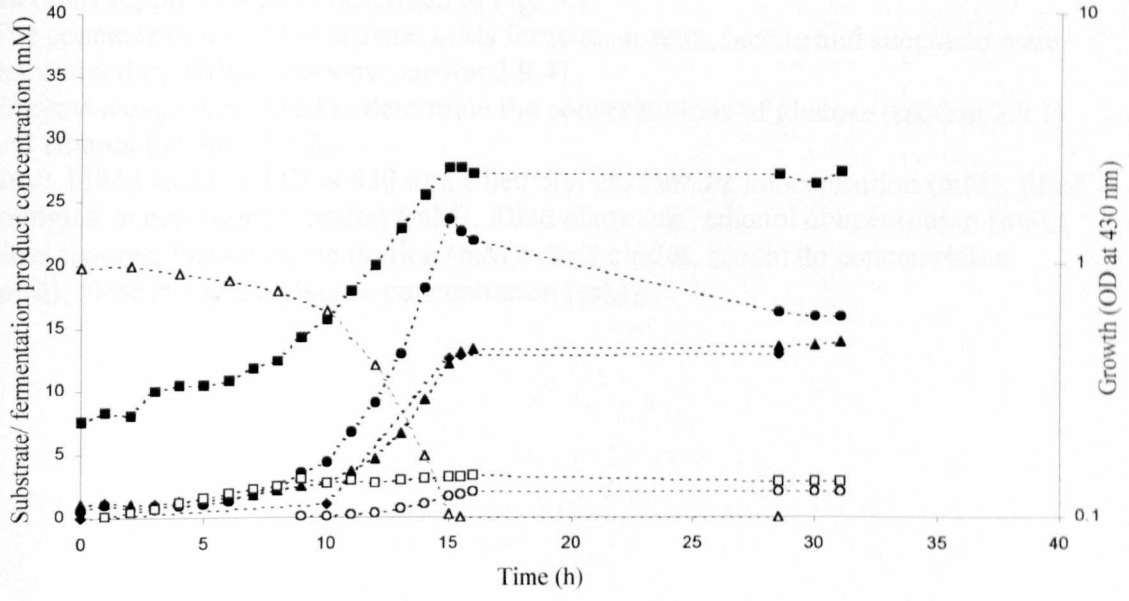


Fig. 5.1 (Page 199). Anaerobic controlled batch cultivation of strains MC4100 (*wt*; A) and JRG3621 (*ΔhyfB-R*; B) in standard minimal medium (section 2.5.3.2) containing glucose (20 mM) with pH maintained at 6.5.
Medium supplemented with nickel chloride (5 μM), ferric citrate (1.6 μM), sodium selenate (1 μM) and sodium molybdate (1 μM).
The concentrations of the organic acids formate, acetate, lactate and succinate were determined by HPLC analysis (section 2.9.4).
Enzyme assays were used to determine the concentrations of glucose (section 2.9.1) and ethanol (section 2.9.2).
Key: Filled squares, OD at 430 nm; filled circles, formate concentration (mM); filled triangles, acetate concentration (mM); filled diamonds, ethanol concentration (mM); clear squares, lactate concentration (mM); clear circles, succinate concentration (mM); clear triangles, glucose concentration (mM).

A MC4100 (*wt*)



B JRG3621 (Δ*hyfB-R*)

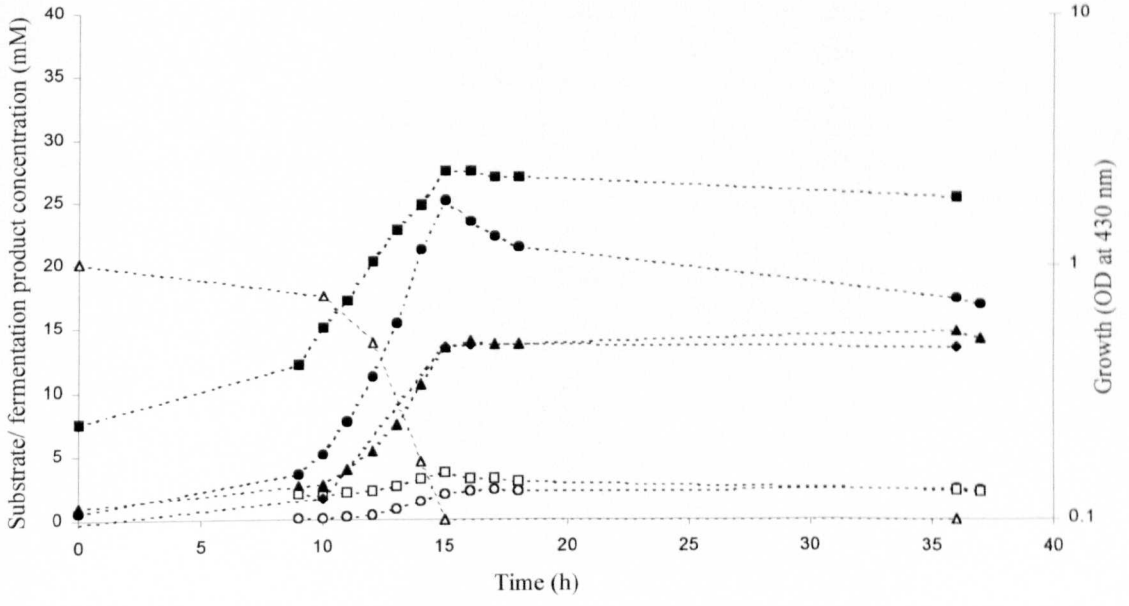


Fig. 5.2 (Page 201). Anaerobic controlled batch cultivation of strains MC4100 (*wt*; A) and JRG3621 (Δ*hyfB-R*; B) in standard minimal medium (section 2.5.3.2) containing glucose (20 mM) with pH maintained at 7.5.
Medium supplemented as described in Fig. 5.1.
The concentrations of the organic acids formate, acetate, lactate and succinate were determined by HPLC analysis (section 2.9.4).
Enzyme assays were used to determine the concentrations of glucose (section 2.9.1) and ethanol (section 2.9.2).
Key: Filled squares, OD at 430 nm; filled circles, formate concentration (mM); filled triangles, acetate concentration (mM); filled diamonds, ethanol concentration (mM); clear squares, lactate concentration (mM); clear circles, succinate concentration (mM); clear triangles, glucose concentration (mM).

Activity or substrate/ fermentation product concentration	pH 6.5		pH 7.5	
	MC4100 (*wt*)	JRG3621 (*ΔhyfB-R*)	MC4100 (wt)	JRG3621 (*ΔhyfB-R*)
μmax	0.13	0.12	0.24	0.23
Final biomass (g [dry weight]/litre)[a]	0.63	0.60	0.65	0.64
Formate[b]	0.81	0.84	1.20	1.25
Formate[c]	0.65	0.65	1.06	1.15
Acetate[b]	0.68	0.68	0.62	0.67
Ethanol[c]	0.61	0.62	0.64	0.67
Lactate[b]	0.19	0.18	0.16	0.19
Succinate[b]	0.12	0.13	0.09	0.10
Glucose consumed[d]	23.05	22.31	19.67	20.06

Table 5.1. Effect of pH on μmax, final biomass and fermentation product distribution for MC4100 (wt) and JRG3621 (*ΔhyfB-R*) during anaerobic controlled batch cultivation in minimal medium (section 2.5.3.2) containing glucose (20 mM).

Medium supplemented as described in Fig. 5.1.

Final biomass estimated and fermentation products measured at the beginning of the stationary phase.

[a] Final biomass estimated from OD at 430 nm (section 5.2).

[b] Concentration in mmoles per mmole of consumed glucose. Concentration of relevant fermentation product determined by HPLC analysis (section 2.9.4). Calculated by dividing the concentration of the relevant fermentation product in the medium at the onset of stationary phase by the quantity of glucose consumed[d].

[c] Concentration in mmoles per mmole of consumed glucose. Concentration of fermentation product determined by enzyme assay (formate, section 2.9.3; ethanol, section 2.9.2). Calculated as for [b].

[d] Concentration in mmoles glucose consumed determined by enzyme assay (section 2.9.1). Calculated by subtracting the concentration of glucose in the medium at the onset of stationary phase from the concentration of glucose at the start of the experiment (0 h).

5.2.1 Effect of *hyfB-R* on *E. coli* anaerobic growth and metabolism

The lag phase, maximum growth rate (μmax) and estimated final biomass of the bacterial cultures grown at both pH 6.5 and 7.5 were essentially unaffected by the *hyfB-R* deletion (Fig. 5.1, Fig. 5.2 & Table 5.1).

The deletion of the *hyfB-R* genes had no effect on the concentration of fermentation products detected at either pH 6.5 or pH 7.5 (Fig. 5.1, Fig. 5.2 & Table 5.1).

Taken together these results suggest that deletion of the *hyfB-R* genes has no effect on *E. coli* growth and metabolism under these conditions.

5.2.2 Effect of pH on *E. coli* anaerobic growth and metabolism

E. coli is a neutrophile, growing optimally at approximately neutral pH (Slonczewski & Foster, 1996). However it can also grow in a moderate acid or base medium regulating its internal pH between 7.4 and 7.8 (Padan *et al.*, 1976; Slonczewski *et al.*, 1981; Zilberstein *et al.*, 1984). Therefore, as expected, the μmax (maximum growth rate) of MC4100 (wt) and JRG3621 (Δ*hyfB-R*) was reduced (almost two fold) by a drop in pH from 7.5 to 6.5 (Table 5.1). The estimated final biomass of MC4100 (wt) was not affected by this drop in pH.

Like the *hyf* operon, expression of the *fdhF* gene and the *hyc* operon is strongly pH dependent, with decreasing extracellular pH resulting in an increase in expression during anaerobic growth (Rossmann *et al.*, 1991). This pH dependence makes sense as the Fhl-1 complex, encoded by the *fdhF* gene and the *hyc* operon, catalyses the non-energy conserving breakdown of formate to H_2 and CO_2 gas reducing acid levels during fermentation. Therefore it is presumed that the 30-40% lower formate concentration at pH 6.5 compared to pH 7.5 is due to increased expression of Fhl-1 (Table 5.1). Also the levels of formate detected for MC4100 (*wt*) at pH 6.5 were reduced 57% in the 11.5 hours after the onset of stationary phase (20 h) (Fig 5.2). This compares to a reduction of just 31% in formate levels detected for MC4100 (*wt*) at pH 7.5 in the 13 hours after the onset of stationary phase (15.5 h) (Fig. 5.1). The concentration of the fermentation products acetate, ethanol and lactate were essentially unaffected by pH. It is surprising that the concentration of lactate was unaffected by the drop in pH from 7.5 to 6.5 as lactate dehydrogenase the

enzyme catalysing the reduction of pyruvate to lactate is induced several fold by acid during anaerobic growth (Mat-Jan *et al.*, 1989).

5.3 Anaerobic controlled batch cultivation of a *hycE* (HD705) deletion strain at pH 6.5

The *hyc* operon encodes hydrogenase-3, a component of the Fhl-1 complex of *E.coli*, which catalyses the non-energy conserving breakdown of formate to H_2 and CO_2. Andrews and co-workers (1997) proposed that the *hyf* operon encoded a fourth hydrogenase in *E. coli*, a component of a second Fhl complex (Fhl-2) catalysing the energy conserving breakdown of formate to H_2 and CO_2 gas. Expression studies presented in this thesis with a *hyfA-lacZ* transcriptional fusion strain (DS5) revealed that optimal *hyf* expression is observed during fermentative growth at acidic pH (pH <6.1) in the presence of formate and the absence of electron acceptors (section 4.9 & section 4.10). Sauter and co-workers (1992) detected no hydrogenase activity at pH 7.0 in the hydrogenase-1, -2 and –3 triple mutant HDJ123 (*Δhya Δhyb ΔhycB-H*) grown under these conditions and concluded that *E. coli* synthesises no further hydrogenases under these growth conditions. If this is the case then the *hyc* encoded Fhl-1 complex is the only active Fhl under these growth conditions. This is further supported by the observation that gas production was lost in strains carrying mutations or deletions in the genes of the *hyc* operon (Pecher *et al.*, 1983; section 3.10). It is possible however that Fdh-H is able to breakdown formate in the cell in the absence of the *hyc* operon encoded hydrogenase-3. Although gas production is lost in a *hycE* mutant, Fdh-H activity is still detectable (Pecher *et* al., 1983; Sauter *et al.*, 1992). It is possible that Fdh-H is able to donate the electrons generated from the oxidation of formate to CO_2 (possibly through the integral membrane proteins HycC and HycD) to anaerobic reductases (section 1.1.4.6). The CO_2 produced may not be detected in the gas production experiments reported (Pecher *et al.*, 1983; section 3.10) as it may be utilised in the carboxylation of phosphoenolpyruvate (Fig 1.1). However during fermentative growth either the required anaerobic reductases would not be synthesised or the required terminal electron acceptors are not present. Similarly, it is unlikely that formate would be removed from the medium by Fdh-N and Fdh-O, as these enzymes are not synthesised under these growth conditions.

Therefore it is not unreasonable to assume that *hycE* strains would be unable to breakdown formate produced during fermentative growth. To investigate the fate of formate in the absence of Fhl-1, a *hycE* deletion strain (HD705) was grown in anaerobic controlled batch culture at pH 6.5 and the concentrations of fermentation products and non-utilised glucose monitored throughout growth (as described in section 5.2). The results are shown in Fig. 5.3 and Table 5.2.

5.3.1 Effect of *hycE* on *E. coli* anaerobic growth at pH 6.5

Surprisingly the *hycE* mutant HD705 had no lag phase when grown in anaerobic controlled batch culture at pH 6.5, compared with a 7 - 8 h lag phase observed for strains MC4100 (wt) and JRG3621 (Δ*hyfB-R*) under these growth conditions (Fig. 5.3). It is hard to conceive that deletion of the *hyc* gene and the subsequent absence of Fhl-1 could have had this effect on the length of the lag phase of the bacterial culture. It seems more likely that observed differences in the length of the lag phase are due to differences in the growth phase of the overnight culture used in inoculation. This is despite attempts made to standardise the inoculum used in these controlled batch cultivation experiments (section 2.6.1.4). Further standardisation is possibly required to avoid the differences in lag phase observed during batch cultivation. Monitoring the optical density of the inoculum during incubation would have more accurately standardised the growth phase at the point of inoculation. However this would have added a considerable and possibly impractical amount of time to the start of the experiment. The maximum growth rate and the estimated final biomass of the bacterial culture was unaffected by the deletion of the *hycE* gene (Tables 5.1 & 5.2).

5.3.2 Effect of *hycE* on *E.coli* anaerobic fermentation product distribution pH 6.5

The levels of formate detected at the beginning of stationary phase in the culture medium of HD705 (Δ*hycE*) grown in anaerobic controlled batch culture at pH 6.5 were increased approximately two fold compared with the levels observed for the wildtype strain (MC4100) grown under identical conditions (Tables 5.1 & 5.2). This was expected as *hycE* encodes the large subunit of hydrogenase-3 a component of the Fhl-1 complex. The Fhl-1 complex catalyses the breakdown of formate to

HD705 (Δ*hycE*)

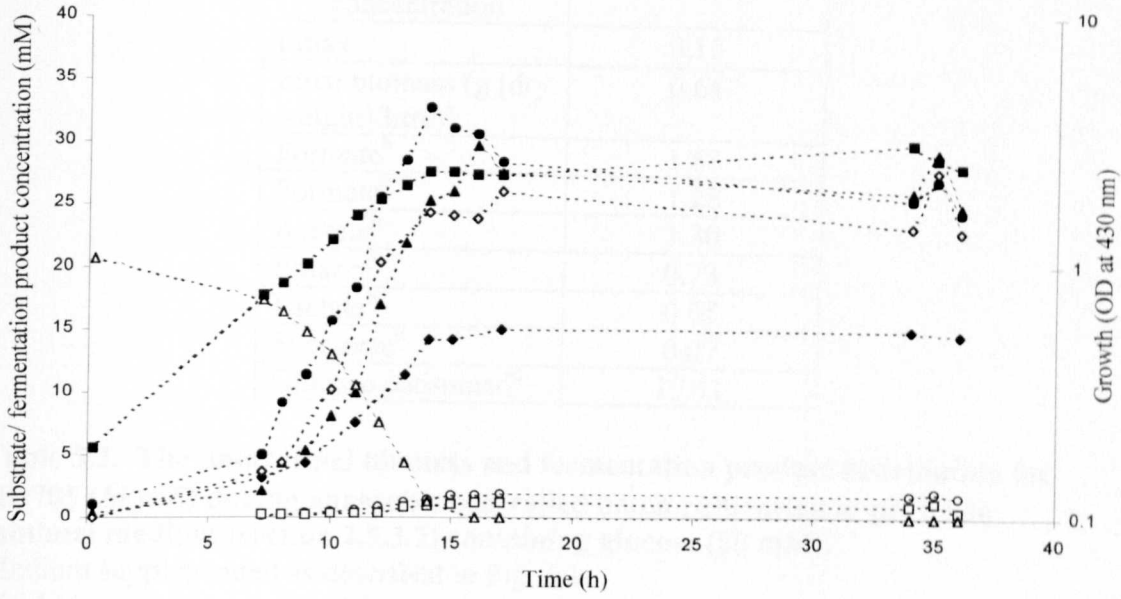


Fig. 5.3. Anaerobic controlled batch cultivation of strain HD705 (Δ*hycE*) in standard minimal medium (section 2.5.3.2) containing glucose (20 mM) with pH maintained at 6.5.

Medium supplemented as described in Fig. 5.1.

The concentrations of the organic acids formate, acetate, lactate and succinate were determined by HPLC analysis (section 2.9.4).

Enzyme assays were used to determine the concentrations of glucose (section 2.9.1), ethanol (section 2.9.2) and formate (section 2.9.3).

Key: Filled squares, OD at 430 nm; filled circles, formate concentration (mM; determined by HPLC analysis); filled triangles, acetate concentration (mM); filled diamonds, ethanol concentration (mM); clear squares, lactate concentration (mM); clear circles, succinate concentration (mM); clear triangles, glucose concentration (mM); clear diamonds, formate concentration (mM; determined by enzyme assay).

Activity or substrate/ fermentation product concentration	pH 6.5
	HD705 (Δ*hycE*)
μmax	0.15
Final biomass (g [dry weight]/litre)[a]	0.64
Formate[b]	1.68
Formate[c]	1.28
Acetate[b]	1.30
Ethanol[c]	0.73
Lactate[b]	0.05
Succinate[b]	0.07
Glucose consumed[d]	19.41

Table 5.2. The μmax, final biomass and fermentation product distribution for HD705 (Δ*hycE*) during anaerobic controlled batch cultivation at pH 6.5 in minimal medium (section 2.5.3.2) containing glucose (20 mM).
Medium supplemented as described in Fig. 5.1.
Final biomass estimated and fermentation products measured at the beginning of the stationary phase.
[a] Final biomass estimated from OD at 430 nm (section 5.2).
[b] Concentration in mmoles per mmole of consumed glucose. Concentration of relevant fermentation product determined by HPLC analysis (section 2.9.4). Calculated as described for Table 5.1.
[c] Concentration in mmoles per mmole of consumed glucose. Concentration of fermentation product determined by enzyme assay (formate, section 2.9.3; ethanol, section 2.9.2). Calculated as for [b].
[d] Concentration in mmoles glucose consumed determined by enzyme assay (section 2.9.1). Calculated as described for Table 5.1.

CO_2 and H_2 during fermentative growth. However, HPLC analysis revealed that formate levels detected in the culture medium of HD705 (Δ*hycE*) grown in anaerobic controlled batch culture at pH 6.5 were reduced 26% in the 22 hours after the onset of stationary phase (h) (Fig. 5.3). This reduction was only 12% when formate levels were detected by enzyme analysis suggesting that this reduction may be exaggerated by or be due to experimental error. An important focus of future work should be to confirm that this drop is significant and to determine whether it exists in a *hyc/hyf* double mutant. The *hycE* deletion strain HD705 produced approximately twice as much acetate and 20% more ethanol by the beginning of stationary phase compared with MC4100 grown under identical conditions, however the levels of both lactate and succinate were reduced approximately 74% and 42% respectively in strain HD705 (Tables 5.1 & 5.2). The reduction in succinate levels is not unexpected as in the absence of a Fhl complex the availability of metabolic CO_2 will limit the amount of succinate production.

5.4 Aerobic glucose-limited chemostat cultivation of the wildtype (MC4100) and a *hyfB-R* (JRG3621) mutant at pH 6.5

All experiments conducted thus far in the study of the *hyf* operon have been based on the proposal of Andrews and co-workers (1997) that the *hyf* operon encoded components of a second Fhl complex in *E. coli* (Fhl-2). It is not unreasonable to assume that Fhl-2 would be active under anaerobic conditions as this is when formate is produced by *E. coli,* and expression studies revealed that optimal *hyf* expression was during anaerobiosis (section 4.9). Bagramyan and co-workers (2000, unpublished) detected Fdh-H activity and H_2 production attributable to the *hyf* operon in *E. coli* grown anaerobically in medium buffered to slightly alkaline pH (pH 7.5). These experiments have been repeated and Fdh-H activity/H_2 production attributable to the *hyf* operon was not detected (section 3.9.3 and section 3.11). All other experiments have not conclusively identified a phenotype for *hyf* mutants. During aerobic growth on glucose expression of the *hyf* operon was reduced (five fold) but the operon was still expressed (P. Golby, unpublished). It is possible that

hyf may have an unexpected or unanticipated aerobic function, which has thus far been unidentified because all previous experiments have investigated an anaerobic function for *hyf* (Andrews *et* al., 1997).

To investigate a possible phenotype for the *hyf* operon under aerobic conditions, the growth and metabolism of strains MC4100 (wt) and JRG3621 (*ΔhyfB-R*) were studied during aerobic glucose-limited chemostat cultivation at pH 6.5 (section 2.6.2). Standard minimal medium was supplemented with glucose (20 mM), nickel chloride (5 μM), ferric citrate (1.6 μM), sodium selenate (1μM) and sodium molybdate (1 μM). The medium was then inoculated with an overnight culture (grown in shake flasks in the same medium) and grown as a controlled batch culture. At exponential/stationary phase of growth the pump adding glucose (20 mM)-limited minimal medium to the bioreactor was switched on and the culture grown at a range of dilution rates (0.10 – 0.25 h^{-1}). Aerobic conditions were achieved by flushing the chemostat with 500 – 800 cm^3/min airflow. Agitation was maintained at 500 rpm. Once the optical density had remained constant for five working volumes the following physiological parameters were measured: biomass concentration (dry weight, section 2.9.5 & section 5.2) and; rate of substrate (glucose) utilisation (q_S, section 2.10.2). The maintenance energy (section 2.11.4) and the maximum growth yield (section 2.11.5) were determined from the plot of substrate utilisation (q_S) against dilution rate (D). The physiological parameters described above were calculated for each strain (section 2.9.5). The results for the glucose limited chemostat cultivation of MC4100 (wt) are shown in Table 5.3 and Fig. 5.4, and JRG3621 shown in Table 5.4 and Fig. 5.5.

The maintenance energy (M_e) and maximum growth yield (Y_{MAX}) of JRG3621 (*ΔhyfB-R*) were increased 27.5% and 7.5% respectively, over levels observed for MC4100 (*wt*) (Fig. 5.4 & 5.5). The increase in maintenance energy means that the *hyfB-R* mutant requires more energy, and therefore metabolises more carbon, to maintain cell viability. If this is the case then the *hyfB-R* mutant would have less remaining carbon available for growth, however as the maximum growth yield is also increased in the *hyfB-R* mutant, experimental error must have contributed to one or both of the differences in maintenance energy or maximum growth yield observed. Further work is required to investigate whether the differences in maintenance energy and maximum growth yield observed between

D (h^{-1})	Dry weight ($g\ l^{-1}$)	q_s (mmol h^{-1} (g dry $weight^{-1}$))
0.10	1.22	1.69
0.16	1.41	2.46
0.21	1.38	3.21
0.25	1.38	3.68

Table 5.3. Physiological parameters of MC4100 (wt) in aerobic glucose-limited chemostat cultivation in standard minimal medium (section 2.5.3.2) with pH maintained at pH 6.5. Medium supplemented as described in Fig. 5.1. Steady states were obtained in glucose (20 mM)-limited chemostat cultures over a range of dilution rates (h^{-1}). Key to abbreviations: D, dilution rate; q_S, rate of substrate (glucose) utilisation.

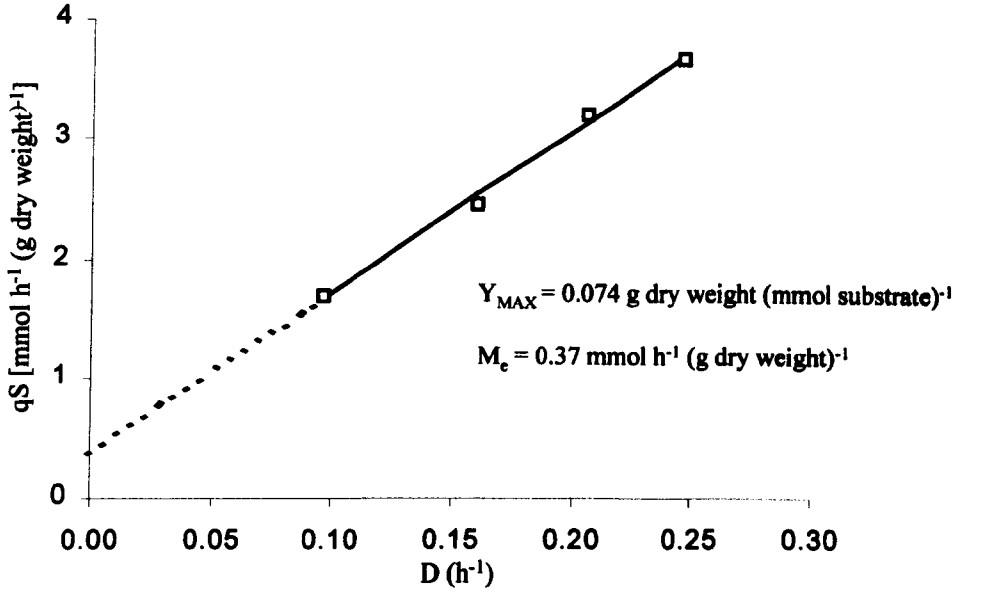


Fig. 5.4. Glucose consumption rates (q_S) of MC4100 (wt) in aerobic glucose-limited chemostat cultivation in standard minimal medium (section 2.5.3.2) with pH maintained at pH 6.5. Medium supplemented as described in Fig. 5.1. Steady states were obtained in glucose (20 mM)-limited chemostat cultures over a range of dilution rates (h^{-1}). Key to abbreviations: D, dilution rate; q_S, rate of substrate (glucose) utilisation; Y_{MAX}, maximum biomass yield; M_e, maintenance energy.

D (h^{-1})	Dry weight ($g\ l^{-1}$)	q_s (mmol h^{-1} (g dry $weight^{-1}$))
0.10	1.34	1.55
0.14	1.26	2.45
0.20	1.36	3.14
0.25	1.52	3.55

Table 5.4. Physiological parameters of JRG3621 (*ΔhyfB-R*) in aerobic glucose-limited chemostat cultivation in standard minimal medium (section 2.5.3.2) with pH maintained at pH 6.5. Medium supplemented as described in Fig. 5.1. Steady states were obtained in glucose (20 mM)-limited chemostat cultures over a range of dilution rates (h^{-1}). Key to abbreviations: D, dilution rate; q_S, rate of substrate (glucose) utilisation.

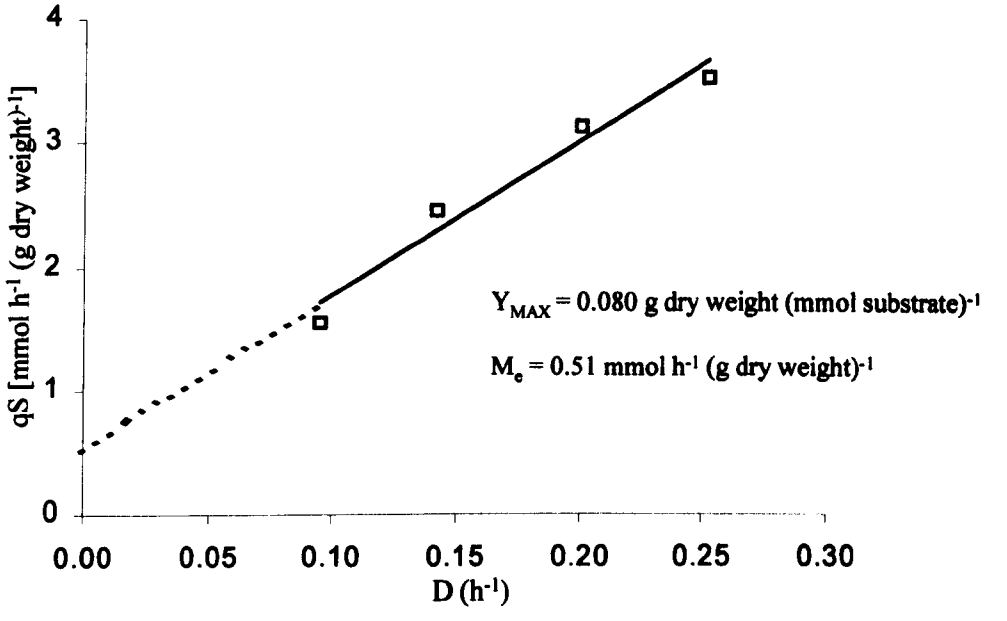


Fig. 5.5. Glucose consumption rates (q_S) of JRG3621 (*ΔhyfB-R*) in aerobic glucose-limited chemostat cultivation in standard minimal medium (section 2.5.3.2) with pH maintained at pH 6.5. Medium supplemented as described in Fig. 5.1. Steady states were obtained in glucose (20 mM)-limited chemostat cultures over a range of dilution rates (h^{-1}). Key to abbreviations: D, dilution rate; q_S, rate of substrate (glucose) utilisation; Y_{MAX}, maximum biomass yield; M_e, maintenance energy.

MC4100 (*wt*) and JRG3621 (*ΔhyfB-R*) are significant or solely due to experimental error.

5.5 Anaerobic glucose-limited chemostat cultivation of the wildtype (MC4100) and *hyfB-R* (JRG3621) mutant at pH 6.5

To investigate whether small differences in growth and metabolism observed between MC4100 (wt) and JRG3621 (*ΔhyfB-R*) during anaerobic batch cultivation (section 5.3.1) are significant, these strains were grown as anaerobic glucose-limited chemostat cultures at a range of dilution rates. Anaerobic conditions were achieved by flushing the chemostat with N_2 (400-500 $cm^3.min^{-1}$) with agitation maintained at 150 rpm. Andrews and co-workers (1997) suggested that the proposed *hyf* encoded Fhl-2 complex would only be active when H_2 was removed as it was produced preventing its build up in the environment. Flushing the bioreactor with N_2 not only serves to create anaerobic conditions during chemostat cultivation but also prevents H_2 build up in the bioreactor. Otherwise glucose (20 mM)-limited chemostat cultivation was as described in section 5.4. Cultivation was in standard minimal medium supplemented with glucose (20 mM), nickel chloride (5 μM), ferric citrate (1.6 μM), sodium selenate (1 μM) and sodium molybdate (1 μM). Due to the poor energy yield of fermentation only approximately 10% of the carbon source is converted to biomass (Böck & Sawers, 1996). To avoid any errors that may be encountered from measuring such small dry weights, the dry weight of the cultures at steady state was estimated from the OD at 430 nm (section 2.6). The results for MC4100 are shown in Table 5.5 and Fig. 5.6, and the results for JRG3621 are shown in Table 5.6 and Fig. 5.7.

Anaerobic glucose (20 mM)-limited chemostat cultivation was originally conducted with agitation maintained at 500 rpm, effectively increasing the distribution of N_2 in the bioreactor. At this increased agitation difficulty was encountered maintaining culture growth and preventing washing out of the culture even at low dilution rates (data not shown). The reason for this is unknown.

The maintenance energy and maximum growth yield of MC4100 (*wt*) and JRG3621 (*ΔhyfB-R*) were very similar, confirming that for anaerobic growth at pH

D (h^{-1})	Dry weight* ($g\ l^{-1}$)	q_s (mmol h^{-1}(g dry weight*$^{-1}$))
0.05	0.38	2.93
0.07	0.34	4.15
0.08	0.39	4.35
0.10	0.40	4.86
0.11	0.40	5.62
0.13	0.39	6.52

Table 5.5. Physiological parameters of MC4100 (wt) in anaerobic glucose-limited chemostat cultivation in standard minimal medium (section 2.5.3.2) with pH maintained at pH 6.5. Medium supplemented as described in Fig. 5.1. Steady states were obtained in glucose (20 mM)-limited chemostat cultures over a range of dilution rates (h^{-1}). Key to abbreviations: D, dilution rate; q_S, rate of substrate (glucose) utilisation. *Dry weight estimated from OD at 430 nm (section 2.6).

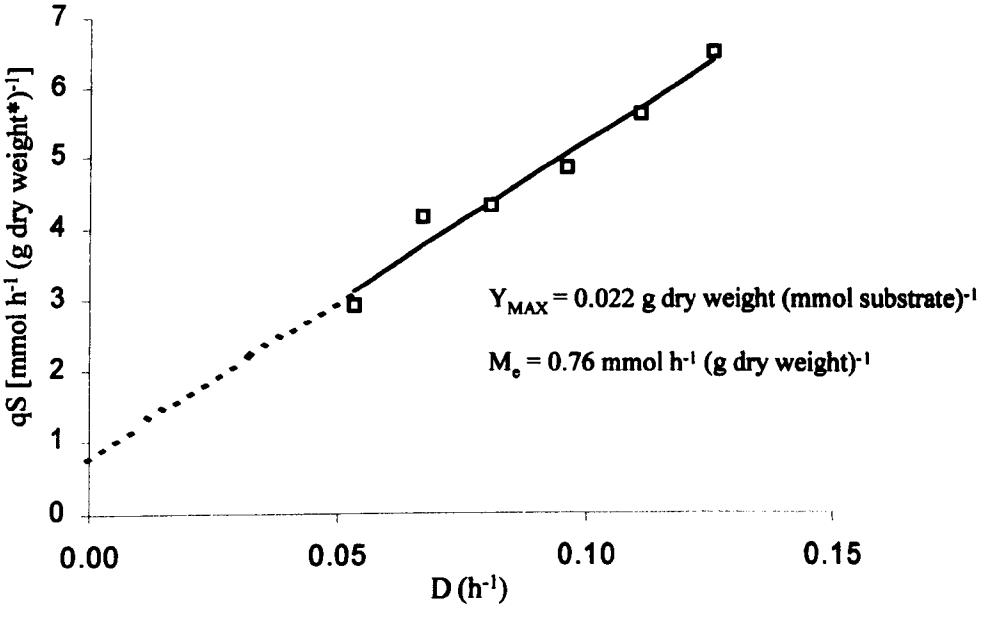


Fig. 5.6. Glucose consumption rates (q_S) of MC4100 (wt) in anaerobic glucose-limited chemostat cultivation in standard minimal medium (section 2.5.3.2) with pH maintained at pH 6.5. Medium supplemented as described in Fig. 5.1. Steady states were obtained in glucose (20 mM)-limited chemostat cultures over a range of dilution rates (h^{-1}). Key to abbreviations: D, dilution rate; q_S, rate of substrate (glucose) utilisation; Y_{MAX}, maximum biomass yield; M_e, maintenance energy. Dry weight estimated from OD at 430 nm (section 2.6).

D (h^{-1})	Dry weight* ($g\ l^{-1}$)	q_s (mmol h^{-1}(g dry weight*$^{-1}$))
0.05	0.34	3.00
0.06	0.35	3.77
0.08	0.34	4.55
0.09	0.38	4.99
0.11	0.39	5.72

Table 5.6. Physiological parameters of JRG3621 (*ΔhyfB-R*) in anaerobic glucose-limited chemostat cultivation in standard minimal medium (section 2.5.3.2) with pH maintained at pH 6.5. Medium supplemented as described in Fig. 5.1. Steady states were obtained in glucose (20 mM)-limited chemostat cultures over a range of dilution rates (h^{-1}). Key to abbreviations: D, dilution rate; q_S, rate of substrate (glucose) utilisation. *Dry weight estimated from OD at 430 nm (section 2.6).

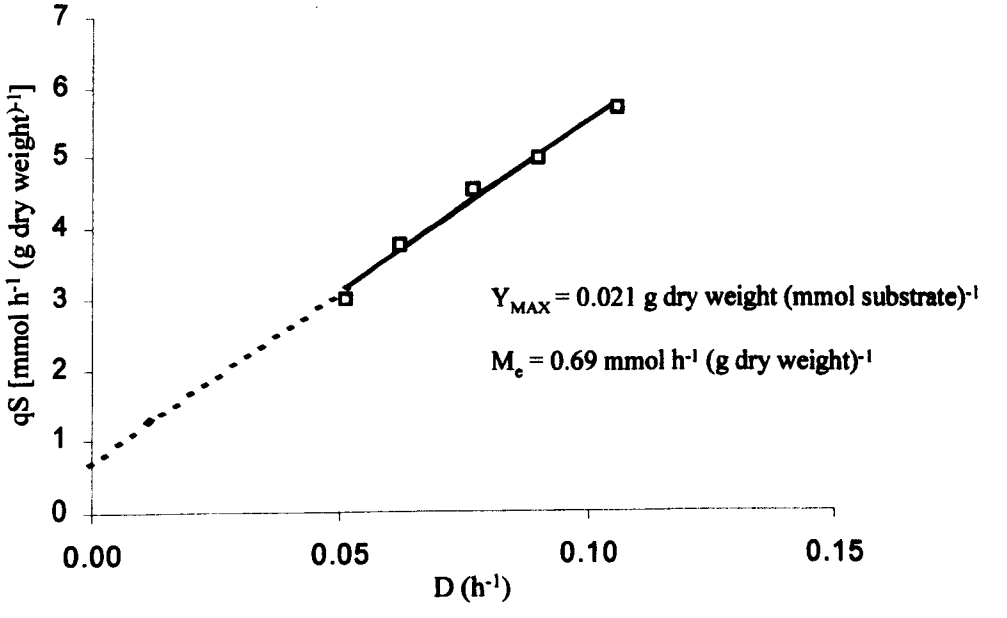


Fig. 5.7. Glucose consumption rates (q_S) of JRG3621 (*ΔhyfB-R*) in anaerobic glucose-limited chemostat cultivation in standard minimal medium (section 2.5.3.2) with pH maintained at pH 6.5. Medium supplemented as described in Fig. 5.1. Steady states were obtained in glucose (20mM)-limited chemostat cultures over a range of dilution rates (h^{-1}). Key to abbreviations: D, dilution rate; q_S, rate of substrate (glucose) utilisation; Y_{MAX}, maximum biomass yield; M_e, maintenance energy. Dry weight estimated from OD at 430 nm (section 2.6).

6.5 the *hyf* operon has no major effect on the growth efficiency of *E. coli* and that the small differences observed between MC4100 (wt) and JRG3621 (*ΔhyfB-R*) during anaerobic batch cultivation (section 5.3.1) were not significant.

The rate of substrate utilisation (q_S; section 2.10.2) is a measure of the amount of carbon substrate required by the bacterial culture to produce unit biomass in unit time. During glucose limited continuous cultivation the rate of substrate utilisation (q_S) for strains MC4100 (*wt*) and JRG3621 (*ΔhyfB-R*) was increased approximately three fold by a shift from aerobic to anaerobic conditions (Tables 5.3, 5.4, 5.5 and 5.6). This expected decrease in growth efficiency during anaerobic growth is due to the lower energy yields of fermentation compared to aerobic respiration.

5.6 Expression of *hyfA-lacZ* during growth in the N_2 flushed bioreactor

To achieve anaerobic conditions during glucose-limited chemostat cultivation the bioreactor was flushed with N_2 (400 – 500 $cm^3.min^{-1}$) and agitation was maintained at 150 rpm. To ensure that the *hyf* operon was expressed under N_2 flushing conditions strain DS5 (MC4100, λ*hyfA-lacZ bla*) was grown in controlled anaerobic batch culture at pH 6.5 (as described in section 5.2). During lag/early exponential phase any fermentation gases produced were left to accumulate in the bioreactor headspace and pressure build up was allowed to escape via a tube running from the bioreactor to an inverted water filled measuring cylinder (gas trap). At early/mid exponential phase the bioreactor was flushed with N_2 (500 $cm^3.min^{-1}$) to recreate conditions observed in the bioreactor during anaerobic glucose-limited chemostat cultivation. Agitation was maintained at 150 rpm throughout the experiment. The culture was sampled every hour during lag and exponential phase and at several points during stationary phase. As a control, strain DS5 (MC4100, λ*hyfA-lacZ bla*) was grown anaerobically in Bijous (8 ml) filled with TYEP (pH 6.6) supplemented with glucose (0.4%) (as described in section 2.6.1.3). Growth was monitored by measuring the OD at 430 nm of diluted culture samples with time. Each sample was also assayed for β-galactosidase activity (as described in section 2.8.5.1). This control was used in the expression experiments reported in chapter four of this thesis,

as a standard against which levels of β-galactosidase activity detected could be compared. In the same way, it has been used in this experiment, so that the levels of β-galactosidase activity detected in the bioreactor can be compared with the levels detected in chapter four of this thesis.

Flushing the bioreactor with N_2 (500 $cm^3.min^{-1}$) immediately halted the increase in *hyfA-lacZ* expression observed (Fig. 5.8). This dip in expression during mid exponential phase was not observed for the expression experiments reported in chapter four of this thesis and must be a direct result of flushing the bioreactor with N_2. Despite this, *hyf* expression was not reduced by N_2 flushing and a further increase in expression was observed at the onset of stationary phase. An approximately two fold increase in β-galactosidase activity was observed for controlled anaerobic batch cultivation in standard minimal medium over anaerobic growth in Bijous (8 ml) filled with TYEP (pH 6.6) plus glucose (0.4%) (control). This two fold increase is consistent with the two fold induction in expression by minimal medium previously described in section 4.10.2. Therefore we can conclude that the *hyf* operon was expressed in the bioreactor whilst it was flushed with N_2 (500 $cm^3.min^{-1}$).

Since the highest levels of *hyf* expression are observed during stationary phase (section 4.9, Fig. 5.8), it could be argued that *hyf* would not be expressed/active during continuous cultivation. However the level of induction of stationary phase markers has been shown to be as high during growth in carbon-limited continuous culture as in carbon starved batch cultures (Notley & Ferenci, 1996). Notley and Ferenci (1996) concluded that cellular metabolism during growth in carbon-limited chemostat continuous culture at low dilution rates (D = 0.05 h^{-1}) is comparable to that observed for carbon starved batch cultures during stationary phase growth.

5.7 Summary

Anaerobic controlled batch cultivation of MC4100 (wt) and JRG3621 (Δ*hyfB-R*) at pH 6.5 and 7.5 revealed no differences in growth and metabolism, which could be attributed to the *hyf* operon (section 5.2). However, deletion of *hycE* (HD705), encoding the large subunit of hydrogenase-3 a component of Fhl-1, had an effect on

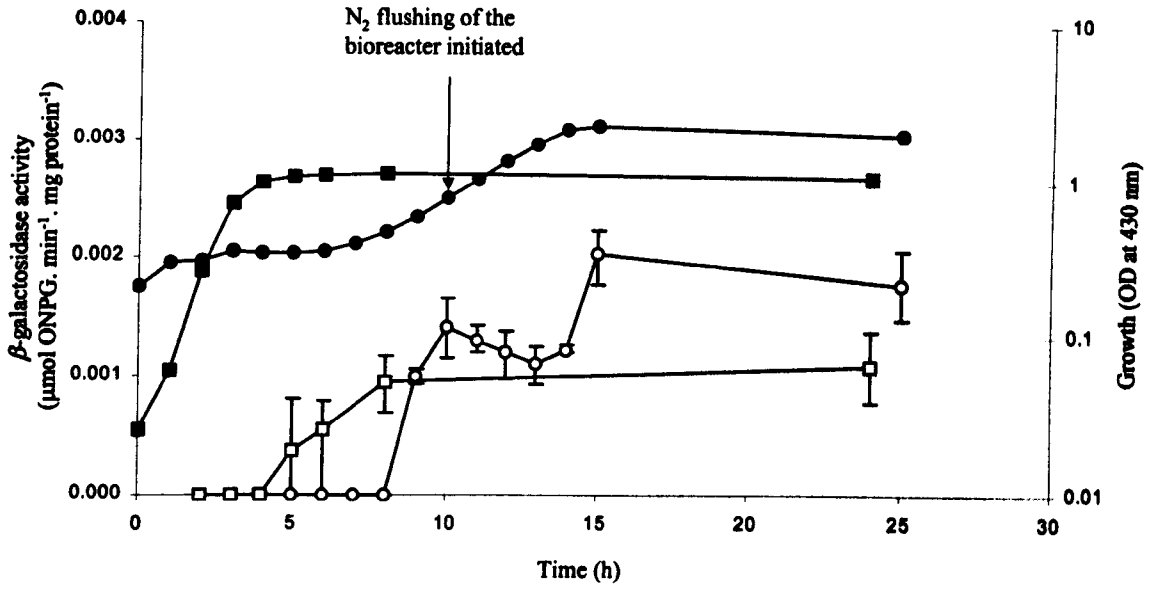


Fig. 5.8. Expression of *hyfA-lacZ* in strain DS5 (*wt*) during anaerobic batch cultivation in standard minimal medium (section 2.5.3.2) containing 20 mM glucose with pH maintained at 6.5.

Medium supplemented as described in Fig. 5.1.

Squares, DS5 (wt) grown anaerobically in 8 ml Bijous filled with TYEP (pH 6.6) plus 0.4% glucose (control); circles, DS5 (wt) grown in anaerobic controlled batch culture in standard minimal medium.

Filled symbols, growth; clear symbols, expression.

Bioreactor (circles) flushed with N_2 (500 $cm^3.min^{-1}$) between 10 h and 25 h.

Agitation maintained at 150 rpm throughout the experiment.

Values plotted are the average of at least two determinations and error bars give the maximal and minimal values.

growth and metabolism at pH 6.5 (section 5.3). The *hycE* mutant (HD705) had no lag phase compared with a 7 – 8 h lag phase observed for strains MC4100 (*wt*) and JRG3621 (Δ*hyfB-R*). I suggest that differences in the growth phase of the overnight inoculum, despite attempts made to standardise it, are likely to have contributed to these differences in the lag phase length. Formate levels detected for the *hycE* mutant (HD705) at the onset of stationary phase were increased two fold over levels observed for the wildtype strain (MC4100), however these levels were reduced during stationary phase despite deletion of the *hycE* gene and no synthesis of Fhl-1 in this strain. Important future work would be to investigate whether this reduction in formate is significant and if so whether it is still observed in a *hyc/hyf* double mutant.

Aerobic and anaerobic glucose limited chemostat cultivation of MC4100 (wt) and JRG3621 (Δ*hyfB-R*) at pH 6.5 revealed no differences in growth and metabolism, which could be attributed to the *hyf* operon (sections 5.4, 5.5 & 5.6).

6. GENERAL DISCUSSION

Sequence analysis of the 55.8-56.0 min region of the *E. coli* genome revealed a twelve-gene operon designated the *hyf* operon (*hyfABCDEFGHIRfocB*) (section 1.3; Andrews *et al.*, 1997). This operon encodes a putative nine-subunit hydrogenase complex (hydrogenase four or Hyf), a potential formate- and σ^{54}-dependent transcriptional activator, HyfR, and a possible formate transporter, FocB (Andrews *et al.*, 1997). Andrews and co-workers (1997) proposed that Hyf together with formate dehydrogenase-H (Fdh-H) formed a proton translocating formate hydrogenlyase (Fhl-2), that HyfR is a formate-dependent regulator of the *hyf* operon and that FocB provides the Fhl-2 complex with external formate as substrate. These proposals have been investigated and the results reported in chapters three, four and five of this thesis.

Bagramyan and co-workers (2000, unpublished) detected ATP dependent Fdh-H activity and H_2 production, attributable to the *hyf* operon in whole cells grown anaerobically, at slightly alkaline pH (pH 7.5) in the absence of formate (section 1.3.6). In contrast to the proposals of Andrews and co-workers (1997), Bagramyan and co-workers (unpublished) proposed that Fdh-H and Hyf combine to form a formate hydrogenlyase (Fhl-2) that is driven by a proton gradient established by F_0F_1-ATPase.

The aim of this study has been to investigate the proposals of Andrews and co-workers and Bagramyan and co-workers by conducting analysis into the function and regulation of the *hyf* operon of *E. coli*.

A *hyfA-lacZ* transcriptional fusion strain was used to study the transcriptional regulation of the *hyf* operon, the results of which are reported in chapter four of this thesis. Optimal transcriptional levels from the *hyf* operon were observed when *E. coli* was grown anaerobically, at acidic pH, in the presence of formate and in the absence of exogenous electron acceptors (sections 4.9 & 4.10.1). This pattern of transcriptional expression is identical to that observed for the *hyc* operon (and other genes of the formate regulon) and like the *hyc* operon induction of expression by all

these factors is channelled via the intracellular concentration of formate and is mediated by the σ^{54}-dependent transcriptional activator FhlA (sections 1.1.4.4 & 4.11). Transcription from the *hyf* operon was also enhanced in minimal medium and repressed by the addition of complex medium components, specifically yeast extract (section 4.10.2). This induction of transcription by minimal medium was absent in *fhlA* mutants suggesting that availability of formate is elevated in this medium (section 4.11). Interestingly expression of the *hyf* operon is also growth phase dependent with optimal transcriptional activation observed at the onset of stationary phase (section 4.9). The formation of the proposed Fhl-2 complex is thought to require the availability, the uptake and the incorporation of nickel, iron, molybdenum and selenium. The presence of nickel chloride in the growth medium did not significantly affect expression of the *hyf* operon, however the presence of the metal ion chelator 2-2'-dipyridyl was found to reduce expression (sections 4.10.3 & 4.10.4). Expression of the *hyc* operon and *fdhF* gene is enhanced by the metal ion molybdate, an induction that is dependent on FhlA (Self *et al.*, 2000). As FhlA regulates the *hyf* operon it is likely that the effect of 2-2'-dipyridyl is due to chelation of the available molybdate in the medium. However, the effect of iron, molybdate and selenium on the transcriptional activation of the *hyf* operon requires further investigation (section 4.10.4).

Deletion of *hyfR*, encoding the proposed σ^{54}-dependent transcriptional activator of the *hyf* operon did not affect *hyf* operon expression under the growth conditions tested. However introduction of a multicopy plasmid encoding *hyfR* under the control of the *lac* promoter increased *hyf* operon expression over a thousand fold (section 4.11). These results suggest that the *hyf* operon is regulated by HyfR but not under the growth conditions employed in this study. Also it is unclear what co-effector(s) is used by HyfR. It is possible that when HyfR is present at low levels, its activity is inhibited by an unknown factor that is titrated out when HyfR levels are higher. HyfR, like FhlA, is σ^{54}-dependent as introduction of the multicopy plasmid encoding *hyfR* failed to increase expression of the *hyf* operon in an *ntrA* mutant (section 4.14). A strongly predicted σ^{54}-dependent promoter located upstream of the transcriptional start site of the *hyf* operon has been confirmed by primer extension (P. Golby, unpublished). Multicopy *hyfR* enhanced *fdhF-lacZ*

transcriptional two fold but had no effect on transcriptional activation of *hycB-lacZ* (section 4.15).

It is possible that expression of the *hyf* operon is considerably increased under conditions that have yet to be identified, and an obvious target for future work would be to screen for these conditions. We must not however lose sight of the difference between transcriptional activation and protein expression. The fact is that mRNA is an intermediate between DNA and protein and increased mRNA levels do not always lead to greater protein expression. The translation of mRNA is influenced by factors such as translational initiation, codon usage, mRNA stability and protein turnover. Immunoblotting with anti-Hyf serum did not detect Hyf polypeptides in extracts from *E. coli* even when the currently recognised optimal transcriptional conditions were further increased 1000 fold by introducing a multicopy plasmid encoding *hyfR* (σ^{54} dependent transcriptional activator of the *hyf* operon) into the cells (section 3.7). If the products of the *hyf* operon are not detectable even when transcriptional levels are artificially increased to this extent then one must speculate whether the *hyf* operon is an evolutionary relic, which is no longer functionally expressed under any conditions. Since transcription from the *hyf* operon shows signs of regulation (activated by FhlA and HyfR) I suggest that the operon is more likely to be cryptic. Genes are termed cryptic when we have not found conditions under which they are expressed (Hinton, 1997; Tamburini & Mastromei, 2000). In many cases these genes are phenotypically silent under the experimental conditions used and expression of these genes is often induced in the natural environment (Tamburini & Mastromei, 2000). Microbiologists studying *E. coli* have tended to focus on what the bacterium needs to thrive in a few specialised niches, with too much emphasis on complex medium at 37 °C (Hinton, 1997). The environmental stimuli required for the expression of particular cryptic genes may be identified by doing extensive studies with a barrage of unusual chemicals, substrates or conditions (Hinton, 1997). An example of a cryptic gene that was subsequently shown to be inducible is *celF*. Originally, expression of the *celF* gene could only be shown at high temperatures, which caused cell death (Droffner & Yamamato, 1992). However, by screening random luciferase fusions, Guzzo and Dubow (1994) showed that *celF* could be induced by nickel. This approach to identify environmental stimuli that induce the *hyf* operon has been attempted but yielded no positive results. The substrates

sorbitol, ethanol, gluconate, glucuronate and acetate did not noticeably induce *hyf* transcription (section 4.10.5). Obviously more intensive studies are required and although this approach could be considered 'hit and miss', it could provide a useful source of clues towards the specific function of the *hyf* operon. Again it must be stressed that although transcriptional analysis may generate clues, other molecular biological approaches are necessary to confirm expression and investigate function.

The results reported in chapter three of this thesis conflict with the work reported by Bagramyan and co-workers (2000, unpublished) as hydrogenase assays, Fdh-H assays, gas production experiments and H_2 production assays, established to investigate the phenotype of the *hyf* operon, detected no activity or gas/H_2 production attributable to the *hyf* operon and the proposed Fhl-2 complex (sections 3.9, 3.10 & 3.11). These experiments were conducted with whole cells grown and assayed under a number of different conditions including those under which Bagramyan and co-workers (2000, unpublished) detected a phenotype for the *hyf* operon and those under which optimal activation of *hyf* operon transcription was observed. Also it appears that not only is *hyf*-attributable activity not detectable but the products of the *hyf* operon are either not synthesised under these conditions or synthesised at very low levels. Immunoblotting experiments with anti-HycE (73% identity to HyfG) and anti-Hyf serum did not detect Hyf polypeptides in extracts from *E. coli* grown under optimal transcriptional activation conditions (sections 3.6 & 3.7). Nickel incorporation experiments did not detect HyfG in extracts from *E. coli* grown under optimal transcriptional activation conditions despite its high similarity to HycE including conserved residues acting as ligands for the [Ni-Fe] centre (section 3.8). This difficulty in detecting Hyf polypeptides is not surprising as expression of the *hyf* operon is very low, even during growth under optimal expression conditions (*hyf* expression approximately 100 fold less than that of the *hyc* operon; estimated from data presented in chapter four). Also, it appears that the *hyf* operon is transcribed as a single transcript (approximately 14000 bp), which would make it prone to fragmentation *in vivo*. Finally, the genes of the *hyf* operon possess codon usages similar to those of very weakly expressed *E. coli* structural genes.

It is difficult to remark on the contradictory nature of these results with those reported by Bagramyan and co-workers (2000, unpublished), but some basic comments can be made. Firstly, despite the fact that optimal transcriptional levels

from the *hyf* operon are observed at acidic pH in the presence of formate the studies of Bagramyan and co-workers (unpublished, 2000) propose that the *hyf* encoded Fhl-2 complex is only functional at slightly alkaline pH and in the absence of exogenous formate. This is surprising considering that even under optimal transcriptional activation conditions expression is too low to detect *hyf*-encoded polypeptides by immunoblotting or nickel incorporation. Secondly, the basis of the work presented by Bagramyan and co-workers (2000, unpublished), is the detection of reduced Fdh-H activity and H_2 production in a *hyf* mutant strain compared with the wildtype strain. They have not, however, reported the detection of any Fhl and/or hydrogenase specific activity in a hydrogenase-1, -2 and –3 triple mutant. The failure of Sauter and co-workers (1992) to detect benzylviologen dependent hydrogenase activity in the Δ*hya* Δ*hyb* Δ*hycB-H* hydrogenase triple mutant (HDJ123) has already been discussed (section 1.3.2) and was attributed to a possible polar effect exerted by the chloramphenicol cassette (contained in the Δ*hycB-H* mutation) on *hycI* expression. HycI is proposed to be required for processing and maturation of the hydrogenase-4 large subunit, HyfG. Also, Bagramyan and co-workers (2000, unpublished) proposed that the *hyf* dependent activity they detected was also dependent on the 16 Fe ferredoxin, HycB, whose gene is deleted in HDJ123. However, no *hyf* dependent Fdh-H activity, hydrogenase activity or gas/H_2 production has been reported for another hydrogenase triple mutant, FTD147 (Δ*hyaB* Δ*hybC* Δ*hycE*), which carries in frame deletions solely in the genes encoding the hydrogenase-1, -2 and –3 large subunits (section 3.9.1 & 3.11).

In chapter five of this thesis bioreactors were used to investigate the phenotype of the *hyf* operon. Anaerobic controlled batch cultivation of MC4100 (*wt*) and JRG3621 (Δ*hyfB-R*) at pH 6.5 and 7.5 revealed no differences in growth and metabolism, which could be attributed to the *hyf* operon (section 5.2). Since transcriptional activation of the *hyf* operon was induced at the onset of stationary phase, use of bioreactors was particularly useful as carbon-limited chemostat continuous culture at low dilution rates ($D = 0.05\ h^{-1}$) is comparable to that observed for carbon starved batch cultures during stationary phase growth (Notley & Ferenci, 1996). However aerobic and anaerobic glucose limited chemostat cultivation of MC4100 (*wt*) and JRG3621 (Δ*hyfB-R*) at pH 6.5 also revealed no differences in

growth and metabolism attributable to the *hyf* operon (sections 5.4 & 5.5). These results are not surprising taking into account that immunoblotting with anti-Hyf serum (section 3.7) did not detect Hyf polypeptides under these growth conditions. However the fact that anaerobic controlled batch cultivations of MC4100 (*wt*) and JRG3621 (Δ*hyfB-R*) at pH 7.5 revealed no differences in growth and metabolism, further questions the results of Bagramyan and co-workers (2000, unpublished). Anaerobic controlled batch cultivation of HD705 (Δ*hycE*) at pH 6.5 detected levels of formate at the onset of stationary phase approximately two fold greater than levels observed for the wildtype strain (MC4100) (section 5.3). This increase was expected as the *hycE* gene encodes the large subunit of hydrogenase-3 and is a component of Fhl-1 which catalyses the disproportionation of formate to carbon dioxide and dihydrogen during fermentation. Therefore it is surprising that these formate levels detected for HD705 (Δ*hycE*) were reduced during the course of stationary phase. Future work is needed to investigate whether this reduction in formate is significant and if so whether it is still observed in a *hyc/hyf* double mutant and therefore attributable to a *hyf* encoded Fhl-2.

To investigate the possible role of FocB in formate transport, a *focB* knockout mutant was created by replacing the *focB* gene with a Δ*focB*::*spc* allele (P. Golby, unpublished). A *focA focB* double mutant was created by transducing this *focB* mutation into a strain carrying a *focA* point mutation (REK701; Suppmann & Sawers, 1994) (P. Golby, unpublished). The growth of these *focA, focB* and *focA focB* mutants was compared to that of the wildtype under fermentative conditions with and without formate. No significant differences in growth or organic acid production were detected for the four strains (P. Golby, unpublished). This is in contrast to the work of Suppmann and Sawers (1994), which detected a reduced level of formate production for the *focA* mutant relative to the wildtype. This discrepancy could be due to the unstable nature of the *focA* point mutation.

Taken together the results presented in this thesis provide no phenotype for the *hyf* mutants. Transcriptional analysis however supports a function of the *hyf* operon in formate metabolism and the proposals of Andrews and co-workers (1997)

that the *hyf* operon, together with *fdhF*, encodes a second Fhl complex (Fhl-2) in *E. coli* which is energy conserving.

'Homologous genes in the same organism whose products perform related but not identical functions' have been termed 'paralogues' (Koonin *et al.*, 1996). *E. coli* possesses a high proportion of paraloguous genes (approximately half the genes in *E. coli* form clusters of paralogues) and this is probably related to its ability to adapt to novel environments (Koonin *et al.*, 1996; Hinton, 1997). In fact many genes which play a role in formate metabolism in *E. coli* are paralogous. Fdh-N and Fdh-O are two homologous formate dehydrogenases possessed by *E. coli* (section 1.1.2). Fdh-N is synthesised during anaerobiosis when nitrate is available, whilst the *fdoGHI-fdhE* operon encoding Fdh-O is less well characterised but transcriptional levels are induced threefold by aerobic growth and about twofold by anaerobic growth in the presence of nitrate (Berg & Stewart, 1990; Abaibou *et al.*, 1995). Similarly *E. coli* possesses two homologous nitrate respiratory systems, Nr-A and Nr-Z (section 1.1.2). Nr-A together with Fdh-N forms the formate-nitrate respiratory pathway and like Fdh-N is synthesised in anaerobiosis when nitrate is available (Stewart, 1982). The *narZYWV* operon encoding Nr-Z is expressed in the early stage of the stationary phase of cell growth and is highly dependent on the stationary phase regulator RpoS (σ^S) (Chang, 1999). It has been proposed that Fdh-O and Nr-Z encode a second formate-nitrate respiratory pathway in *E. coli*, which serves to allow the cell to harness nitrate more efficiently as an alternative electron acceptor under aerobic, stress-associated conditions (Chang *et al.*, 1999). An alternative function for this second formate-nitrate respiratory pathway has been proposed in facilitating rapid adaption to anaerobic conditions pending the induction of Fdh-N and Nr-A (Abaibou *et al.*, 1995). Either way the genes encoding these formate-nitrate respiratory pathways are examples of the paralogues *E. coli* possesses which allow it to adapt to novel environments. The uptake hydrogenases of *E. coli* can also be considered to be paralogues (section 1.1.5). Hydrogenase-2 has a respiratory function and is the principle uptake hydrogenase during growth on non-fermentable carbon sources such as fumarate (Ballantine & Boxer, 1985; Sawers *et al.*, 1985; Menon *et al.*, 1994). However transcriptional activation of the *hyaABCDEF* operon encoding hydrogenase-1 is induced by anaerobiosis and is growth phase dependent (expression increased during stationary phase) and dependent on the stationary phase

regulator RpoS (σ^S) (Atlung *et al.*, 1997). The function of hydrogenase-1 remains to be resolved definitively, but roles in recycling the hydrogen produced by Fhl-1 and maintaining the proton potential of the cytoplasmic membrane in response to stress have been proposed (Sawers *et al.*, 1985, 1986; King & Przybyla, 1999). The genes of the *hyc* and *hyf* operons of *E. coli* can similarly be considered paralogues. However unlike the other paralogues discussed here, no phenotype has been identified for the *hyf* operon and regulation of both operons appears to be very similar. The pioneering work of August Böck and co-workers which characterised the function and regulation of the *hyc* operon provided a template for this study into the *hyf* operon around which many of the core experiments reported in this thesis were based. Prehaps if future work took a different approach and focused on finding differences between these two operons, more clues into the function of the *hyf* operon would be revealed.

The *hyc* operon was identified by mapping and sequencing a Mu *d* (Ap *lac*) insertion mutation in Fhl pathway synthesis (Pecher *et al.*, 1993; Bohm *et al.*, 1990). The *hyf* operon in contrast was identified from sequence analysis of the 55.8 min region of the *E. coli* linkage map and its function can only be speculated upon (Andrews *et al.*, 1991; 1997). Now that the complete *E.coli* genome has been sequenced and the completion of many other microbial genome sequences is continuing apace (Lucchini *et al.*, 2001, report the completion of 37 genome sequences with 142 still in progress) new approaches to molecular biology have arisen to facilitate the study of this genetic information (Blattner *et al.*, 1997). The terms 'functional genomics', 'transcriptomics' and 'proteomics' have arisen to describe the large scale application of mass mutagenesis, gene expression profiling and global protein analysis (Wasinger *et al.*, 1995; Velculescu *et al.*, 1997; Luccini *et al.*, 2001). Studying transcriptional activation at the genomic scale has been achieved with DNA microarrays. These are glass slides containing the entire genome as an ordered mosaic of either oligonucleotides or PCR products representing individual genes (Lucchini *et al.*, 2001). Sampling of mRNA after subjecting bacteria to certain environmental conditions, followed by hybridisation to the microarray produces a 'gene expression profile' or 'signature' for the microbe under the conditions studied (Hinton, 1997; Lucchini *et al.*, 2001). Microarrays can be used to gain clues to gene function by comparing the 'gene expression profile' in

wildtype and knockout mutant strains (Lucchini *et al.*, 2001). This approach is particularly useful for predicted regulatory genes and could quickly lead to the identification of the members of the regulon. Clues to the specific function of the *hyf* operon may be gained by studying the gene expression profile of a *hyfR* knockout mutant or for a wildtype strain transformed with the multicopy plasmid carrying the *hyfR* gene. This genome-wide approach is likely to reveal the functions of many genes, including those of the *hyf* operon, which have been missed by more conventional approaches (Lucchini *et al.*, 2001).

REFERENCES

Abaibou, H., Pommier, J., Benoit, S., Giordano, G. & Mandrand-Berthelot, M.-A. (1995). Expression and characterisation of the *Escherichia coli fdo* locus and a possible physiological role for aerobic formate dehydrogenase. *Journal of Bacteriology* **177**, 7141-7149.

Abdel-Hamid, A. M., Attwood, M. M. & Guest, J. R. (2001). Pyruvate oxidase contributes to the aerobic growth efficiency of *Escherichia coli*. *Microbiology* **147**, 1483-1498.

Altuvia, S., Weinstein-Fischer, D., Zhang, A., Postow, L. & Storz, G. (1997). A small, stable RNA induced by oxidative stress: role as a pleiotrophic regulator and antimutator. *Cell* **90**, 43-53.

Altuvia, S., Zhang, A., Argaman, L., Tiwari, A. & Storz, G. (1998). The *Escherichia coli* OxyS regulatory RNA represses *fhlA* translation by blocking ribosome binding. *The EMBO Journal* **17**, 6069-6075.

Andrews, S. C., Berks, B. C., McClay, J., Andrew, A., Quail, M. A., Golby, P. & Guest, J. R. (1997). A 12-cistron *Escherichia coli* operon (*hyf*) encoding a putative proton-translocating formate hydrogenlyase system. *Microbiology* **143**, 3633-3647.

Argaman, L. & Altuvia, S. (2000). *fhlA* repression by OxyS RNA: Kissing complex formation at two sites results in a stable antisense-target RNA complex. *Journal of Molecular Biology* **300**, 1101-1112.

Atlung, T., Knudsen, K., Heerfordt, L. & Brondsted, L. (1997). Effects of σ^S and the transcriptional activator AppY on induction of the *Escherichia coli hya* and *cbdAB-appA* operons in response to carbon and phosphate starvation. *Journal of Bacteriology* **179**, 2141-2146.

Atlung, T., Nielsen, A. & Hansen, F. G. (1989). Isolation charcterisation and nucleotide sequence of *appY*, a regulatory gene for growth-phase-dependent gene expression in *Escherichia coli*. *Journal of Bacteriology* **171**, 1683-1691.

Augier, V., Asso, M., Guigliarelli, B., More, C., Bertand, P., Santini, C. L., Blasco, F., Chippaux, M. & Giordano, G. (1993b). Removal of the high potential (4Fe-4s) centre of the b-subunit from *Escherichia coli* nitrate reductase. Physiological, biochemical, and EPR chacterisation of site-directed mutated enzymes. *Biochemistry* **32**, 5099-5108.

Augier, V., Guigliarelli, B., Asso, M., Bertrand, P., Frixon, C., Giordano, G., Chippaux, M. B. & Blasco, F. (1993a). Site directed mutagenesis of conserved cysteine residues within the b subunit of *Escherichia coli* nitrate reductase. Physiological, biochemical, and EPR characterisation of the mutated enzymes. *Biochemistry* **32**, 2013-2023.

Ausubel (1989). Preparation of genomic DNA from bacteria. In *Current Protocols in Molecular Biology*, pp. Unit 2.4. Edited by F. K. Ausubel & others. New York: John Wiley and Sons.

Axley, M. J., Grahame, D. A. & Stadtman, T. C. (1990). *Escherichia coli* Formate-Hydrogen Lyase: purification and properties of the selenium-dependent formate dehydrogenase component. *Journal of Biological Chemistry* **265**, 18213-18218.

Bagramyan, K., Mnatsakanyan, N., Poladian, A., Vassilian, A. & Trchounian, A. (unpublished). The roles of hydrogenases 3 and 4, and the F_0F_1-ATPase, in H_2 production by *Escherichia coli* at alkaline and acidic pH. .

Bagramyan, K., Vassilian, A., Mnatsakanyan, N. & Trchounian, A. (2000). Participation of hydrogenase 4, encoded by hyf operon, in liberation of molecular hydrogen and proton-potassium exchange by *Escherichia coli*. *Biologicheskie Membrany* **17**, 604-615.

Ballard, A. L. & Ferguson, S. J. (1988). Respiartory nitrate reductase from *Paracoccus denitrificans*. Evidence for two *b*-type haems in the g subunit and properties of a water soluble enzyme containing a and b subunits. *European Journal of Biochemistry* **174**, 207-212.

Ballantine, S. P. & Boxer, D. H. (1985). Nickel-containing hydrogenase isoenzymes from anaerobically grown *Escherichia coli* K-12. *Journal of Bacteriology* **163**, 901-908.

Ballantine, S. P. & Boxer, D. H. (1986). Isolation and charcterisation of a soluble active fragment of hydrogenase isoenzyme 2 from the membranes of anaerobically grown *Escherichia coli*. *European Journal of Biochemistry* **156**, 277-284.

Bartolome, B., Jubete, Y., Martinez, E. & De La Cruz, F. (1991). Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene* **102**, 75-78.

Begg, Y. A., Whyte, J. N. & Haddock, B. A. (1977). The identification of mutants of *Escherichia coli* deficient in formate dehydrogenase and nitrate reductase activities using dye indicator plates. *FEMS Microbiology Letters* **2**, 47-50.

Benoit, S., Abaibou, H. & Mandrand-Berthelot, M.-A. (1998). Topological analysis of the Aerobic Membrane-Bound Formate Dehydrogenase of *Escherichia coli*. *Journal of Bacteriology* **180**, 6625-6634.

Berg, B. L., Li, J., Heider, J. & Stewart, V. (1991). Nitrate inducible formate dehydroegnase in *Escherichia coli* K-12: nucleotide sequence of the *fdnGHI* operon and evidence that opal (UGA) encodes selenocysteine. *Journal of Biological Chemistry* **266**, 22380-22385.

Berg, B. L. & Stewart, V. (1990). Structual genes for nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. *Genetics* **125**, 691-702.

Bergmeyer, H.-U. & Bernt, E. (1965). α-D-glucose. Determination with Glucose Oxidase and Peroxidase. In *Methods of Enzymic Analysis*, pp. 123-130. Edited by H.-U. Bergmeyer. New York and London: Verlag Chemie Academic Press, Inc.

Berks, B. C. (1996). A common export pathway for proteins binding complex redox cofactors? *Molecular Microbiology* **22**, 393-404.

Berks, B. C., Page, M. D., Richardson, D. J., Reilly, A., Cavill, A., Outen, F. & Ferguson, S. J. (1995). Sequence analysis of subunits of the membrane-bound nitrate reductase from a denitrifying bacterium: the integral membrane subunit provides a prototype for the dihaem electron-carrying arm of a redox loop. *Molecular Microbiology* **15**, 319-331.

Beutler, H.-O. (1984). In *Methods of Enzymic Analysis*, pp. 598-606. Edited by H.-U. Bergmeyer. Verlag chemie, Weinheim, Deerfield Beach/ Florida, Basel.

Bilous, P. T., Cole, S. T., Anderson, W. F. & Weiner, J. H. (1988). Nucleotide sequence of the *dmsABC* operon encoding the anaerobic DMSO reductase of *Escherichia coli*. *Molecular Microbiology* **2**, 785-796.

Bilous, P. T. & Weiner, J. H. (1985). Proton translocation coupled to dimethyl sulfoxide reduction in anaerobically grown *Escherichia coli* HB101. *Journal of Bacteriology* **163**, 369-375.

Bilous, P. T. & Weiner, J. H. (1988). Molecular cloning and expression of the *Escherichia coli* dimethyl sulphoxide reductase operon. *Journal of Bacteriology* **170**, 1511-1518.

Birkmann, A. & Böck, A. (1989). Characterization of a cis regulatory DNA element necessary for formate induction of the formate dehydrogenase gene (*fdhF*) of *Echerichia coli*. *Molecular Microbiology* **3**, 187-195.

Birkmann, A., Hennecke, H. & Böck, A. (1989). Construction of chimaeric promoter regions by exchange of the upstream regulatory sequences from *fdh* and *nif* genes. *Molecular Microbiology* **3**, 697-703.

Birkmann, A., Sawer, R. G. & Böck, A. (1987b). Involvement of the *ntrA* gene product in the anaerobic metabolism of *Escherichia coli*. *Molecular and General Genetics* **210**, 535-542.

Birkmann, A., Zinoni, F., Sawers, R. G. & Böck, A. (1987a). Factors affecting transcriptional regulation of the formate-hydrogen-lyase pathway of *Escherichia coli*. *Archives of Microbiology* **148**, 44-51.

Blasco, F., Dos Santos, J.-P., Magalon, A., Frixon, C., Guigliarelli, B., Santini, C.-L. & Giordano, G. (1998). NarJ is a specific chaperone required for molybdenum cofactor assembly in nitrate reductase A of *Escherichia coli*. *Molecular Microbiology* **28**, 435-447.

Blasco, F., Iobbi, C., Giordano, G., Chippaux, M. & Bonnefoy, V. (1989). Nitrate reductase of *Escherichia coli*: Completion of the nucleotide sequence of the *nar* operon and reassessment of the role of the alpha and beta subunits in iron binding and electron transfer. *Molecular and General Genetics* **218**, 249-256.

Blasco, F., Iobbi, C., Ratouchniak, J., Bonnefoy, V. & Chippaux, M. (1990). Nitrate reductases of *Escherichia coli*: Sequence of the second nitrate reductase and comparison with that encoded by the *narGHJI* operon. *Molecular and General Genetics* **222**, 104-111.

Blattner, F. R., Plunkett, I., G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shau, Y. (1997). The complete genome sequence of *Escherichia coli*. *Science* **277**, 1453-1462.

Blokesch, M., Magalon, A. & Böck, A. (2001). Interplay between the specific chaperone-like proteins HybG and HypC in maturation of hydrogenases 1, 2 and 3 from *Escherichia coli*. *Journal of Bacteriology* **183**, 2817-2822.

Böck, A. & Sawers, G. (1996). Fermentation. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, pp. 262-282. Edited by F. C. Neidhardt & others. Washington DC: American Society for Microbiology.

Bohm, R., Sauter, M. & Böck, A. (1990). Nucleotide sequence and expression of an operon in *Escherichia coli* coding for formate hydrogenlyase components. *Molecular Microbiology* **4**, 231-243.

Bonnefoy, V., Burini, J. F., Giordano, G., Pascal, M. C. & Chippaux, M. (1987). Presence in the 'silient' terminus region of the *Escherichia coli* K12 chromosome of cryptic gene(s) encoding a new nitrate reductase. *Molecular Microbiology* **1**, 143-150.

Bonnefoy, V. & DeMoss, J. A. (1994). Nitrate reductases in *Escherichia coli*. *Antonie van Leeuwenhoek* **66**, 47-56.

Bonnefoy-Orth, V., Lepelletier, M., Pascal, M. C. & Chippaux, M. (1981). Nitrate reductase and cytochrome $b_{\text{nitrate reductase}}$ structural genes as parts of the nitrate reductase operon. *Molecular and General Genetics* **181**, 535-540.

Bott, M. & Thauer, R. K. (1989). Proton translocation coupled to the oxidation of carbon monoxide to CO_2 and H_2 in *Methanosarchina barkeri*. *European Journal of Biochemistry* **179**, 469-472.

Boyington, J. C., Gladyshev, V. N., Khangulov, S. V., Stadtman, T. C. & Sun, P. D. (1997). Crystal structure of formate dehydrogenase H: catalysis involving Mo, molybdopterin, selenocysteine, and an Fe4S4 cluster. *Science* **275**, 1305-1308.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254.

Braun, M., Bungert, S. & Friedrich, T. (1998). Characterisation of the overproduced NADH dehydrogenase fragment of the NADH: Ubiquinone oxidoreductase (complex I) from *Escherichia coli. Biochemistry* **37**, 1861-1867.

Breton, J., Berks, B. C., Reilly, A., Thomson, A. J., Ferguson, S. J. & Richardson, D. J. (1994). Characterisation of the paramagnetic iron-containing redox centres of *Thiosphaera pantotropha* periplasmic nitrate reductase. *FEBS Letters* **345**, 76-80.

Brondsted, L. & Atlung, T. (1994). Anaerobic regulation of the hydrogenase 1 (*hya*) operon of *Escherichia coli. Journal of Bacteriology* **176**, 5423-5428.

Brooke, A. G., Watling, E. M., Attwood, M. M. & Tempest, D. W. (1989). Environmental control of metabolic fluxes in thermotolerant methylotrophic *Bacillus* strains. *Archives of Microbiology* **151**, 268-273.

Casadaban, M. J. & Cohen, S. N. (1979). Lactose genes fused to exogenous promoters in one step using a Mu-*lac* bacteriophage: in vivo probe for transcriptional control sequences. *Proceedings of the National Academy of Sciences USA* **76**, 4530-4533.

Casalot, L. & Rousset, M. (2001). Maturation of the [Ni-Fe] hydrogenases. *Trends in Microbiology* **9**, 228-237.

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

Chang, L. (1999). Expression of the *Escherichia coli* Nr-Z nitrate reductase is highly growth phase dependent and is controlled by RpoS, the alternative vegetative sigma factor. *Molecular Microbiology* **34**, 756-766.

Chaudhry, G. R. & MacGregor, C. H. (1983a). *Escherichia coli* nitrate reductase subunit A: its role as the catalytic site and evidence for its modification. *Journal of Bacteriology* **154**, 387-394.

Chaudhry, G. R. & MacGregor, C. H. (1983b). Cytochrome *b* from *Escherichia coli* nitrate reductase. Its properties and association with enzyme complex. *Journal of Biological Chemistry* **258**, 5819-5827.

Clark, D. P. (1989). The fermentation pathways of *Escherichia coli. FEMS Microbiological Reviews* **63**, 223-234.

Cotter, P. A. & Gunsalus, R. P. (1989). Oxygen, nitrate, and molybdenum regulation of *dmsABC* gene expression in *Escherichia coli. Journal of Bacteriology* **171**, 3817-3823.

Cox, J. C., Edwards, E. S. & DeMoss, J. A. (1981). Resolution of distinct selenium-containing formate dehydrogenases from *Escherichia coli*. *Journal of Bacteriology* **145**, 1317-1324.

Darwin, A., Hussain, H., Griffiths, L., Grove, J., Sambongi, Y., Busby, S. & Cole, J. (1993a). Regulation and sequence of the structural gene for cytochrome c552 from *Escherichia coli*: not a hexahaem but a 50 kDa tetrahaem nitrite reductase. *Molecular Microbiology* **9**, 1255-1265.

Darwin, A., Tormay, P., Page, L., Griffiths, L. & Cole, J. (1993b). Identification of the formate dehydrogenases and genetic determinants of the formate-dependent nitrite reduction by *Escherichia coli* K12. *Journal of General Microbiology* **139**, 1829-1840.

De Pina, K., Desjardin, V., Mandrand-Berthelot, M.-A., Giordano, G. & Wu, L.-F. (1999). Isolation and characterisation of the *nikR* gene encoding a nickel-responsive regulator in *Escherichia coli*. *Journal of Bacteriology* **181**, 670-674.

De Pina, K., Navarro, C., McWalter, L., Boxer, D. H., Price, N. C., Kelly, S. M., Mandrand-Berthelot, M.-A. & Wu, L.-F. (1995). Purification and characterisation of the periplasmic nickel-binding protein NikA of *Escherichia coli* K12. *European Journal of Biochemistry* **227**, 857-865.

Drapal, N. & Böck, A. (1998). Interaction of the hydrogenase accessory protein HypC with hycE, the large subunit of *Escherichia coli* hydrogenase 3 during maturation. *Biochemistry* **37**, 2941-2948.

Droffner, M. L. & Yamamoto, N. (1992). Demonstration of cell operon expression of *Escherichia coli*, *Salmonella typhimurium* and *Pseudomonas aeruginosa* at elevated temperatures refactory to their growth. *Applied Environmental Microbiology* **58**, 1784-1785.

Dross, F., Geisler, V., Lenger, R., Theis, F., Kraft, T., Fahrenholz, F., Kojro, E., Duchene, A., Tripier, D., Juvenal, K. & Kroger, A. (1992). The quinone-reactive Ni/Fe hydrogenase of *Wolinella succinogenes*. *European Journal of Biochemistry* **206**, 93-102.

Earhart, C. F. (1996). Uptake and Metabolism of Iron and Molybdenum. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, pp. 1075-1089. Edited by F. C. Neidhardt & others. Washington DC: American Society for Microbiology.

Eaves, D. J., Grove, J., Staudenmann, W., James, P., Poole, R. K., White, S. A., Griffiths, I. & Cole, J. A. (1998). Involvement of the products of the *nrfEFG* genes in the covalent attachment of haem *c* to a novel cysteine-lysine motif in the cytochrome c_{552} nitrite reductase from *Escherichia coli*. *Molecular Microbiology* **28**, 205-216.

Edwards, E. S., Rondeau, S. S. & DeMoss, J. A. (1983). *chlC(nar)* operon of *Escherichia coli* includes structural genes for alpha and beta subunits of nitrate reductase. *Journal of Bacteriology* **153**, 1513-1520.

Enoch, H. G. & Lester, R. L. (1974). The role of a novel cytochrome b-containing nitrate reductase and quinone in the *in vitro* reconstruction of formate-nitrate reductase activity of *E. coli*. *Biochemical and Biophysical Research Communications* **61**, 1234-1241.

Enoch, H. G. & Lester, R. L. (1975). The purification and properties of formate dehydrogenase and nitrate reductase from *Escherichia coli*. *The Journal of Biological Chemistry* **250**, 6693-6705.

Falk-Krzesinski, H. J. & Wolfe, A. J. (1998). Genetic analysis of the *nuo* locus, which encodes the proton-translocating NADH dehydrogenase in *Escherichia coli*. *Journal of Bacteriology* **180**, 1174-1184.

Fearnley, I. M. & Walker, J. E. (1992). Conservation of sequences of subunits of mitochondrial complex I and their relationships with other proteins. *Biochimica et Biophysica Acta* **1140**, 105-134.

Forchhammer, K., Boesmiller, K. & Böck, A. (1991a). The function of selenocysteine synthase and SELB in the synthesis and incorporation of selenocysteine. *Biochemie* **73**, 1481-1486.

Forchhammer, K., Leinfelder, W. & Böck, A. (1989). Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. *Nature* **342**, 453-456.

Forchhammer, K., Leinfelder, W., Boesmiller, K., Veprek, B. & Böck, A. (1991b). Selenocysteine synthase from *Escherichia coli*: nucleotide sequence of the gene (*selA*) and purification of the protein. *Journal of Biological Chemistry* **266**, 6318-6323.

Friedmann, D. H. (1988). Integration Host Factor- A protein for all reasons. *Cell* **55**, 545-554.

Friedrich, T., Steinmuller, K. & Weiss, H. (1995). The proton-pumping complex I of bacteria and mitochondria and its homologue in chloroplasts. *FEBS Letters* **367**, 107-111.

Friedrich, T., Weidner, U., Nehls, U., Fecke, W., Schneider, R. & Weiss, H. (1993). Attempts to define distinct parts of NADH:ubiquinone reductase. *Journal of Bioenergetics and Biomembranes* **25**, 331-337.

Gennis, R. B. & Stewart, V. (1996). Respiration. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, pp. 217-261. Edited by F. C. Neidhardt. Washington DC: American Society for Microbiology.

Giordano, G., Medani, C.-L., Mandrand-Berthelot, M.-A. & Boxer, D. H. (1983). Formate dehydrogenases from *Escherichia coli*. *FEMS Microbiological Letters* **17**, 171-171.

Golby, P., Davies, S., Kelly, D. J., Guest, J. R. & Andrews, S. C. (1999). Identification and characterisation of a two-component sensor-kinase and response-regulator system (DcuS-DcuR) controlling gene expression in response to C4-dicarboxylates in *Escherichia coli*. *Journal of Bacteriology* **181**, 1238-1248.

Gon, S., Patte, J.-C., Mejean, V. & Iobbi-Nivol, C. (2000). The *torYZ* (*yecK bisZ*) operon encodes a third respiratory trimethylamine *N*-oxide reductase in *Escherichia coli*. *Journal of Bacteriology* **182**, 5779-5786.

Grunden, A. M., Ramesh, M. R., Rosentel, J. K., Healy, F. G. & Shanmugam, K. T. (1996). Repression of the *Escherichia coli modABCD* (molybdate transport) operon by ModE. *Journal of Bacteriology* **178**, 735-744.

Grunden, A. M., Ray, R. M., Rosentel, J. K., Healy, F. G. & Shanmugam, K. T. (1996). Repression of the *Escherichia coli modABCD* (molybdate transport) operon by ModE. *Journal of Bacteriology* **178**, 735-744.

Grunden, A. M. & Shanmugam, K. T. (1997). Molybdate transport and regulation in bacteria. *Arch. Microbiol* **168**, 345-354.

Guest, J. R., Angier, S. J. & Russell, G. C. (1989). Structure, expression and protein engineering of the pyruvate dehydrogenase complex of *Escherichia coli*. *Annals of the New York Academy of Sciences* **573**, 76-99.

Guzzo, A. & DuBow, M. S. (1994). A *luxAB* transcriptional fusion to the cryptic *celF* gene of *Escherichia coli* displays increased luminescence in the presence of nickel. *Molecular and General Genetics* **242**, 455-460.

Hamamoto, T., Hashimoto, M., Hino, M., Kitada, M., Seto, Y., Kudo, T. & Horikoshi, K. (1994). Characterisation of a gene responsible for the Na^+/H^+ antiporter system of alkalophilic *Bacillus* species strain C-125. *Molecular Microbiology* **14**, 939-946.

Hamilton, C. M., Aldea, M., Washburn, B. K., Babitzke, P. & Kusher, S. R. (1989). New method for generating deletions and gene replacements in *Escherichia coli*. *Journal of Bacteriology* **171**, 4617-4622.

Hanahan, D. (1985). In *DNA cloning - A practical approach*, pp. 109-135. Edited by D. M. Glover. Oxford: IRL Press.

Happe, R. P., Roseboom, W., Pierik, A. J. & Albracht, S. P. J. (1997). Biological activation of hydrogen. *Nature* **385**, 126.

Hasona, A., Ray, R. M. & Shanmugam, K. T. (1998b). Physiological and genetic analysis leading to identification of a biochemical role for the *moeA* (molybdate

metabolism) gene product in *Escherichia coli*. *Journal of Bacteriology* **180**, 1466-1472.

Hasona, A., Self, W. T., Ramesh, M. R. & Shanmugam, K. T. (1998). Molybdate-dependent transcription of *hyc* and *nar* operons of *Escherichia coli* requires MoeA protein and ModE-molybdate. *FEMS Microbiological Letters* **169**, 111-116.

Hasona, A., Self, W. T., Ray, R. M. & Shanmugam, K. T. (1998a). Molybdate-dependent transcription of *hyc* and *nar* operons of *Escherichia coli* requires MoeA protein and ModE-molybdate. *FEMS Microbiological Letters* **169**, 111-116.

Hengge-Aronis, R. (1996). Regulation of gene expression during entry into stationary phase. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, pp. 1497-1512. Edited by F. C. Neidhardt. Washington DC: American Society for Microbiology Press.

Hinton, J. C. D. (1997). The *Escherichia coli* genome sequence: the end of an era or the start of the FUN? *Molecular Microbiology* **26**, 417-422.

Hiramatsu, T., Kodama, K., Kuroda, T., Mizushima, T. & Tsuchiya, T. (1998). A putative multisubunit Na^+/H^+ antiporter from *Staphylococcus aureus*. *Journal of Bacteriology* **180**, 6642-6648.

Hirschmann, J., Wang, P. K., Sei, K., Keener, J. & Kustu, S. (1985). Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription *in vitro*: evidence that the *ntrA* product is a σ factor. *Proceedings of the National Academy of Sciences USA* **82**, 7525-7529.

Höpner, T. & Knappe, J. (1974). In *Methods of Enzymic Analysis*, pp. 1551-1555. Edited by H.-U. Bergmeyer. New York and London: Verlag Chemie, Weinheim/Academic press, Inc.

Hopper, S., Babst, M., Schlensog, V., Fischer, H.-M., Hennecke, H. & Böck, A. (1994). Regulated expression *in vitro* of genes coding for formate hydrogenlyase components of *Escherichia coli*. *Journal of Biological Chemistry* **269**, 19597-19604.

Hopper, S. & Böck, A. (1995). Effector-mediated stimulation of ATPase activity by the σ54-dependent transcriptional activator FhlA from *Escherichia coli*. *Journal of Bacteriology* **177**, 2798-2803.

Hopper, S., Korsa, I. & Böck, A. (1996). The nucleotide concentration determines the specificity of in vitro transcription activation by the σ54-dependent activator FhlA. *Journal of Bacteriology* **178**, 199-203.

Hunt, T. P. & Magasanik, B. (1985). Transcription of *glnA* by purified *E. coli* components: RNA polymerase and the products of *glnF*, *glnG* and *glnL*. *Proceedings of the National Academy of Sciences USA* **82**, 8453-8457.

Hussain, H., Grove, J., Griffiths, L., Busby, S. & Cole, J. (1994). A seven-gene operon essential for formate-dependent nitrite reduction to ammonia by enteric bacteria. *Molecular Microbiology* **12**, 153-163.

Iobbi-Nivol, C., Santini, C.-L., Blasco, F. & Giordano, G. (1990). Purification and further characterisation of the second nitrate reductase of *Escherichia coli* K12. *European Journal of Biochemistry* **188**, 679-687.

Ito, M., Guffanti, A. A., Oudega, B. & Krulwich, T. A. (1999). *mrp*, a multigene, multifunctional locus in *Bacillus subtilis* with roles in resistance to cholate and to Na^+ and in pH homeostasis. *Journal of Bacteriology* **181**, 2394-2402.

Iuchi, S. & Lin, E. C. C. (1987). The *narL* gene product activates the nitrate reductase operon and represses the fumarate reductase and trimethylamine *N*-oxide reductase operons in *Escherichia coli*. *Proceedings of the National Academy of Sciences USA* **84**, 3901-3905.

Jacobi, A., Rossmann, R. & Böck, A. (1992). The *hyp* operon gene products are required for the maturation of catalytically active hydrogenase isoenzymes in *Escherichia coli*. *Archives of Microbiology* **158**, 444-451.

Johann, S. & Hinton, S. M. (1987). Cloning and nucleotide sequence of the *chlD* locus. *Journal of Bacteriology* **169**, 1911-1916.

Johnson, R. C., Ball, C. A., Pfeffer, D. & Simon, M. I. (1988). Isolation of the gene encoding the Hin recombinational enhancer binding protein. *Proceedings of the National Academy of Sciences USA* **85** 3484-3488.

Jones, H. M. & Gunsalus, R. P. (1985). Transcription of the *Escherichia coli* fumarate reductase genes (*frdABCD*) and their coordinate regulation by oxygen, nitrate and fumarate. *Journal of Bacteriology* **164**, 1100-1109.

Jones, H. M. & Gunsalus, R. P. (1987). Regulation of *Escherichia coli* fumarate reductase (*frdABCD*) operon expression by respiratory electron acceptors and the *fnr* gene product. *Journal of Bacteriology* **169**, 3340-3349.

Jones, R. W. (1980a). Proton translocation by the membrane-bound formate dehydrogenase of *Escherichia coli*. *FEMS Microbiology Letters* **8**, 167-171.

Jones, R. W. (1980b). The role of membrane-bound hydrogenase in the energy-conserving oxidation of molecular hydrogen by *Escherichia coli*. *Biochemical Journal* **188**, 345-350.

Jones, R. W. & Garland, P. B. (1977). Sites and specificity of the reaction of bipyridilium compounds with anaerobic respiratory systems of *Escherichia coli*. Effects of permeability barriers imposed by the cytoplasmic membrane. *Biochemical Journal* **164**, 199-211.

Jones, R. W., Lamont, W. A. & Garland, P. B. (1980). The mechanism of proton translocation driven by the respiratory nitrate reductase complex in *Escherichia coli*. *Biochemical Journal* **190**, 79-94.

Jourlin, C., Bengrine, A., Chippaux, M. & Mejean, V. (1996a). An unorthodox sensor protein (TorS) mediates the induction of the *tor* structural genes in response to trimethylamine *N*-oxide in *Escherichia coli*. *Molecular Microbiology* **20**, 1297-1306.

Jourlin, C., Simon, G., Pommier, J., Chippaux, M. & Mejean, V. (1996b). The periplasmic TorT protein is required for trimethylamine *N*-oxide reductase gene induction in *Escherichia coli*. *Journal of Bacteriology* **178**, 1219-1223.

Kaiser, M. & Sawers, G. (1994). Pyruvate formate-lyase is not essential for nitrate respiration by *Escherichia coli*. *FEMS Microbiology Letters* **117**, 163-168.

Kerby, R. L., Ludden, P. W. & Roberts, G. P. (1995). Carbon monoxide-dependent growth of *Rhodospirillum rubrum*. *Journal of Bacteriology* **177**, 2241-2244.

Kessler, D. & Knappe, J. (1996). Anaerobic dissimilation of pyruvate. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, pp. 199-205. Edited by F. C. Neidhardt & others. Washington DC: American Society for Microbiology.

King, P. W. & Przybyla, A. E. (1999). Response of *hya* expression to external pH in *Escherichia coli*. *Journal of Bacteriology* **181**, 5250-5256.

Knappe, J. & Sawers, G. (1990). A radical-chemical route to acetyl-CoA: the anaerobically induced pyruvate formate-lyase system of *Escherichia coli*. *FEMS Microbiology Reviews* **75**, 383-398.

Kolesnikow, T., Schroder, I. & Gunsalus, R. P. (1992). Regulation of *narK* gene expression in *Escherichia coli* in response to anaerobiosis, nitrate, iron, and molybdenum. *Journal of Bacteriology* **174**, 7104-7111.

Koonin, E. V., Tatusov, R. L. & Rudd, K. E. (1996). Protein sequence comparison at genomic scale. *Methods Enzymol* **266**, 295-322.

Korsa, I. & Böck, A. (1997). Characterisation of *fhlA* mutations resulting in ligand-independent transcriptional activation and ATP hydrolysis. *Journal of Bacteriology* **179**, 41-45.

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

Lambden, P. R. & Guest, J. R. (1976). Mutants of *Escherichia coli* K12 unable to use fumarate as an anaerobic electron acceptor. *Journal of General Microbiology* **97**, 145-160.

Leif, H., Sled, V. D., Ohnishi, T., Weiss, H. & Friedrich, T. (1995). Isolation and charcterisation of the proton-translocating NADH: ubiquinone oxidoreductase from *Escherichia coli*. *European Journal of Biochemistry* **230**, 538-548.

Leinfelder, W., Forchhammer, K., Zinoni, F., Sawers, G., Mandrand-Berthelot, M.-A. & Böck, A. (1988a). *Escherichia coli* genes whose products are involved in selenium metabolism. *Journal of Bacteriology* **170**, 540-546.

Leinfelder, W., Zehelein, E., Mandrand-Berthelot, M.-A. & Böck, A. (1988b). Gene for a novel tRNA species that accepts L-serine and cotanslationally inserts selenocysteine. *Nature* **331**, 723-725.

Leinfelder, W. K., Forchhammer, B., Veprek, E., Zehelein, E. & Böck, A. (1990). *In vitro* synthesis of selenocysteinyl-$tRNA_{UCA}$ from seryl-$tRNA_{UCA}$: involvement and characterisation of the *selD* gene product. *Proceedings of the National Academy of Sciences USA* **87**, 543-547.

Lemire, B. D., Robinson, J. J., Bradley, R. D., Scraba, D. G. & Weiner, J. H. (1983). Structure of fumarate reductase on the cytoplasmic membrane of *Escherichia coli*. *Journal of Bacteriology* **155**, 391-397.

Lennox, E. S. (1955). Transduction of linked genetic characters of the host by bacteriophage P1. *Virology* **1**, 190-206.

Leonhartsberger, S., Ehrenreich, A. & Böck, A. (2000). Analysis of the domain structure and the DNA binding site of the transcriptional activator FhlA. *European Journal of Biochemistry* **267**, 3672-3684.

Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951). Protein measurements with the Folin phenol reagent. *Journal of Biological Chemistry* **193**, 265-275.

Lucchini, S., Thompson, A. & Hinton, J. C. D. (2001). Microarrays for microbiologists. *Microbiology* **147**, 1403-1414.

Lutz, S., Bohm, R., Beier, A. & Böck, A. (1990). Characterisation of divergent NtrA-dependent promoters in the anaerobically expressed gene cluster coding for hydrogenase 3 components of *Escherichia coli*. *Molecular Microbiology* **4**, 13-20.

Lutz, S., Jacobi, A., Schlensog, V., Bohm, R., Sawers, G. & Böck, A. (1991). Molecular characterisation of an operon (*hyp*) necessary for the activity of the three hydrogenase isoenzyme isoenzyme in *Escherichia coli*. *Molecular Microbiology* **5**, 123-135.

Magalon, A. & Böck, A. (2000a). Analysis of the hypC-HycE complex, a key intermediate in the assembly of the metal center of the *Escherichia coli* hydrogenase 3. *Journal of Biological Chemistry* **275**, 21114-21120.

Magalon, A. & Böck, A. (2000b). Dissection of the maturation reactions of the [Ni-Fe] hydrogenase 3 from *Escherichia coli* taking after nickel incorporation. *FEBS Letters* **473**, 254-258.

Magalon, A., Lemesle-Meunier, D., Rothery, R. A., Frixton, C., Weiner, J. H. & Blasco, F. (1997). Heme axial ligation by the highly conserved His residues in helix II of cytochrome *b* (NarI) of *Escherichia coli* nitrate reductase A (NarGHI). *Journal of Bacteriology* **272**, 25652-25658.

Maier, T., Binder, U. & Böck, A. (1996). Analysis of the *hydA* locus of *Escherichia coli*: two genes (*hydN* and *hypF*) involved in formate and hydrogen metabolism. *Archives of Microbiology* **165**, 333-341.

Maier, T., Jacobi, A., Sauter, M. & Böck, A. (1993). The product of the *hypB* gene, which is required for nickel incorporation into hydrogenases, is a novel guanine nucleotide-binding protein. *Journal of Bacteriology* **175**, 630-635.

Maier, T., Lottspeich, F. & Böck, A. (1995). GTP hydrolysis by HypB is essential for nickel insertion into hydrogenases of *Escherichia coli*. *European Journal of Biochemistry* **230**, 133-138.

Mandrand-Berthelot, M.-A., Couchoux-Luthaud, G., Santini, C.-L. & Giordano, G. (1988). Mutants of *Escherichia coli* specifically deficient in respiratory formate dehydroegnase activity. *Journal of General Microbiology* **134**, 3129-3139.

Mat-Jan, F., Alam, K. Y. & Clark, D. P. (1989). Mutants of *Escherichia coli* deficient in the fermentative lactate dehydrogenase. *Journal of Bacteriology* **171**, 342-348.

Maupin, J. A. & Shanmugam, K. T. (1990). Genetic regulation of formate hydrogenlyase of *Escherichia coli*: role of the *fhlA* gene product as a transcriptional activator for a new regulatory gene, *fhlB*. *Journal of Bacteriology* **172**, 4798-4806.

Mauplin-Furlow, J. A., Rosentel, J. K., Lee, J. H., Deppenmeier, V., Gunsalus, R. P. & Shanmugam, K. T. (1995). Genetic analysis of the ModABCD (molybdate transport) operon of *Escherichia coli*. *Journal of Bacteriology* **177**, 4851-4856.

McClellard, M., Florea, L., Sanderson, K., Clifton, S. W., Parkhill, J., Churder, C., Dougan, G., Wilson, R. K. & Miller, W. (2000). Comparison of *Escherichia coli* K-12 genome with sampled genomes of a *Pneumoniae* and three *Salmonella enterica* serovars, Typhimurium, Typhi and Paratyphi. *Nucleic Acids Research* **28**, 4974-4986.

McNicholas, P. M., Chiang, R. C. & Gunsalus, R. P. (1998). Anaerobic regulation of the *Escherichia coli dmsABC* operon requires the molybdate-responsive regulator ModE. *Molecular Microbiology* **27**, 197-208.

Mejean, V., Iobbi-Nivol, C., Lepelletier, M., Giordano, G., Chippaux, M. & Pascal, M.-C. (1994). TMAO anaerobic respiration in *Escherichia coli*: involvement of the *tor* operon. *Molecular Microbiology* **11**, 1169-1179.

Menon, N. K., Chatelus, C. Y., Devartanian, M., Wendt, J. C., Shanmugam, K. T., Peck, J. H. D. & Przybyla, A. E. (1994). Cloning, sequencing and mutational analysis of the *hyb* operon encoding *Escherichia coli* hydrogenase 2. *Journal of Bacteriology* **176**, 4416-4423.

Menon, N. K., Robbins, J., Peck, H. D., Chatelus, C. Y., Choi, E. S. & Przybyla, A. E. (1990). Cloning and sequencing of a putative Escherichia coli (NiFe) hydrogenase-1 operon containing six open reading frames. *Journal of Bacteriology* **172**, 1969-1977.

Menon, N. K., Robbins, J., Wendt, J. C., Shanmugam, K. T. & Przybyla, A. E. (1991). Mutational analysis and characterisation of the Escherichia coli *hya* operon, which encodes (NiFe) hydrogenase 1. *Journal of Bacteriology* **173**, 4851-4861.

Meuer, J., Bartoschek, S., Koch, J., Kunkel, A. & Hedderich, R. (1999). Purification and catalytic properties of Ech hydrogenase from *Methanosarcina barkeri*. *European Journal of Biochemistry* **265**, 325-335.

Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbour Laboratory, New York: Cold Spring Harbour Laboratory Press.

Morpeth, F. F. & Boxer, D. H. (1985). Kinetic analysis of respiratory nitrate reductase fron *Escherichia coli* K12. *Biochemistry* **24**, 40-46.

Motteram, P. A. S., McGarthy, J. E. G., Ferguson, S. J., Jackson, J. B. & Cole, J. A. (1981). Energy conservation during the formate-dependent reduction of nitrite by *Escherichia coli*. *FEMS Microbiology Letters* **12**, 317-320.

Nandi, R. & Sengupta, S. (1996). Involvement of anaerobic reductases in the spontaneous lysis of formate by immobilized cells of *Escherichia coli*. *Enzyme and Microbial Technology* **19**, 20-25.

Navarro, C., Wu, L.-F. & Mandrand-Berthelot, M.-A. (1993). The *nik* operon of *Escherichia coli* encodes a periplasmic binding-protein-dependent transport system for nickel. *Molecular Microbiology* **9**, 1181-1191.

Nohno, T., Kasai, Y. & Saito, T. (1988). Cloning and sequencing of the *Escherichia coli chlEN* operon involved in molybdopterin biosynthesis. *Journal of Bacteriology* **170**, 4097-4102.

Notley, L. & Ferenci, T. (1996). Introduction of RpoS-dependent functions in glucose-limited continuous culture: what level of nutrient limitation induces the stationary phase of *Escherichia coli*. *Journal of Bacteriology* **178**, 1465-1468.

Padan, E., Zilberstein, D. & Rottenberg, H. (1976). The proton electrochemical gradient in *Echerichia coli*. *European Journal of Biochemistry* **63**, 533-541.

Page, L., Griffiths, L. & Cole, J. A. (1990). Different physiological roles for two independent pathways for nitrite reduction to ammonia by enteric bacteria. *Archives of Microbiology* **154**, 349-354.

Palmer, T., Vasishta, A., Whitty, P. W. & Boxer, D. H. (1994). Isolation of protein FA, a product of the *mob* locus required for molybdenum cofactor biosynthesis in *Escherichia coli*. *European Journal of Biochemistry* **222**, 687-692.

Park, M. H., Wong, B. B. & Lusk, J. E. (1976). Mutants in three genes affecting transport of magnesium in *Escherichia coli* physiology and genetics. *Journal of Bacteriology* **126**, 1096-1103.

Pascal, M. C., Burini, J. F. & Chippaux, M. (1984). Regulation of trimethylamine *N*-oxide (TMAO) reductase of *Escherichia coli* K-12: analysis of *tor::*MudI fusion. *Molecular and General Genetics* **195**, 351-355.

Paveglio, M. T., Tang, J. S., Unger, R. E. & Barrett, E. L. (1988). Formate-nitrate respiration in *Salmonella typhimurium*: studies of two *rha*-linked *fdn* genes. *Journal of Bacteriology* **170**, 213-217.

Pecher, A., Zinoni, F. & Böck, A. (1985). The seleno-polypeptide of formic dehydrogenase (formate hydrogen-lyase linked) from *Escherichia coli*: genetic analysis. *Archives of Microbiology* **141**, 359-363.

Pecher, A., Zinoni, F., Jatisatienr, C., Wirth, R., Hennecke, H. & Böck, A. (1983). On the redox control of synthesis of anaerobically induced enzymes in enterobactericeae. *Archives of Microbiology* **136**, 131-136.

Peck, H. D. & Gest, H. (1957). Formic dehydrogenase and the hydrogenlyase enzyme complex in coli-aerogenes bacteria. *Biochemical Journal* **57**, 10-16.

Phillips-Jones, M. K., Watson, F. J. & Martin, R. (1993). The 3' codon context on UAG suppressor tRNA is different in *Escherichia coli* and human cells. *Journal of Molecular Biology* **233**, 1-6.

Pitterle, D. M., Johnson, J. L. & Rajagopalan, K. V. (1993). *In vitro* synthesis of molybdopterin from precursor Z using purified converting factor. Role of protein-bound sulfur in formation of the dithiolene. *Journal of Biological Chemistry* **268**, 13506-13509.

Plunkett, G., Burland, V., Daniels, D. L. & Blattner, F. R. (1993). Analysis of the *Escherichia coli* genome. III. DNA sequence of the region from 87.2 to 89.2 minutes. *Nucleic Acids Research* **21**, 3391-3398.

Pommier, J., Mandrand, M. A., Holt, S. E., Boxer, D. H. & Giordano, G. (1992). A second phenazine methosulphate-linked formate dehydrogenase isoenzyme in *Escherichia coli*. *Biochimica et Biophysica Acta* **1107**, 305-313.

Przybyla, A. E., Menon, N. K., Robbins, J., DerVartanian, L. & Peck, H. D. (1991). Further characterisation of the *hya* and *hyb* operons of *Escherichia coli*. In *3rd International Conference on the Molecular Biology of Hydrogenase*. Troia, Portugal.

Przybyla, A. E., Robbins, J., Menon, N. & Peck, H. D. (1992). Structure-function relationships among nickel-containing hydrogenases. *FEMS Microbiological Reviews* **8**, 109-135.

Rabin, R. S. & Stewart, V. (1993). Dual reponse regulators (NarL and NarP) interact with dual sensors (NarX and NarQ) to control nitrate- and nitrite-regulated gene expression in *Escherichia coli* K-12. *Journal of Bacteriology* **175**, 3259-3268.

Rajagopalan, K. V. (1996). Biosynthesis of the molybdenum cofactor. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, pp. 674-679. Edited by F. C. Neidhardt & others. Washington DC: American Society for Microbiology.

Rhode, M., Furstenau, V., Mayer, F., Przybyla, A. E., Peck, H. D., LeGall, J., Choi, E.-S. & Menon, N. K. (1989). Localisation of membrane-associated (NiFe) and (NiFeSe) hydrogenases of *Desulfovibrio vulgaris* using immunoelectron microscope procedures. *European Journal of Biochemistry* **180**, 421-427.

Richard, D. J., Sawers, G., Sargent, F., McWalter, L. & Boxer, D. H. (1999). Transcriptional regulation in response to oxygen and nitrate of the operons encoding the [NiFe] hydrogenases 1 and 2 of *Escherichia coli*. *Microbiology* **145**, 2903-2912.

Rodrigue, A., Boxer, D. H., Mandrand-Berthelot, M. A. & Wu, L.-F. (1996). Requirement for nickel of the transmembrane translocation of NiFe-hydrogenase 2 in *Escherichia coli*. *FEBS Letters* **392**, 81-86.

Rossmann, R., Maier, T., Lottspeich, F. & Böck, A. (1995). Characterisation of a protease from *Escherichia coli* involved in hydrogenase maturation. *European Journal of Biochemistry* **227**, 545-550.

Rossmann, R., Sauter, M., Lottspeich, F. & Böck, A. (1994). Maturation of the large subunit (HycE) of *Escherichia coli* hydrogenase 3 requires nickel incorporation followed by C-terminal processing at Arg537. *Journal of Biochemistry* **220**.

Rossmann, R., Sawers, G. & Böck, A. (1991). Mechanism of regulation of the formate-hydrogenlyase pathway by oxygen, nitrate, and pH: definition of the formate regulon. *Molecular Microbiology* **5**, 2807-2814.

Rothery, R. A. & Weiner, J. H. (1993). Topological characterisation of *Escherichia coli* DMSO reductase by electron paramagnetic resonance spectroscopy of an engineered [3Fe-4S] cluster. *Biochemistry* **32**, 5855-5861.

Rowe, J. J., Ubbink-Kok, T., Molenaar, D., Konings, W. N. & Driessen, J. M. (1994). NarK is a nitrite-extrusion system involved in anaerobic nitrate respiration by *Escherichia coli*. *Molecular Microbiology* **12**, 579-586.

Sambasivarao, D., Scraba, D. G., Trieber, C. & Weiner, J. H. (1990). Organisation of dimethyl sulfoxide reductase in the plasma membrane of *Escherichia coli*. *Journal of Bacteriology* **172**, 5938-5948.

Sambasivarao, D. & Weiner, J. H. (1991). Differentiation of the multiple *S*-oxide and *N*-oxide-reducing activities of *Escherichia coli*. *Current Microbiology* **23**, 105-110.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). RNase (DNase free). In *Molecular cloning : a laboratory manual*. Edited by J. Sambrook, E. F. Fritsch & T. Maniatis. Cold Spring Harbour, New York.

Sargent, F., Ballantine, P., Rugman, P. A., Palmer, T. & Boxer, D. H. (1998). Reassignment of the gene encoding the *Escherichia coli* hydrogenase 2 small subunit: Identification of a soluble precursor of the small subunit in a *hypB* mutant. *European Journal of Biochemistry* **255**, 746-754.

Sauter, M., Bohm, R. & Böck, A. (1992). Mutational analysis of the operon (*hyc*) determining hydrogenase 3 formation in *Escherichia coli*. *Molecular Microbiology* **6**, 1523-1532.

Sawers, G. (1994). The hydroegnases and formate dehydrogenases of *Escherichia coli*. *Antonie Leeuwenhoek* **66**, 57-88.

Sawers, G. & Böck, A. (1988). Anaerobic regulation of Pyruvate Formate-Lyase from *Escherichia coli* K-12. *Journal of Bacteriology* **170**, 5330-5336.

Sawers, G. & Böck, A. (1989). Novel transcriptional control of the pyruvate formate-lyase gene: upstream regulatory sequences and multiple promoters regulate anaerobic expression. *Journal of Bacteriology* **171**, 2485-2498.

Sawers, G., Heider, J., Zehelein, E. & Böck, A. (1991). Expression and operon structure of the *sel* genes of *Escherichia coli* and identification of a third selenium-containing formate dehydrogenase isoenzyme. *Journal of Bacteriology* **173**, 4983-4993.

Sawers, G. & Suppmann, B. (1992). Anaerobic induction of pyruvate formate-lyase gene expression is mediated by the ArcA and FNR proteins. *Journal of Bacteriology* **174**, 3474-3478.

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. *Journal of Bacteriology* **164**, 1324-1331.

Sawers, R. G. & Boxer, D. H. (1986). Purification and properties of membrane-bound hydrogenase isoenzyme 1 from anaerobically grown *Escherichia coli* K12. *European Journal of Biochemistry* **156**, 265-275.

Sawers, R. G., Jamieson, D. J., Higins, C. F. & Boxer, D. H. (1986). Characterisation and physiological roles of membrane-bound hydrogenase isoenzymes from *Escherichia coli*. *Journal of Bacteriology* **168**, 398-404.

Schaller, K.-H. & Triebig, G. (1984). In *Methods of Enzymic Analysis*, pp. 668-672. Edited by H.-U. Bergmeyer. Verlag Chemie, Weinheim, Deerfield Beach/ Florida, Basel.

Schlensog, V., Birkmann, A. & Böck, A. (1989). Mutations in trans which affect the anaerobic expression of a formate dehydrogenase (*fdhF*) structual gene. *Archives in Microbiology* **152**, 83-89.

Schlensog, V. & Böck, A. (1990). Identification and sequence analysis of the gene encoding the transcriptional activator of the formate hydrogenlyase system of *Escherichia coli*. *Molecular Microbiology* **4**, 1319-1327.

Schlensog, V., Lutz, S. & Böck, A. (1994). Purification and DNA-binding properties of FhlA, the transcriptional activator of the formate hydrogenlyase system from *Escherichia coli*. *Journal of Biological Chemistry* **269**, 19590-19596.

Schlindwein, C., Giordano, G., Santini, C.-L. & Mandrand-Berthelot, M.-A. (1990). Identification and expression of the *Escherichia coli fdhD* and *fdhE* gene involved in respiratory formate dehydrogenase formation in *Escherichia coli* K-12. *Gene* **97**, 147-147.

Self, W. T., Grunden, A. M., Hasona, A. & Shanmugam, K. T. (1999). Transcriptional regulation of molybdoenzyme synthesis in *Escherichia coli* in response to molybdenum: ModE-molybdate, a repressor of the *modABCD* (molybdate transport) operon is a secondary transcriptional activator of the *hyc* and *nar* operons. *Microbiology* **145**, 41-55.

Self, W. T. & Shanmugam, K. T. (2000). Isolation and characterisation of mutated FhlA proteins which activate transcription of the *hyc* operon (formate hydrogenlyase) of *Escherichia coli* in the absence of molybdate. *FEMS Microbiology Letters* **184**, 47-52.

Shanmugam, K. T., Stewart, V., Gunsalus, R. P., Boxer, D. H., Cole, J. A., Chippaux, M., DeMoss, J. A., Giordano, G., Lin, E. C. C. & Rajagopalan, K. V. (1992). Proposed nomenclature for the genes involved in molybdenum metabolism in *Escherichia coli* and *Salmonella typhimurium*. *Molecular Microbiology* **6**, 3452-3454.

Shingler (1996). Signal sensing by σ54-dependent regulators: derepression as a control mechanism. *Molecular Microbiology* **19**, 409-416.

Silva, P. J., van den Ban, E. C. D., Wassink, H., Haaker, H., de Castro, B., Robb, F. T. & Hagen, W. R. (2000). Enzymes of hydrogen metabolism in *Pyrococcus furiosus*. *European Journal of Biochemistry* **267**, 6541-6551.

Silvestro, A. J., Pommier, J. & Giordano, G. (1989). The inducible trimethylamine *N*-oxide reductase of *Escherichia coli* K12: its localisation and inducers. *Biochimica et Biophysica Acta* **999**, 208-216.

Simon, G., Mejean, V., Jourlin, C., Chippaux, M. & Pascal, M.-C. (1994). The *torR* gene of *Escherichia coli* encodes a response regulator protein involved in the expression of the trimethylamine *N*-oxide reductase genes. *Journal of Bacteriology* **176**, 5601-5606.

Simons, R. W., Houman, F. & Kleckner, N. (1987). Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene* **53**, 85-96.

Slonczewski, J. L. & Foster, J. W. (1996). pH-regulated genes and survival at extreme pH. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, pp. 1539-1549. Edited by F. C. Neidhardt & others. Washington DC: American Society for Microbiology.

Slonczewski, J. L., Rosen, B. P., Alger, J. R. & Macnab, R. M. (1981). pH homeostasis in *Escherichia coli*: measurement by ^{31}P nuclear magnetic resonance of methyl phosphonate and phosphate. *Proceedings of the National Academy of Sciences* **78**, 6271-6275.

Sodergren, E. J. & DeMoss, J. A. (1988). *narI* region of the *Escherichia coli* nitrate reductase (*nar*) operon contains two genes. *Journal of Bacteriology* **170**, 1721-1729.

Spiro, S. & Guest, J. R. (1987). Regulation and over-expression of the *fnr* gene of *Escherichia coli*. *Journal of General Microbiology* **133**, 3279-3288.

Stewart, V. (1988). Nitrate respiration in relation to facultative metabolism in enterobacteria. *Microbiological Reviews* **52**, 190-232.

Stewart, V. (1993). Nitrate regulation and anaerobic respiratory gene expression in *Escherichia coli*. *Molecular Microbiology* **9**, 425-434.

Stewart, V. & MacGregor, C. H. (1982). Nitrate reductase in *Escherichia coli* K12: involvement of *chlC, chlE* and *chlG* loci. *Journal of Bacteriology* **151**, 788-799.

Stoker, K., Reijnder, W. N. M., Oltmann, L. F. & Stouthamer, A. H. (1980). Initial cloning and sequencing of *hydGH*, an operon homologous to *nirBC* and regulating the labile hydrogenase activity in *Escherichia coli* K-12. *Journal of Bacteriology* **171**, 4448-4456.

Suppmann, B. & Sawers, G. (1994). Isolation and characterisation of hyophosphite-resistant mutants of *Escherichia coli*: identification of the FocA protein, encoded by the *pfl* operon, as a putative formate transporter. *Molecular Microbiology* **11**, 965-982.

Takagi, M., Tsuchiya, T. & Ishimoto, M. (1981). Proton translocation coupled to trimethylamine *N*-oxide reduction in anaerobically grown *Escherichia coli*. *Journal of Bacteriology* **148**, 762-768.

Tamburini, E. & Mastromei, G. (2000). Do bacterial cryptic genes really exist? *Research in Microbiology* **181**, 179-182.

Thauer, R. K., Hedderich, R. & Fischer, R. (1993). Reactions and enzymes involved in methanogenesis from CO_2 and H_2. In *Methanogenesis*. Edited by J. G. Ferry. New York: Chapman & Hall.

Theodoratou, E., Paschos, A., Magalon, A., Frische, E., Huber, R. & Böck, A. (2000). Nickel serves as a substrate recognition motif for the endopeptidase involved in hydrogenase maturation. *European Journal of Biochemistry* **267**, 1995-1999.

Tomiyama, M., Shiotani, M., Sode, K., Tamiya, E. & Karube, I. (1991). Nucleotide sequence analysis and expression control of *hydA* in *Escherichia coli*. In *Third International Conference on the Molecular Biology of Hydrogenases*. Troia, Portugal.

Tyson, K. L., Cole, J. A. & Busby, J. W. (1994). Nitrite and nitrate regulation at the promoters of two *Escherichia coli* operons encoding nitrite reductase: identification of common target heptamers for both NarP- and NarL-dependent regulation. *Molecular Microbiology* **13**, 1045-1055.

Velculescu, V. E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M. A., Bassett, D. E. J., Hieter, P., Vogelstein, B. & Kinzler, K. W. (1997). Characterisation of the yeast transcriptome. *Cell* **88**, 243-251.

Vishniac, W. & Santer, M. (1957). The thiobacilli. *Bacteriological Reviews* **21**, 195-209.

Volbeda, A., Charon, M.-H., Piras, C., Hatchikian, E. C., Frey, M. & Fontecilla-Camps, J. C. (1995). Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*. *Nature* **373**, 580-587.

Wackwitz, B., Bongaerts, J., Goodman, S. D. & Unden, G. (1999). Growth phase-dependent regulation of *nuoA-N* expression in *Escherichia coli* K-12 by the Fis protein: upstream binding sites and bioenergetic significance. *Molecular and general genetics* **262**, 876-883.

Walkenhorst, H. M., Hemschemeier, S. K. & Eichenlaub, R. (1995). Molecular analysis of the molybdate uptake operon, *modABCD*, of *Escherichia coli* and *modR*, a regulatory gene. *Microbiology Research* **150**, 347-361.

Wallace, B. J. & Young, I. G. (1977). Role of quinones in electron transport to oxygen and nitrate in *Escherichia coli*. Studies with a *ubiA menA* double quinone mutant. *Biochimica et Biophysica Acta* **461**, 84-100.

Wang, H. & Gunsalus, R. P. (2000). The *nrfA* and *nirB* nitrite reductase operon in *Escherichia coli* are expressed differently in response to nitrate than to nitrite. *Journal of Bacteriology* **182**, 5813-5822.

Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A. & authors, o. (1995). Progress with gene product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis* **16**, 1090-1094.

Weidner, U., Geier, S., Ptock, A., Freichrich, T., Leif, H. & Weiss, H. (1993). The gene locus of the proton-translocating NADH: Ubiquinone oxidoreductase in *Escherichia coli*. Organisation of the 14 genes and relationship between the derived proteins and subunits of mitochondrial complex I. *European Journal of Biochemistry* **233**, 109-122.

Weiss, D. S., Batut, J., Klose, K. E., Keener, J. & Kustu, S. (1991b). The phosphorylated form of the enhancer-binding protein NTRC has ATPase activity that is essential for activation of transcription. *Cell* **67**, 155-167.

Weiss, H., Friedrich, T., Hofhaus, G. & Preis, D. (1991a). The respiratory-chain NADH dehydrogenase (complex I) of mitochondria. *European Journal of Biochemistry* **197**, 563-579.

Wimpenny, J. W. T. & Cole, J. A. (1967). The regulation of metabolism in facultative bacteria. III. The effect of nitrate. *Biochimica et Biophysica Acta* **148**, 133-242.

Wissenbach, U., Ternes, D. & Unden, G. (1992). An *Escherichia coli* mutant containing only demethylmenaquinone, but no menaquinone: effects on fumarate, dimethylsulfoxide, trimethylamine N-oxide and nitrate respiration. *Archives of Microbiology* **158**, 68-73.

Wu, L.-F. & Mandrand-Berthelot, M.-A. (1986). Genetic and physiological characterisation of new *Escherichia coli* mutants impaired in hydrogenase activity. *Biochimie* **68**, 167-179.

Wu, L.-F., Mandrand-Berthelot, M. A., Waugh, R., Edmonds, C. J., Holt, S. E. & Boxer, D. H. (1989). Nickel deficiency gives rise to the defective hydrogenase phenotype of *hydC* and *fnr* mutants in *Escherichia coli*. *Molecular Microbiology* **3**, 1709-1718.

Yamamoto, I. & Ishimoto, M. (1977). Anaerobic growth of *Escherichia coli* on formate by reduction of nitrate, fumarate and trimethylamine N-oxide. *Zeitschrift fur Allgemeine Mikrobiologie* **17**, 235-242.

Zientz, E., Bongaerts, J. & Unden, G. (1998). Fumarate regulation of gene expression in *Escherichia coli* by the DcuSR (*dcuSR* genes) two-component regulatory system. *Journal of Bacteriology* **180**, 5421-5425.

Zilberstein, D., Agmon, V., Schuldiner, S. & Padan, E. (1984). *Escherichia coli* intracellular pH, membrane potential, and cell growth. *Journal of Bacteriology* **158**, 246-252.

Zinoni, F., Beier, A., Pecher, A., Wirth, R. & Böck, A. (1984). Regulation of the synthesis of hydrogenase (formate hydrogen-lyase linked) of *E. coli*. *Archives of Microbiology* **139**, 299-304.

Zinoni, F., Birkmann, A., Leinfelder, W. & Böck, A. (1987). Cotranslational insertion of selenocysteine into formate dehydrogenase from *Escherichia coli* deirected by a UGA codon. *Proceedings of the National Academy of Sciences USA* **84**, 3156-3160.

Zinoni, F., Birkmann, A., Stadtman, T. C. & Böck, A. (1986). Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from *Escherichia coli*. *Proceedings of the National Academy of Sciences USA* **83**, 4650-4654.

Appendix I

Nucleotide sequence of a 14837 bp fragment in the 55-57 min region of the *E. coli* chromosome showing translations of the 12 genes of the *hyf* operon and the two upstream open reading frames, *o177* and *bcp*.

Translation initiation sites (bold) and stop sites (asterisks) are indicated, as are the sequences of PCR primers used in this study (underlined) (Table 2.3). The EMBL accession number for the nucleotide sequence is M37689 (Andrews *et al.*, 1997).
To create strain JRG3615 (*ΔhyfA-B::spc*) the region between base position 1572 and 1979 was deleted and replaced with a spectinomycin cassette. To create strain JRG3618 (*ΔhyfR::spc*) the region between base position 12001 and 13583 was deleted and replaced with a spectinomycin cassette. To create strain JRG3621 (*ΔhyfB-R::spc*) the region between base position 3156 and 13874 was deleted and replaced with a spectinomycin cassette.

o177

```
1     ttg ggt gcc gat cgc cct gga att gtg aac acc atc acc cgt cat   45
1      L   G   A   D   R   P   G   I   V   N   T   I   T   R   H   15

46    gtc agt agt tgc ggc tgt aat att gaa gac agt cgc ctg gcg atg   90
16     V   S   S   C   G   C   N   I   E   D   S   R   L   A   M   30

91    ctg gga gaa gag ttc acg ttt att atg ctg ctt tcc ggt tca tgg   135
31     L   G   E   E   F   T   F   I   M   L   L   S   G   S   W   45

136   aat gcc att act ctg att gaa tca acg tta ccg ttg aaa ggt gcc   180
46     N   A   I   T   L   I   E   S   T   L   P   L   K   G   A   60

181   gaa ctg gat ctt tta atc gtg atg aag cgc acg acg gcg cgt ccg   225
61     E   L   D   L   L   I   V   M   K   R   T   T   A   R   P   75

226   cgt ccg cca atg cca gca tct gtc tgg gtt cag gtc gat gtg gca   270
76     R   P   P   M   P   A   S   V   W   V   Q   V   D   V   A   90

271   gac tcc ccg cat tta att gaa cgc ttc aca gca ctt ttc gac gcg   315
91     D   S   P   H   L   I   E   R   F   T   A   L   F   D   A   105

316   cat cat atg aac att gcg gag ctg gtg tcg cgc acg caa cct gct   360
106    H   H   M   N   I   A   E   L   V   S   R   T   Q   P   A   120

361   gaa aat gaa cgg gct gcg cag ttg cat att cag ata acc gcc cac   405
121    E   N   E   R   A   A   Q   L   H   I   Q   I   T   A   H   135

406   agc ccc gca tct cgg acg cag caa ata ttg agc aag cgt tca aag   450
136    S   P   A   S   R   T   Q   Q   I   L   S   K   R   S   K   150

451   ccc tat gta cag aac tca atg cac aag gca gta tta acg tcg tca   495
151    P   Y   V   Q   N   S   M   H   K   A   V   L   T   S   S   165

                                                    bcp
496   att att ccc aac atg atg aac agg atg gag tta agt aat gaa tcc   540
166    I   I   P   N   M   M   N   R   M   E   L   S   N   E   S   180
                                                        M   N   P   3
```

```
541   act gaa agc cgg tga tat cgc acc gaa att tag ctt gcc gga tcaa  586
181    T   E   S   R   *
4       L   K   A   G   D   I   A   P   K   F   S   L   P   D   Q   18

587   gac gga gaa caa gtt aat ttg acc gac ttc cag gga cag cgt gtt   631
19     D   G   E   Q   V   N   L   T   D   F   Q   G   Q   R   V   33

632   ctg gtt tat ttc tac ccg aaa gcc atg acc ccc ggc tgt acc gta   676
34     L   V   Y   F   Y   P   K   A   M   T   P   G   C   T   V   48
677   cag gcc tgc ggc tta cgc gat aac atg gat gag ttg aaa aaa gcg   721
49     Q   A   C   G   L   R   D   N   M   D   E   L   K   K   A   63

722   ggc gtt gat gtg ctg ggt atc agc acc gat aaa ccc gaa aaa ctc   766
64     G   V   D   V   L   G   I   S   T   D   K   P   E   K   L   78

767   tcc cgt ttt gcg gaa aaa gag ctg ctt aac ttt acg ctc ctg tct   811
79     S   R   F   A   E   K   E   L   L   N   F   T   L   L   S   93

812   gat gag gac cac cag gtg tgc gaa caa ttc ggc gtc tgg ggt gaa   856
94     D   E   D   H   Q   V   C   E   Q   F   G   V   W   G   E   108

857   aag tcc ttc atg ggc aaa acc tac gat ggc att cat cgc atc agc   901
109    K   S   F   M   G   K   T   Y   D   G   I   H   R   I   S   123

902   ttc ctg att gac gct gat ggc aaa atc gaa cat gtc ttt gac gat   946
124    F   L   I   D   A   D   G   K   I   E   H   V   F   D   D   138

                                                    hyfRB-1F
947   ttc aaa acc agc aat cac cac gac gtt gtg ctg aac tgg ctg aaa   991
139    F   K   T   S   N   H   H   D   V   V   L   N   W   L   K   153

992   gaa cac gcc tga tta ctt tgc tcc att ccg tgc tgg ctg cgc ttg  1036
154    E   H   A   *

1037  cgg cca gca tac ctc act tct cgt gat caa gat cac att ctc gct  1081

1082  ttc ccc tgc gac acg ggt gtc gaa tcc att ttt tgc tga acg tta  1126

1127  atg acc atc att ttt gta ccg ttc aga atc cag tta ata cat aac  1171

1172  tta ttg aat ata ttg agt taa tca gaa tgg cat cct tta tgc aat  1216

                                                          hyfA
1217  atg aaa tgc aat gtt tca tat cat ttt caa gga gcc gac atg aac  1261
                                                          M   N   2

1262  cgc ttt gtg gtg gcc gaa cca ctg tgg tgt aca gga tgt aat acc  1306
3      R   F   V   V   A   E   P   L   W   C   T   G   C   N   T   17

1307  tgt ctc gct gcc tgt tcg gac gtg cat aaa acg caa ggt tta cag  1351
18     C   L   A   A   C   S   D   V   H   K   T   Q   G   L   Q   32

1352  caa cac ccg cgc ctg gcc ctg gcg aag acg tca aca atc act gcc  1396
33     Q   H   P   R   L   A   L   A   K   T   S   T   I   T   A   47

1397  cct gtc gtg tgt cat cac tgt gag gaa gcc cct tgc ctg cag gtc  1441
48     P   V   V   C   H   H   C   E   E   A   P   C   L   Q   V   62

1442  tgc ccg gtc aat gcc atc tct cag agg gat gat gcg atc caa ctc  1486
63     C   P   V   N   A   I   S   Q   R   D   D   A   I   Q   L   77

1487  aac gaa agc ctc tgt att ggc tgc aag ctt tgc gcc gtg gtc tgc  1531
78     N   E   S   L   C   I   G   C   K   L   C   A   V   V   C   92
```

```
1532  cca ttt ggc gca atc agc gct tca gga agc cgt ccg gtg aat gcc  1576
93     P   F   G   A   I   S   A   S   G   S   R   P   V   N   A   107

1577  cat gcg caa tat gtt ttt cag gct gaa ggc tca ctc aaa gac ggc  1621
108    H   A   Q   Y   V   F   Q   A   E   G   S   L   K   D   G   122

1622  gaa gaa aac gcg cca aca caa cat gct ttg ctg cgc tgg gaa cct  1666
123    E   E   N   A   P   T   Q   H   A   L   L   R   W   E   P   137

1667  ggt gtc cag acc gtc gcg gtg aaa tgc gac ctg tgt gat ttc ttg  1711
138    G   V   Q   T   V   A   V   K   C   D   L   C   D   F   L   152

1712  cca gaa ggt ccg gcc tgc gtt cgc gct tgc ccg aat cag gcg tta  1756
153    P   E   G   P   A   C   V   R   A   C   P   N   Q   A   L   167

1757  cgg ctg atc acc ggt gat agc ctg caa cgt cag atg aaa gaa aaa  1801
168    R   L   I   T   G   D   S   L   Q   R   Q   M   K   E   K   182

1802  cag cgc ctt gcc gca agc tgg ttt gcc aat ggc ggg gag gat ccc  1846
183    Q   R   L   A   A   S   W   F   A   N   G   G   E   D   P   197
```

hyfB

```
1847  ctt tcc ctc act cag gag caa cgc taa tgg atg ccc tgc aat ta   1890
198    L   S   L   T   Q   E   Q   R   *
                                       M   D   A   L   Q   L      6

1891  tta acc tgg tcg ctg att ctc tat ctg ttt gct agt ctg gct tcg  1935
7      L   T   W   S   L   I   L   Y   L   F   A   S   L   A   S   21

1936  ctg ttt tta ctc ggt ctg gac aga ctg gct att aag ctt tcc ggc  1980
22     L   F   L   L   G   L   D   R   L   A   I   K   L   S   G   36

1981  atc aca tcg ctg gtg ggc ggc gtg att ggc atc atc agc gga att  2025
37     I   T   S   L   V   G   G   V   I   G   I   I   S   G   I   51

2026  acg caa tta cat gct ggt gta act tta gtc gcc cgt ttt gcc ccc  2070
52     T   Q   L   H   A   G   V   T   L   V   A   R   F   A   P   66

2071  cct ttt gaa ttt gcc gat tta acc ctg cga atg gat agc ctc tcg  2115
67     P   F   E   F   A   D   L   T   L   R   M   D   S   L   S   81

2116  gca ttt atg gtg ctg gtt atc tcc ttg ctg gtg gtg gtt tgt tct  2160
82     A   F   M   V   L   V   I   S   L   L   V   V   V   C   S   96

2161  ctc tat tca ttg act tat atg cgc gaa tac gag ggc aaa ggc gcg  2205
97     L   Y   S   L   T   Y   M   R   E   Y   E   G   K   G   A   111

2206  gcg gcg atg ggc ttc ttt atg aat att ttc atc gca tcg atg gtt  2250
112    A   A   M   G   F   F   M   N   I   F   I   A   S   M   V   126

2251  gcc ctg ctg gtg atg gac aac gct ttt tgg ttc atc gtg ctg ttt  2295
127    A   L   L   V   M   D   N   A   F   W   F   I   V   L   F   141

2296  gaa atg atg tcg ctg tct tcc tgg ttt ctg gtc att gcc agg cag  2340
142    E   M   M   S   L   S   S   W   F   L   V   I   A   R   Q   156

2341  gat aaa acg tcg atc aac gct ggc atg ctc tac ttt ttt atc gcc  2385
157    D   K   T   S   I   N   A   G   M   L   Y   F   F   I   A   171

2386  cac gcc gga tcg gtg ctg ata atg atc gcc ttc ttg ctg atg ggg  2430
172    H   A   G   S   V   L   I   M   I   A   F   L   L   M   G   186

2431  cgc gaa agc ggc agc ctc gat ttt gcc agt ttc cgc acg ctt tca  2475
187    R   E   S   G   S   L   D   F   A   S   F   R   T   L   S   201

2476  ctt tct ccg ggg ctg gcg tcg gcg gtg ttc ctg ctg gcc ttt ttc  2520
202    L   S   P   G   L   A   S   A   V   F   L   L   A   F   F   216
```

```
2521  ggt ttt ggc gcg aaa gcc ggg atg atg ccg ttg cac agc tgg ttg   2565
217    G   F   G   A   K   A   G   M   M   P   L   H   S   W   L   231

2566  ccg cgc gct cac cct gcc gca cca tcg cac gct tcg gcg ttg atg   2610
232    P   R   A   H   P   A   A   P   S   H   A   S   A   L   M   246

2611  tct ggc gta atg gtc aaa ata ggt att ttc ggc atc ctg aaa gta   2655
247    S   G   V   M   V   K   I   G   I   F   G   I   L   K   V   261

2656  gcg atg gat ctg ctg gcg caa acg ggt ttg cct ctg tgg tgg ggc   2700
262    A   M   D   L   L   A   Q   T   G   L   P   L   W   W   G   276

2701  att ctg gtg atg gcg atc ggc gca atc tcc gcg ctc ctg ggc gtg   2745
277    I   L   V   M   A   I   G   A   I   S   A   L   L   G   V   291
2746  cta tat gcg ctg gcg gaa cag gat atc aaa cgg ctg ctg gcc tgg   2790
292    L   Y   A   L   A   E   Q   D   I   K   R   L   L   A   W   306

2791  agt acc gtc gaa aac gtc ggc att att ttg ctg gca gtc ggt gtg   2835
307    S   T   V   E   N   V   G   I   I   L   L   A   V   G   V   321

2836  gcg atg gtc ggt ctg tca ctg cac gac ccg ctg ctc acc gtg gtt   2880
322    A   M   V   G   L   S   L   H   D   P   L   L   T   V   V   336

2881  gga ctg ctc ggc gca ctg ttt cat ctg ctc aac cat gcg ctg ttc   2925
337    G   L   L   G   A   L   F   H   L   L   N   H   A   L   F   351

2926  aaa ggg ctg cta ttt ctc ggc gcg gga gcg att att tcg cgt ttg   2970
352    K   G   L   L   F   L   G   A   G   A   I   I   S   R   L   366

2971  cat acc cac gac atg gaa aaa atg ggg gca cta gcg aaa cgg atg   3015
367    H   T   H   D   M   E   K   M   G   A   L   A   K   R   M   381

3016  ccg tgg aca gcc gca gca tgc ctg att ggt tgc ctc gcg ata tca   3060
382    P   W   T   A   A   A   C   L   I   G   C   L   A   I   S   396

3061  gcc att cct ccg ctg aat ggt ttt atc agc gaa tgg tac acc tgg   3105
397    A   I   P   P   L   N   G   F   I   S   E   W   Y   T   W   411

                 hyfBR-F1
3106  cag tcg ctg ttc tca cta agt cgt gtg gaa gcc gta gcg cta caa   3150
412    Q   S   L   F   S   L   S   R   V   E   A   V   A   L   Q   426

3151  ctt gcg ggt cct att gct atg gta atg ctg gca gtc act ggt ggg   3195
427    L   A   G   P   I   A   M   V   M   L   A   V   T   G   G   441

3196  ctg gca gta atg tgc ttc gta aaa atg tac ggt att act ttc tgt   3240
442    L   A   V   M   C   F   V   K   M   Y   G   I   T   F   C   456

3241  ggt gcg ccg cgc agt aca cac gct gaa gag gca cag gaa gtg cca   3285
457    G   A   P   R   S   T   H   A   E   E   A   Q   E   V   P   471

3286  aat acg atg atc gtc gcc atg cta ctg ctc gcg gca ctc tgc gta   3330
472    N   T   M   I   V   A   M   L   L   L   A   A   L   C   V   486

3331  tta att gcg ctt agt gcc agt tgg ctg gca ccg aag ata atg cat   3375
487    L   I   A   L   S   A   S   W   L   A   P   K   I   M   H   501

3376  att gcc cat gcg ttt acc aat acc cct ccc gcc act gtc gcc agc   3420
502    I   A   H   A   F   T   N   T   P   P   A   T   V   A   S   516

3421  gga ata gca ctt gta ccc ggc acg ttt cat aca cag gtc acc ccc   3465
517    G   I   A   L   V   P   G   T   F   H   T   Q   V   T   P   531

3466  tca tta ctg ttg ctg tta cta ctg gcg atg cct ttg ctg cct ggc   3510
532    S   L   L   L   L   L   L   L   A   M   P   L   L   P   G   546

3511  ctt tac tgg ctg tgg tgt cgt tcg cgc cgc gca gcg ttt cgt cgc   3555
547    L   Y   W   L   W   C   R   S   R   R   A   A   F   R   R   561
```

```
3556  aca gga gat gcc tgg gca tgc ggc tac ggc tgg gaa aat gcg atg   3600
562    T   G   D   A   W   A   C   G   Y   G   W   E   N   A   M    576

3601  gcc ccg tca ggc aat ggc gtg atg cag ccg ctg cgt gtg gtc ttt   3645
577    A   P   S   G   N   G   V   M   Q   P   L   R   V   V   F    591

3646  tct gcg cta ttt cgt cta cga caa cag ctc gac cct acg ctg agg   3690
592    S   A   L   F   R   L   R   Q   Q   L   D   P   T   L   R    606

3691  cta aat aaa ggt ctt gcg cac gtc acc gcc agg gct cag agc aca   3735
607    L   N   K   G   L   A   H   V   T   A   R   A   Q   S   T    621

3736  gaa ccc ttc tgg gat gag cgg gtg atc cgc ccc atc gtg agc gcc   3780
622    E   P   F   W   D   E   R   V   I   R   P   I   V   S   A    636

3781  acc caa cgg ctg gcc aaa gaa ata cag cat ctg caa agc ggc gac   3825
637    T   Q   R   L   A   K   E   I   Q   H   L   Q   S   G   D    651
3826  ttt cgt ctc tat tgc ctg tat gtg gtc gcc gca ctg gtt gtg ctg   3870
652    F   R   L   Y   C   L   Y   V   V   A   A   L   V   V   L    666

                                                    hyfC
3871  cta atc gct att gcc gtc taa gga aat cac cat gag aca aac tctt   3916
667    L   I   A   I   A   V   *
                                                    M   R   Q   T   L    5

3917  tgc gac gga tat ctg gtc att ttt gcg tta gca cag gcc gtg att   3961
6      C   D   G   Y   L   V   I   F   A   L   A   Q   A   V   I    20

3962  ctg ctg atg cta acc cca ctt ttt acg ggt att tcc cgg cag ata   4006
21     L   L   M   L   T   P   L   F   T   G   I   S   R   Q   I    35

4007  cgc gcg cgt atg cac tcc cgc cgc ggg ccg ggg atc tgg cag gat   4051
36     R   A   R   M   H   S   R   R   G   P   G   I   W   Q   D    50

4052  tat cgc gat atc cac aaa ctg ttt aaa cgc cag gaa gtt gcg ccg   4096
51     Y   R   D   I   H   K   L   F   K   R   Q   E   V   A   P    65

4097  aca tct tca ggt ctg atg ttc cgc ctg atg ccg tgg gta tta atc   4141
66     T   S   S   G   L   M   F   R   L   M   P   W   V   L   I    80

4142  agc agc atg ctg gtg ctg gcg atg gcc tta cca ctg ttt att acc   4186
81     S   S   M   L   V   L   A   M   A   L   P   L   F   I   T    95

4187  gtt tcc cct ttt gcg ggc ggc ggc gat ctg atc acc ctt atc tat   4231
96     V   S   P   F   A   G   G   G   D   L   I   T   L   I   Y   110

4232  ctt ctt gcc ctg ttt cgt ttt ttc ttt gct ctt tcc ggg ctg gat   4276
111    L   L   A   L   F   R   F   F   F   A   L   S   G   L   D   125

4277  acc gga agt ccg ttt gcg gga gtc ggt gcc agt cgc gag ttg acg   4321
126    T   G   S   P   F   A   G   V   G   A   S   R   E   L   T   140

4322  ctc ggc att ctg gtc gaa cca atg ctt att ctc tca ctg ctg gta   4366
141    L   G   I   L   V   E   P   M   L   I   L   S   L   L   V   155

4367  ttg gcg ctg ata gca ggt tcc acg cat atc gag atg atc agc aat   4411
156    L   A   L   I   A   G   S   T   H   I   E   M   I   S   N   170

4412  acg ctg gcg atg ggc tgg aac tcg ccg cta acc acc gta ctg gcg   4456
171    T   L   A   M   G   W   N   S   P   L   T   T   V   L   A   185

4457  tta ctg gcc tgt ggt ttt gcc tgc ttc att gag atg gga aaa att   4501
186    L   L   A   C   G   F   A   C   F   I   E   M   G   K   I   200

4502  ccc ttt gat gtt gct gaa gca gaa cag gaa tta cag gaa ggc ccg   4546
201    P   F   D   V   A   E   A   E   Q   E   L   Q   E   G   P   215
```

```
4547  ctg acc gaa tat tcc ggt gcc ggg ctg gcg cta gcg aaa tgg ggg   4591
216    L   T   E   Y   S   G   A   G   L   A   L   A   K   W   G    230

4592  ctg ggg ctg aaa cag gtc gtg atg gca tca ctg ttt gtg gcc ctg   4636
231    L   G   L   K   Q   V   V   M   A   S   L   F   V   A   L    245

4637  ttt ctg ccc ttt ggg cgc gcg caa gaa ctt tct ctc gcc tgc ctg   4681
246    F   L   P   F   G   R   A   Q   E   L   S   L   A   C   L    260

4682  ctg act tca ctt gtc gtt acg ctg ctc aag gtt ttg ctg att ttt   4726
261    L   T   S   L   V   V   T   L   L   K   V   L   L   I   F    275

4727  gta ctg gcc tca atc gca gaa aac acg ctg gca cgc ggg cgt ttt   4771
276    V   L   A   S   I   A   E   N   T   L   A   R   G   R   F    290

4772  tta ctc att cac cat gtg acc tgg ctt ggc ttc agc ctt gct gcg   4816
291    L   L   I   H   H   V   T   W   L   G   F   S   L   A   A    305

4817  ctt gca tgg gtc ttc tgg tta acc ggt ctg taa gga gca ctg acgg  4862
306    L   A   W   V   F   W   L   T   G   L   *
```

hyfD

```
4863  aat atg gaa aat ctt gct ctg acg acg tta ttg ctg cct ttt atc   4907
        M   E   N   L   A   L   T   T   L   L   L   P   F   I     14

4908  ggc gca ctg gtc gtt tcg ttt tcg cca caa cgt cgg gcc gcc gaa   4952
15     G   A   L   V   V   S   F   S   P   Q   R   R   A   A   E    29

4953  tgg ggg gtt ttg ttc gcc gcg ctg acc acg ctg tgc atg ttg tca   4997
30     W   G   V   L   F   A   A   L   T   T   L   C   M   L   S    44

4998  ctg atc tcc gcg ttt tat cag gcc gat aaa gtt gcc gtc acg ttg   5042
45     L   I   S   A   F   Y   Q   A   D   K   V   A   V   T   L    59

5043  acg ttg gtc aac gtg ggg gat gtg gcg ttg ttt ggc ctg gtc att   5087
60     T   L   V   N   V   G   D   V   A   L   F   G   L   V   I    74

5088  gat cgc gtg agt acg ctg att ctg ttt gtg gtg gtg ttt ctc ggt   5132
75     D   R   V   S   T   L   I   L   F   V   V   V   F   L   G    89

5133  ttg ctg gtc acg atc tac tcc acg ggt tat ctg acg gat aaa aat   5177
90    L   L   V   T   I   Y   S   T   G   Y   L   T   D   K   N    104

5178  cgc gaa cac ccg cat aac ggc acg aat cgt tat tac gca ttt tta   5222
105    R   E   H   P   H   N   G   T   N   R   Y   Y   A   F   L    119

5223  ctg gtg ttt atc ggc gcg atg gcg gga ctg gta ctc tcc tcg acg   5267
120    L   V   F   I   G   A   M   A   G   L   V   L   S   S   T    134

5268  ctg ctc ggt cag ttg ttg ttt ttt gaa att aca ggc ggc tgc tcc   5312
135    L   L   G   Q   L   L   F   F   E   I   T   G   G   C   S    149

5313  tgg gcg ttg atc agt tat tac cag agc gat aaa gcg cag cgt tca   5357
150    W   A   L   I   S   Y   Y   Q   S   D   K   A   Q   R   S    164

5358  gca cta aaa gcg tta ctt atc act cat atc ggc tcg ttg ggg ttg   5402
165    A   L   K   A   L   L   I   T   H   I   G   S   L   G   L    179

5403  tat ctt gcc gcc gcc acg ctg ttt ttg cag acc gga acg ttt gcg   5447
180    Y   L   A   A   A   T   L   F   L   Q   T   G   T   F   A    194

5448  ctt agc gcg atg agc gag tta cac ggc gac gca cgt tat ctg gtt   5492
195    L   S   A   M   S   E   L   H   G   D   A   R   Y   L   V    209

5493  tat ggc ggc atc ctg ttt gcc gcg tgg ggg aaa tcg gcc cag cta   5537
210    Y   G   G   I   L   F   A   A   W   G   K   S   A   Q   L    224
```

```
5538  ccg atg caa gcg tgg cta ccg gac gca atg gaa gcg cca aca ccg   5582
225    P   M   Q   A   W   L   P   D   A   M   E   A   P   T   P   239

5583  atc agc gcc tat ctc cac gcc gca tcg atg gtg aaa gtg ggc gtt   5627
240    I   S   A   Y   L   H   A   A   S   M   V   K   V   G   V   254

5628  tac att ttt gcc cgc gct att atc gac ggc ggc aat atc ccg cat   5672
255    Y   I   F   A   R   A   I   I   D   G   G   N   I   P   H   269

5673  gtg att ggc ggc gtt ggc atg gtc atg gca ctg gtc acc att ctt   5717
270    V   I   G   G   V   G   M   V   M   A   L   V   T   I   L   284

5718  tat ggc ttt ctg atg tat ttg cca cag cag gat atg aag cgg ttg   5762
285    Y   G   F   L   M   Y   L   P   Q   Q   D   M   K   R   L   299

5763  cta gcc tgg tcg acc atc act caa ctt ggc tgg atg ttc ttc ggc   5807
300    L   A   W   S   T   I   T   Q   L   G   W   M   F   F   G   314

5808  ttg tcg ctc tcc atc ttc ggc tcg cgg ctg gcg ctg gag ggc agc   5852
315    L   S   L   S   I   F   G   S   R   L   A   L   E   G   S   329

5853  atc gcc tac atc gtc aac cac gcg ttc gct aaa agc ctg ttt ttc   5897
330    I   A   Y   I   V   N   H   A   F   A   K   S   L   F   F   344

5898  ctt gta gca ggt gcg ctg agt tac agc tgc ggc acg cgc ttg ttg   5942
345    L   V   A   G   A   L   S   Y   S   C   G   T   R   L   L   359

5943  ccg cgt ctg cgt ggc gta ttg cac acc ctg ccg ttg cca ggc gtg   5987
360    P   R   L   R   G   V   L   H   T   L   P   L   P   G   V   374

5988  ggt ttc tgc gtg gca gcg ctg gcg att acc ggc gtg ccg ccg ttc   6032
375    G   F   C   V   A   A   L   A   I   T   G   V   P   P   F   389

6033  aac ggc ttc ttc agt aaa ttc ccg ctg ttt gct gcc ggt ttt gcg   6077
390    N   G   F   F   S   K   F   P   L   F   A   A   G   F   A   404

6078  ttg tca gtg gag tac tgg atc ctg ctg ccc gcc atg att ctt ctg   6122
405    L   S   V   E   Y   W   I   L   L   P   A   M   I   L   L   419

6123  atg att gaa tcg gtc gcc agt ttc gcc tgg ttt att cgc tgg ttt   6167
420    M   I   E   S   V   A   S   F   A   W   F   I   R   W   F   434

6168  ggt cgc gtt gtg cct ggc aaa ccg agc gag gcc gtc gcc gat gcc   6212
435    G   R   V   V   P   G   K   P   S   E   A   V   A   D   A   449

6213  gca ccg ctg cca gga tca atg cgc ctg gtg ttg att gta ctg att   6257
450    A   P   L   P   G   S   M   R   L   V   L   I   V   L   I   464

6258  gtg atg tcg ctg att tcc agc gta atc gcc gcg acc tgg ttg cag   6302
465    V   M   S   L   I   S   S   V   I   A   A   T   W   L   Q   479
                                                          t
                          hyfE
6303  taa gga gat gat gaa tga ccg gtt cta tga tcg taa ata atc tg    6346
480    *
                          M   T   G   S   M   I   V   N   N   L    10

6347  gcg gga ctg atg atg ctg aca tcg ctg ttt gtg att agc gtc aaa   6391
11     A   G   L   M   M   L   T   S   L   F   V   I   S   V   K   25

6392  agc tat cgc ctg tca tgc gga ttt tac gcc tgc cag tca ctg gtg   6436
26     S   Y   R   L   S   C   G   F   Y   A   C   Q   S   L   V   40

6437  ctg gtg tct att ttc gcc act ctc tcg tgc ctg ttc gcc gca gag   6481
41     L   V   S   I   F   A   T   L   S   C   L   F   A   A   E   55

6482  caa ctg ctg atc tgg tcc gcc agc gcc ttt atc acc aaa gtg ctg   6526
56     Q   L   L   I   W   S   A   S   A   F   I   T   K   V   L   70
```

```
6527  ctg gta ccg tta atc atg act tac gct gca cga aat att ccc cag   6571
71     L   V   P   L   I   M   T   Y   A   A   R   N   I   P   Q    85

6572  aac atc ccg gaa aaa gcg tta ttc ggt ccg gca atg atg gca ctg   6616
86     N   I   P   E   K   A   L   F   G   P   A   M   M   A   L   100

6617  ctc gcg gcg tta att gtc ctg ctt tgc gca ttt gtc gtt cag ccc   6661
101    L   A   A   L   I   V   L   L   C   A   F   V   V   Q   P   115

6662  gtg aag cta ccg atg gct acc ggg ctg aaa ccg gcg ctg gcg gta   6706
116    V   K   L   P   M   A   T   G   L   K   P   A   L   A   V   130

6707  gcg tta ggt cat ttt ctg ctt ggc ctg ctg tgc att gtc agc cag   6751
131    A   L   G   H   F   L   L   G   L   L   C   I   V   S   Q   145

6752  cgc aat atc ctg cgg caa att ttt ggt tac tgc ctg atg gaa aac   6796
146    R   N   I   L   R   Q   I   F   G   Y   C   L   M   E   N   160

6797  ggc tcc cat ctg gtg ctg gcg ctt ctt gcc tgg cga gca ccg gaa   6841
161    G   S   H   L   V   L   A   L   L   A   W   R   A   P   E   175

6842  ctg gtg gaa ata ggt atc gct acc gac gcc atc ttc gcc gtc att   6886
176    L   V   E   I   G   I   A   T   D   A   I   F   A   V   I   190

6887  gtg atg gtg tta ctg gca aga aaa ata tgg cgt acc cac ggc acg   6931
191    V   M   V   L   L   A   R   K   I   W   R   T   H   G   T   205

                                                          hyfF
6932  ctg gac gtg aac aac ttg acc gcg ctg aag gga taa tga gat gagt   6977
206    L   D   V   N   N   L   T   A   L   K   G   *
                                                          M   S    2

6978  tat tct gtg atg ttc gct tta ctc ctg ctc acg ccg ctg ctt ttt   7022
3      Y   S   V   M   F   A   L   L   L   L   T   P   L   L   F    17

7023  tcg ctg ctc tgt ttt gcc tgc cgg aaa cgg aga ctt tct gcg act   7067
18     S   L   L   C   F   A   C   R   K   R   R   L   S   A   T    32

7068  cgc acg gtg acc gta tta cat agc tta ggg atc aca ctg ctg ctg   7112
33     R   T   V   T   V   L   H   S   L   G   I   T   L   L   L    47

7113  att ctg gca ctc tgg gtg gtc caa act gcc gct gat gca gga gaa   7157
48     I   L   A   L   W   V   V   Q   T   A   A   D   A   G   E    62

7158  ata ttc gct gcg gga ctg tgg ctt cat att gat ggt ctg ggc ggt   7202
63     I   F   A   A   G   L   W   L   H   I   D   G   L   G   G    77

7203  ttg ttc ctc gcc att ctt ggt gtg att ggc ttt ctc acc ggt att   7247
78     L   F   L   A   I   L   G   V   I   G   F   L   T   G   I    92

7248  tac tcg att ggc tac atg cgt cat gaa gtg gca cac ggc gag ctt   7292
93     Y   S   I   G   Y   M   R   H   E   V   A   H   G   E   L   107

7293  tca ccc gtt acg ctg tgc gat tac tac ggt ttc ttc cat ctg ttt   7337
108    S   P   V   T   L   C   D   Y   Y   G   F   F   H   L   F   122

7338  ttg ttc acc atg ctg ctg gtt gtt acc agc aat aac ctg att gtg   7382
123    L   F   T   M   L   L   V   V   T   S   N   N   L   I   V   137

7383  atg tgg gcg gcg atc gaa gcc acc acc tta agc tcg gcg ttt ctg   7427
138    M   W   A   A   I   E   A   T   T   L   S   S   A   F   L   152

7428  gta ggc att tac ggt cag cgt tca tcg ctg gaa gct gca tgg aag   7472
153    V   G   I   Y   G   Q   R   S   S   L   E   A   A   W   K   167

7473  tac atc att att tgt act gtt ggt gtc gct ttt ggt ctg ttc ggt   7517
168    Y   I   I   I   C   T   V   G   V   A   F   G   L   F   G   182
```

```
7518  acc gtg ctg gta tac gcc aac gcc gcc agc gtt atg ccg cag gca  7562
183    T   V   L   V   Y   A   N   A   A   S   V   M   P   Q   A   197

7563  gaa atg gcg ata ttc tgg agc gag gtt ctt aag caa tcg tcc ttg  7607
198    E   M   A   I   F   W   S   E   V   L   K   Q   S   S   L   212

7608  ctt gac cca aca tta atg ctg ttg gcc ttt gtg ttt ttg cta att  7652
213    L   D   P   T   L   M   L   L   A   F   V   F   L   L   I   227

7653  ggc ttt ggt acc aaa acc ggg cta ttt ccc atg cac gcc tgg ctg  7697
228    G   F   G   T   K   T   G   L   F   P   M   H   A   W   L   242

7698  ccg gat gct cac agt gaa gcg ccg agt ccg gtc agc gcc ctg ctc  7742
243    P   D   A   H   S   E   A   P   S   P   V   S   A   L   L   257

7743  tcc gcc gta ttg ctg aac tgc gcg ctg ttg gtg ctg att cgc tat  7787
258    S   A   V   L   L   N   C   A   L   L   V   L   I   R   Y   272

7788  tac atc att att tgc caa gcc atc ggc agc gat ttc ccc aac cgg  7832
273    Y   I   I   I   C   Q   A   I   G   S   D   F   P   N   R   287

7833  ttg ttg ctc atc ttc ggc atg ttg tcg gtt gcc gtg gcg gca ttt  7877
288    L   L   L   I   F   G   M   L   S   V   A   V   A   A   F   302

7878  ttc att ctg gta cag cgg gac att aag cgt ctg ctg gcg tac tcc  7922
303    F   I   L   V   Q   R   D   I   K   R   L   L   A   Y   S   317

7923  agc gtg gag aac atg ggg ctg gtc gcg gtg gag cta ggc att ggc  7967
318    S   V   E   N   M   G   L   V   A   V   E   L   G   I   G   332
7968  ggg ccg ctg gga att ttt gcc gcg ctg ctg cac atc tta aac cac  8012
333    G   P   L   G   I   F   A   A   L   L   H   I   L   N   H   347

8013  agt ctg gca aaa acg ctg ctg ttc tgc ggt tcc ggc aat gta ctg  8057
348    S   L   A   K   T   L   L   F   C   G   S   G   N   V   L   362

8058  ctc aag tac ggc acg cgc gat ctc aac gtc gtc tgt ggg atg ctc  8102
363    L   K   Y   G   T   R   D   L   N   V   V   C   G   M   L   377

8103  aaa atc atg cca ttt acc gcc gtg ctg ttt ggc ggc ggt gcg ctg  8147
378    K   I   M   P   F   T   A   V   L   F   G   G   G   A   L   392

8148  gcg ctg gca ggg atg ccg ccc ttc aac att ttt ctt agc gaa ttt  8192
393    A   L   A   G   M   P   P   F   N   I   F   L   S   E   F   407

8193  atg acc att acc gcc gga ctg gca cgt aat cac ctg ctg att atc  8237
408    M   T   I   T   A   G   L   A   R   N   H   L   L   I   I   422

8238  gtc ctg ctg tta ttg ctg tta acg ctg gtg ctg gcg ggc ctg gta  8282
423    V   L   L   L   L   L   L   T   L   V   L   A   G   L   V   437

8283  cgg atg gct gcg cgg gtg tta atg gcg aaa ccg ccg cag gcc gtt  8327
438    R   M   A   A   R   V   L   M   A   K   P   P   Q   A   V   452

8328  aac cgg ggt gat ctc ggc tgg ttg acc acc tcg cca atg gtg att  8372
453    N   R   G   D   L   G   W   L   T   T   S   P   M   V   I   467

8373  ctg ctg gtc atg atg ctg gcg atg gga acg cat att cca caa cct  8417
468    L   L   V   M   M   L   A   M   G   T   H   I   P   Q   P   482

8418  gtc atc agg atc ctg gcg ggc gct tcc act ata gtc ctc tca ggg  8462
483    V   I   R   I   L   A   G   A   S   T   I   V   L   S   G   497

8463  acg cac gat ctg cct gca caa cgt agc acc tgg cat gat ttt ttg  8507
498    T   H   D   L   P   A   Q   R   S   T   W   H   D   F   L   512
```

hyfG

```
8508  cct tca ggc acc gca tct gtt tcg gag aaa cac agt gaa cgt t     8550
513    P   S   G   T   A   S   V   S   E   K   H   S   E   R       526
                                                  M   N   V       3

8551  aat tca tcg tca aat cgt ggc gaa gcg att ctc gcc gcc ctg aaa   8595
526   *
4      N   S   S   S   N   R   G   E   A   I   L   A   A   L   K   18

8596  acg cag ttc ccc ggc gcg gtg ctg gat gaa gag cga caa acg cct   8640
19     T   Q   F   P   G   A   V   L   D   E   E   R   Q   T   P   33

8641  gaa cag gtc acc att acg gtg aaa atc aat ctg ctg cct gac gtt   8685
34     E   Q   V   T   I   T   V   K   I   N   L   L   P   D   V   48

8686  gta cag tat ctt tat tat caa cat gat ggc tgg ctt ccg gtc ctg   8730
49     V   Q   Y   L   Y   Y   Q   H   D   G   W   L   P   V   L   63

8731  ttt ggc aac gac gag cgg aca ctt aac ggt cat tac gcg gtt tat   8775
64     F   G   N   D   E   R   T   L   N   G   H   Y   A   V   Y   78

8776  tat gcc ctt tca atg gaa ggg gcc gaa aaa tgc tgg att gtg gtg   8820
79     Y   A   L   S   M   E   G   A   E   K   C   W   I   V   V   93

8821  aag gcg ctg gtc gat gcc gac agt cgg gag ttt ccg tca gtc aca   8865
94     K   A   L   V   D   A   D   S   R   E   F   P   S   V   T   108

8866  ccg cgc gtc cct gcc gcg gtc tgg ggc gag cga gaa att cgc gat   8910
109    P   R   V   P   A   A   V   W   G   E   R   E   I   R   D   123

8911  atg tac ggg ctg att ccg gtt ggc ctg ccg gat cag cgt cgc ctg   8955
124    M   Y   G   L   I   P   V   G   L   P   D   Q   R   R   L   138

8956  gtg ttg ccc gat gac tgg ccg gaa gat atg cat ccg ctg cgc aaa   9000
139    V   L   P   D   D   W   P   E   D   M   H   P   L   R   K   153

9001  gat gcg atg gat tat cga ctg cgc cct gaa ccg acg act gat tcc   9045
154    D   A   M   D   Y   R   L   R   P   E   P   T   T   D   S   168

9046  gaa acg tat ccg ttt atc aat gag ggc aac agc gat gcg cgg gtg   9090
169    E   T   Y   P   F   I   N   E   G   N   S   D   A   R   V   183

9091  atc cct gtc ggc ccg ctg cat atc acc tcc gat gaa ccg ggt cac   9135
184    I   P   V   G   P   L   H   I   T   S   D   E   P   G   H   198

9136  ttc cgc ttg ttt gtg gat ggc gag caa att gtc gat gct gat tac   9180
199    F   R   L   F   V   D   G   E   Q   I   V   D   A   D   Y   213

9181  cgc ctg ttt tat gtc cat cgc ggc atg gag aaa ctg gca gaa acg   9225
214    R   L   F   Y   V   H   R   G   M   E   K   L   A   E   T   228

9226  cgg atg ggc tac aac gaa gtg acc ttc tta tcg gac cgc gtg tgt   9270
229    R   M   G   Y   N   E   V   T   F   L   S   D   R   V   C   243

9271  ggg att tgc ggt ttt gcc cac agt gtg gcc tat acc aat tcg gtt   9315
244    G   I   C   G   F   A   H   S   V   A   Y   T   N   S   V   258

9316  gaa aat gca ctg ggg att gag gtg ccg caa cga gca cat act att   9360
259    E   N   A   L   G   I   E   V   P   Q   R   A   H   T   I   273

9361  cgc tcg att ctg ctg gaa gtc gaa cgg cta cac agt cat ttg ctt   9405
274    R   S   I   L   L   E   V   E   R   L   H   S   H   L   L   288

9406  aac ctt ggc ctc tcc tgc cat ttc gtt ggt ttt gat acc ggc ttt   9450
289    N   L   G   L   S   C   H   F   V   G   F   D   T   G   F   303

9451  atg caa ttt ttc cgc gtg cgg gaa aag tcg atg acg atg gcg gaa   9495
304    M   Q   F   F   R   V   R   E   K   S   M   T   M   A   E   318
```

```
9496  ttg ctg atc ggg tcg cgt aaa acc tac ggt ctg aat ctg att ggt   9540
319    L   L   I   G   S   R   K   T   Y   G   L   N   L   I   G   333

9541  ggt gtt cgc cgc gat att ctc aaa gag caa cgt ctg caa acg ctg   9585
334    G   V   R   R   D   I   L   K   E   Q   R   L   Q   T   L   348

9586  aaa ctg gtg cgc gag atg cgc gcc gac gtg tcg gag ctg gta gag   9630
349    K   L   V   R   E   M   R   A   D   V   S   E   L   V   E   363

9631  atg ctg ctt gct acg ccg aat atg gaa caa cgc act cag ggc att   9675
364    M   L   L   A   T   P   N   M   E   Q   R   T   Q   G   I   378

9676  ggc att ctc gac cga caa atc gcc cgt gat ttg cgc ttt gat cac   9720
379    G   I   L   D   R   Q   I   A   R   D   L   R   F   D   H   393

9721  ccc tac gcc gac tac ggc aat att cca aaa aca ctg ttt acc ttt   9765
394    P   Y   A   D   Y   G   N   I   P   K   T   L   F   T   F   408

9766  acc ggc ggc gat gtt ttc tcc cgc gtg atg gtc cgt gtc aaa gag   9810
409    T   G   G   D   V   F   S   R   V   M   V   R   V   K   E   423

9811  acg ttt gat tcg ctg gca atg ctg gaa ttt gcc ctc gac aac atg   9855
424    T   F   D   S   L   A   M   L   E   F   A   L   D   N   M   438

9856  ccg gat acc cca ctg ctg acc gaa ggc ttt agc tat aaa cct cac   9900
439    P   D   T   P   L   L   T   E   G   F   S   Y   K   P   H   453

9901  gca ttc gcg ctg ggc ttt gtt gaa gcg cca cgc ggt gaa gac gtg   9945
454    A   F   A   L   G   F   V   E   A   P   R   G   E   D   V   468

9946  cac tgg agc atg ctc ggt gat aac caa aaa ttg ttc cgc tgg cgc   9990
469    H   W   S   M   L   G   D   N   Q   K   L   F   R   W   R   483

9991  tgc cgt gcc gcc acc tac gcc aac tgg ccg gtg ttg cgt tac atg   10035
484    C   R   A   A   T   Y   A   N   W   P   V   L   R   Y   M   498
10036 ctg cgc ggc aat acc gtt tct gac gca ccg ctg att atc ggt agc   10080
499    L   R   G   N   T   V   S   D   A   P   L   I   I   G   S   513

10081 ctt gat ccc tgc tac tcc tgt acc gac cgt gtg acg ctg gta gat   10125
514    L   D   P   C   Y   S   C   T   D   R   V   T   L   V   D   528

10126 gtg cgc aag cgc cag tca aaa acc gtg ccg tat aaa gag atc gaa   10170
529    V   R   K   R   Q   S   K   T   V   P   Y   K   E   I   E   543

10171 cgc tac ggc att gat cgt aac cgt tcg ccg ctg aag taa gga cag   10215
544    R   Y   G   I   D   R   N   R   S   P   L   K   *
```

hyfH

```
10216 aag atg ctg aag tta ctg aaa act att atg cgc gcc gga acc gcg   10260
          M   L   K   L   L   K   T   I   M   R   A   G   T   A   14

10261 acg gtg aaa tat ccc ttc gcg cca ctg gag gtc agc cct ggc ttt   10305
15     T   V   K   Y   P   F   A   P   L   E   V   S   P   G   F   29

10306 cgc gga aaa ccg gac ctg atg ccc agc caa tgt att gcc tgc ggt   10350
30     R   G   K   P   D   L   M   P   S   Q   C   I   A   C   G   44

10351 gcc tgc gcc tgt gct tgt ccg gca aat gcg ctg act atc cag acc   10395
45     A   C   A   C   A   C   P   A   N   A   L   T   I   Q   T   59

10396 gac gac cag caa aat tcg cgc acc tgg cag ctc tat ctg ggg cgt   10440
60     D   D   Q   Q   N   S   R   T   W   Q   L   Y   L   G   R   74

10441 tgt att tac tgc gga cgt tgt gaa gaa gtg tgc ccg acc aga gcc   10485
75     C   I   Y   C   G   R   C   E   E   V   C   P   T   R   A   89

10486 atc cag ctt acc aat aac ttt gaa ctg acc gtc acc aat aaa gcc   10530
90     I   Q   L   T   N   N   F   E   L   T   V   T   N   K   A   104
```

```
10531 gat ctc tat acc cgc gcg acg ttc cat cta caa cgt tgc agc cgt   10575
105    D   L   Y   T   R   A   T   F   H   L   Q   R   C   S   R    119

10576 tgc gaa cgc ccg ttt gcc ccg caa aaa acc atc gca ctg gct gct   10620
120    C   E   R   P   F   A   P   Q   K   T   I   A   L   A   A    134

10621 gaa ttg tta gca cag caa caa aat gcg cca caa aac cgc gaa atg   10665
135    E   L   L   A   Q   Q   Q   N   A   P   Q   N   R   E   M    149

10666 ttg tgg gcg caa gcg agc gtc tgc ccg gaa tgc aaa caa cgc gcg   10710
150    L   W   A   Q   A   S   V   C   P   E   C   K   Q   R   A    164

10711 acg ctg atc aac gac gat aca gat gta ctg ctg gtg gct aag gag   10755
165    T   L   I   N   D   D   T   D   V   L   L   V   A   K   E    179

                hyfI
10756 cag cta tga gtc cag tgc tta cac aac atg tca gcc agc cca tcacg 10802
180    Q   L   *
                M   S   P   V   L   T   Q   H   V   S   Q   P   I   T 14

10803 ctg gac gag caa acg caa aag atg aag cgg cat ttg cta cag gat   10847
15     L   D   E   Q   T   Q   K   M   K   R   H   L   L   Q   D    29

10848 atc cgt cgc tcg gct tac gtt tat cgc gtc gat tgc ggc ggc tgc   10892
30     I   R   R   S   A   Y   V   Y   R   V   D   C   G   G   C    44

10893 aac gcc tgt gaa atc gaa att ttt gct gcc att aca cca gta ttc   10937
45     N   A   C   E   I   E   I   F   A   A   I   T   P   V   F    59

10938 gac gca gaa cgt ttt ggc att aag gtt gtt tca tca ccg cgt cac   10982
60     D   A   E   R   F   G   I   K   V   V   S   S   P   R   H    74

10983 gcc gat att ttg tta ttt act ggc gca gtc acc cgg gcg atg cgt   11027
75     A   D   I   L   L   F   T   G   A   V   T   R   A   M   R    89


11028 atg cct gca ctt cgg gcg tat gag tct gcc ccc gat cat aaa att   11072
90     M   P   A   L   R   A   Y   E   S   A   P   D   H   K   I   104

11073 tgt gtt tcc tac ggc gcg tgc ggt gtc ggc ggc ggt att ttc cac   11117
105    C   V   S   Y   G   A   C   G   V   G   G   G   I   F   H   119

11118 gat ctc tac agc gtc tgg ggc ggt agc gac acc att gtc ccc att   11162
120    D   L   Y   S   V   W   G   G   S   D   T   I   V   P   I   134

11163 gat gtt tgg atc ccc ggc tgc ccg cca aca ccg gcc gcc acc att   11207
135    D   V   W   I   P   G   C   P   P   T   P   A   A   T   I   149

11208 cac ggt ttc gcc gtg gcg ctc ggt ttg ctg caa cag aag att cac   11252
150    H   G   F   A   V   A   L   G   L   L   Q   Q   K   I   H   164

11253 gct gtg gat tat cgc gat ccc acc ggg gtg act atg caa ccg ttg   11297
165    A   V   D   Y   R   D   P   T   G   V   T   M   Q   P   L   179

11298 tgg ccg cag atc ccg cca tca cag cgt atc gcc att gag cga gaa   11342
180    W   P   Q   I   P   P   S   Q   R   I   A   I   E   R   E   194

11343 gcg cgg cgg ctg gcg ggc tat cgt cag ggg cga gaa att tgc gat   11387
195    A   R   R   L   A   G   Y   R   Q   G   R   E   I   C   D   209

11388 cgg ctc ctg cgc cat tta agc gac gat cct aca gga aat cgg gtt   11432
210    R   L   L   R   H   L   S   D   D   P   T   G   N   R   V   224

11433 aac acc tgg ttg cgc gat gcc gac gat cca cgt ctc aat agt atc   11477
225    N   T   W   L   R   D   A   D   D   P   R   L   N   S   I   239
```

hyfJ

```
11478 gtt cag caa ctc ttt cgc gta ctc cgg ggg tta cat gac tga a     11520
240    V   Q   Q   L   F   R   V   L   R   G   L   H   D   *
                                                   M   T   E       3

11521 gag tgc ggg gaa att gtt ttc tgg acg ctg cga aaa aag ttt gtc   11565
4      E   C   G   E   I   V   F   W   T   L   R   K   K   F   V    18

11566 gcc agt agc gac gag atg ccg gaa cac agc tct cag gta atg tat   11610
19     A   S   S   D   E   M   P   E   H   S   S   Q   V   M   Y    33

11611 tac tcg cta gct atc ggc cat cac gtt ggc gtg att gat tgt ctg   11655
34     Y   S   L   A   I   G   H   H   V   G   V   I   D   C   L    48

11656 aat gtc gcc ttc cgc tgc cca ctg acg gaa tac gaa gat tgg ctt   11700
49     N   V   A   F   R   C   P   L   T   E   Y   E   D   W   L    63

11701 gca ctg gtc gaa gag gag caa gcc cga cgt aag atg ctg ggg gtg   11745
64     A   L   V   E   E   E   Q   A   R   R   K   M   L   G   V    78

11746 atg act ttt ggt gag att gtt att gac gcc agc cac acc gcc ctg   11790
79     M   T   F   G   E   I   V   I   D   A   S   H   T   A   L    93

11791 ttg acc cgg gca ttc gcg cca ctg gcg gat gac gcg acg tct gtg   11835
94     L   T   R   A   F   A   P   L   A   D   D   A   T   S   V   108

11836 tgg cag gcg cgt agc att caa ttc att cat ctg ttg gat gaa att   11880
109    W   Q   A   R   S   I   Q   F   I   H   L   L   D   E   I   123

11881 gtg cag gaa ccg gcc atc tat ctg atg gcc aga aaa att gcg tgaga 11927
124    V   Q   E   P   A   I   Y   L   M   A   R   K   I   A   *
```

hyfR

hyfR-F1

```
11928 agg att tct cat taa taa gga ctg ttg atg gct atg tca gac gag   11972
                                      M   A   M   S   D   E       6

11973 gcg atg ttt gcc ccg cca caa gga ata aca att gaa gcg gta aac   12017
7      A   M   F   A   P   P   Q   G   I   T   I   E   A   V   N    21

12018 gga atg ctc gcg gag cgg tta gca cag aaa cac ggc aag gcg tct   12062
22     G   M   L   A   E   R   L   A   Q   K   H   G   K   A   S    36

12063 tta tta cgc gcc ttc atc ccg ctg ccg ccg ccg ttc agc ccg gta   12107
37     L   L   R   A   F   I   P   L   P   P   P   F   S   P   V    51

12108 caa ctt att gaa ctg cat gtt ctc aaa agc aac ttc tat tac cgc   12152
52     Q   L   I   E   L   H   V   L   K   S   N   F   Y   Y   R    66

12153 tac cat gat gat ggc agc gat gtg acg gca aca aca gag tat cag   12197
67     Y   H   D   D   G   S   D   V   T   A   T   T   E   Y   Q    81

12198 ggc gag atg gtc gat tat tcg cgt cac gcc gtc ctt ctc ggc agt   12242
82     G   E   M   V   D   Y   S   R   H   A   V   L   L   G   S    96

12243 agt gga atg gcg gag cta cgc ttt att cgc acc cac ggc agt cgt   12287
97     S   G   M   A   E   L   R   F   I   R   T   H   G   S   R   111

12288 ttt act tcc cag gat tgc aca ctg ttt aac tgg ctg gcg cgg ata   12332
112    F   T   S   Q   D   C   T   L   F   N   W   L   A   R   I   126

12333 atc acc ccg gtt ctg caa tca tgg ctc aat gat gaa gaa cag cag   12377
127    I   T   P   V   L   Q   S   W   L   N   D   E   E   Q   Q   141

12378 gtg gcg ctg cgt ttg ctg gag aaa gat cgc gat cat cat cgg gta   12422
142    V   A   L   R   L   L   E   K   D   R   D   H   H   R   V   156
```

```
12423 ctg gtt gat att act aat gca gtg ctg tca cat ctt gat ctc gac   12467
157    L   V   D   I   T   N   A   V   L   S   H   L   D   L   D    171

12468 gat ctg atc gct gac gtc gct cgt gag atc cat cat ttt ttc ggt   12512
172    D   L   I   A   D   V   A   R   E   I   H   H   F   F   G    186

12513 ctg gct tca gtc agt atg gta ctg ggc gat cat cga aag aac gag   12557
187    L   A   S   V   S   M   V   L   G   D   H   R   K   N   E    201

12558 aag ttc agc ctg tgg tgc agc gat ctt tct gcc tca cat tgt gcg   12602
202    K   F   S   L   W   C   S   D   L   S   A   S   H   C   A    216

12603 tgt ctg cca cgc tgt atg cct ggc gaa agt gta ttg ctg aca caa   12647
217    C   L   P   R   C   M   P   G   E   S   V   L   L   T   Q    231

12648 acg cta caa acc cga caa ccg acc ttg acg cac cgt gca gat gat   12692
232    T   L   Q   T   R   Q   P   T   L   T   H   R   A   D   D    246

12693 ctg ttt ctc tgg caa cgc gac ccg tta tta ctc tta ctt gca tct   12737
247    L   F   L   W   Q   R   D   P   L   L   L   L   L   A   S    261

12738 aac ggc tgc gaa tct gcg ctc ctt ata ccg ctt acc ttt ggc aac   12782
262    N   G   C   E   S   A   L   L   I   P   L   T   F   G   N    276

12783 cat aca ccg ggt gca ttg ttg ctg gcg cat acc tct tcc act ctc   12827
277    H   T   P   G   A   L   L   L   A   H   T   S   S   T   L    291

12828 ttt agt gag gaa aac tgc cag cta cta caa cac ata gcc gat cgc   12872
292    F   S   E   E   N   C   Q   L   L   Q   H   I   A   D   R    306

12873 atc gct att gcc gtt ggc aat gcc gat gcc tgg cgt agc atg acc   12917
307    I   A   I   A   V   G   N   A   D   A   W   R   S   M   T    321

12918 gat ttg cag gaa agt ttg cag caa gaa aac cac cag ctt agc gag   12962
322    D   L   Q   E   S   L   Q   Q   E   N   H   Q   L   S   E    336

12963 cag ctc ctt tcg aat ctg ggc atc ggt gac att atc tat caa agc   13007
337    Q   L   L   S   N   L   G   I   G   D   I   I   Y   Q   S    351

13008 cag gca atg gaa gac ctg ctc cag cag gta gat att gtg gcg aag   13052
352    Q   A   M   E   D   L   L   Q   Q   V   D   I   V   A   K    366

13053 agc gac agt acg gtg ttg att tgt ggt gaa acc gga act ggc aaa   13097
367    S   D   S   T   V   L   I   C   G   E   T   G   T   G   K    381
13098 gag gtg atc gcc aga gcg atc cat caa ctt agc ccg cga cgc gac   13142
382    E   V   I   A   R   A   I   H   Q   L   S   P   R   R   D    396

13143 aag ccg ctg gtc aaa atc aac tgc gct gcc atc ccc gcc agt ctt   13187
397    K   P   L   V   K   I   N   C   A   A   I   P   A   S   L    411

13188 ctg gaa agt gag tta ttc ggt cat gac aaa ggg gcg ttt act ggt   13232
412    L   E   S   E   L   F   G   H   D   K   G   A   F   T   G    426

13233 gcg att aat acc cat cgt ggt cgt ttt gaa att gcc gat ggc ggc   13277
427    A   I   N   T   H   R   G   R   F   E   I   A   D   G   G    441

13278 acg ttg ttt ctc gat gaa att ggc gat ctg ccg tta gaa ctt cag   13322
442    T   L   F   L   D   E   I   G   D   L   P   L   E   L   Q    456

13323 cct aaa ctg ctg cgc gta ttg cag gag cgg gag att gag cgt ctc   13367
457    P   K   L   L   R   V   L   Q   E   R   E   I   E   R   L    471

13368 ggc ggg agt aga acg atc ccg gtg aat gtc aga gtc att gcc gcc   13412
472    G   G   S   R   T   I   P   V   N   V   R   V   I   A   A    486

13413 acc aac cgt gat ttg tgg caa atg gtt gaa gat cgc cag ttt cgc   13457
487    T   N   R   D   L   W   Q   M   V   E   D   R   Q   F   R    501
```

```
13458 agc gat ctc ttt tat cgc ctg aat gtc ttc cca ctg gaa ttg ccg   13502
502    S   D   L   F   Y   R   L   N   V   F   P   L   E   L   P    516

13503 ccg cta cgc gac cgt ccg gaa gat atc cct ctt tta gca aaa cat   13547
517    P   L   R   D   R   P   E   D   I   P   L   L   A   K   H    531

13548 ttc acg caa aaa atg gcg cgc cat atg aat cgc gca att gac gcc   13592
532    F   T   Q   K   M   A   R   H   M   N   R   A   I   D   A    546

13593 atc ccg acc gag gca cta cgc cag ttg atg tcg tgg gat tgg ccg   13637
547    I   P   T   E   A   L   R   Q   L   M   S   W   D   W   P    561

                                                        hyfR-R1
13638 ggc aac gtg cgc gag ctg gaa aac gtg att gag cgg gcg gta ctg   13682
562    G   N   V   R   E   L   E   N   V   I   E   R   A   V   L    576

13683 ttg act cgt ggt aac agt ctg aat tta cat cta aat gtc cga caa   13727
577    L   T   R   G   N   S   L   N   L   H   L   N   V   R   Q    591

13728 agc cgt tta ctg ccg acg cta aat gaa gat tca gcg ctt cgc agt   13772
592    S   R   L   L   P   T   L   N   E   D   S   A   L   R   S    606

13773 tca atg gcg cag tta ctg cac ccg acg acg cca gag aat gac gaa   13817
607    S   M   A   Q   L   L   H   P   T   T   P   E   N   D   E    621

13818 gaa gaa cgt cag cgc att gtt cag gta ttg cga gaa acc aat ggc   13862
622    E   E   R   Q   R   I   V   Q   V   L   R   E   T   N   G    636

13863 att gtt gcc ggg ccc cgt ggc gca gcg acg cga tta ggg atg aag   13907
637    I   V   A   G   P   R   G   A   A   T   R   L   G   M   K    651

                                                        hyfBR-R1
13908 cgc acc acg ctg ctg tca cga atg cag cgg ctg ggg atc tcg gtt   13952
652    R   T   T   L   L   S   R   M   Q   R   L   G   I   S   V    666

                                                        focB
13953 cgc gag gtg ttg taa tct gct ttt gca gga gta tgc atg aga aac   13997
667    R   E   V   L   *                                 M   R   N    3

13998 aaa ctc tct ttc gac ttg cag ttg agc gcc aga aaa gcg gca atc   14042
4      K   L   S   F   D   L   Q   L   S   A   R   K   A   A   I    18

14043 gct gaa cgg att gcc gcc cat aaa att gcc cgc agt aaa gtg tcg   14087
19     A   E   R   I   A   A   H   K   I   A   R   S   K   V   S    33

14088 gtc ttt tta atg gcg atg tcc gct ggc gtg ttt atg gcg atc gga   14132
34     V   F   L   M   A   M   S   A   G   V   F   M   A   I   G    48

14133 ttt act ttt tac ctt tcc gtt atc gcc gat gcc ccg tct tca cag   14177
49     F   T   F   Y   L   S   V   I   A   D   A   P   S   S   Q    63

14178 gca tta acc cat ctg gtg ggc ggc ctt tgc ttt aca ctc ggc ttt   14222
64     A   L   T   H   L   V   G   G   L   C   F   T   L   G   F    78

14223 att ttg ctg gcg gtt tgc ggc acc agc ctg ttc acc tcg tcg gta   14267
79     I   L   L   A   V   C   G   T   S   L   F   T   S   S   V    93

14268 atg acg gtg atg gca aaa agt cgg ggc gtt att agt tgg cga act   14312
94     M   T   V   M   A   K   S   R   G   V   I   S   W   R   T   108

14313 tgg ctg att aac gca ctt ctg gtg gcc tgc ggt aat ctg gca ggt   14357
109    W   L   I   N   A   L   L   V   A   C   G   N   L   A   G   123

14358 att gcc tgt ttc agt ttg tta atc tgg ttt tcc ggg ctg gtg atg   14402
124    I   A   C   F   S   L   L   I   W   F   S   G   L   V   M   138
```

```
14403 agt gaa aac gcg atg tgg gga gtc gcg gtt tta cac tgc gcc gag   14447
139    S   E   N   A   M   W   G   V   A   V   L   H   C   A   E    153

14448 ggc aaa atg cat cat aca ttt act gaa tct gtc agc ctc ggc att   14492
154    G   K   M   H   H   T   F   T   E   S   V   S   L   G   I    168

14493 atg tgc aat ctg atg gtt tgc ctg gcg ctg tgg atg agt tat tgc   14537
169    M   C   N   L   M   V   C   L   A   L   W   M   S   Y   C    183

14538 ggg cgt tcg tta tgc gac aaa atc gtc gcc atg att ttg ccc atc   14582
184    G   R   S   L   C   D   K   I   V   A   M   I   L   P   I    198

14583 acc ctg ttt gtc gcc agt ggc ttt gag cac tgt atc gcc aat ttg   14627
199    T   L   F   V   A   S   G   F   E   H   C   I   A   N   L    213

14628 ttt gtg att ccg ttc gcc att gcc att cgc cat ttc gcc cct ccc   14672
214    F   V   I   P   F   A   I   A   I   R   H   F   A   P   P    228

14673 ccc ttc tgg cag ctg gcg cac agt agc gca gac aat ttt ccg gca   14717
229    P   F   W   Q   L   A   H   S   S   A   D   N   F   P   A    243

14718 ctg acg gtc agc cat ttt att acc gcc aat ctg ctc ccg gtg atg   14762
244    L   T   V   S   H   F   I   T   A   N   L   L   P   V   M    258

14763 ctg ggt aat att atc ggc ggt gcg gtg ctg gtg agt atg tgt tat   14807
259    L   G   N   I   I   G   G   A   V   L   V   S   M   C   Y    273

14808 cgg gct att tat tta cgt cag gaa ccc tga
274    R   A   I   Y   L   R   Q   E   P   *
```

Appendix II

Nucleotide sequence of the *hyc* operon of *E. coli* and translations of the nine genes.

Translation initiation sites (bold) and stop sites (asterisks) are indicated, as are the sequences of PCR primers used in this study (underlined) (Table 2.3). The EMBL accession number for the nucleotide sequence is X17506 (Bohm *et al.*, 1990; Rossmann *et al.*, 1995).

```
1     tcg cct ccc att aac tat tgc cag cta caa gca ata att gtg cca   45

46    gtg ttg att atc cct gcg gtg aat aat gtc gat gat gtc gaa atg   90

91    aca cgt cga cac ggc gac gaa att cat ctt tag ctt aaa aat ctc   135

136   ttt aat aac aat aaa tta aaa gtt ggc aca aaa aat gct taa agc   180

                                              hycA          hycA-L
181   tgg cat ctc tgt taa acg ggt aac ctg aca atg act att tgg gaa   225
                                              M   T   I   W   E     5
                                    hyc-F
226   ata agc gag aaa gcc gat tac atc gca cag cgg cat cgt cgc cta   270
6      I   S   E   K   A   D   Y   I   A   Q   R   H   R   R   L    20

271   cag gac cag tgg cac atc tac tgc aat tcg ctg gtt cag ggg atc   315
21     Q   D   Q   W   H   I   Y   C   N   S   L   V   Q   G   I    35

316   acg tta tcg aaa gcg cgc ctg cat cac gcc atg agc tgc gcg ccg   360
36     T   L   S   K   A   R   L   H   H   A   M   S   C   A   P    50

361   gac aaa gaa ctc tgt ttc gtc ctt ttt gaa cat ttt cgc att tac   405
51     D   K   E   L   C   F   V   L   F   E   H   F   R   I   Y    65

406   gtc acc ctg gcg gat ggc ttt aac agc cac acc atc gag tat tac   450
66     V   T   L   A   D   G   F   N   S   H   T   I   E   Y   Y    80

451   gtc gaa aca aaa gat ggc gaa gac aaa cag cgg att gcg cag gcg   495
81     V   E   T   K   D   G   E   D   K   Q   R   I   A   Q   A    95

496   caa ctg agc att gac ggc atg att gat ggc aag gtc aac atc cgc   540
96     Q   L   S   I   D   G   M   I   D   G   K   V   N   I   R   110

541   gat cgc gaa cag gtt ctg gaa cac tat ctc gaa aaa atc gct ggc   585
111    D   R   E   Q   V   L   E   H   Y   L   E   K   I   A   G   125

586   gtt tac gac agc tta tac acc gct att gaa aac aat gtg ccg gtg   630
126    V   Y   D   S   L   Y   T   A   I   E   N   N   V   P   V   140

631   aat tta agc caa ctg gta aag gga caa agc ccg gca gca tga gct   675
141    N   L   S   Q   L   V   K   G   Q   S   P   A   A   *

676   gag gct ttg ccc gtt ttg cag gcg tta cgc ctg ttt ggg gat ggg   720
```

hycA-R

721 cgt gtc gat gag tgt cga aaa tga cat ttc atc ggc atg ttt tcgt 766

hycB

767 caa aaa tga caa tca cct gag gaa tgc ctg **gtg** aat cgt ttt gta 811
M N R F V 5

812 att gct gac tcc acg ctc tgt atc ggc tgc cac act tgt gag gcc 856
6 I A D S T L C I G C H T C E A 20

857 gcc tgt tca gag acg cat cgc cag cac ggc ctg caa tca atg ccg 901
21 A C S E T H R Q H G L Q S M P 35

902 cgc ctg aga gtg atg ctg aat gaa aaa gaa tct gcg ccg cag ctc 946
36 R L R V M L N E K E S A P Q L 50

947 tgt cac cac tgt gaa gat gca ccc tgc gcg gtg gtc tgc ccg gtt 991
51 C H H C E D A P C A V V C P V 65

992 aac gcc atc acc cgc gtc gat ggg gcc gtg cag ttg aat gaa agc 1036
66 N A I T R V D G A V Q L N E S 80

1037 ctg tgc gta agc tgc aag ctg tgc ggc atc gcc tgc ccg ttt ggc 1081
81 L C V S C K L C G I A C P F G 95

1082 gca att gaa ttt tcc ggc agc cgt ccg ctg gat att ccg gca aac 1126
96 A I E F S G S R P L D I P A N 110

1127 gcc aat acc ccg aaa gcg cca ccg gca ccg cct gct ccg gcg cgt 1171
111 A N T P K A P P A P P P A P A R 125

1172 gtc agc aca ttg ctt gac tgg gtg cca ggt att cgc gcg atc gcc 1216
126 V S T L L D W V P G I R A I A 140

1217 gtc aaa tgt gac ctt tgt agc ttt gat gaa caa ggt ccg gcc tgc 1261
141 V K C D L C S F D E Q G P A C 155

1262 gcg cgg atg tgc ccg act aaa gcc ctg cat ctg gtg gat aac acc 1306
156 A R M C P T K A L H L V D N T 170

1307 gat atc gcc cgc gtc agc aaa cgt aag cgt gag ctg acc ttt aac 1351
171 D I A R V S K R K R E L T F N 185

1352 acg gac ttt ggc gat ctc acc ttg ttt cag cag gct caa agt gga 1396
186 T D F G D L T L F Q Q A Q S G 200

hycC

1397 gag gct aaa **tga** gcg caa ttt ccc tga tca ata gcg gcg tgg cg 1440
201 E A K *
M S A I S L I N S G V A 12

1441 tgg ttt gtc gcc gcc gct gtt ctg gca ttt ctc ttt tct ttt caa 1485
13 W F V A A A V L A F L F S F Q 27

1486 aaa gcg tta agt ggc tgg ata gct gga att ggc ggc gcg gtt ggt 1530
28 K A L S G W I A G I G G A V G 42

1531 agt ctg tat acg gca gcc gcg ggc ttc act gta ctg act ggc gcg 1575
43 S L Y T A A A G F T V L T G A 57

1576 gtt ggc gtg agc ggt gcg ctg tcg ctg gta agc tac gat gtg caa 1620
58 V G V S G A L S L V S Y D V Q 72

1621 atc tct ccg ctt aac gcg att tgg ctg att acg ctc ggt ctg tgc 1665
73 I S P L N A I W L I T L G L C 87

```
1666  ggt ctg ttt gtc agc ctc tac aac att gac tgg cat cgc cac gcg   1710
88     G   L   F   V   S   L   Y   N   I   D   W   H   R   H   A    102

1711  cag gtg aag tgc aac ggc ttg cag atc aat atg ttg atg gct gcc   1755
103    Q   V   K   C   N   G   L   Q   I   N   M   L   M   A   A    117

1756  gcc gtc tgc gcc gtc att gcc agc aac ctc ggc atg ttc gtg gta   1800
118    A   V   C   A   V   I   A   S   N   L   G   M   F   V   V    132

1801  atg gcc gaa atc atg gcc ctg tgc gcg gtg ttc ctc acc agc aac   1845
133    M   A   E   I   M   A   L   C   A   V   F   L   T   S   N    147

1846  agc aaa gag ggc aaa ctg tgg ttt gcg ctg ggg cgt ctt ggc act   1890
148    S   K   E   G   K   L   W   F   A   L   G   R   L   G   T    162
1891  ctg ctg ctg gcg att gct tgc tgg ctg ctg tgg cag cgt tac ggc   1935
163    L   L   L   A   I   A   C   W   L   L   W   Q   R   Y   G    177

1936  acg ctg gat ctg cgc ctg ctg gat atg cgt atg caa cag ctg ccg   1980
178    T   L   D   L   R   L   L   D   M   R   M   Q   Q   L   P    192

1981  ctc ggt tcc gat atc tgg ctg ctc gga gtg att ggc ttt ggc ctg   2025
193    L   G   S   D   I   W   L   L   G   V   I   G   F   G   L    207

2026  ctg gcc ggg att att ccg ctg cac ggc tgg gtg ccg cag gca cat   2070
208    L   A   G   I   I   P   L   H   G   W   V   P   Q   A   H    222

2071  gcg aac gcc tct aca cca gct gcc gcg ttg ttt tct acg gta gtc   2115
223    A   N   A   S   T   P   A   A   A   L   F   S   T   V   V    237

2116  atg aaa att ggc ctg ctg ggc att tta acc ctg tca ctg ctg ggc   2160
238    M   K   I   G   L   L   G   I   L   T   L   S   L   L   G    252

2161  ggt aat gca ccg ctg tgg tgg ggg atc gcg ctg ctg gtg ctc ggc   2205
253    G   N   A   P   L   W   W   G   I   A   L   L   V   L   G    267

2206  atg atc acc gcg ttt gtc ggt ggt ctg tat gcg ctg gtg gag cac   2250
268    M   I   T   A   F   V   G   G   L   Y   A   L   V   E   H    282

2251  aac atc cag cgc ctg ctg gct tac cac acc ctg gaa aat atc ggc   2295
283    N   I   Q   R   L   L   A   Y   H   T   L   E   N   I   G    297

2296  atc atc ctg ctg ggg ctg ggc gct ggc gta acg ggt atc gcg ctc   2340
298    I   I   L   L   G   L   G   A   G   V   T   G   I   A   L    312

2341  gaa caa ccg gcg ctg att gct ctt ggc ctg gtc ggt ggt ctg tac   2385
313    E   Q   P   A   L   I   A   L   G   L   V   G   G   L   Y    327

2386  cat ctg ctt aac cat agc ctg ttc aaa agc gta ctg ttc ctc ggg   2430
328    H   L   L   N   H   S   L   F   K   S   V   L   F   L   G    342

2431  gcg ggg agc gtc tgg ttc cgt acc ggt cat cgc gat atc gaa aaa   2475
343    A   G   S   V   W   F   R   T   G   H   R   D   I   E   K    357

2476  ctc ggt ggt att ggc aag aaa atg ccg gtt atc tcc atc gcc atg   2520
358    L   G   G   I   G   K   K   M   P   V   I   S   I   A   M    372

2521  tta gtc ggg ctg atg gca atg gct gcg ctg ccg ccg ctg aat ggt   2565
373    L   V   G   L   M   A   M   A   A   L   P   P   L   N   G    387

2566  ttt gcc ggg gaa tgg gtt atc tat caa tca ttt ttc aaa ctg agc   2610
388    F   A   G   E   W   V   I   Y   Q   S   F   F   K   L   S    402

2611  aat agt ggc gcg ttt gtt gcc cgt ctg ctg ggg ccg ctg ctc gct   2655
403    N   S   G   A   F   V   A   R   L   L   G   P   L   L   A    417

2656  gtg ggg ctg gca att acc ggt gcg ctg gcg gtg atg tgt atg gcg   2700
418    V   G   L   A   I   T   G   A   L   A   V   M   C   M   A    432
```

```
2701  aaa gtc tat ggc gtc acg ttc ctc ggc gcg ccg cgc acc aaa gaa   2745
433    K   V   Y   G   V   T   F   L   G   A   P   R   T   K   E    447

2746  gcc gaa aac gcc acc tgt gcg ccg ctc ctg atg agc gta agc gta   2790
448    A   E   N   A   T   C   A   P   L   L   M   S   V   S   V    462

2791  gtg gca ctg gcg att tgc tgc gta att ggc ggt gtt gct gcg ccg   2835
463    V   A   L   A   I   C   C   V   I   G   G   V   A   A   P    477

2836  tgg cta ctg ccg atg ctc tct gct gct gta cct ctg ccg ctg gag   2880
478    W   L   L   P   M   L   S   A   A   V   P   L   P   L   E    492

2881  cct gct aac acc acc gtt tct caa ccg atg atc acg ttg ctg ctg   2925
493    P   A   N   T   T   V   S   Q   P   M   I   T   L   L   L    507

2926  att gcc tgc ccg ctg ctg cca ttc atc att atg gcg att tgc aaa   2970
508    I   A   C   P   L   L   P   F   I   I   M   A   I   C   K    522
2971  ggc gat cgt ttg cca tcg cgt tcc cgc ggt gcg gcc tgg gtg tgc   3015
523    G   D   R   L   P   S   R   S   R   G   A   A   W   V   C    537

3016  ggt tac gac cac gaa aaa tca atg gtg att acc gct cac ggt ttt   3060
538    G   Y   D   H   E   K   S   M   V   I   T   A   H   G   F    552

3061  gcc atg ccg gtg aaa cag gcg ttt gcg ccg gtg ctg aaa cta cgc   3105
553    A   M   P   V   K   Q   A   F   A   P   V   L   K   L   R    567

3106  aaa tgg ctg aat ccg gtg tct ctg gtg ccg ggc tgg cag tgc gag   3150
568    K   W   L   N   P   V   S   L   V   P   G   W   Q   C   E    582

3151  ggg agt gcg ttg ctg ttc cgc cgg atg gcg ctg gtt gaa ctg gcg   3195
583    G   S   A   L   L   F   R   R   M   A   L   V   E   L   A    597

                                                      hycD
3196  gta ctg gtg gtg att att gtt tca cga gga gcc tga gaa tga gtgtt 3242
598    V   L   V   V   I   I   V   S   R   G   A   *
                                                          M   S   V 2

3243  tta tat ccg tta att cag gcg ctg gtg tta ttt gcc gtt gcg ccg   3287
3      L   Y   P   L   I   Q   A   L   V   L   F   A   V   A   P    17

3288  ctg ctc tcc ggt ata acc cgc gtg gcg cgc gcc cgc ttg cat aac   3332
18     L   L   S   G   I   T   R   V   A   R   A   R   L   H   N    32

3333  cgt cgc ggg ccg ggc gtg ttg cag gag tat cgc gac att atc aaa   3377
33     R   R   G   P   G   V   L   Q   E   Y   R   D   I   I   K    47

3378  ctg ctg ggg cgt cag agc gtc ggc ccg gat gcc tcc ggc tgg gtg   3422
48     L   L   G   R   Q   S   V   G   P   D   A   S   G   W   V    62

3423  ttc cgc ctg acg ccg tat gtg atg gtg ggc gtc atg ctg act atc   3467
63     F   R   L   T   P   Y   V   M   V   G   V   M   L   T   I    77

3468  gct act gcg ctg ccg gtg gtg acc gtc ggt tct ccg ctg ccg caa   3512
78     A   T   A   L   P   V   V   T   V   G   S   P   L   P   Q    92

3513  ctg ggt gat ttg atc acc tta ctg tat ctc ttt gcc atc gcg cgt   3557
93     L   G   D   L   I   T   L   L   Y   L   F   A   I   A   R   107

3558  ttc ttc ttt gcc att tct ggt ctg gat acc ggt agc ccg ttt acc   3602
108    F   F   F   A   I   S   G   L   D   T   G   S   P   F   T   122

3603  gct atc ggc gcg agc cgt gaa gcg atg ctt ggc gtg ctg gtc gaa   3647
123    A   I   G   A   S   R   E   A   M   L   G   V   L   V   E   137

3648  ccg atg ctg ctg ctt ggt ctg tgg gtt gcc gca cag gtt gcc ggt   3692
138    P   M   L   L   L   G   L   W   V   A   A   Q   V   A   G   152
```

```
3693  tcc acc aac atc agc aac atc acc gac acc gtt tat cac tgg ccg   3737
153    S   T   N   I   S   N   I   T   D   T   V   Y   H   W   P   167

3738  ctg agc cag agc atc ccg ctg gta ctg gcg ctt tgt gcc tgt gcg   3782
168    L   S   Q   S   I   P   L   V   L   A   L   C   A   C   A   182

3783  ttc gcc acc ttt atc gaa atg ggc aaa ctg ccg ttc gac ctg gcg   3827
183    F   A   T   F   I   E   M   G   K   L   P   F   D   L   A   197

3828  gaa gcc gag cag gag ctg cag gaa ggc ccg ctc tct gaa tac agc   3872
198    E   A   E   Q   E   L   Q   E   G   P   L   S   E   Y   S   212

3873  ggc agc ggc ttt ggc gtc atg aaa tgg ggt atc agc ctg aaa cag   3917
213    G   S   G   F   G   V   M   K   W   G   I   S   L   K   Q   227

3918  ctg gtg gtg ttg cag atg ttc gtc ggg gtg ttt att ccg tgg gga   3962
228    L   V   V   L   Q   M   F   V   G   V   F   I   P   W   G   242

3963  caa atg gaa acc ttc acc gcc ggt gga ctg ctg ctg gcg ctg gtg   4007
243    Q   M   E   T   F   T   A   G   G   L   L   L   A   L   V   257
4008  att gcc atc gta aaa ctg gtg gtc ggc gtc ctg gtt atc gcg ctg   4052
258    I   A   I   V   K   L   V   V   G   V   L   V   I   A   L   272

4053  ttc gaa aac agc atg gcc cgt ctg cgt ctt gat att act ccg cgc   4097
273    F   E   N   S   M   A   R   L   R   L   D   I   T   P   R   287

4098  att acc tgg gct ggg ttt ggc ttt gca ttt tta gcg ttc gtc tcc   4142
288    I   T   W   A   G   F   G   F   A   F   L   A   F   V   S   302
```

hycE

```
4143  ttg ctg gcg gcg tga tta aag aga gtt tga gca tgt ctg aag aa    4186
303    L   L   A   A   *
                                      M   S   E   E     4

4187  aaa tta ggt caa cat tat ctc gcc gcg ctg aat gag gca ttt ccg   4231
5      K   L   G   Q   H   Y   L   A   A   L   N   E   A   F   P   19

4232  ggc gtc gtg ctg gac cac gcc tgg cag acc aaa gat cag ctg act   4276
20     G   V   V   L   D   H   A   W   Q   T   K   D   Q   L   T   34

4277  gtc acc gta aag gtg aac tac ctg ccg gaa gtg gtg gag ttt ctt   4321
35     V   T   V   K   V   N   Y   L   P   E   V   V   E   F   L   49

4322  tac tac aaa cag ggt ggc tgg ctg tcg gtg ctg ttt ggt aac gac   4366
50     Y   Y   K   Q   G   G   W   L   S   V   L   F   G   N   D   64

4367  gaa cgc aaa ctg aat ggt cat tac gcc gtt tac tac gtg ctg tcg   4411
65     E   R   K   L   N   G   H   Y   A   V   Y   Y   V   L   S   79

4412  atg gag aag ggc act aag tgt tgg att acg gtt cgc gtc gaa gtt   4456
80     M   E   K   G   T   K   C   W   I   T   V   R   V   E   V   94

4457  gac gcc aac aaa ccg gaa tat ccg tcc gtg acg ccg cgc gtt ccg   4501
95     D   A   N   K   P   E   Y   P   S   V   T   P   R   V   P   109

4502  gcg gcg gtg tgg ggc gag cgt gaa gtg cgc gat atg tac ggt ttg   4546
110    A   A   V   W   G   E   R   E   V   R   D   M   Y   G   L   124

4547  att ccg gtt ggt ctg ccg gat gaa cgt cgt ctg gtg ctg ccg gat   4591
125    I   P   V   G   L   P   D   E   R   R   L   V   L   P   D   139

4592  gac tgg ccg gat gaa ctt tat ccg ctg cgt aaa gac agc atg gat   4636
140    D   W   P   D   E   L   Y   P   L   R   K   D   S   M   D   154

4637  tat cgt cag cgt ccg gca ccg acc acc gat gct gaa acc tac gag   4681
155    Y   R   Q   R   P   A   P   T   T   D   A   E   T   Y   E   169
```

```
4682  ttc atc aac gaa ctg ggc gac aag aaa aac aac gtc gtg ccg att   4726
170    F   I   N   E   L   G   D   K   K   N   N   V   V   P   I    184

4727  ggt ccg ctg cac gtc act tct gat gaa ccg ggc cac ttc cgt ctg   4771
185    G   P   L   H   V   T   S   D   E   P   G   H   F   R   L    199

4772  ttc gtc gat ggc gaa aac att atc gac gcc gac tac cgt ctg ttc   4816
200    F   V   D   G   E   N   I   I   D   A   D   Y   R   L   F    214

4817  tac gtc cat cgc ggc atg gaa aaa ctg gcg gaa acc cgt atg ggt   4861
215    Y   V   H   R   G   M   E   K   L   A   E   T   R   M   G    229

4862  tat aac gaa gtg acc ttc ctc tct gac cgt gtg tgc ggg atc tgc   4906
230    Y   N   E   V   T   F   L   S   D   R   V   C   G   I   C    244

4907  ggc ttt gcc cac agc acc gcc tac acc acg tcg gtg gaa aac gcg   4951
245    G   F   A   H   S   T   A   Y   T   T   S   V   E   N   A    259

4952  atg ggt att cag gtg cca gaa cgt gcg cag atg atc cgc gcc att   4996
260    M   G   I   Q   V   P   E   R   A   Q   M   I   R   A   I    274

4997  ctg ctg gag gta gaa cgc ttg cac tcg cat ctg ctc aac ctt ggc   5041
275    L   L   E   V   E   R   L   H   S   H   L   L   N   L   G    289
5042  ctg gcc tgt cac ttt acc ggc ttc gac tcc ggc ttt atg cag ttc   5086
290    L   A   C   H   F   T   G   F   D   S   G   F   M   Q   F    304

5087  ttc cgc gtg cgt gaa acc tcc atg aaa atg gca gag atc ctt acc   5131
305    F   R   V   R   E   T   S   M   K   M   A   E   I   L   T    319

5132  ggt gcg cgt aaa acc tac ggc ctg aac ttg atc ggc ggg att cgt   5176
320    G   A   R   K   T   Y   G   L   N   L   I   G   G   I   R    334

5177  cgc gat ctg ctg aaa gac gac atg atc cag acc cgc cag ctg gca   5221
335    R   D   L   L   K   D   D   M   I   Q   T   R   Q   L   A    349

5222  caa cag atg cgt cgt gaa gtg cag gag ctg gtg gat gtg ctg ctg   5266
350    Q   Q   M   R   R   E   V   Q   E   L   V   D   V   L   L    364

5267  agc act ccg aac atg gaa cag cgc act gtc ggc att ggt cgt ctg   5311
365    S   T   P   N   M   E   Q   R   T   V   G   I   G   R   L    379

5312  gac ccg gaa atc gct cgc gac ttc agt aac gtc ggc ccg atg gtc   5356
380    D   P   E   I   A   R   D   F   S   N   V   G   P   M   V    394

5357  cgt gcc agc ggt cac gcc cgt gat acc cgc gcc gat cac ccg ttt   5401
395    R   A   S   G   H   A   R   D   T   R   A   D   H   P   F    409

5402  gtc ggc tat ggc ctg ctg cca atg gaa gtc cac agc gag cag ggc   5446
410    V   G   Y   G   L   L   P   M   E   V   H   S   E   Q   G    424

5447  tgc gac gtt att tcc cgt ctg aaa gtg cgt atc aac gaa gtc tat   5491
425    C   D   V   I   S   R   L   K   V   R   I   N   E   V   Y    439

5492  acc gcg ctg aac atg atc gac tac ggt ctg gat aac ctg ccg ggt   5536
440    T   A   L   N   M   I   D   Y   G   L   D   N   L   P   G    454

5537  ggc cca ctg atg gtg gaa ggc ttt acc tac att ccg cac cgc ttt   5581
455    G   P   L   M   V   E   G   F   T   Y   I   P   H   R   F    469

5582  gcg ctg ggc ttt gcc gaa gcg ccg cgc ggc gat gat atc cac tgg   5626
470    A   L   G   F   A   E   A   P   R   G   D   D   I   H   W    484

5627  agc atg acc ggc gac aac cag aag ctg tac cgc tgg cgc tgc cgt   5671
485    S   M   T   G   D   N   Q   K   L   Y   R   W   R   C   R    499

5672  gcc gcg acc tac gcg aac tgg ccg acc ctg cgc tac atg ctg cgc   5716
500    A   A   T   Y   A   N   W   P   T   L   R   Y   M   L   R    514
```

```
5717  ggc aac acc gtt tcc gat gcg ccg ctg att atc ggt agc ctc gac   5761
515    G   N   T   V   S   D   A   P   L   I   I   G   S   L   D    529

5762  cct tgc tac tcc tgt acc gac cgc atg acc gtg gtc gat gtg cgt   5806
530    P   C   Y   S   C   T   D   R   M   T   V   V   D   V   R    544

5807  aag aag aag agc aaa gtg gtg ccg tac aaa gaa ctc gag cgt tac   5851
545    K   K   K   S   K   V   V   P   Y   K   E   L   E   R   Y    559

                                                          hycF
5852  agc att gag cgt aaa aac tcg ccg ctg aaa taa gga atc gcc atg   5896
560    S   I   E   R   K   N   S   P   L   K   *                   M    1

5897  ttt acc ttt atc aaa aaa gtc atc aaa acc ggc acg gcg acc tcg   5941
2      F   T   F   I   K   K   V   I   K   T   G   T   A   T   S    16

5942  tct tat ccg ctg gag ccg att gcg gtt gat aaa aac ttc cgt ggt   5986
17     S   Y   P   L   E   P   I   A   V   D   K   N   F   R   G    31

5987  aag cca gag cag aac ccg cag cag tgc atc ggc tgc gcg gcc tgc   6031
32     K   P   E   Q   N   P   Q   Q   C   I   G   C   A   A   C    46

6032  gtc aat gcc tgc ccg tca aac gcc tta acg gtt gaa act gac ctc   6076
47     V   N   A   C   P   S   N   A   L   T   V   E   T   D   L    61
6077  gcc aca gga gag ctt gcc tgg gag ttt aat ctt ggg cac tgc atc   6121
62     A   T   G   E   L   A   W   E   F   N   L   G   H   C   I    76

6122  ttc tgt gga cgc tgc gaa gaa gtc tgc ccg acg gcg gcg atc aaa   6166
77     F   C   G   R   C   E   E   V   C   P   T   A   A   I   K    91

6167  ctg tcg caa gag tac gaa ctg gcg gtg tgg aag aaa gaa gac ttc   6211
92     L   S   Q   E   Y   E   L   A   V   W   K   K   E   D   F   106

6212  ctg caa cag tcc cgc ttc gcg ctg tgc aac tgc cgc gtc tgc aat   6256
107    L   Q   Q   S   R   F   A   L   C   N   C   R   V   C   N   121

6257  cgt cct ttc gcc gtc cag aaa gag atc gac tac gcc att gcg ctg   6301
122    R   P   F   A   V   Q   K   E   I   D   Y   A   I   A   L   136

6302  ctt aag cac aac ggc gac agc cgc gcg gaa aac cac cgc gaa agc   6346
137    L   K   H   N   G   D   S   R   A   E   N   H   R   E   S   151

6347  ttt gag act tgc ccg gaa tgt aag cgc cag aaa tgc ctg gtg ccg   6391
152    F   E   T   C   P   E   C   K   R   Q   K   C   L   V   P   166

6392  tcc gac cgt att gaa ctg act cgc cat atg aaa gag gcc atc tg    6435
167    S   D   R   I   E   L   T   R   H   M   K   E   A   I   *

      hycG
6436  atg agc aat tta tta ggc ccc cgt gac gcc aac ggc att ccg gtc   6480
1      M   S   N   L   L   G   P   R   D   A   N   G   I   P   V    15

6481  ccc atg acg gtg gat gaa tcc atc gcc agc atg aag gcg tcg tta   6525
16     P   M   T   V   D   E   S   I   A   S   M   K   A   S   L    30

6526  ctg aaa aaa atc aaa cgt tct gcc tat gtt tac cgc gtg gac tgc   6570
31     L   K   K   I   K   R   S   A   Y   V   Y   R   V   D   C    45

6571  ggc ggc tgc aac ggt tgc gaa atc gaa att ttc ggc acg ctt tcg   6615
46     G   G   C   N   G   C   E   I   E   I   F   G   T   L   S    60

6616  ccg ctg ttt gat gca gaa cgc ttc ggc att aaa gtc gtt cct tca   6660
61     P   L   F   D   A   E   R   F   G   I   K   V   V   P   S    75

6661  ccg cgt cat gcg gat att tta ctg ttt acc ggc gcg gtc acc cgt   6705
76     P   R   H   A   D   I   L   L   F   T   G   A   V   T   R    90
```

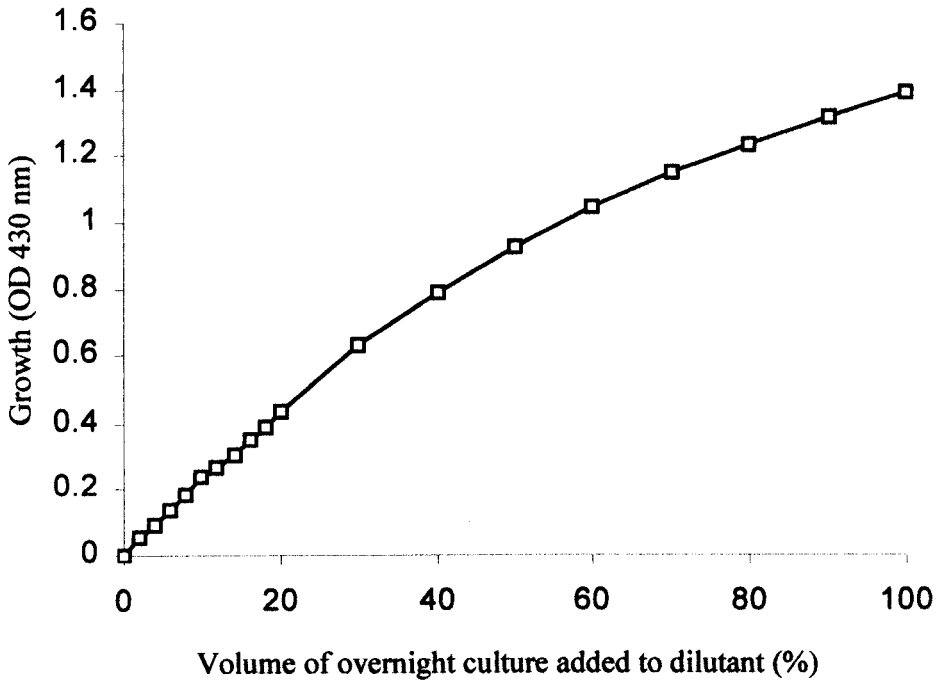
```
6706  gca atg cga tcc cct gcg ctg cgt gcg tgg cag tcc gcg ccg gac   6750
91     A   M   R   S   P   A   L   R   A   W   Q   S   A   P   D    105

6751  ccg aaa att tgt atc tcc tac ggt gcc tgc ggt aac agt ggc ggg   6795
106    P   K   I   C   I   S   Y   G   A   C   G   N   S   G   G    120

6796  atc ttc cac gat ctc tac tgc gtg tgg ggc ggt acg gat aaa att   6840
121    I   F   H   D   L   Y   C   V   W   G   G   T   D   K   I    135

6841  gtc cct gtg gat gtt tat atc cct ggc tgc ccg cca acg cct gcc   6885
136    V   P   V   D   V   Y   I   P   G   C   P   P   T   P   A    150

6886  gcc acg ctg tac ggc ttt gca atg gcg ctc ggc ctg ctg gag cag   6930
151    A   T   L   Y   G   F   A   M   A   L   G   L   L   E   Q    165

6931  aaa att cac gcc cgt ggg ccg ggt gaa ctg gat gaa caa ccg gcg   6975
166    K   I   H   A   R   G   P   G   E   L   D   E   Q   P   A    180

6976  gag atc ctg cat ggt gat atg gtg cag ccg ctg cgc gtg aaa gtg   7020
181    E   I   L   H   G   D   M   V   Q   P   L   R   V   K   V    195

7021  gat cgc gaa gca cgt cgc ctg gcg ggt tat cgt tac ggt cgt cag   7065
196    D   R   E   A   R   R   L   A   G   Y   R   Y   G   R   Q    210

7066  att gcc gat gat tac ctt aca cag tta ggg cag ggc gaa gaa cag   7110
211    I   A   D   D   Y   L   T   Q   L   G   Q   G   E   E   Q    225

7111  gtt gca cgc tgg ctg gaa gcg gaa aac gat ccg cgt ctg aac gag   7155
226    V   A   R   W   L   E   A   E   N   D   P   R   L   N   E    240

                                                          hycH
7156  att gtc agc cat ctg aat cat gtt gtt gaa gag gcg cgt atc cgatg 7202
241    I   V   S   H   L   N   H   V   V   E   E   A   R   I   R  *  256
                                                              M      0

7203  agt gaa aag gtg gtg ttc agt caa ctg agc cgt aaa ttt att gat   7247
1      S   E   K   V   V   F   S   Q   L   S   R   K   F   I   D    15

7248  gag aac gat gcc acg ccc gcc gag gcg cag cag gtg gtc tat tac   7292
16     E   N   D   A   T   P   A   E   A   Q   Q   V   V   Y   Y    30

7293  agc ctg gcg att ggt cac cac ctt ggg gtt atc gat tgc ctg gaa   7337
31     S   L   A   I   G   H   H   L   G   V   I   D   C   L   E    45

7338  gcg gcg ctc acc tgc ccg tgg gat gaa tat ctg gca tgg att gcc   7382
46     A   A   L   T   C   P   W   D   E   Y   L   A   W   I   A    60

7383  act ctg gag gca ggc agt gaa gcc cgc cgc aaa atg gaa ggc gtg   7427
61     T   L   E   A   G   S   E   A   R   R   K   M   E   G   V    75

7428  ccg aaa tat ggt gag atc gtc atc gac att aac cat gtg ccg atg   7472
76     P   K   Y   G   E   I   V   I   D   I   N   H   V   P   M    90

7473  ctg gcc aac gca ttc gat aaa gcc cgg gca gcg caa act tcg cag   7517
91     L   A   N   A   F   D   K   A   R   A   A   Q   T   S   Q   105

7518  cag cag gaa tgg agt aca atg ctg tta agt atg ctg cat gat att   7562
106    Q   Q   E   W   S   T   M   L   L   S   M   L   H   D   I   120

                                                          hycI
7563  cat cag gaa aac gcc atc tat ttg atg gtg agg aga ctg cgt g     7605
121    H   Q   E   N   A   I   Y   L   M   V   R   R   L   R        134
                                                              M      1

7606  act gac gtt tta ctc tgt gtt ggc aat agc atg atg ggc gat gat   7650
135   D   *
2      T   D   V   L   L   C   V   G   N   S   M   M   G   D   D    16
```

```
7651  ggc gca ggt ccg ctg ctg gcg gaa aag tgc gcc gcc gcg ccg aaa   7695
17     G   A   G   P   L   L   A   E   K   C   A   A   A   P   K    31

7696  ggt aac tgg gtg gtg att gac ggc ggt agc gca ccg gaa aac gac   7740
32     G   N   W   V   V   I   D   G   G   S   A   P   E   N   D    46

                                                  hycA-R
7741  atc gtc gct atc cgt gaa ctg cgc ccg aca cga ctg ctg att gtc   7785
47     I   V   A   I   R   E   L   R   P   T   R   L   L   I   V    61

7786  gac gcc acg gat atg ggg cta aac ccc ggc gag atc cgc atc atc   7830
62     D   A   T   D   M   G   L   N   P   G   E   I   R   I   I    76

7831  gac ccg gat gat atc gcc gag atg ttt atg atg act acc cat aac   7875
77     D   P   D   D   I   A   E   M   F   M   M   T   T   H   N    91

7876  atg ccg ttg aat tac ctt atc gac cag ttg aaa gaa gat att ggc   7920
92     M   P   L   N   Y   L   I   D   Q   L   K   E   D   I   G   106

7921  gaa gtg att ttc ctc ggc att cag ccg gat atc gtc ggc ttt tac   7965
107    E   V   I   F   L   G   I   Q   P   D   I   V   G   F   Y   121

7966  tac ccg atg acc cag ccg att aaa gat gcg gta gaa acc gtt tat   8010
122    Y   P   M   T   Q   P   I   K   D   A   V   E   T   V   Y   136

8011  caa cga ctg gaa ggc tgg gaa gga aat ggc ggc ttc gcg cag tta   8055
137    Q   R   L   E   G   W   E   G   N   G   G   F   A   Q   L   151

8056  gcg gtg gaa gaa gag tag ttt ttc att aag gaa tca gga cag gga   8100
152    A   V   E   E   E   *

8101  tgt tct tga tgg ggt gaa cca gct ctg atg cca aat gct aaa ttg   8145

8146  ccc gat gcg ctg cgc tta tcg ggc ctt cat ggt tcg tgc gac atg   8190

8191  tag gcc gga taa ggc gtt cac gcc gca tcc ggc act gtt acc tac   8235

8236  tct aaa tct
```

Appendix III

Estimation of bacterial cell density and biomass from OD at 430 nm

A plot of OD at 430 nm, as a function of bacterial culture concentration.
Values are the averages of at least two measurements.



A plot of OD at 430 nm, as a function of bacterial culture dry weight ($g.l^{-1}$).

Values are the average of at least two measurements.