

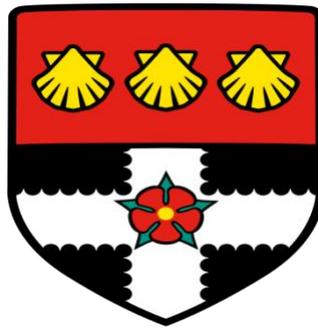
**The University of Reading**

**Molecular-genetic analysis of the *hyf* gene-cluster  
encoding the Hydrogenase-4 system of *E. coli*; role in  
hydrogen production**

A thesis submitted to the Faculty of Life Sciences in Partial Fulfilment of  
the Requirements for the Degree of Doctor of Philosophy

**by**

**Aida Al Lawati**



**Division of Microbiology**

**School of Biological Sciences**

**The University of Reading**

**Reading**

**UK**

**April 2018**

## **Declaration**

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged

Aida Al Lawati

Signed:

Date:

## Abstract

In *E. coli*, four types of hydrogenases are recognized. Hydrogenases 1 (Hyd-1) and 2 (Hyd-2) are responsible for H<sub>2</sub> consumption via respiration. Hydrogenase-3 (Hyd-3), which is encoded by the *hycABCDEFGH* operon, combines with formate dehydrogenase H (Fdh-H) to form the formate hydrogen lyase (FHL) complex, which is responsible for formate disposal and H<sub>2</sub> evolution during anaerobic fermentation under acidic pH. The *hyfABCDEFGHIJR-focB* operon of *E. coli* encodes a fourth potential hydrogenase (Hyd-4) closely related to Hyd-3. The physiological purpose of the Hyf system is unclear although it is speculated to form a second, energy conserving, FHL with Fdh-H. However, current evidence suggests that the *hyf* operon is not expressed and may thus be cryptic. The aim here is to discover whether the Hyf/Hyd-4 system is indeed capable of hydrogenase activity and can thus be assigned a fourth hydrogenase. Analysis of the NCBI database showed that *hyf* is conserved in all *E. coli/Shigella* strains, except for the most basal phylogroup (B2) which is associated with avian- and uro-pathogenic strains; this might suggest a role for Hyf in colonisation of the mammalian intestine. To determine if *hyf*-encoded FocB acts as a second formate channel, a *focA* mutant was complemented with *focB* carried by pBAD<sub>rha</sub>. Rhamnose induction of *focB* restored sensitivity to hypophosphite (formate analogue) supporting a role for FocB as a second formate channel. Similarly, induction of pBAD<sub>rha</sub>-encoded *focA* also restored hypophosphite sensitivity. FocA supported H<sub>2</sub> production at low pH, but had little impact at high pH. Low levels of FocB could compensate for absence of FocA. Plasmid induced *focA* restored the levels of formate released in a *focA* mutant at pH 6.5, to those seen in the wildtype, whereas plasmid induced *focB* increased such levels more modestly, although both FocA and FocB supported fermentative growth at pH 6.5. Expression of *focB* appeared toxic at pH 7.5. Thus, the results generally support a role for FocB in formate

export and suggest roles in delivery of exogenous formate to FHL-1 indicating a role in formate import also under acidic pH conditions.

When a *hycE* mutant strain was complemented with *hyf* carried by pBAD<sub>rha</sub> at pH 6.5, H<sub>2</sub> production was restored, and this was accompanied by restoration of formate consumption; a similar effect was seen upon plasmid-borne *hyc* induction. This indicates that *hyf* specifies an active hydrogen-evolving hydrogenase. Complementation of a  $\Delta fdhF \Delta hycE$  double mutant with both *hyf* and *fdhF*, carried by pBAD<sub>rha</sub> and pBAD<sub>ara</sub>, respectively, re-enabled H<sub>2</sub> production and formate consumption, thus showing that Hyf activity is Fdh-H dependent indicating that Hyf and Fdh-H combine to form a second FHL (FHL-2). FHL-2 activity was formate concentration dependent, with increased H<sub>2</sub> production with increased formate provided. Induced-Hyf H<sub>2</sub>-production activity was enhanced by provision of nickel at up to 0.25 mM (but was inhibited at 0.5 mM nickel) supporting the presence of Ni in the proposed [NiFe] centre of the Hyf-hydrogenase. Molybdenum and selenium were also shown to be required for high Hyf-FdhH (FHL-2) H<sub>2</sub>-production activity. In conclusion, this thesis provides strong evidence showing that the *hyf-focB* locus encodes a second formate channel and a fourth hydrogenase (Hyd-4) in *E. coli*, with the Hyd-4 enzyme acting as an H<sub>2</sub> producing and formate dependent enzyme that forms a second FdhF-dependent FHL complex. Further work is required to establish the environmental conditions under which the Hyf system is active, and offers an advantage, in the wildtype.

## **POSTER/ ORAL PRESENTATION**

### **Oral presentations:**

1. The postgraduate symposium, University of Reading, 09 June 2015.
2. The Annual Doctoral Research Conference, University of Reading, 23 June 2016.
3. The postgraduate symposium, University of Reading, 26 May 2016.
4. Microbiology research day, University of Reading, 18 August 2016.

### **Poster presentation:**

1. Microbiology research away day, London road campus, University of Reading, 19 July 2016.
2. Microbiology research day, University of Reading, 18 August 2016.
3. 11<sup>th</sup> Hydrogenase Conference, Marseille, France, 10-14 July, 2016.

## **Acknowledgements**

Firstly, I would like to express my sincere gratitude to my advisor prof Simon Andrews for the continuous support of my PhD study, for his patience, motivation and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Rob Jackson for his insightful comments and encouragement, and Dr. Shiela for her guidance and precise advice as an internal examiner.

My sincere thanks also go to Mr. Simon Fiest, Mr. Kevan and Mr. Jordan, all CSS family who gave access to the laboratory and research facilities. Without their precious support it would not be possible to conduct this research.

I would also like to thank all the current and previous members of Prof. Simon Andrew's iron research lab, especially; Rana Hilmi, Rabab Makharita, Christopher Moon, Abdul-Ameer Ghareeb, Fawzi Issa, Nancy Elhalfawy, Salem Al-aidy, and Dhama, Arvind, Afrah Salman, Marwa, Fawzih Al-Zahrani, Ameena, Noora, Kang, Shayma, Lwes, Zara, Fawziyea and my best friend Muzna for their scientific criticized discussions and being incessantly supporting and helping me to overcome difficulties during my PhD.

# Table of Contents

	<b>Page</b>
<b>Abstract</b> .....	iii
 <b>Chapter 1: Literature review</b>	
1. Introduction.....	1
1.1. Biofuel.....	1
1.1.1. Hydrogen as biofuel.....	5
1.2. Biotechnological production of hydrogen.....	7
1.3. Energy conversion in bacteria.....	10
1.3.1. Respiration.....	10
1.3.2. Fermentation.....	12
1.4. Hydrogenase enzyme in bacteria.....	14
1.4.1. Introduction.....	14
1.4.2. [Ni-Fe]-Hydrogenase.....	16
1.4.3. [Fe-Fe]-Hydrogenase.....	18
1.4.4. [Fe]-Hydrogenase.....	20
1.5. Hydrogen production and consumption in <i>E. coli</i> .....	22
1.5.1. Hydrogen consuming enzymes- Hyd-1 and Hyd-2.....	22
1.5.2. Hydrogen producing enzymes- Hyd-3.....	23
1.5.3. Hydrogen producing enzymes- Hyd-4.....	24
1.5.3.1. Comparison between Hyf and Hyc.....	25
1.5.4. Regulation of hydrogenase in <i>E. coli</i> .....	28
1.5.4.1. Other hydrogenase regulators.....	30
1.6. Some functional studies of Hyf.....	33
1.7. Formate transport and its role in formate hydrogen lyase (FhlA).....	34
1.7.1. <i>E. coli</i> is hypophosphite sensitive under fermentative growth.....	36
1.8. Systems required for formate dehydrogenase and hydrogenase formation.....	37
1.8.1. Types of formate dehydrogenase.....	37
1.8.2. Nickel uptake and its incorporation in <i>E. coli</i> .....	38
1.8.3. Selenium and molybdenum uptake and their incorporation.....	40
1.8.4. Accessory genes ( <i>hyp</i> genes).....	41
1.9. Aims and objectives of this study.....	43
 <b>Chapter 2: Materials and Methods</b>	
2.1. Chemicals.....	45
2.2. Enzymes.....	45
2.3. The size marker.....	45
2.4. Antibiotics.....	46

2.5.	Microbiological media.....	46
2.5.1.	Lysogeny broth (LB) and agar.....	46
2.5.2.	Werkmann Minimal media (WM).....	47
2.5.2.1.	Trace elements.....	47
2.5.3.	Rich media.....	48
2.5.4.	TGYEP media.....	48
2.6.	Media sterilisation.....	48
2.7.	Equipment.....	49
2.8.	Strains.....	49
2.9.	Plasmids.....	53
2.10.	DNA manipulation and analysis methods.....	54
2.10.1.	Extraction of genomic DNA.....	54
2.10.2.	Polymerase chain reaction (PCR).....	55
2.10.2.1.	Colony PCR.....	55
2.10.3.	Purification of PCR products.....	56
2.10.4.	Restriction digestion.....	57
2.10.5.	DNA concentration determination.....	57
2.10.6.	Ligation reactions.....	57
2.10.7.	Plasmid ‘miniprep’ isolation.....	57
2.10.8.	Gel electrophoresis.....	59
2.10.8.1.	Agarose gel electrophoresis.....	59
2.10.8.2.	Native Polyacrylamide gel electrophoresis.....	59
2.11.	Cloning into plasmids.....	60
2.11.1.	Primers designing for gene amplification using PCR..	60
2.11.2.	Preparation of plasmid DNA.....	65
2.12.	Preparation of competent cells for transformation.....	66
2.12.1.	Transformation into chemically competent cells and plating of the cells.....	66
2.12.1.1.	Heat-shock transformation.....	66
2.12.1.2.	Electroporation transformation.....	67
2.12.2.	Gene cloning into vectors.....	67
2.12.2.1.	In-Fusion® HD cloning.....	67
2.12.2.2.	Primers designing for In-Fusion® HD cloning.....	68
2.12.2.3.	In-Fusion cloning protocol.....	68
2.12.2.4.	In-Fusion transformation procedure using Stellar™ competent cells.....	68
2.13.	Elimination of the Kanamycin resistance cassette.....	69
2.14.	Gene inactivation procedure “The knockout”.....	70
2.14.1.	Lambda (λ) Red disruption system.....	70
2.14.2.	Primer design.....	70
2.14.3.	PCR amplification of CAT cassette.....	71
2.14.4.	Induction and preparation of host cells.....	71
2.14.5.	Electroporation with linear DNA.....	72
2.15.	Formate-hypophosphite inhibition test.....	72
2.16.	Hydrogen gas production assay.....	73
2.17.	Formate assay.....	74
2.18.	Protein cell extract and hydrogenase activity staining.....	75
2.18.1.	Protein sample preparation.....	75
2.18.2.	Determination of protein concentration.....	75

2.18.3.	Hydrogenase activity staining.....	76
2.19.	Phenotypic studies.....	76
2.19.1.	Phenotypic analysis of $\Delta focA$ , $\Delta focB$ , $\Delta hyc$ and $\Delta hyf$ ..	76

### Chapter 3: Bioinformatics analysis

3.1.	Bioinformatic data collection.....	78
3.2.	Phylogenetic tree.....	79
3.3.	Physical map.....	83
3.4.	Detailed comparison of the <i>hyf</i> locus in <i>hyf</i> <sup>+</sup> and <i>hyf</i> <sup>-</sup> <i>E. coli</i> strains.....	90
3.5.	Discussion and conclusion.....	93

### Chapter 4: Analysis of *focB* function

4.1.	Introduction.....	97
4.2.	Comparison between FocA and FocB amino acid sequences.....	99
4.3.	Cloning of <i>focA</i> and <i>focB</i> into inducible vectors.....	100
4.3.1.	Genomic DNA extraction of wildtype (MG1655).....	101
4.3.2.	Isolation and digestion of pBAD <sub>rha</sub> .....	101
4.3.3.	PCR amplification of <i>focA</i> and <i>focB</i> .....	103
4.3.4.	Cloning of <i>focA</i> and <i>focB</i> into the vector, pBAD <sub>rha</sub> ....	103
4.4.	Complementation of the <i>focA</i> mutant phenotypic by <i>focB</i> .....	106
4.4.1.	Hypophosphite anaerobic-growth inhibition is FocA dependent.....	106
4.4.2.	Can formate protect against hypophosphite toxicity?.....	110
4.4.3.	Can the resistance of the <i>focA</i> mutant to hypophosphite be reversed by complementation with <i>focA</i> and/or <i>focB</i> ?.....	114
4.5.	The effect of pH and formate on hypophosphite sensitivity upon induction of <i>focA</i> or <i>focB</i> .....	118
4.6.	Effect of <i>focA/focB</i> status on formate production and consumption, and H <sub>2</sub> production .....	126
4.6.1.	The <i>focA</i> and <i>focB</i> mutants.....	126
4.6.2.	Complementation with pBAD <sub>rha</sub> expression <i>focA</i> and <i>focB</i> .....	128
4.7.	Discussion.....	135

### Chapter 5: Phenotypic analysis of the effect of *hyf* operon induction in *E. coli* K-12

5.1.	Introduction .....	141
5.2.	Cloning of <i>hyc</i> and <i>hyf</i> into inducible vectors.....	143

5.2.1.	Extraction and amplification of <i>hyc</i> and <i>hyf</i> from the wildtype (MG1655) .....	143
5.2.2.	Cloning of <i>hyc</i> and <i>hyf</i> into the pBAD <sub>rha</sub> vector .....	145
5.3.	Can pBAD <sub>rha</sub> -encoded <i>hyf</i> produce hydrogen in a <i>hyc</i> mutant background? .....	148
5.3.1.	The effect of pH and formate on hydrogen production in the wildtype, $\Delta$ <i>hyf</i> and $\Delta$ <i>hyc</i> mutants. ....	148
5.3.2.	The effect of <i>hyc</i> or <i>hyf</i> induction on H <sub>2</sub> production in the wildtype and $\Delta$ <i>hycE</i> mutant. ....	153
5.3.2.1.	Is <i>hyf</i> complementation of the $\Delta$ <i>hycE</i> mutant achieved by <i>hyfG</i> replacement of <i>hycE</i> function?.....	157
5.4.	Does formate stimulate H <sub>2</sub> production by Hyf? .....	161
5.5.	Effect of <i>hyc/hyf</i> status on formate production and consumption .....	166
5.5.1.	The <i>hycE</i> mutant .....	166
5.5.2.	Effect on formate consumption of complementation of the <i>hyc</i> mutant with pBAD <sub>rha</sub> - <i>hyc</i> or - <i>hyf</i> .....	169
5.6.	Identification of Hyd-4 activity by gel activity staining .....	174
5.7.	Can <i>hyf</i> -expressing strains use formate as an energy source? .....	176
5.8.	Discussion .....	179

## Chapter 6: Role of FdhH in Hyf activity

6.1.	Introduction .....	182
6.1.1.	Hyd-3 as part of the FHL complex .....	182
6.1.2.	Formate dehydrogenase-H .....	182
6.1.3.	Formate dehydrogenase-H and hydrogenase-4 .....	185
6.2.	Cloning of <i>fdhF</i> into an inducible vector .....	186
6.2.1.	Isolation and digestion of pBAD <sub>ara</sub> .....	187
6.2.2.	Extraction, amplification and cloning of <i>fdhF</i> .....	187
6.2.3.	Cloning of <i>fdhF</i> into the vector, pBAD <sub>ara</sub> .....	188
6.3.	Complementation of <i>fdhF</i> and $\Delta$ <i>hycE</i> mutant phenotypes by <i>hyf</i> and <i>fdhF</i> .....	190
6.3.1.	Confirmation that <i>fdhF</i> is required for production of hydrogen .....	190
6.3.2.	Complementation of the $\Delta$ <i>fdhF</i> mutant with pBAD <sub>ara</sub> - <i>fdhF</i> .....	192
6.3.3.	Effect of introduction of pBAD <sub>rha</sub> - <i>hyc</i> and - <i>hyf</i> into the $\Delta$ <i>fdhF</i> mutant on H <sub>2</sub> production .....	193
6.4.	Inactivation of <i>fdhF</i> in the $\Delta$ <i>hycE</i> mutant by knockout using the $\lambda$ red procedure.....	194
6.4.1.	Amplification of pKD3 to get the Cm <sup>R</sup> cassette .....	196
6.4.2.	Confirmation of the ( $\Delta$ <i>fdhF</i> ):: <i>cat</i> substitution mutation and $\Delta$ <i>hycE</i> mutation in the $\Delta$ <i>hycE</i> mutant strain .....	197
6.5.	Effect of <i>fdhF</i> and <i>hyf</i> status on hydrogen production in the $\Delta$ <i>hycE</i> $\Delta$ <i>fdhF</i> mutant .....	201
6.5.1.	Complementation of the double mutant with <i>fdhF</i> or <i>hyf</i> .....	201

6.5.2.	Complementation of the $\Delta hycE \Delta fdhF$ mutant with pBAD <sub>ara/rha</sub> -expressed <i>fdhF</i> and <i>hyf</i> .....	202
6.6.	Increased Hyf-FdhF-dependent H <sub>2</sub> production in the complemented $\Delta hycE \Delta fdhF$ double mutant .....	205
6.6.1.	Effect of inducer levels on H <sub>2</sub> production .....	205
6.6.2.	Effect of formate addition on H <sub>2</sub> production .....	206
6.7.	Effect of the plasmid-specified Hyf-FdhF and Hyc-FdhF, FHL complexes on formate production and consumption in the $\Delta hycE \Delta fdhF$ mutant .....	207
6.8.	Discussion .....	211

## Chapter 7: Effect of metals and accessory genes on the activity of Hyf

7.1.	Introduction .....	214
7.2.	Accessory genes .....	214
7.2.1.	<i>hyp</i> operon .....	214
7.2.2.	<i>nik</i> operon .....	218
7.3.	Metals required in hydrogenase activity .....	218
7.3.1.	Nickel .....	218
7.3.2.	Iron .....	219
7.3.3.	Molybdate .....	219
7.3.4.	Selenium .....	219
7.4.	Is there any increase in H <sub>2</sub> production with addition of nickel to wildtype <i>E. coli</i> with pBAD <sub>rha-hyc</sub> or <i>-hyf</i> ? ..	220
7.5.	Effect of metals on H <sub>2</sub> production and bacterial growth	223
7.6.	Discussion .....	232

## Chapter 8. General discussion

8.1.	Introduction .....	234
8.2.	<i>hyf</i> is not fully conserved in <i>E. coli</i> strains .....	235
8.3.	FocB as a formate transporter .....	236
8.4.	Hyf is an active H <sub>2</sub> -producing (and formate-consuming) hydrogenase, and can thus be designated as the fourth hydrogenase (Hyd-4) of <i>E. coli</i> .....	238
8.5.	FdhH is required for Hyf activity .....	240
8.6.	Hyf (FHL-2) activity is Ni, Se and Mo dependent .....	241
8.7.	Conclusion .....	242
8.8.	Future work .....	243

<b>References</b>	245
-------------------	-----

<b>Appendix</b>	269
-----------------	-----

## List of Figures

		<b>Page</b>
<b>Figure 1.1.</b>	The best-known biofuels (bioethanol and biodiesel), production trends from 2005 to 2017.....	4
<b>Figure 1.2.</b>	Schematic mechanism of the photosynthesis in photoautotrophic microorganisms.....	6
<b>Figure 1.3.</b>	Different approaches in biological hydrogen production. ....	9
<b>Figure 1.4.</b>	Mixed acid fermentation of glucose by <i>E. coli</i> . ....	14
<b>Figure 1.5.</b>	The heterodimer [Ni-Fe] structure in <i>D. fructosovorans</i> . ....	17
<b>Figure 1.6.</b>	Molecular structure of [Fe-Fe]-hydrogenase enzyme. ....	19
<b>Figure 1.7.</b>	Structure of [Fe]-hydrogenase enzyme. ....	21
<b>Figure 1.8.</b>	Organisation and function of FHL complex. ....	24
<b>Figure 1.9.</b>	A schematic map of Hyc, Hyf and colmplex1. ....	27
<b>Figure 1.10.</b>	Functional model for the proposed ‘Hyf:Fdh-H’ complex (FHL-2). ....	27
<b>Figure 1.11.</b>	Schematic diagram of FNR regulation and its activation and inactivation by Oxygen. ....	28
<b>Figure 1.12.</b>	ArcA and ArcB component signal transduction systems, quinone dependent. ....	30
<b>Figure 1.13.</b>	The four domains of sigma factor 54 ( $\sigma^{54}$ ).....	32
<b>Figure 1.14.</b>	Structure of FocA. ....	35
<b>Figure 1.15.</b>	Diagram with sequence of each transmembrane segments and their connections. ....	36
<b>Figure 1.16.</b>	Schematic illustration of membrane-bound formate dehydrogenase in <i>E. coli</i> . ....	38
<b>Figure 1.17</b>	Uptake of nickel across cytoplasmic membrane in <i>E. coli</i> , and the regulation of nickel ion uptake by components of <i>nik</i> operon .....	40
<b>Figure 1.18.</b>	Genetic organisation of <i>hyc</i> and <i>hyp</i> operon on the <i>E. coli</i> chromosome. ....	42
<b>Figure 2.1.</b>	DNA marker. ....	45
<b>Figure 2.2.</b>	Diagram describing the process of primers design and PCR product generation for use in generating single gene knockout strain. ....	71
<b>Figure 2.3.</b>	Hydrogen assay using syringes.....	73

<b>Figure 3.1.</b>	Physical map and gene organisation of <i>hyf</i> operon in the <i>E. coli</i> K-12 chromosome .....	78
<b>Figure 3.2.</b>	Phylogenetic tree of all <i>E. coli</i> strains and close relatives within the complete microbial chromosome database of NCBI .....	82
<b>Figure 3.3.</b>	Physical map of the <i>hyf</i> region of <i>E. coli</i> (W3110) and adjacent genes. ....	83
<b>Figure 3.4.</b>	<i>E. coli</i> strains with some missing <i>hyf</i> genes. ....	86
<b>Figure 3.5.</b>	Translated sequences of the two fragments of <i>hyfD</i> in <i>Shigella sonnei</i> (53G). ....	87
<b>Figure 3.6.</b>	Translated sequences of the two fragments of <i>hyfC</i> in <i>E. coli</i> (55989). ....	87
<b>Figure 3.7.</b>	Translated sequences of the two fragments of <i>hyfC</i> in <i>E. coli</i> (55989). ....	88
<b>Figure 3.8.</b>	Different <i>E. coli</i> (SE15) strain with no <i>hyf</i> operon. ....	89
<b>Figure 3.9.</b>	<i>E. coli</i> (MG1655) with full <i>hyf</i> region. ....	89
<b>Figure 3.10.</b>	Nucleotide sequence alignments of the upstream <i>hyf</i> region of three <i>hyf</i> <sup>+</sup> <i>E. coli</i> strains .....	91
<b>Figure 3.11</b>	Nucleotide sequence alignments of the downstream <i>hyf</i> region of three <i>hyf</i> <sup>+</sup> <i>E. coli</i> strains .....	92
<b>Figure 3.12</b>	Schematic map showing the organisation of the <i>hyf</i> region in <i>hyf</i> <sup>+</sup> and <i>hyf</i> <sup>-</sup> <i>E. coli</i> strains. ....	92
<b>Figure 4.1.</b>	Sequence alignment of FocA from <i>E. coli</i> (FocAE) and <i>Salmonella</i> (FocAS) with FocB from <i>E. coli</i> . ....	100
<b>Figure 4.2.</b>	Agarose gel (0.7%) electrophoretic analysis of genomic DNA from <i>E. coli</i> MG1655. ....	101
<b>Figure 4.3.</b>	Agarose gel (0.7%) electrophoretic analysis of undigested pBAD <sub>rha</sub> . ....	102
<b>Figure 4.4.</b>	Agarose gel (0.7%) electrophoretic analysis of pBAD <sub>rha</sub> following restriction digestion with <i>Nde</i> I and <i>Bam</i> HI. ....	102
<b>Figure 4.5.</b>	Gel electrophoretic analysis of the <i>focA</i> and <i>focB</i> PCR products. ....	103
<b>Figure 4.6.</b>	Gel electrophoretic analysis of pBAD <sub>rha</sub> - <i>focA</i> or <i>focB</i> . ....	104
<b>Figure 4.7.</b>	Restriction map of pBAD- <i>focA</i> and pBAD- <i>focB</i> . ....	105
<b>Figure 4.8.</b>	Effect of hypophosphite on anaerobic growth and the impact of <i>focA/focB</i> status. ....	108
<b>Figure 4.9.</b>	Effect of hypophosphite under aerobic and anaerobic growth...	110
<b>Figure 4.10.</b>	The effect of formate on inhibition by 0.5 mM hypophosphite at pH 6.5. ....	112
<b>Figure 4.11.</b>	The effect of pH on HP growth inhibition under anaerobic fermentative conditions. ....	113
<b>Figure 4.12.</b>	The effect of <i>focA</i> and <i>focB</i> complementation of hypophosphite sensitivity. ....	116

<b>Figure 4.13.</b>	Effect of pH on the toxicity of hypophosphite upon induction of <i>focA</i> or <i>focB</i> in a wildtype or <i>focA</i> background during anaerobic growth. ....	122
<b>Figure 4.14.</b>	Effect of pH on the toxicity of hypophosphite upon induction of <i>focA</i> or <i>focB</i> in a wildtype or <i>focA</i> background during aerobic growth. ....	124
<b>Figure 4.15.</b>	The effect of <i>focA</i> mutation on formate concentration during fermentative growth compared to wt in WM-medium at pH 6.5. ....	131
<b>Figure 4.16.</b>	Hydrogen production in <i>focA</i> mutant compared to wildtype at pH 6.5. ....	131
<b>Figure 4.17.</b>	Effect of <i>focA</i> or <i>focB</i> induction on formate production and consumption in a wildtype or <i>focA</i> mutant at pH 6.5. ....	132
<b>Figure 4.18.</b>	Effect of <i>focA</i> or <i>focB</i> induction on formate production and consumption in a wildtype or <i>focA</i> mutant at pH 7.5. ....	133
<b>Figure 4.19.</b>	Hydrogen production upon induction of <i>focA</i> or <i>focB</i> in a wildtype or $\Delta$ <i>focA</i> background. ....	134
<b>Figure 4.20.</b>	A view of FocA channel in <i>Salmonella enterica</i> serovar Typhimurium. ....	136
<b>Figure 5.1</b>	Agarose gel (0.7%) electrophoretic analysis of purified <i>hyc</i> operon DNA ....	144
<b>Figure 5.2</b>	Agarose gel (0.7%) electrophoretic analysis of purified <i>hyf</i> operon DNA ....	144
<b>Figure 5.3.</b>	Map of the pBAD <sub>rha</sub> - <i>hyc</i> and pBAD <sub>rha</sub> - <i>hyf</i> plasmids. Relevant restriction sites. ....	146
<b>Figure 5.4.</b>	Agarose gel (0.7%) electrophoretic analysis of the pBAD- <i>hyc</i> constructs. ....	147
<b>Figure 5.5.</b>	Gel electrophoretic analysis of pBAD <sub>rha</sub> - <i>hyf</i> constructs. ....	147
<b>Figure 5.6.</b>	Effect of $\Delta$ <i>hycE</i> or $\Delta$ <i>hyfG</i> mutation on hydrogen production at different pH levels (pH 6, 6.5, 7 and 7.5). ....	151
<b>Figure 5.7.</b>	Effect of formate on hydrogen production in <i>E. coli</i> wildtype and $\Delta$ <i>hycE</i> mutant at pH 6-7.5. ....	152
<b>Figure 5.8.</b>	Plasmid induced <i>hyc</i> or <i>hyf</i> enables production of hydrogen gas in a $\Delta$ <i>hycE</i> mutant. ....	156
<b>Figure 5.9.</b>	Agarose gel (0.7%) electrophoretic analysis of <i>hyfG</i> DNA....	159
<b>Figure 5.10.</b>	Hydrogen production upon <i>hyc</i> and <i>hyfG</i> induction along with vector control in $\Delta$ <i>hycE</i> mutant.....	159
<b>Figure 5.11.</b>	Hydrogen production upon induction of <i>hyc</i> and <i>hyf</i> along with vector control in $\Delta$ <i>hycE</i> and $\Delta$ <i>hycA-E</i> mutants.....	161
<b>Figure 5.12.</b>	Hydrogen production upon induction of <i>hyc</i> or <i>hyf</i> in a wildtype or $\Delta$ <i>hycE</i> background, at pH 5.5. ....	162
<b>Figure 5.13.</b>	Hydrogen production upon induction of <i>hyc</i> or <i>hyf</i> in a wildtype or $\Delta$ <i>hycE</i> background at pH 6.5. ....	164
<b>Figure 5.14.</b>	Hydrogen production upon induction of <i>hyc</i> or <i>hyf</i> in a wildtype or $\Delta$ <i>hycE</i> background at pH 7.5. ....	166

<b>Figure 5.15.</b>	The effect of <i>hycE</i> mutation on formate levels during fermentative growth in WM-medium at pH 6.5. ....	168
<b>Figure 5.16.</b>	Hydrogen production in <i>hycE</i> mutant compared to wildtype at pH 6.5 .....	168
<b>Figure 5.17.</b>	Effect of complementation with <i>hyc</i> or <i>hyf</i> on formate production and consumption in a wildtype or <i>hycE</i> mutant at pH 6.5. ....	171
<b>Figure 5.18.</b>	Effect of <i>hyc</i> or <i>hyf</i> induction on formate production and consumption in a wildtype or <i>hycE</i> mutant at pH 7.5. ....	173
<b>Figure 5.19.</b>	Analysis of Hyd-4 hydrogenase activity by activity-staining and native PAGE in Tris-glycine buffer .....	175
<b>Figure 5.20.</b>	Analysis of Hyd-4 hydrogenase activity by activity-staining and native PAGE in Tris-barbitone buffer. ....	176
<b>Figure 5.21.</b>	Effect of <i>hyf</i> and <i>hyc</i> induction on growth with 10 mM formate. ....	178
<b>Figure 6.1.</b>	The structure of hydrogenase-3 and FdhF to form the FHL-1 complex. ....	184
<b>Figure 6.2.</b>	Crystal structure of formate dehydrogenase-H ( <i>fdhF</i> ) of <i>E. coli</i> . ....	184
<b>Figure 6.3.</b>	The structure of hydrogenase-4 and FdhF to form the FHL-2 complex. ....	186
<b>Figure 6.4.</b>	Agarose gel (0.7%) electrophoretic analysis of pBAD <sub>ara</sub> . ....	187
<b>Figure 6.5.</b>	Gel electrophoretic analysis of <i>fdhF</i> amplification from MG1655 with two different sets of primers. ....	188
<b>Figure 6.6.</b>	Gel electrophoretic analysis of cloned <i>fdhF</i> in pBAD <sub>ara</sub> . ....	189
<b>Figure 6.7.</b>	Map of pBAD <sub>ara</sub> - <i>fdhF</i> . ....	190
<b>Figure 6.8.</b>	Effect of <i>fdhF</i> mutation of H <sub>2</sub> production with and without formate. ....	192
<b>Figure 6.9.</b>	Hydrogen production upon induction of <i>fdhF</i> in the $\Delta$ <i>fdhF</i> and $\Delta$ <i>hycE</i> mutants. ....	193
<b>Figure 6.10.</b>	A schematic diagram of the knockout strategy using $\lambda$ Red system. ....	196
<b>Figure 6.11.</b>	Agarose gel (0.7%) electrophoretic analysis of the PCR amplified Cm <sup>R</sup> cassette from pKD3 .....	197
<b>Figure 6.12.</b>	Agarose gel (0.7%) electrophoretic analysis of colony PCR amplification of the <i>fdhF</i> locus in two suspected ( $\Delta$ <i>fdhF</i> ):: <i>cat</i> mutants. ....	198
<b>Figure 6.13.</b>	Agarose gel (0.7%) electrophoretic analysis of pJET1.2-( $\Delta$ <i>fdhF</i> ):: <i>cat</i> plasmids. ....	199
<b>Figure 6.14.</b>	Agarose gel (0.7%) electrophoretic analysis of the <i>hycE</i> PCR product of the putative $\Delta$ <i>hycE</i> ( $\Delta$ <i>fdhF</i> ):: <i>cat</i> strain. ....	200
<b>Figure 6.15.</b>	Agarose gel (0.7%) electrophoretic analysis of the <i>hycE</i> PCR product of the wildtype strain (contain <i>hycE</i> ). ....	200
<b>Figure 6.16.</b>	Expression of <i>hyf</i> fails to restore H <sub>2</sub> production in the $\Delta$ <i>fdhF</i> $\Delta$ <i>hycE</i> double mutant. ....	202

<b>Figure 6.17.</b>	Restoration of H <sub>2</sub> production in the $\Delta hycE \Delta fdhF$ strain by complementation with <i>hyf</i> and <i>fdhF</i> expressing plasmids or with <i>hyc</i> and <i>fdhF</i> expressing plasmids. ....	204
<b>Figure 6.18.</b>	Growth of $\Delta hycE \Delta fdhF$ strain by complementation with <i>hyf</i> and <i>fdhF</i> expressing plasmids or with <i>hyc</i> and <i>fdhF</i> expressing plasmids. ....	204
<b>Figure 6.19.</b>	Effect of formate concentration on hydrogen production by the plasmid encoded Hyf-FdhF and Hyc-FdhF FHL complexes in the $\Delta hycE \Delta fdhF$ strain. ....	207
<b>Figure 6.20.</b>	Effect of plasmid-specified <i>fdhF</i> and <i>hyf</i> , or plasmid-specified <i>fdhF</i> and <i>hyc</i> induction on formate production and consumption in the $\Delta hycE \Delta fdhF$ double mutant at pH 6.5.....	210
<b>Figure 7.1.</b>	The position of <i>hypF</i> close to the <i>hyc</i> operon in <i>E. coli</i> . ....	216
<b>Figure 7.2.</b>	The organisation of <i>hyp</i> locus in <i>E. coli</i> . ....	216
<b>Figure 7.3.</b>	The biosynthesis of the first part of the [NiFe] cofactor. ....	217
<b>Figure 7.4.</b>	The biosynthesis of second half of the [NiFe] cofactor. ....	217
<b>Figure 7.5.</b>	Effect of different concentrations of nickel (II) on hydrogen production in <i>E. coli</i> BW25113 with pBAD <sub>rha-hyc</sub> induction..	221
<b>Figure 7.6.</b>	Effect of different concentrations of nickel (II) on hydrogen production in <i>E. coli</i> BW25113 with pBAD <sub>rha-hyf</sub> induction...	223
<b>Figure 7.7.</b>	Effect of iron and nickel on bacterial growth and hydrogen production in <i>wildtype</i> strains complimented with pBAD <sub>rha-hyc</sub> , pBAD <sub>rha-hyf</sub> and the control. ....	226
<b>Figure 7.8.</b>	Effect of iron and nickel on bacterial growth and hydrogen production in <i>hycE</i> mutant strains complimented with pBAD <sub>rha-hyc</sub> , pBAD <sub>rha-hyf</sub> and the control. ....	227
<b>Figure 7.9.</b>	Effect of Molybdenum and Selenium on bacterial growth and hydrogen production in <i>wildtype</i> strains upon induction of pBAD <sub>rha-hyc</sub> and pBAD <sub>rha-hyf</sub> and the control. ....	230
<b>Figure 7.10.</b>	Effect of Molybdenum and Selenium on bacterial growth and hydrogen production in <i>hycE</i> mutant strains upon induction of pBAD <sub>rha-hyc</sub> and pBAD <sub>rha-hyf</sub> and the control. ....	231

## List of Tables

	Page
<b>Table 2.1.</b> Antibiotics used in this study.....	46
<b>Table 2.2.</b> Strains used in this study. ....	49
<b>Table 2.3.</b> Plasmids used in this study .....	53
<b>Table 2.4.</b> Primers designed for PCR amplification for <i>focA</i> and <i>focB</i> cloning into pBAD <sub>rha</sub> and their sequence primers used in this study .....	60
<b>Table 2.5.</b> Primers designed for PCR amplification of <i>hyf</i> operon of <i>E. coli</i> MG1655 for In-Fusion cloning and its sequence after the cloning into pBAD <sub>rha</sub> .....	61
<b>Table 2.6.</b> Primers designed for PCR amplification of <i>hyc</i> operon of <i>E. coli</i> MG1655 for In-Fusion cloning and its sequence after the cloning into pBAD <sub>rha</sub> .....	62
<b>Table 2.7.</b> primers designed for PCR amplification of <i>fdhF</i> of <i>E. coli</i> MG1655 for In-Fusion cloning into pBAD <sub>ara</sub> .....	64
<b>Table 2.8.</b> Primers designed for PCR amplification of <i>fdhF</i> and <i>hycE</i> in <i>E. coli</i> MG1655 for post transduction and <i>kan</i> cassette removal confirmation .....	64
<b>Table 2.9.</b> Primers designed for PCR amplification of pKD3 as chloramphenicol template for <i>fdhF</i> knockout in the $\Delta$ <i>hycE</i> strain .....	65
<b>Table 2.10.</b> Primers designed for PCR amplification of <i>hyfG</i> in <i>E. coli</i> MG1655 for cloning into pBAD <sub>rha</sub> in this study.....	65
<b>Table 3.1.</b> The presence of <i>hyfD</i> , <i>hyfE</i> and <i>hyfF</i> in <i>E. coli</i> strains and close relatives with percentage match indicated. ....	80
<b>Table 4.1.</b> pH values before and after growth related to Fig. 4.12 .....	113
<b>Table 4.2.</b> pH values difference before and after growth related to fig. 4.13. ....	114
<b>Table 5.1.</b> pH values difference before and after growth related to Fig. 5.8. ....	157

## Abbreviation

<b>Amp</b>	<b>Ampicillin</b>
<b>Amp<sup>R</sup></b>	Ampicillin resistance
<b>APS</b>	Ammonium persulphate
<b>Cm<sup>R</sup></b>	Chloramphenicol resistant
<b>CoA</b>	Co- enzyme A
<b>dNTP</b>	Deoxyribonucleotide triphosphate
<b>EDTA</b>	Ethylene diaminetetra acetic acid
<b>H<sup>+</sup></b>	Proton
<b>h</b>	Hour(s)
<b>IPTG</b>	Irsopropyl-β-D-thiogalactoside
<b>Kan</b>	Kanamycin
<b>kDa</b>	KiloDaltons
<b>LA</b>	Luria-Bertani agar medium
<b>LB</b>	Lysogeny broth medium
<b>M</b>	Molar
<b>min</b>	Minute(s)
<b>NAD</b>	Nicotinamide adenine dinucleotide (oxidised)
<b>NADH</b>	Nicotinamide adenine dinucleotide (reduced)
<b>OD</b>	Optical density
<b>OD<sub>600</sub></b>	Optical density at 600 nanometer
<b>PAGE</b>	Polyacrylamide gel electrophoresis
<b>PCR</b>	polymerase chain reaction
<b>rpm</b>	Revolutions per minute
<b>SDS</b>	Sodium dodecyl sulphate
<b>TM</b>	Trans membrane
<b>TEMED</b>	N, N, N',N', Tetramethyl ethylene diamine
<b>Tris</b>	Tris (hydroxymethyl) amino methane
<b>°C</b>	Degrees centigrade

## **Chapter 1: Literature review**

### **1. Introduction**

#### **1.1 Biofuel**

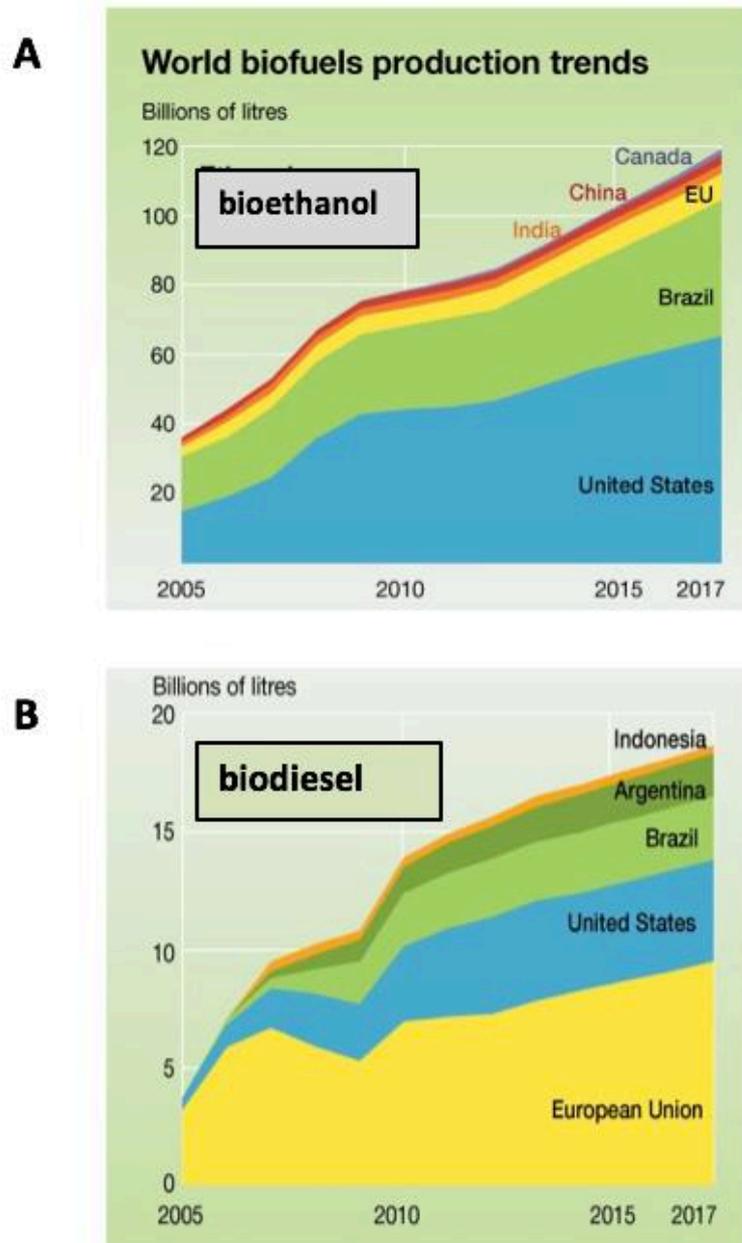
With the expected increase of global energy requirements, there is a growing concern about energy supply security, air pollution and worldwide climate change. Today's energy systems are based mostly on fossil fuels. Oil, natural gas and coal are examples of fossil fuels, which are used to fuel cars and airplanes, and heat our homes (Urbaniec *et al.*, 2010). However, the amount of resource is limited, because fossil fuels are formed by natural processes whereby buried dead organisms (plants or animals) are decomposed over millions of years to form coal, petroleum, or natural gas. In addition, fossil fuels cause damage to the environment when used for energy by humans as they represent the largest source of 'unnatural' emissions of carbon dioxide (Miyake *et al.*, 1999). Indeed, CO<sub>2</sub> is a greenhouse gas that contributes to global warming. The burning of fossil fuels produces around 21.3 billion tons of CO<sub>2</sub> per year. But it is estimated that natural processes can only absorb about half of that amount, so there is a net increase of 10.65 billion tons of atmospheric carbon dioxide per year (International conference 2018-2019). Beside CO<sub>2</sub>, burning fossil fuels releases over 100 toxic chemicals that directly affect human health. One of the main toxic chemicals that is emitted the PAHs (polyaromatic hydrocarbons), which cause irritation, nausea, vomiting and diarrhea during short-term exposure. However, the long-term symptoms are cataracts, kidney and liver damage and jaundice (Illinois Department of Public Health, 2009). Furthermore, burning of fossil fuel has resulted in the accumulation of the heavy metals lead and mercury, and derived compounds, in the environment, which can damage human organs and causes neurodevelopmental damage (Martin and Griswold, 2009). In addition, benzene is released which is carcinogenic (Hennig, 2017). Thus, development of renewable energy

in place of fossil fuels is an urgent mission and an increasing focus on clean energy as an alternative for satisfying growing energy demand is required (Das and Veziroğlu, 2001). Biofuels are fuels produced from organic matter in a relatively short period of time (Savaliya *et al.*, 2013). Biofuel history started when man first discovered fire. Wood, dung and charcoal were the first forms of fuel (biofuel) used for cooking and heating (Demirbas and Balat, 2006). Furthermore, whale oil was widely used in the mid-1700s and early 1800s and was the fuel of choice for lighting houses (O'Connor, 2010). In 1834, the first US patent was awarded to S. Casey, who used alcohol derived from whale oil as a lamp fuel (Kovarik, 1998).

Biofuels are classified into three generations (Fig. 1.1). These generations differ in the sources (feed stock) from which the fuel is derived. First generation of biofuels are produced directly from feedstock that also can be consumed as human food, such as corn, sugarcane, wheat, starch, sugar, vegetable oil and animal fats (Lee and Lavoie, 2013). Such biofuels include biodiesel (produced from vegetable oil derived from seeds of soybean or sunflower), bioethanol (fermentation product of glucose, a simple sugar from sugarcane or starch crops) and biogas (anaerobic fermentation of organic food and crop residues). The second-generation biofuels are known as 'advanced biofuels' and they differ in not being produced from edible food crops or crop products not useful for consumption (Oliver and King, 2009). They are produced from grasses, waste from agricultural residues and other solid wastes (human waste, woody crops and yard clippings) which are difficult to extract biofuel from (Demirbas, 2009). The second generation biofuels include bioethanol produced from cellulose and methanol which derived from lignocellulose and municipal wastes. The third-generation biofuels refer to those derived from algae. Previously, algae were grouped with second-generation biofuels. However, when it became apparent that algae are capable of much higher yields

with lower resource inputs than other feedstock, it was suggested that they be moved to their own category (BC Bioenergy Strategy, 2010). Production of biofuels from algae usually depends on the lipid content of the microorganisms. *Chlorella* species are usually targeted because of their high lipid content (60 to 70%). Many biofuel types can be produced from algae, such as biodiesel, butanol, gasoline, methane, bioethanol, vegetable oil and jet fuel (Liang *et al.*, 2009; Behera *et al.*, 2014).

The most well-known biofuels are bioethanol and biodiesel. They are the most widely used globally, because they are simple and easy to produce. In 2012, the UNEP (United Nations Environment Programme) report indicated that the supply of both bioethanol and biodiesel is steadily increasing (Fig. 1.1). Markets for both are increasing, not only in the European Union, Brazil and the United States, but also in countries such as China, India and Argentina. (UNEP report, 2012).



**Figure 1.1. The best-known biofuels (bioethanol and biodiesel), production trends from 2005 to 2017. (a) Bioethanol production trend (b) Biodiesel production. As shown by (UNEP report, 2012).**

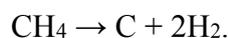
### 1.1.1 Hydrogen as biofuel

Hydrogen is the most abundant element in the universe; it makes about three-quarters of all matter. The atmosphere contains about 0.07% hydrogen, whereas the earth's crust contains 0.14% of hydrogen (Dharmaraj and AdishKumar, 2012). It may be considered a perfect biofuel, because hydrogen produces no pollutants or greenhouse gases when it is used. The only emission it makes is water vapour when it reacts with oxygen in air, so it is environmentally friendly. Additionally, hydrogen's physical and chemical characteristics make it a good candidate as a fuel for cars (Kotay and Das, 2008). On the other hand, many technical challenges must be overcome for major deployment of hydrogen as a fuel. For example, hydrogen production is expensive and it is difficult to transport by pipelines (Hallenbeck and Ghosh, 2012), because hydrogen has an active electron, which can easily migrate into the crystal structure of most metals. Thus, hydrogen pipes have to resist such corrosion.

Hydrogen can be produced by a number of different processes such as, electrolysis of water; thermo-catalytic reformation of hydrogen-rich organic compounds and biological processes (Levin *et al.*, 2004). Water electrolysis is usually dependent on enormous electrical currents required to break the water molecules into oxygen and hydrogen as shown in this equation:



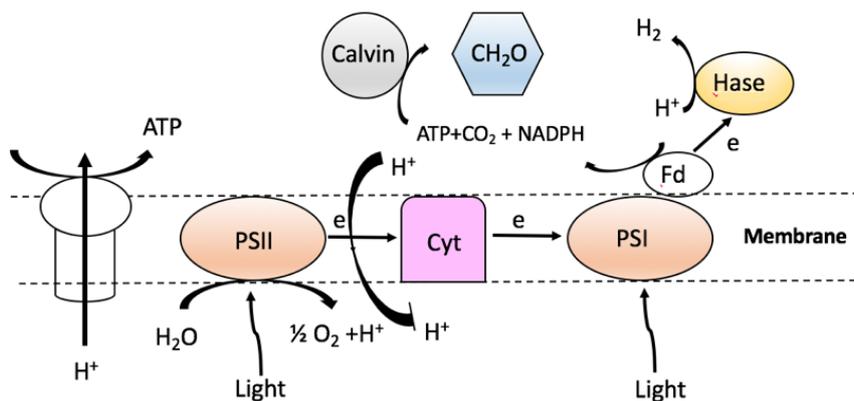
The other approach to produce hydrogen without forming CO<sub>2</sub> is thermos-catalytic reformation of hydrogen-rich organic compounds, which is achieved by decomposition of methane (CH<sub>4</sub>) and production of solid carbon and pure hydrogen steam from natural gas:



Because this reaction is mildly endothermic, the temperature must be more than 600 °C

for the reaction to proceed. Hydrogen production via this process has been achieved using solar radiation, a molten metal bath, and thermal reactors with no catalyst (Dunker *et al.*, 2006).

The third process in hydrogen production involves the use of microorganisms (Levin *et al.*, 2004). Hydrogen production using microorganisms is mainly achieved using photosynthetic microorganisms such as cyanobacteria and green algae. These have the ability to use energy from sun light through photosynthesis to produce hydrogen (Yu and Takahashi, 2007). Cyanobacteria and green algae use chlorophyll a and other pigments, as well as photosynthetic systems (PSI and PSII), to capture sun light to perform plant-like photosynthesis, and at the end of the process they produce hydrogen as a by-product (section 1.2; Fig 1.2). Hydrogen produced by microalgae and bacteria is called ‘biohydrogen’. The first biohydrogen production from fermentative bacteria was reported in the 1930s, followed by light-dependent hydrogen production from microalgae and photosynthetic bacteria (Benemann, 1998).



**Figure 1.2. Schematic mechanism of the photosynthesis in photoautotrophic microorganisms.** CH<sub>2</sub>O, carbohydrates; Calvin, Calvin cycle; Hase, Hydrogenase; Fd, Ferredoxin; ATP, Adenosine tri-phosphate (energy source) and PS, Photosystem. As shown by (Yu and Takahashi, 2007).

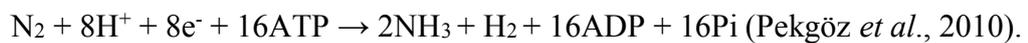
## 1.2. Biotechnological production of hydrogen

In order to foster the reduction of fossil fuels and their environmental effects, hydrogen has been suggested as an alternative form of energy. About 50% of fertilizers and 37% of petroleum products use hydrogen for manufacturer, so hydrogen production has an importance beyond as a direct fuel. In fertilizers, hydrogen is used to make ammonia (as a source of nitrogen for plants). However, petroleum industries also use hydrogen in processing of petroleum products to break down crude oil into fuel oil (Kane, 2006). Globally, 40% of hydrogen is produced from natural gas, 30% from heavy oils and Naphtha, 18% from coal and 4% from electrolysis (Das and Veziroglu, 2008). For economic use of hydrogen as a fuel, the development of efficient, large-scale and sustainable H<sub>2</sub> production systems are a necessity. Basic research in the biotechnological hydrogen production field started only in the late 1920s with bacterial H<sub>2</sub> production and in the 1940s with micro-algal hydrogen production. The first attempts highlighted the importance of anaerobic conditions in H<sub>2</sub> production with some bacteria (Hallenbeck and Benemann, 2002).

The processes of biological hydrogen production can be classified into two distinct types, light dependent (direct bio-photolysis, indirect bio-photolysis, photo-fermentations) and light-independent (dark-fermentation) (Fig. 1.2). Direct bio-photolysis is the process of algal photosynthesis, where water is oxidized and oxygen is produced. Light energy is absorbed by photosystem II (PSII), which generates electrons and transferred to ferredoxin (iron-sulphur proteins that mediate electron transfer) using the light energy absorbed by photosystem I (PSI). A reversible hydrogenase accepts electrons directly from the reduced ferredoxin to generate hydrogen (Fig. 1.3a) (Lee *et al.*, 2010). The indirect approach is the use of bi-directional hydrogenases with the help of specific cell-types in cyanobacteria called resting cells or 'heterocysts', which perform nitrogen

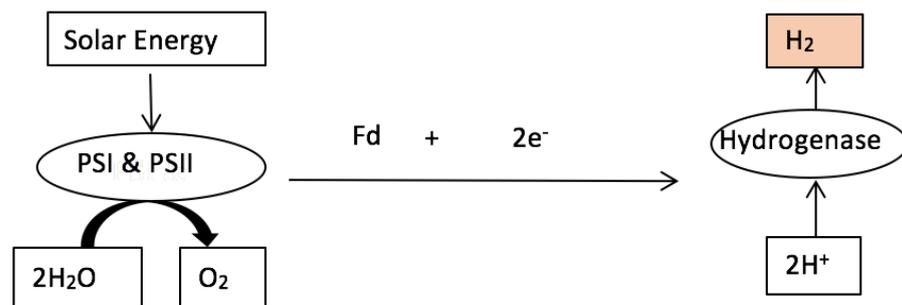
fixation, and have the ability to oxidize and synthesize hydrogen (Fig. 1.3b) (Kirtay, 2011).

In photo-fermentation, photosynthetic bacteria produce hydrogen from different organic acids and agricultural wastes (Hallenbeck and Benemann, 2002) (Fig. 1.3c). The photosynthetic non-sulphur (PNS) bacteria are an example of photosynthetic bacteria used to produce hydrogen from volatile fatty acids (VFA) under anaerobic conditions. PNS bacteria also have the ability to produce hydrogen from carbon sources such as glucose or sucrose via a photo-fermentation pathway (Argun and Kargi, 2011). Both hydrogenase and nitrogenase enzymes in PNS photo-fermentative bacteria have a role in hydrogen production (Hellenbeck and Benemann, 2002). *Rhodobacter capsulatus*, an example of a PNS bacterium, produces biohydrogen via photo-fermentation. It has two kinds of nitrogenase enzymes [Mo-nitrogenase] and [Fe-nitrogenase] that are responsible for hydrogen production. Nitrogenase enzymes normally reduce nitrogen to ammonia, and also produce hydrogen as a by-product:

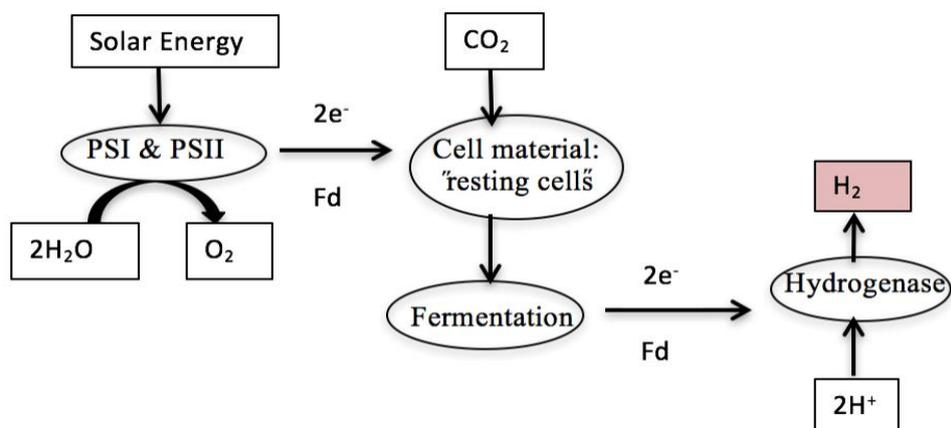


However, in dark-fermentation, anaerobic bacteria, which are grown on carbohydrate-rich substrates, can produce hydrogen by fermenting these carbohydrate-rich substrates to hydrogen and other products, such as acids (lactic, acetic, butyric, etc) and alcohols (ethanol, butanol, etc) (Hallenbeck and Ghosh, 2009) (Fig.1.3d). Species of the *Enterobacter* genus, *Bacillus coagulans* (IIT-BT-S1) isolated from anaerobically sewage sludge and *Clostridium*, produce hydrogen in the dark. Different amounts of hydrogen per mole of glucose are produced, depending on the fermentation pathway and the fermentation end products (section 1.3.2) (Kotay and Das, 2008). Theoretically, in strict anaerobic bacteria such as *Clostridium* species, a maximum of 4 moles of hydrogen per

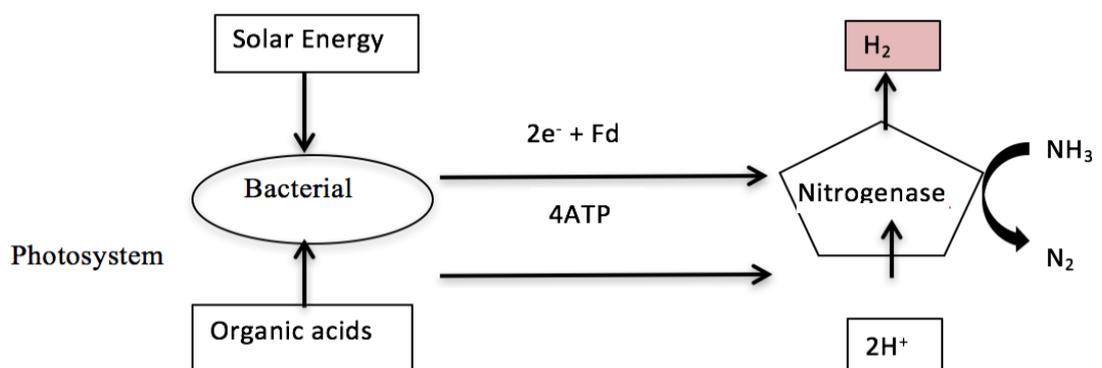
one mole of glucose are obtained, however, in facultative anaerobes such as *Escherichia coli*, a maximum of 2 moles of hydrogen per mole of glucose can be produced (Manish and Banerjee, 2008). Hydrogen production by these bacteria needs a bioreactor (batch or continuous) for good hydrogen yield (Levin *et al.*, 2004).



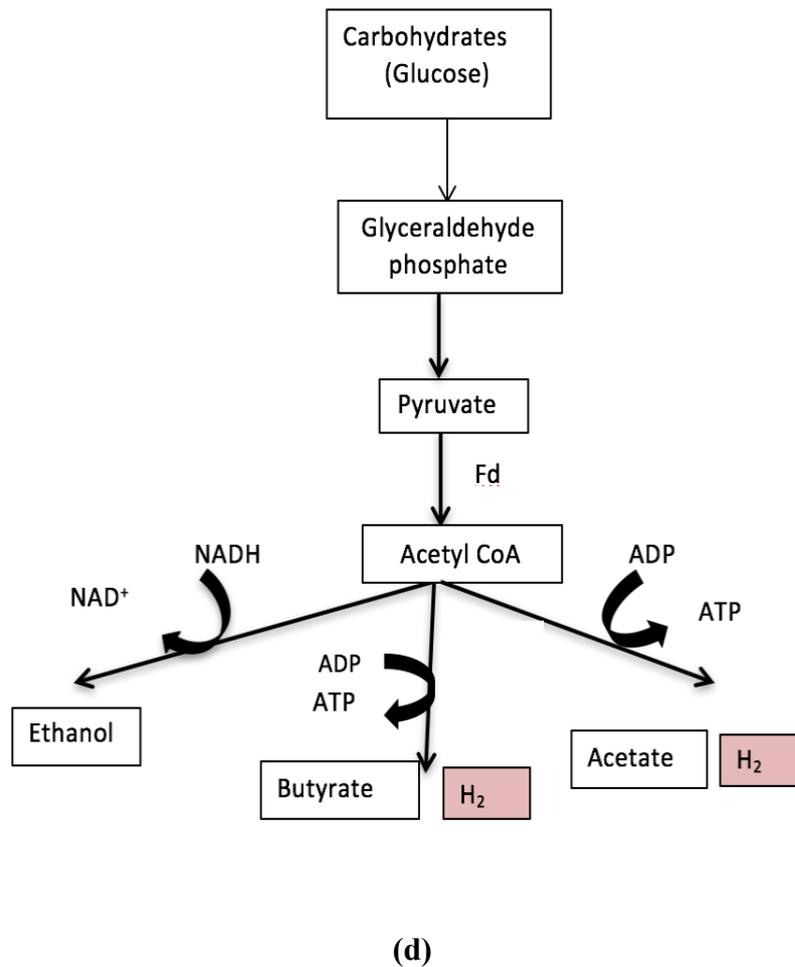
(a)



(b)



(c)



**Figure 1.3. Different approaches in biological hydrogen production.** (a) Direct photolysis (b) Indirect photolysis (c) Photo-fermentation (d) Dark fermentation. Fd (ferredoxins) are iron-sulphur proteins that mediate electron transfer in a range of metabolic reactions. As shown by (Lee *et al.*, 2010).

### 1.3. Energy conversion in bacteria

#### 1.3.1 Respiration

Respiration is a process by which chemical energy is released and partially captured in the form of adenosine tri-phosphate (ATP). Carbohydrates, fats and proteins can all be used as fuels in respiration, but glucose is most commonly used as an example to study the reactions and pathways involved.

The overall mechanism of respiration in many cases involves three processes: glycolysis, the Krebs' cycle and electron transport system. Glycolysis is a multistep metabolic pathway that occurs in the cytoplasm of microbial cells. The six-carbon glucose molecule is converted into intermediary compounds and then split into two three-carbon molecules to form pyruvic acid at the end of the process. During glycolysis, four ATP molecules are generated; two are consumed in the substrate phosphorylation process, giving a net energy output of two ATP molecules. Furthermore, during glycolysis, another energy source (two molecules of NADH) is produced, and since each molecule of NADH can generate three ATP molecules, a total of six ATPs is generated by glucose glycolysis from reducing equivalents associated with NADH. Following glycolysis, another multistep process is involved, the Krebs' cycle. Before the cycle begins, two molecules of pyruvic acid (which are produced from one glucose molecule by glycolysis) undergo conversion to acetyl-CoA which yields two high-energy molecules of NADH (thus 6 ATPs). Acetyl-CoA enters the cycle to eventually yield oxaloacetic acid. During the Krebs' cycle, a total of 2 ATP, 2 FADH (where each FADH can give 2 molecules of ATP) and 6 NADH molecules are produced per two acetyl-CoA molecules. At the end of this process, the total energy, which is produced starting with glycolysis and ending up with the Krebs' cycle, is 38 ATP molecules. The electrons that are released from the breakdown during glycolysis and the Krebs' cycle are collected by the electron carriers, NADH and FADH, in the form of hydride ions (Verhees *et al.*, 2003). During respiration, these electrons are passed from NADH and FADH to membrane-bound electron carriers, which are then passed to other electron carriers until they are finally associated with a terminal carrier (e.g. oxygen, resulting in the production of water). This electron transfer process releases energy which is harnessed for translocation of protons across the cytoplasmic membrane (prokaryotes) or inner mitochondrial membrane (eukaryotes) creating a

concentration gradient of protons which can be used to drive ATP synthase to produce ATP from ADP (Haddock and Jones, 1977). However, under anaerobiosis, other alternative electron acceptors may be utilized, such as nitrate, fumarate, nitrite, dimethylsulphoxide (DMSO) and trimethylamine N-oxide (TMAO) (Lin and Luchi, 1991). If these are not available, then energy can be obtained from glucose by fermentation.

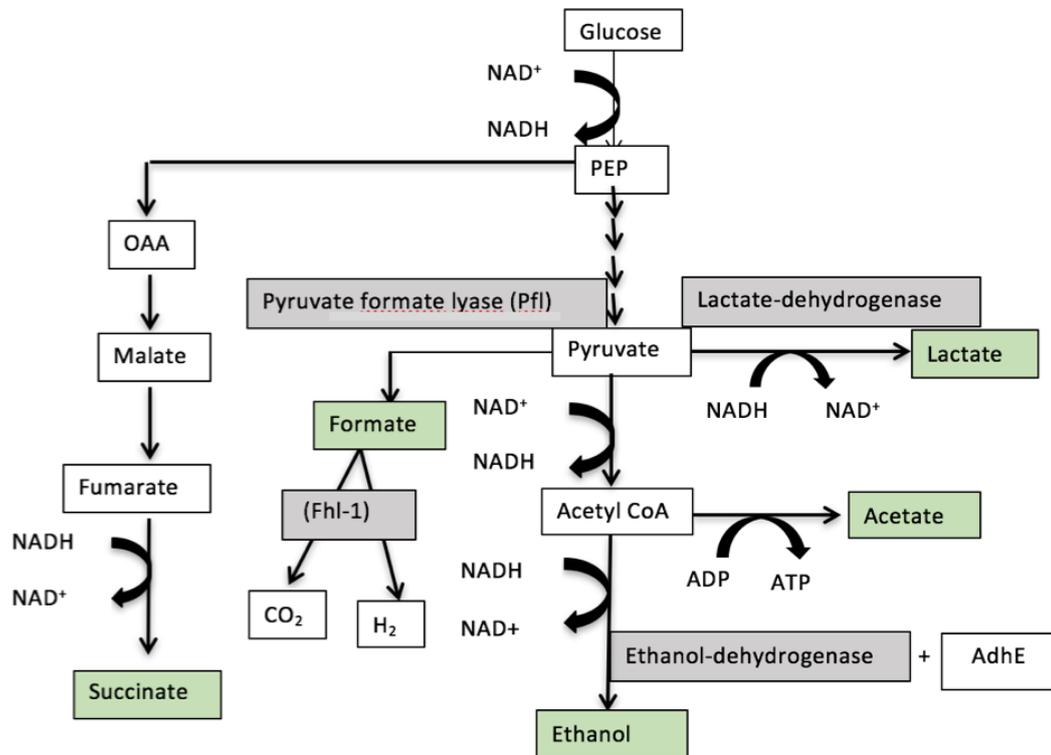
*Escherichia coli* is one of the preferred bacteria to study the energy and regulation of respiration (Unden and Bongaerts, 1997). *E. coli* shows flexibility in both electron input and output. In this case there is no cytochrome *bc*<sub>1</sub> complex, which is the third complex in mitochondrial electron transport chain. However, a number of electron acceptors can be used instead of oxygen (Bekker *et al.*, 2009).

The respiratory chain of *E. coli* is categorized by a highly variable composition of constituents and pathways compared to the mitochondria in eukaryotes. In *E. coli*, three types of quinone are involved, menaquinone (MK); demethylmenaquinone (DMK) and ubiquinone (UQ). When *E. coli* is grown on substrates that are catabolized via glycolysis and the Krebs' cycle, NADH is the major electron carrier. In this process, electrons are assumed to be transferred directly to UQ, (or DMK or MK, which are used in anaerobic respiratory pathways). The genome of *E. coli* encodes three oxygen-reducing terminal oxidases, cytochrome *bo* oxidase; cytochrome *bd* I oxidase and cytochrome *bd* II oxidase. All three terminal oxidases transfer electrons from ubiquinone to oxygen (Kim *et al.*, 2008).

### **1.3.2. Fermentation**

Fermentation is a type of anaerobic redox process where energy for ATP synthesis is provided from the breakdown of carbohydrates (Hjersing, 2011). Fermentation is used

by the cell to recycle NADH into NAD<sup>+</sup> and not to generate energy directly; so, glycolysis can continue anaerobically, as long as an appropriate substrate is available. When glucose and other carbohydrates are oxidized, the energy released is used to generate adenosine triphosphate (ATP) (Pollak *et al.*, 2007). The ATP yield from a fermentation process is dependent on the pathway used and it can range from 0.3 to 4 moles of ATP per a mole of glucose. This output is considerably smaller by, 9 fold, from that obtained during aerobic respiration. Therefore, most bacteria prefer to avoid fermentation if a respiratory electron acceptor is present (Müller, 2001). *E. coli* uses a type of fermentation called 'mixed acid fermentation' where pyruvate dehydrogenase is no longer synthesized. Instead, during fermentation pyruvate formate-lyase (PFL) is used in facultative anaerobic *E. coli*, or pyruvate ferridoxin oxido reductase (Pfor) in the case of strict anaerobes (Hallenbeck, 2009). Both enzymes are active in the absence of oxygen and both are used in breaking-down pyruvate to two products, formate, which is secreted as formic acid, and acetyl-CoA. Some pyruvate is converted to lactate via lactate dehydrogenase (Becker *et al.*, 1999). In addition, some acetyl CoA is converted to ethanol by ethanol dehydrogenase and the rest is converted to acetate which allows ATP generation. Most phosphoenolpyruvate (PEP) is converted to pyruvate but some is converted to succinate through three steps and NAD<sup>+</sup> is generated (Fig. 1.4). Subsequently, during mixed acid fermentation, four additional enzymes are required as discussed earlier, pyruvate formate lyase; lactate dehydrogenase; ethanol dehydrogenase and formate hydrogen lyase (Fhl-1), which cleaves formate to CO<sub>2</sub> and H<sub>2</sub> (Zhou *et al.*, 2003).



**Figure 1.4. Mixed acid fermentation of glucose by *E. coli*.** PEP: Phosphoenolpyruvate. OAA: Oxalo-acetic acid, AdhE: Alcohol dehydrogenase, Fhl-1: Formate hydrogen lyase -A system found in *hyc* operon. As illustrated by (Ruiz *et al.*, 2012; Thakker *et al.*, 2012).

## 1.4. Hydrogenase enzyme in bacteria

### 1.4.1. Introduction

Hydrogenases are metallo-enzymes that can be found in many microorganisms. These metallo-enzymes contain nickel and iron metals, which are earth-abundant metals, and contain organometallic ligands such as carbon monoxide and cyanide, which were previously unknown to exist in biological systems (Shafaat *et al.*, 2013). They are used in energy metabolism and are responsible for catalysing the reversible cleavage of hydrogen into protons and electrons (Hjersing, 2011; Shomura *et al.*, 2011). This

enzymatic activity was first stated in 1931, when it was shown that *E. coli* produce hydrogen during their growth at lower pH under anaerobic conditions (stated in Heinekey, 2009). Hydrogenases can be located in either the cytoplasm, the periplasm of bacteria or bound to the plasma membrane. The cytoplasmic hydrogenases are soluble, monomeric hydrogenases and found mainly in strict anaerobic bacteria such as *Clostridium pasteurianum*, *Megasphaera elsdenii* and *Cyanobacteria*. They have the ability in catalyse both H<sub>2</sub> uptake and evolution. The periplasmic hydrogenases are heterodimeric, as found in *Desulfovibrio* spp, which can be active under aerobic conditions. However, the membrane-bound hydrogenases are called energy-producing hydrogenases, which couple H<sub>2</sub> oxidation to electron dependent energy-generation via membrane-bound subunits and they can be synthesized under either aerobic or anaerobic conditions (Benoit *et al.*, 1998; Casalot and Rousset, 2001).

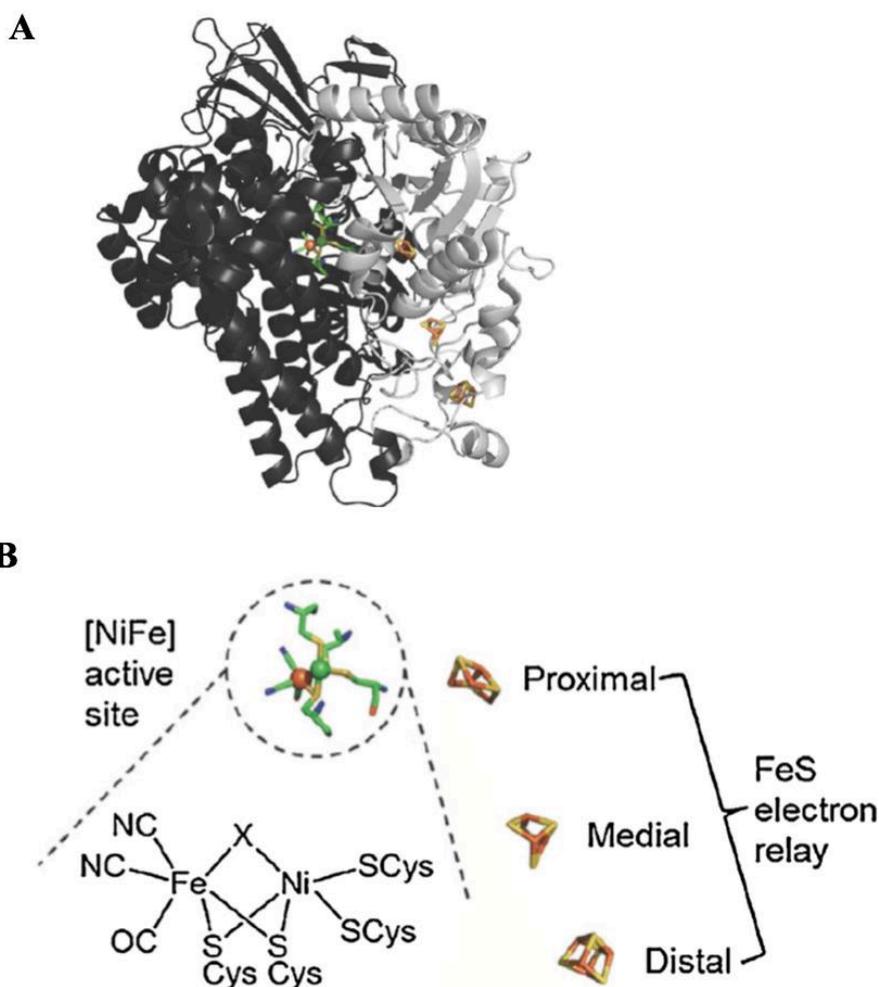
Hydrogenases are classified, according to their active site metal ion content, into three different classes: nickel-iron [Ni-Fe]-, di-iron [Fe-Fe]-, and mono-iron (iron-only) [Fe]-hydrogenases (Lukey *et al.*, 2010). In general, hydrogenases are sensitive to oxygen and this is applicable to *E. coli* hydrogenases (Abou Hamdan *et al.*, 2012). In *E. coli*, there are four [Ni-Fe] hydrogenases. These consist of a small subunit containing iron-sulphur clusters, used in electron transfer, and a large subunit containing the nickel-iron active site. Commonly, hydrogenase 1 (Hyd-1) and hydrogenase 2 (Hyd-2) oxidize hydrogen, while, hydrogenase 3 (Hyd-3) is a part of the Formate Hydrogen Lyase (FHL) complex, which produces hydrogen from formate. In *E. coli*, under anaerobic conditions, synthesis of Hyd-1 and Hyd-2 occurs during fermentation and these enzymes are generally associated with the appearance of the third hydrogenase, Hyd-3, which forms a part of the hydrogen evolving activity of the FHL (formate hydrogenlyase) complex (Soboh *et al.*, 2011). However, hydrogenase 4 (Hyd-4) appears to be inactive since the genes

encoding Hyd-4 are not expressed at significant levels (Sanchez-Torres *et al.*, 2013). However, Hyd-4 shows similar sequence in its majority subunits to Hyd-3 and is more similar to complex I than Hyd-3, and is also more complicated (nine rather than six subunits) (Andrews *et al.*, 1997).

#### 1.4.2. [Ni-Fe]-Hydrogenase

[Ni-Fe]-hydrogenase catalyses the reversible oxidation of H<sub>2</sub> to protons in prokaryotes. In 1995, the structure of [Ni-Fe]-hydrogenase was obtained from x-ray crystallography studies of five different sulphate-reducing bacteria: *D. vulgaris* Miyazaki F, *D. gigas*, *D. fructosovorans*, *D. desulfuricans* and *Desulfomicrobium baculatum* and *D. vulgaris* (Sawers, 1995). The [Ni-Fe]-hydrogenase in *D. fructosovorans* is a heterodimeric protein consisting of larger subunit of 62.5 kDa and carries the nickel-iron [Ni-Fe] at the active site (Ogata *et al.*, 2009). The active site nickel is coordinated by four cysteine-S ligands; two of them are also linked to an iron atom, which is coordinated by three diatomic ligands (two cyanides and one carbonyl) (Lukey *et al.*, 2010). The smaller subunit has a molecular mass of 28.8 kDa, and it contains three iron-sulphur [Fe-S] clusters (Fig. 1.5). These [Fe-S] clusters serve as an electron transport chain of the electrons, which are produced from the cleavage of H<sub>2</sub> at the [Ni-Fe] redox active site, to the electron acceptor cytochrome *c*<sub>3</sub> (Ogata *et al.*, 2009; Pinske and Sawers, 2012). [Ni-Fe]-hydrogenases function in hydrogen oxidation, as well as the transfer of electrons from H<sub>2</sub> to a redox partner (cytochrome P450). This activation of hydrogen has a potential application in energy generation. Generally, [Ni-Fe]-hydrogenases are sensitive to oxygen. Oxygen reacts with the active site and the result is an inactive state, which is Ni-A (inactive unready, oxidized state) and Ni-B (inactive ready, oxidized state) (Fritsch *et al.*, 2011). If the partial pressure of oxygen is high, it results in high Ni-B, whereas under low O<sub>2</sub>,

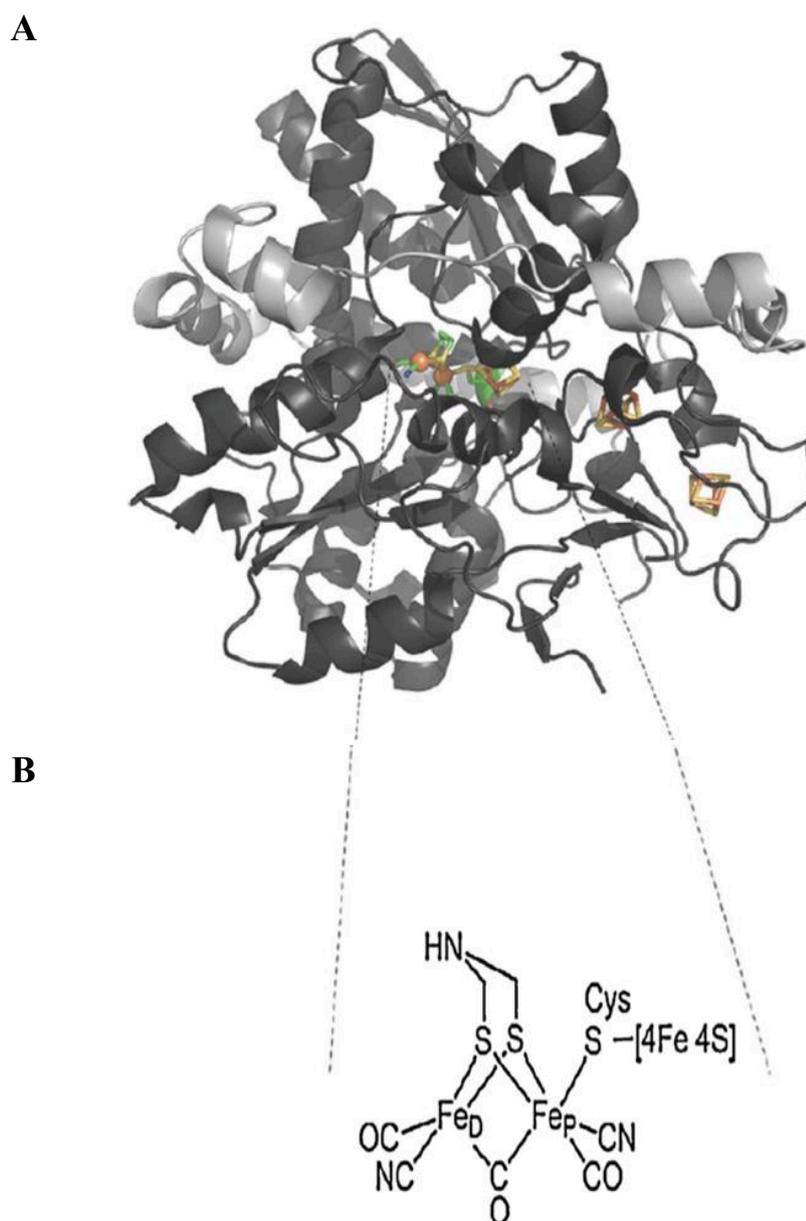
more Ni-A is formed. The interaction of O<sub>2</sub> with the enzyme and the inactive state was explained with the use of the isotope O<sup>17</sup>, which showed that O<sub>2</sub> must be close to the Ni atom and form a bridging ligand with the active centre, in order to be inactive (Ogata *et al.*, 2005; Shafaat *et al.*, 2013).



**Figure 1.5.** The heterodimer [Ni-Fe] structure in *D. fructosovorans*. **(A)** Ribbon representation of [Ni-Fe]-hydrogenase, which consisted of two subunits: the large subunit ([Ni-Fe] active site) (dark grey) and the small subunit (light grey) consist of three [Fe-S] clusters (P: proximal, M: medial and D: distal) as shown in (B). **(B)** The details of the metallic centre in the active site of [Ni-Fe] in its oxidised form. Green ball; Ni, small dark red ball; oxygen, blue; nitrogen, yellow; sulphur, green sticks; hydrocarbon, orange ball and sticks; iron (Ogata *et al.*, 2009).

### 1.4.3. [Fe-Fe] Hydrogenase

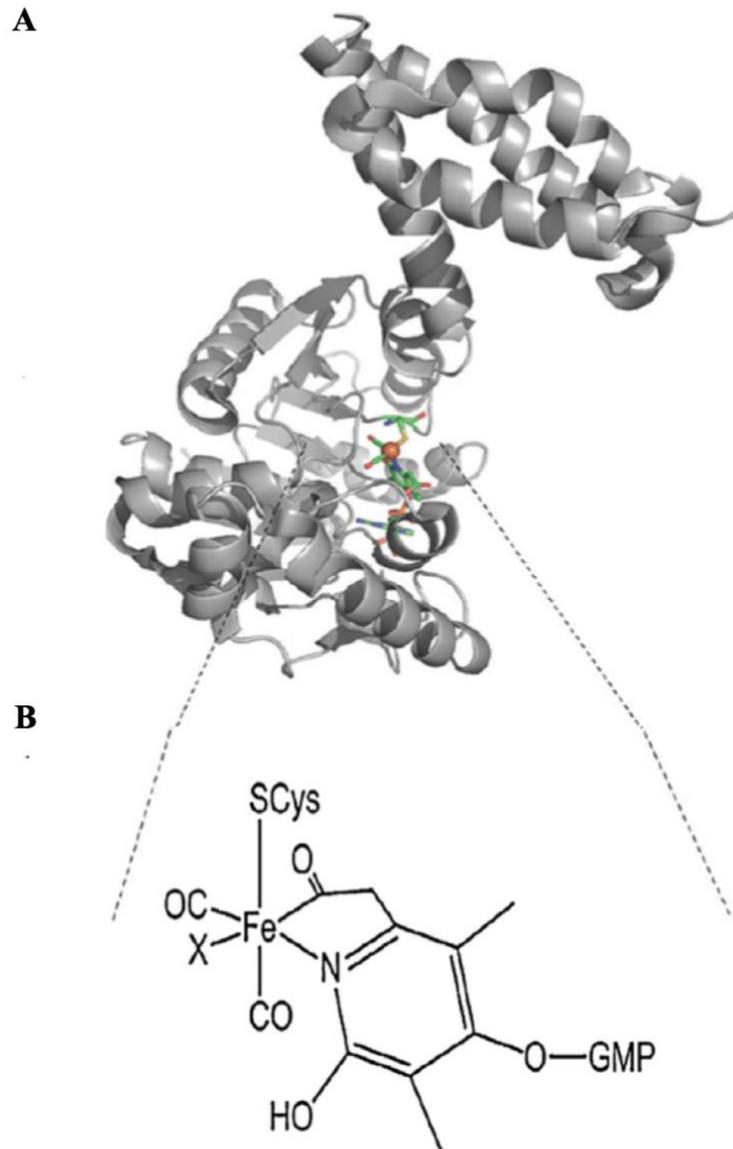
[Fe-Fe]-hydrogenases are usually found in strictly anaerobic bacteria, but are also present in green algae and some aerobic cyanobacteria. [Fe-Fe] enzymes consist of a di-nuclear iron site in the active centre of the enzyme (Berggren *et al.*, 2013). Two well-characterized examples are the monomeric, cytoplasmic [Fe-Fe] hydrogenase (61 kDa) in *Clostridium pasteurianum* and the dimeric periplasmic [Fe-Fe]-hydrogenase with a large subunit (43 kDa) and small subunit (10 kDa), as in *Desulfovibrio desulfuricans*. All [Fe-Fe] hydrogenases have only iron as the metal in their active site, which is named the 'H-cluster' (Adams, 1990). The H-cluster is composed of a [4Fe-4S] cubic cluster bridged to two Fe atoms coordinated by putative CO, CN<sup>-</sup> and di-thiolate ligands (Fig. 1.6). [Fe-Fe]- and [Ni-Fe]-hydrogenases catalyse the activation of H<sub>2</sub> through a reversible reaction (Fontecilla-Camps *et al.*, 2007 and Mulder *et al.*, 2010) although they are not similar in their genetic sequence (Vignais *et al.*, 2001; Vignais and Billoud, 2007), both have analogous overall structures as the active site is located within a large subunit and electrons are delivered to this centre through iron-sulphur clusters (small subunit) (Forzi and Sawers, 2007). Furthermore, both include the presence of iron bound to CO and CN ligands at the active site, which is uncommon in nature (Vignais and Colbeau, 2004). [Fe-Fe] enzymes are more active than [Ni-Fe]-hydrogenases by 10<sup>2</sup> fold in hydrogen production (Holtz *et al.*, 1991). The *D. desulfuricans* hydrogenase is able to produce hydrogen at up to 8700 molecules per second (Glick *et al.*, 1980 and Hatchikian *et al.*, 1992). However, it is more sensitive to oxygen during catalysis than the [Ni-Fe] enzymes (Pinske *et al.*, 2012). The catalytic activity of [Fe-Fe] is irreversibly inhibited once by O<sub>2</sub>. However, [Ni-Fe] is reversibly inhibited by O<sub>2</sub> (Przybyla *et al.*, 1992).



**Figure 1.6. Molecular structure of [Fe-Fe]-hydrogenase enzyme from *Clostridium pasteurianum*.** (A) The active site of [Fe-Fe] hydrogenase structure (orange ball). (B) The expanded view of the active site ‘H cluster’ as ball and other accessory clusters, stated in the text. As shown by (Shepard *et al.*, 2010; Mulder *et al.*, 2011).

#### 1.4.4. [Fe]-hydrogenase

The third type of hydrogenases are the [Fe]-only hydrogenase enzymes that contain a mononuclear Fe in its active site, neither Ni nor Fe-S (Nicolet *et al.*, 2002). [Fe]-hydrogenases are enzymes unlike the other two types, found only in some hydrogenotrophic methanogenic Archaea, which function in methanogenic pathways. It has different enzymatic redox mechanisms regarding how electrons are delivered to the active site with respect to the [Ni-Fe] and [Fe-Fe] enzymes. The electrons are directly transported to the active site over a short distance, where a cofactor methenyl-H<sub>4</sub>MPT<sup>+</sup> directly accepts the hydride from H<sub>2</sub>, so the hydrogenation of a methenyl-H<sub>4</sub>MPT<sup>+</sup> occurs by [Fe]-hydrogenase instead of oxidation/production of hydrogen as in the other two types (Hiromoto *et al.*, 2009; Salomone-Stagnia *et al.*, 2010). For the [Fe-Fe] and [Ni-Fe] hydrogenases, the electrons travel through a series of metallogenic clusters over a long distance and the structure of the active site remain as it is through the whole process. This third type contains a unique iron-guanylyl-pyridinol (FeGP) cofactor, in which the iron atom is attached by one cysteine-sulphur atom, two CO groups, one GMP molecule, and an sp<sup>2</sup>-hybridized nitrogen atom and an acyl carbon atom from the pyridinol ring (Shima *et al.*, 2008; Shima and Ermler, 2010). [Fe]-only hydrogenases are similar to the [Fe-Fe] type in that both are extremely sensitive to oxygen, unlike the [Ni-Fe] hydrogenases.



**Figure 1.7. Structure of [Fe]-hydrogenase enzyme from *Methanocaldococcus jannaschii*.** (A) The active site of [Fe]-hydrogenase is illustrated in monomeric protein subunit (grey). It zoomed in (B) (B) FeGP cofactor (shown as yellowish ball in (A)). GMP; Guanosine monophosphate molecule; N, hybrid nitrogen and SCys is cysteine sulphur atom. As shown by (Shima and Ermler, 2010)

## 1.5. Hydrogen production and consumption in *E. coli*

### 1.5.1. Hydrogen consuming enzymes – Hyd-1 and Hyd-2

In *E. coli*, four different types of hydrogenases are recognized, all of which are of the [Ni-Fe] type and membrane bound. Hydrogenase 1 (Hyd-1) and hydrogenase 2 (Hyd-2) are responsible for hydrogen uptake and oxidation (Trchounian and Trchounian, 2009). Hyd-1 is encoded by the six gene *hya* operon, *hyaABCDEF* (Vignais *et al.*, 2001). It is a heterodimer with a large subunit ( $\alpha$ -subunit, 60 kDa) encoded by *hyaB* carrying the [Ni-Fe] active centre and the small subunit ( $\beta$  subunit, 30 kDa) encoded by *hyaA*, which contains [Fe-S] clusters: a distal [4Fe-4S] cluster; a medial [3Fe-4S] cluster; and proximal [4Fe-3S] cluster that is essential for oxygen tolerance (Volbeda *et al.*, 2012, Dross *et al.*, 1992).

Hydrogenase 2 (Hyd-2) is encoded by an eight genes operon, *hybOABCDEFG*, with *hybC* coding for the large 58 kDa subunit ( $\alpha$ -subunit) and *hybO* coding for the small 30 kDa subunit ( $\beta$ -subunit) (Menon *et al.*, 1990).

The physiological role of Hyd-1 and 2 is to couple oxidation of hydrogen with the reduction of the quinone pool in the inner membrane (Pinske *et al.*, 2012). It was found that Hyd-1 is tolerant to oxygen, while Hyd-2 is a standard oxygen sensitive enzyme, as a result, both are repressed under aerobic environments (Lukey *et al.*, 2010). Hyd-1 and 2 are repressed also by nitrate or phosphate availability (Richard *et al.*, 1999). Hyd-1 is expressed highly under acidic condition, in the presence of formate and at the stationary phase (Brøndsted and Atlung, 1994). On the other hand, Hyd-1 activity is reduced in pyruvate formate-lyase mutants where formate cannot be produced. Unlike Hyd-1, Hyd-2 is inhibited by formate and under alkaline pH condition (Ballantine *et al.*, 1986) and expressed most strongly during the exponential phase of growth (Richard *et al.*, 1999). Hyd-2 activity is enhanced by glycerol and fumarate or hydrogen (Brøndsted and Atlung,

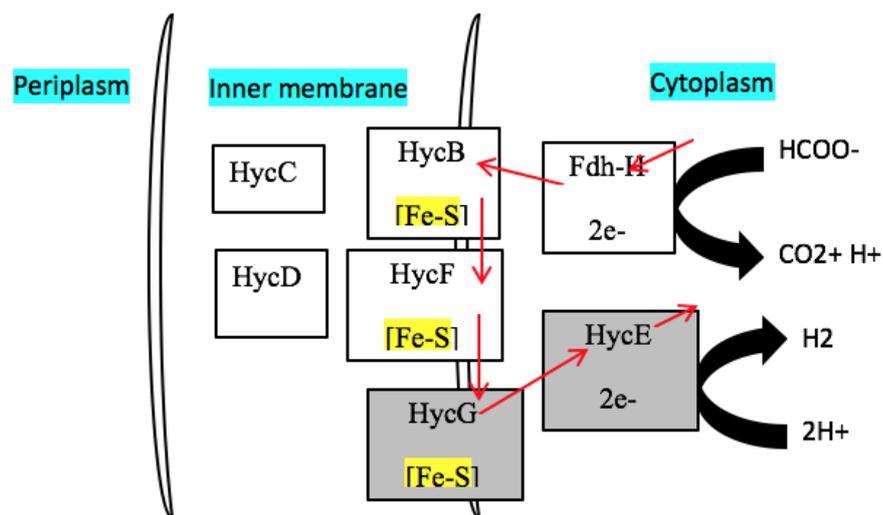
1994) and its activity in the presence of glycerol under neutral or slightly alkaline conditions (Trchounian and Trchounian, 2009).

In 2010, Kim and his co-workers showed that recombinant expression of Hyd-1 makes it able to produce hydrogen under micro-aerobic condition. The production was about 12.5 ml/h in 400 ml minimal media. However, there was no hydrogen production in the wildtype even with the addition of formate as a substrate of formate hydrogen lyase pathway. Sanchez-Torres *et al.* (2013) found, similar to previous studies, that Hyd-2 has a major impact on hydrogen production and maintenance of redox balance under anaerobic conditions in glycerol minimal media at pH 7.5, but had no effect on H<sub>2</sub> production at pH 6.5 or in complex medium. Loss of Hyd-1 and Hyd-4 had little impact on H<sub>2</sub> production.

### 1.5.2. Hydrogen producing enzymes – Hyd-3

Hydrogenase-3 is the third hydrogenase enzyme found in *E. coli*, which is encoded by the *hyc* genes (*hycBCDEFGHI*) (Bagramyan *et al.*, 2002). Hyd-3 contains large and small subunits that are characteristic of 'standard' [Ni-Fe] hydrogenases, and carries [Fe-S] clusters located in the hydrophilic part of the complex, which may form the electron transport pathway (Fig. 1.8) (Rossmann *et al.*, 1994). The large 65 kDa subunit of Hyd-3 is encoded by *hycE*, while the small 28 kDa subunit is encoded by *hycG*, which is tightly attach to the membrane and functions in electron transfer within the Fhl-1 complex. In order to produce hydrogen, Fhl-1 is the main enzyme needed, and is a complex of Hyd-3 (*hyc*) and 'formate dehydrogenase H' (FdhH) which is encoded by the *fdhF* gene. The Fhl-1 complex is responsible for the cleavage of formic acid to carbon dioxide and hydrogen during anaerobic fermentation under acidic pH condition (Sawers, 1995). The expression of the *hyc* operon is repressed by the *hycA* gene product, which encodes the

FhlA, the activator. It is also repressed by oxygen and nitrate, and induced by formate under fermentative growth conditions (Pecher *et al.*, 1982). The 22 kDa HycB and 20 kDa HycF subunits are membrane associated ferredoxin-like electron transport proteins, which act as electron carriers of the FhlA complex, where HycB is also the small subunit of FdhH (Sauter *et al.*, 1992). On the other hand, HycC (64 kDa) and HycD (33 kDa) are integral membrane-spanning proteins, which are anchored to the cytoplasmic side of the membrane and they are similar to Hyf components, and have homology to NADH-ubiquinone-oxidoreductase components (McDowall *et al.*, 2014) (Fig 1.8).



**Figure 1.8. Organisation and function of FHL complex.** FHL subunits are illustrated with electron transport pathway. [Fe-S], Iron-sulphur cluster; Fdh-H, Formate dehydrogenase-H. Red arrow describe the electron pathway. As shown by (Leonhartsberger, 2002).

### 1.5.3. Hydrogen producing enzymes – Hyd-4

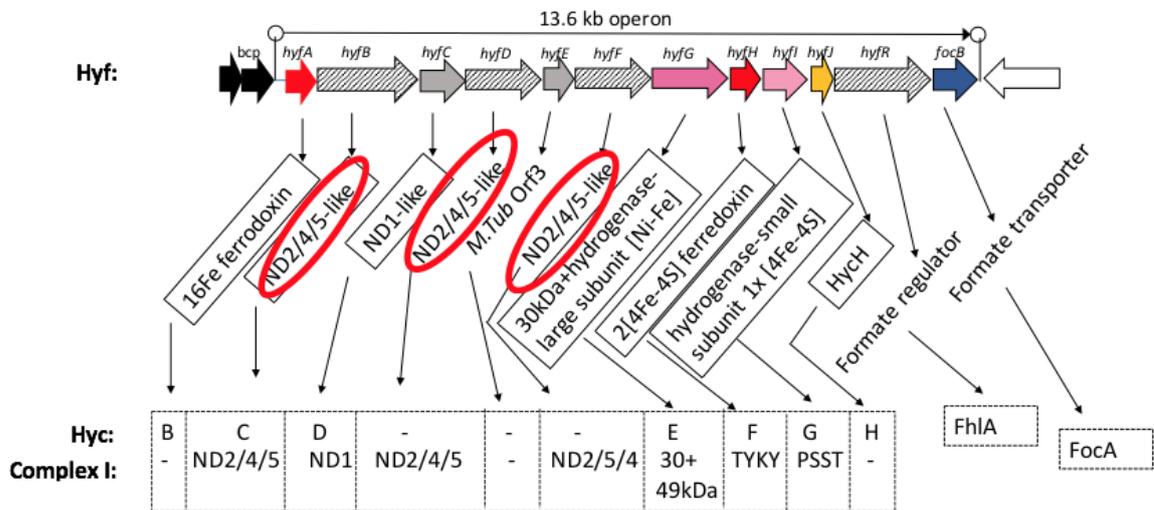
The *hyf* operon in *E. coli*, encoding the fourth hydrogenase isoenzyme (Hyd-4), is a twelve-gene operon (*hyfABCDEFGHIIR-focB*). The *hyf* operon is located from 55.8' to 56.0' in the genome and encodes a predicted ten-subunit hydrogenase complex, plus a formate- and sigma 54-dependent transcriptional activator (HyfR homologous with

FhlA), and a putative formate transporter channel (FocB homologous to FocA) (Andrews *et al.*, 1997). HyfA is a member of the ferredoxin family of bacterial proteins and is likely to be involved in electron transfer between Fdh-H and hydrogenase-4 (Sauter *et al.*, 1992). HyfB, HyfD and HyfF are not closely related to each other but they are polytopic membrane proteins and they are similar to the subunits (ND2, ND4 and ND5) of the proton-translocating NADH-ubiquinone oxidoreductase (complex I of the mitochondrial electron transport chain and Nuo in *E. coli*). Moreover, HyfD includes eight predicted transmembrane helices. However, HyfE is not homologous to complex I or Nuo components, but it is predicted to be found as a polytopic membrane protein. HyfG and HyfI are particularly closely related (73% and 63% identity) to the large and small subunits of hydrogenase-3 (HycE and HycG, respectively). HyfH is a member of the 8 Fe ferredoxin family. HyfJ is homologous to HycH of Hyd-3, which has an unknown function. Recently, it was found to act as a chaperone of HycH and its homologue, HyfJ, plays a dual- role in improving the assembly of formate hydrogen-lyase complex, although HycH is not a part of the final structure of the FHL complex (Lindenstrauß *et al.*, 2017). On the other hand, HyfR regulates transcription of the *hyf* operon in the presence of formate and it has similarity to FhlA (regulator of Hyd-3). FocB is closely resembles the formate transporter FocA, therefore it appears to be a second putative formate transporter (Andrews *et al.*, 1997) (Fig 1.9).

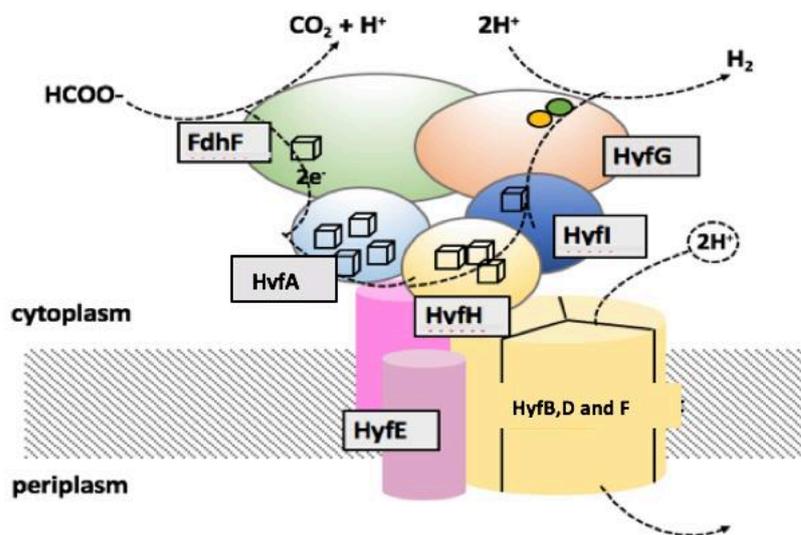
### 1.5.3.1 Comparison between Hyf and Hyc

There are seven products of *hyf* (*hyf**ABCGHIJ*) homologous to seven Hyc subunits; where HyfA is homologous to HycB, HyfB to HycC, HyfC to HycD, HyfG to HycE, HyfH to HycF, HyfI to HycG and HyfJ to HycH (Skibinski *et al.*, 2002). Furthermore, HyfR is related closely to FhlA, in that both are formate-responsive  $\sigma^{54}$  dependent transcriptional

regulators, where they regulate the transcription process of the *hyf* and *hyc* operons, respectively. Moreover, FocB is very similar to FocA, the formate transport system of *E. coli*. According to sequence similarities, FocB is considered to be a second putative formate transporter (Fig 1.9) (Suppmann and Sawers, 1994; Andrews *et al.*, 1997). On the other hand, there are three *hyf* genes that can be found only in the *hyf* operon and not in the *hyc* operon, *hyfD*; *hyfE* and *hyfF* (Andrews *et al.*, 1997), on which two of the subunits (HyfD and HyfF) and HyfB are counted as a part of complex I. Andrews *et al.*, (1997) suggested that these additional subunits have a function associated with the respective subunits of complex I as proton translocators, which are not found in Hyd-3. They concluded that Hyf hydrogenase, together with FdhH, couples the oxidation of formate to proton translocation and it forms a putative 'Fhl-2 complex', that is similar to the 'Fhl-1 complex' formed by Hyd-3 and FdhH. A schematic model of the subunits arrangement of the proposed FdhF-Hyf complex, and known structural features of complex I, are shown in Fig 1.10.



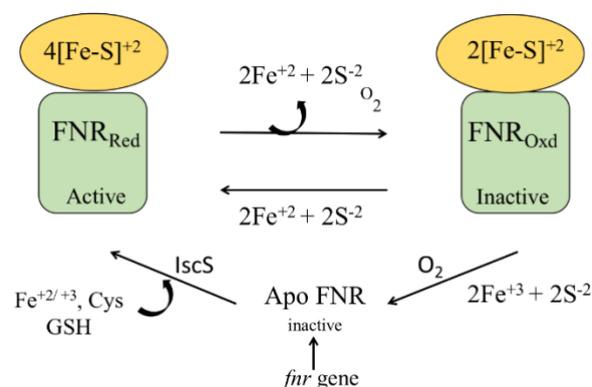
**Figure 1.9. A schematic map of Hyc, Hyf and colcomplex1.** Seven of Hyc components match those of Hyf. Two additional subunits from Hyf resemble complex I components (in red circle). As shown by (Skibinski *et al.*, 2002).



**Figure 1.10. Functional model for the proposed 'Hyf:Fdh-H' complex (FHL-2).** This model was based on the Hyc:FdhH complex, showing proton translocation (absent in Hyc:FdhH). The yellow big cylinder indicates the proton-translocating subunit where proton movement occur and energy is produced. This subunit is not found in Hyc. As listed by (Skibinski *et al.*, 2002).

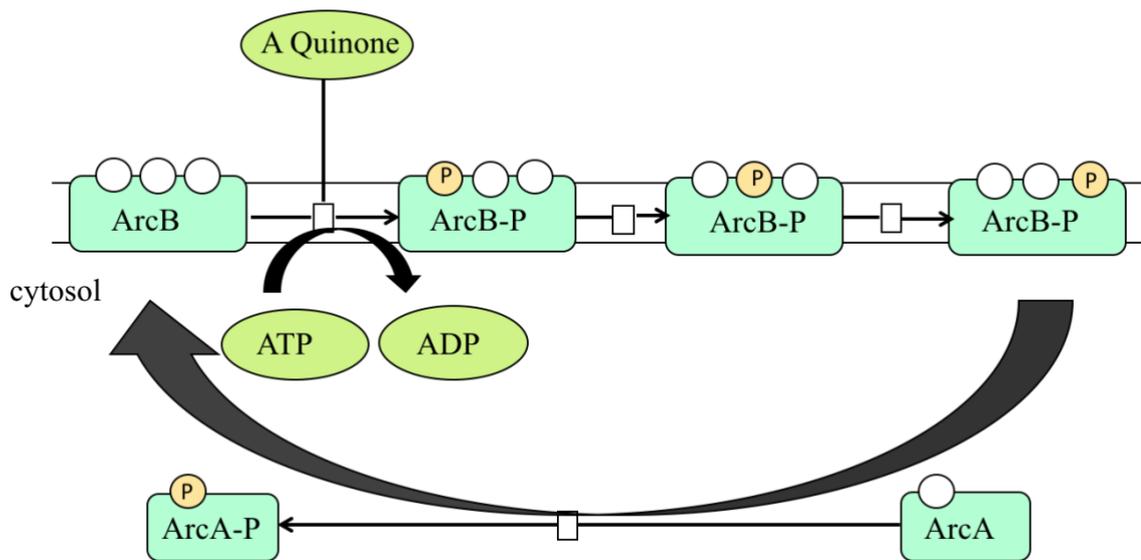
### 1.5.4. Regulation of hydrogenases in *E. coli*

The global regulatory systems of hydrogenase enzymes in *E. coli* are FNR (fumarate and nitrate regulator), which is encoded by the *fnr* (fumarate and nitrate reductase) gene, and the ArcAB (Anoxic Respiratory Control A and B) regulator system. FNR controls the expression of a number of genes involved in strict anaerobic respiratory pathways (Levanon *et al.*, 2005). Under anaerobic conditions, FNR is a dimer with two  $[4\text{Fe-4S}]^{2+}$  clusters which each bind to four cysteine residues. In the presence of  $\text{O}_2$ , FNR is converted to a  $[2\text{Fe-2S}]$  form, which is not active in gene regulation, and if the oxygen exposure is extensive FNR is converted to apoFNR, which is not active (Flint *et al.*, 1996). The interconversion of active and inactive FNR is a reversible reaction. The cysteine desulfurase (IscS) enzyme is one of the most important components for  $[4\text{Fe-4S}]$  formation in apoFNR *in vivo*. It provides sulphur for Fe-S cluster synthesis *in vitro* (Schwartz *et al.*, 2000). The reaction needs cysteine, which provides a bound sulfhydryl group for Fe-S cluster formation. Also, glutathione (GSH) acts as a reductant and can be used in conversion of  $[2\text{Fe-2S}]$  to  $[4\text{Fe-4S}]$  (Schwartz *et al.*, 2000; Uden *et al.*, 2002) (Fig 1.11).



**Figure 1.11. Schematic diagram of FNR regulation and its activation and inactivation by Oxygen.** Red, reduction; Oxd, oxidation. The arrows show the interconversion process of FNR from active to inactive state and the opposite. As shown by (Guest, 1992).

In addition, the Arc (Anoxic respiratory control) system is a two-component response-regulatory system in *E. coli* that controls the transcription of genes under anaerobic conditions. ArcA is a cytoplasmic regulatory protein, while ArcB is a histidine kinase transmembrane protein sensor (Drury and Buxton, 1985; Iuchi *et al.*, 1990). Under anaerobic condition, the quinone pools are in reducing form, ArcB senses the redox state of the cell and undergoes phosphorylation. The phosphate is transferred onto further signalling domains (located at the amino terminal of ArcB) and then on to ArcA, which is a response regulator and it becomes functional (Malpica *et al.*, 2004; Taylor and Zhulin, 1999). The phosphorylated ArcA is then able to act as an activator or repressor of different metabolic reactions (Fig 1.12). For example, when oxygen is low, genes encoding proteins that use oxygen and are active under aerobic conditions are repressed, however, genes that do not require aerobic conditions for their further metabolic processes will be upregulated (Liu and De Wulf, 2004; Salmon *et al.*, 2005). Mutation in *arcA* or *arcB* affects more than 30 operons. Most of them are flavo-protein dehydrogenases, and/or Krebs' cycle, glyoxylate cycle and fatty acid degradation pathways enzymes, which are repressed under anaerobic environment. However, two of them were shown to be activated by ArcA; cytochrome d oxidase and pyruvate-formate lyase. The organization between FNR and ArcAB allows the cell to adapt itself in different oxygen environments (Compan and Touati, 1994).



**Figure 1.12. ArcA and ArcB component signal transduction systems, quinone dependent.** Under anaerobic conditions, ArcB autophosphorylates and transphosphorylates ArcA via four steps of phosphorylation. In an aerobic environment, the oxidised forms of quinone predominates, which causes inhibition of ArcB autophosphorylation, and ArcB can then act as a specific ArcA-P phosphatase. As illustrated by (Malpica *et al.*, 2004).

#### 1.5.4.1. Other hydrogenase regulators

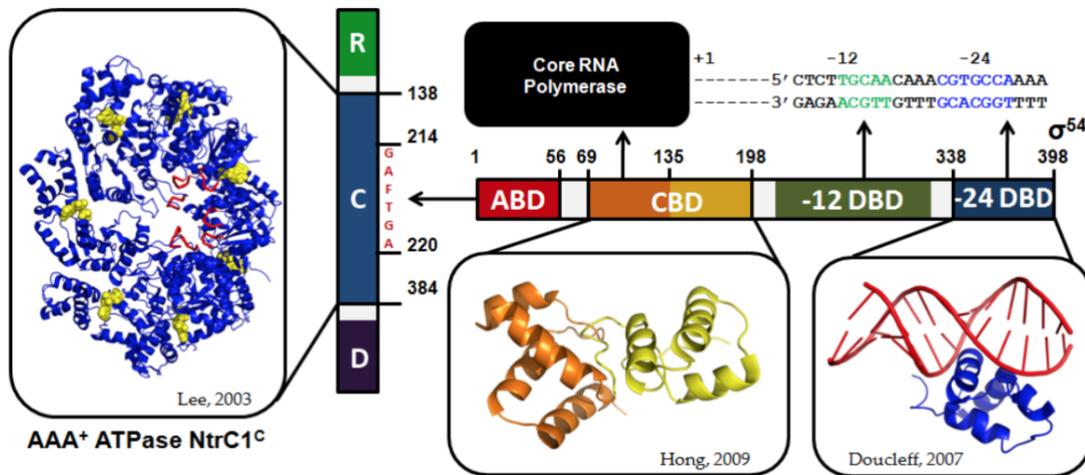
Formate is the primary environmental signal that mediates induction of the formate regulon. Formate-hydrogen lyase A (FhlA, encoded by *fhlA*) is the formate transcriptional activator that directly controls the formate regulon under low pH. FhlA binds to an upstream activating sequence located about 100 bp upstream of the transcriptional start site of the *hyc* operon and in the presence of formate, it activates the transcription process at the  $\sigma^{54}$ -dependent promoter (Sawers, 1994; Self and Shanmugam, 2000). It was originally defined as a regulator of genes involved in nitrogen metabolism and assimilation under nitrogen limiting conditions (Hunt and Magasanik, 1985).

Sigma-54 is a class of sigma factor that requires additional ATP-dependent activation provided by ATPase transcriptional activators (such as FhlA) in-order to melt double

strand DNA and initiate transcription (Wedel and Kustu, 1995). Sigma-54 has four domains (Fig 1.13):

- The N-terminal Activator Bind Domain (ABD), at that interacts with a conserved motif (GAFTGA) located in the middle of the ATPase domain. This interaction leads to a conformational change that allows  $\sigma^{54}$  to switch from promoters from the closed to open complex whereby double stranded DNA is melted around the +1 site.
- The linker region, which links between the N-terminal domain and the core. This second domain is the core binding domain (CBD), which is a seven-helix domain that interacts with the core RNA polymerase to promote a conformational change that is responsible for transferring the activator binding and ATP hydrolysis into larger conformational changes at the C- terminal region of  $\sigma^{54}$  to promote transcription initiation (Hong *et al.*, 2009).
- The third domain is the -12 DNA binding domain (-12DBD). This domain interacts with the -12 DNA element, which is 12 bp upstream of the transcriptional binding site. Thus, once sigma is activated, this domain is needed in melting DNA to initiate transcription (Wong *et al.*, 1994; Barrios *et al.*, 1999).
- The last domain is the -24 DNA binding domain (-24DBD). This C- terminal domain the binds to a conserved -24 DNA element upstream of the transcription start site. Unlike the -12 DBD, it helps to position and anchor the inactive  $\sigma^{54}$  factor at the proper location along the DNA (Douceff *et al.*, 2007).

$\sigma^{54}$  plays a significant role in the transcription of the FhlA/HyfR regulated genes required for Hyd-3 and Hyd-4.



**Figure 1.13. The four domains of sigma factor 54 ( $\sigma^{54}$ ).** ABD: Activator binding domain. CBD: Core binding domain. The -12 DBD is the -12 DNA-binding domain and -24 DBD is -24 DNA-binding domain. As posted by (Doucleff *et al.*, 2007; Hong *et al.*, 2009).

Fhl-1 complex, which is encoded by *fdhF* and *hyc*, is a transcriptional regulator of Fhl (formate hydrogen lyase) complex gene expression; this complex catalysis the breakdown of formate to hydrogen and carbon dioxide under fermentative conditions (Fig 1.8) (Sauter *et al.*, 1992). The FhlA protein has three main domains:

- The C-terminal DNA-binding domain;
- The central region, which is responsible for formate- and DNA-dependent ATP hydrolysis and where the interaction with  $\sigma^{54}$ -takes place and the initiation of transcription occurs; and
- The N-terminal region, where the formate is bound (Self *et al.*, 2001).

From the sequence similarity between FhlA and HyfR (the regulator of the Hyd-4 genes), their central (69% identity) and C-terminal regions (69% identity) show strong similarity, but they are less similar in their N-terminal regions which in FhlA recognise formate.

HyfR is another type of  $\sigma^{54}$ -transcriptional regulator, that regulates the transcription of the *hyf* operon (Skibinski *et al.*, 2002). Over-production of HyfR increases *hyf* expression by up to 1000 fold. This high expression depends on growth conditions, such as anaerobiosis, low pH and post exponential growth phase. This induction was observed with a *hyfA-lacZ* reporter strain when *hyfR* was placed under the control of the *P<sub>lac</sub>* promoter in the presence of  $\sigma^{54}$ . Moreover, this induction was not affected by *fhIA* mutation, which indicates that FhIA does not participate in the activation of *hyf* operon by HyfR (Skibinski *et al.*, 2002). Self *et al.* (2004) found similar results in that HyfR activated a *hyfA-lacZ* fusion. This result confirms that transcription of *hyf* depends on HyfR. However, under aerobic condition, HyfR does not activate the transcription of *hyf-lacZ*.

### 1.6. Some functional studies on Hyf

Andrews *et al.* (1997) suggested that Hyf in combination with formate hydrogenase H (Fdh-H) forms a second formate hydrogenlyase system (Fhl-2) that, unlike Fhl-1, is an energy conserving, proton translocating system. A further study by Skibinski *et al.* (2002), identified the transcriptional start site ( $\sigma^{54}$ ) of the *hyf* operon by reverse transcriptase (RT) mediated primer extension. A cDNA product was observed which initiates 30 bp upstream of *hyfA* start codon. However, no cDNA products were detected with primers to *hyfR* and *focB*, which provides evidence that *hyfR* and *focB* promoters are absent. Furthermore, they used RT-PCR to determine all of the *hyf* cistrons, except for *hyfJ-hyfR*, are cotranscribed. This suggests that *hyfR-focB* may be independently transcribed from the rest of *hyf* operon and *hyfR-focB* are not part of *hyf* operon (despite the lack of a RT mediated primer extension product for *hyfR*). Mnatsakanyan *et al.* (2004) measured hydrogen gas production from Hyc- or Hyf-dependent FHL during

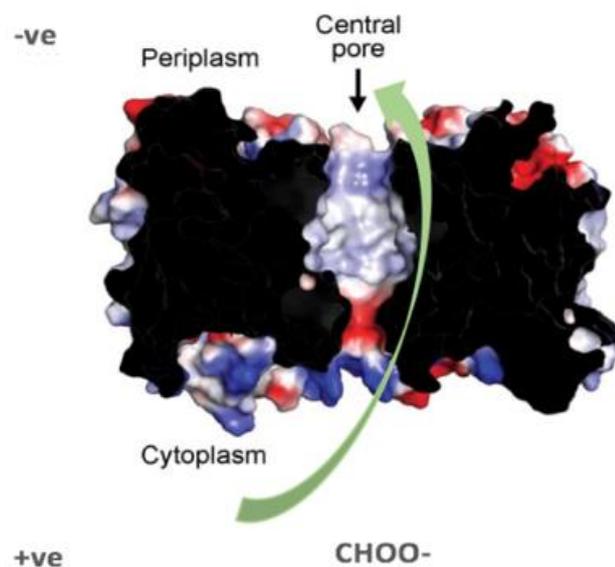
fermentation growth of *E. coli* under alkaline and acidic pH, with/ or without formate. The hydrogen gas was measured using a pair of the oxidation-reduction, platinum and titanium-silicate electrodes. The influence of ATPase was determined using N,N'-dicyclohexylcarbodiimide (DCCD), which inhibit F<sub>0</sub>F<sub>1</sub>-ATPase and other proton translocating membrane proteins. The results indicated that at slightly alkaline pH (pH 7.5), hydrogen production is dependent on Hyd-4 and F<sub>0</sub>F<sub>1</sub>-ATPase, however, addition of external formate increased Hyd-3 activity, which indicated that Hyd-3, but not Hyd-4, depend on FdhH. On the other hand, at acidic pH (pH 5.5), hydrogen production was dependent on Hyd-3, but was unresponsive to DCCD. It should be noted that other workers have been unable to find evidence for any Hyd-4 dependent H<sub>2</sub> production (Skibinski *et al.*, 2002; Self *et al.*, 2004).

Additional studies on *E. coli* K-12 hydrogen production (Trchounian and Trchounian, 2013) indicated that Hyd-4 is able to produce hydrogen even in glycerol fermentation media and the amount of the hydrogen produced was the same as in glucose at pH 6.5. Also, Hyd-2 was reported to produce hydrogen in glycerol medium, while with glucose only hydrogen consumption was observed. Kuniyoshi *et al.* (2015) reported that the expression of proteorhodopsin (PR) play a role in enhancing hydrogen production in *E. coli* when HyfR was overexpressed in the presence of a light source. It was suggested that protons generated by PR activity could only enhance hydrogen production by Hyd-4 with associated proton translocating subunits but not Hyd-3.

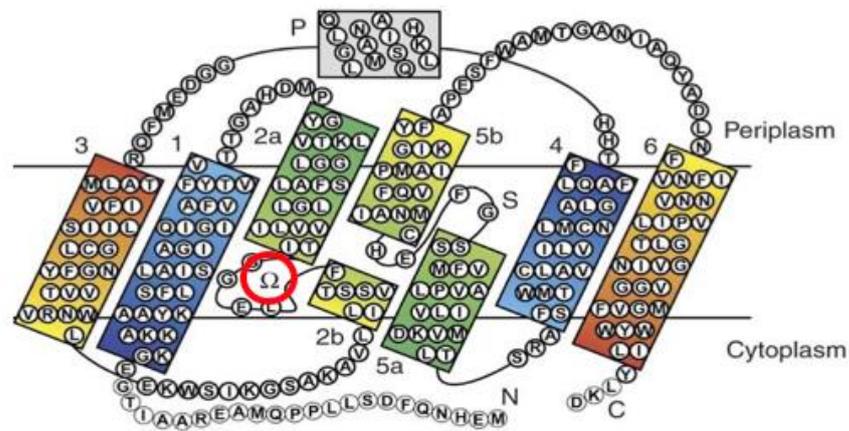
### **1.7. Formate transport and its role in formate hydrogen lyase (Fhl-1)**

Formate is a mixed acid fermentative product of the pyruvate cleavage by pyruvate formate lyase (Pfl) and it is translocated across the cytoplasmic membrane by the pentameric ion channel/transporter 'FocA', which together with the nitrite channel NirC,

is a member of the formate/nitrite membrane transporter (FNT) family (Lü *et al.*, 2011). Formate is transported across the inner membrane by the protein FocA, encoded by *focA*. Expression of the *focA-pfl* operon is induced during anaerobic growth and in the presence of formate. FocA provides a channel with a narrow, positively charged entrance to the pore that is just wide enough for formate to pass through (Fig 1.14). FocA forms a symmetric pentamer, with each protomer consisting of six transmembrane (TM) segments (Wang *et al.*, 2009). TM1, TM3, TM4 and TM6 are each a single  $\alpha$  helix. However, TM2 and TM5 consist of two  $\alpha$  helices connected by a loop  $\Omega$ . The formate ions travel through the pore, and the negatively charged surface on the periplasmic side repels the formate out of the channel into the cytoplasm (Fig 1.15) (Waight *et al.*, 2010).



**Figure 1.14. Structure of FocA.** The FocA pentamer got a central pore. (-ve) indicates negatively charged periplasmic surface area and (+ve) indicates the cytoplasmic positive charge. The green arrow indicates CHOO<sup>-</sup> formate movement through the central pore. As illustrated by (Wang *et al.*, 2009).



**Figure 1.15. Diagram with sequence of each transmembrane segments and their connections.** Numbers indicate the segments. It shows clearly that TM2 and TM5 are broken in the middle of the membrane and are connected by a loop (shown in red circle). This figure as shown by (Hunger *et al.*, 2014).

### 1.7.1. *E. coli* is hypophosphite sensitive during fermentative growth

Hypophosphite is a formate toxic analogue, as it is a strong inhibitor of the pyruvate formate lyase (pfl) enzyme. It inhibits *E. coli* growth completely when it is added (about 75 mM) to media under anaerobic conditions. However, under aerobic conditions, there is no effect of hypophosphite on bacterial growth. In a *focA* mutant strain, *E. coli* is more resistance to hypophosphite compared to the wildtype. This indicates that FocA is responsible for hypophosphite transport (Suppmann and Sawers, 1994). Moreover, under fermentative conditions, FocA transports the fermentation product depending on pH of the cytoplasm. At high pH ( $\geq 6.8$ ), FocA acts as passive facilitator, however, at lower pH ( $< 6.8$ ) it switches to be as an active proton symporter (Lü *et al.*, 2012).

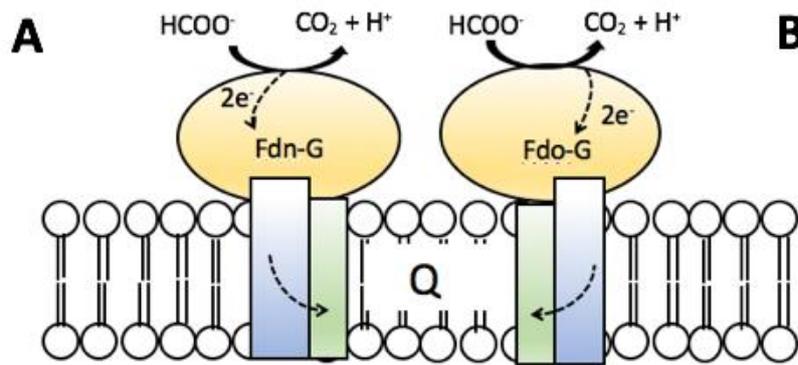
## 1.8. Systems required for formate dehydrogenase and hydrogenase formation

### 1.8.1. Types of formate dehydrogenases

There are three formate dehydrogenases in *E. coli*, formate dehydrogenase-N (Fdh-N); formate dehydrogenase-O (Fdh-O) and formate dehydrogenase-H (Fdh-H). Fdh-N consists of three subunits,  $\alpha$  (FdnG),  $\beta$  (FdnH) and  $\gamma$  (FbnI) with molecular weights of 110 kDa, 32 kDa and 20 kDa, respectively (Enoch and Lester, 1975). The large  $\alpha$  subunit is a seleno-molybdo protein, which forms a catalytic site for formate oxidation (Berg *et al.*, 1991). FdnH binds to four iron-sulphur clusters to mediate electron transfer between  $\alpha$  and  $\gamma$ . The  $\gamma$  subunit transfers electron from  $\beta$  to quinone. The *fdnGHI* operon is located at 32 min on the *E. coli* chromosome. The expression of this operon requires anaerobic conditions with the presence of nitrate for optimal transcription (Berg and Stewart, 1990). The expression is regulated by FNR (section 1.5.4). Two accessory genes are required in synthesis of active Fdh-N, encoded by *fdhD* and *fdhE*. FdhD functions as a sulphur transferase required for sulphur transfer from IscS to the Mo located in the Fdh active centre for its maturation (Thome *et al.*, 2012), while FdhE functions in the processing of Se (Lüke *et al.*, 2008; Fig 1.16A).

Fdh-O differs from Fdh-N, since it catalyse the oxidation of formate using oxygen as a terminal electron acceptor (Sawers *et al.*, 1991; Gennis and Stewart, 1996). It is composed of three subunits:  $\alpha$  (FdoG),  $\beta$  (FdoH) and  $\gamma$  (FdoI). The functions of FdoG and FdoH are similar to the Fdh-N equivalents. However, FdoI acts to mediate the located localisation of the Fdh-O complex to the inner surface of the cytoplasmic membrane (Benoit *et al.*, 1998). Fdh-O is encoded by *fdoGHI* operon at 88 min on the *E. coli* chromosome (Plunkett *et al.*, 1993). This operon is expressed under aerobic conditions. In contrast to Fdh-N, the expression of the Fdh-O operon is not regulated by FNR. The

main regulators are the histone-like H-NS protein and the cyclic AMP receptor protein (CRP), which induce Fdh-O expression (Abaibou *et al.*, 1995) (Fig 1.16B). Note that Fdh-H was described in detail in section 1.5.2.



**Figure 1.16. Schematic illustration of membrane-bound formate dehydrogenase in *E. coli*.** A: Fdh-N and B: Fdh-O. Both enzymes are located at cytoplasmic side. Electron flow within the dehydrogenases is indicated by the dotted thin arrows. Q indicates quinone. As shown by (Hartwig *et al.*, 2015)

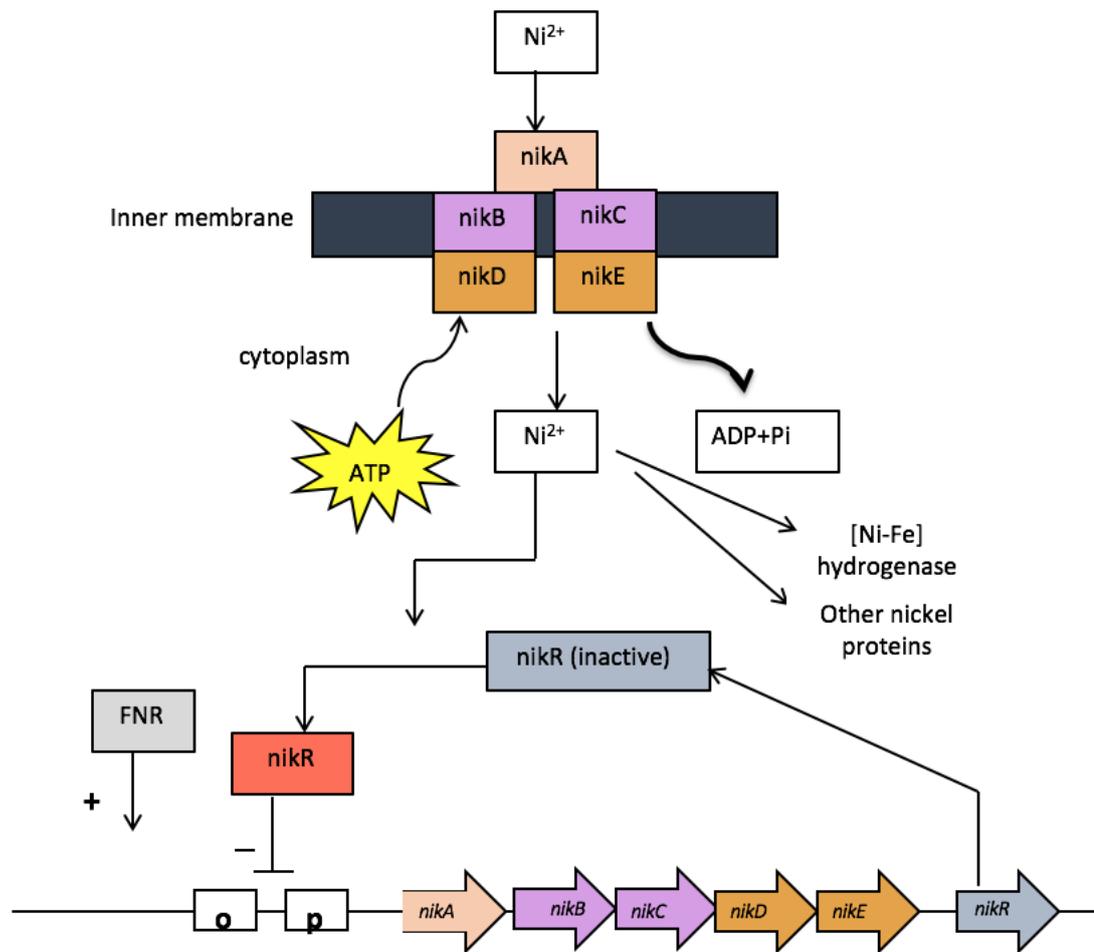
The formation of formate dehydrogenase and hydrogenase enzymes in *E. coli* depends on the availability of a number of metals. These include nickel, which is required in hydrogenase synthesis, and selenium and molybdenum, which are a must in formate dehydrogenase synthesis (Bock and Sawers, 1996). In addition, the hydrogenases requires some accessory genes (e.g. *hyp* genes), which are necessary for the maturation and synthesis of all hydrogenases.

### 1.8.2. Nickel uptake and its incorporation in *E. coli*

Nickel is found in all hydrogenase enzymes in *E. coli*. The synthesis of the [Ni-Fe] cluster depends on nickel supply (Rodrigue *et al.*, 1996; Maier and Triplett, 1996). Nickel (Ni) is taken up into prokaryotic cells by two types of high-affinity transport system. The

first method involves the ATP-Binding Cassette (ABC-type) transporters. The best-characterized ABC-type transporter for  $\text{Ni}^{2+}$  is found in *E. coli* (Eitinger and Mandrand-Berthelot, 2000). The second mechanism makes use of permeases, which is found to be in other bacteria and fission yeast (Komeda *et al.*, 1997).

The multicomponent of ABC-type transport system consists of five proteins, NikABCDE encoded by the *nikABCDE* operon that carries out the ATP-dependent transport of  $\text{Ni}^{2+}$ . NikA is a soluble, periplasmic-binding protein; NikB and NikC form two integral transmembrane forming proteins as a channel for  $\text{Ni}^{2+}$  and NikD and NikE hydrolyse ATP and couple this energy to power Ni-transport (Fig. 1.17) (Mulrooney and Hausinger, 2003). Expression of *nikABCDE* is induced by FNR under anaerobic condition and repressed by NikR. At high nickel concentrations (0.3 mM), as nickel ions are also toxic to cells, synthesis of Nik is completely repressed by the transcriptional regulator NikR, which is encoded by *nikR* located immediately downstream of the *nikABCDE* operon (Eitinger and Mandrand-Berthelot, 2000).



**Figure 1.17. Uptake of nickel across cytoplasmic membrane in *E. coli*, and the regulation of nickel ion uptake by components of *nik* operon.** (o): operator of the operon, (p) promoter of the operon. The (-) sign indicates the repression of *nikABCDE* operon by NikR transcription factor and (+) indicates the expression of *nik* operon and its regulation by FNR. As illustrated by (Navarro *et al.*, 1993).

### 1.8.3. Selenium and molybdenum uptake and their incorporation

Selenium is incorporated into the large  $\alpha$  subunit of the three formate dehydrogenase isoenzymes in *E. coli* in the form of a single selenocysteine residue (Bock and Sawers, 1996). The *selAB* operon, the *selC* and the *selD* genes are required for the biosynthesis and incorporation of the selenocysteine residue into the  $\alpha$  subunit of Fdh (Leinfelder *et*

*al.*, 1988). The seryl residue is converted to selenocysteine on a unique *selC*-encoded tRNA, through the action of selenophosphate synthase (SelD) and selenocysteine synthase (SelA). SelB is a special elongation factor that binds to selenocystyl-tRNA and delivers it to the translating ribosome.

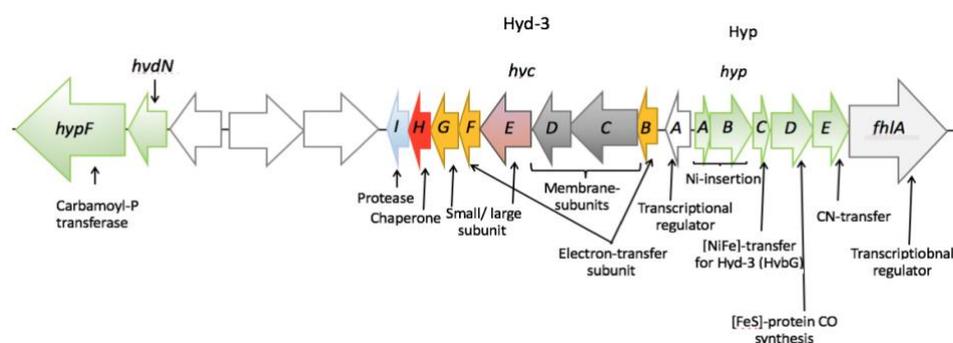
Likewise, molybdenum is an essential component in Fdh-enzymes in *E. coli*. The uptake of molybdenum is as the molybdate ion,  $\text{MoO}_4^{2-}$ , which is carried by the ModABCD high-affinity molybdate-specific transporter. ModABCD is an ABC-transporter, and consists of a molybdate-specific periplasmic-binding protein (ModA), two integral-membrane channel-forming proteins (ModB and ModD) and an ATP-binding protein (ModC) (Maupin-Furlow *et al.*, 1995; Earhart, 1996). ModABCD is encoded by the *modABCD* operon located at 17 min on the *E. coli* chromosome. Expression of the operon is repressed by ModE in response to high concentrations of molybdenum.

#### **1.8.4. Accessory genes (*hyp* genes)**

The Hyp (hydrogenase pleotropic) proteins, encoded by the *hypABCDE* operon and *hypF*, are located upstream of *hyc* operon and downstream of *fhfA* (Fig. 1.18). The *hyp* genes are involved in hydrogenases synthesis through the maturation of the large subunits and [Ni-Fe] metal centre synthesis (Jacobi *et al.*, 1992). HypBDE and HypF are generally involved in maturation of Hyd-1, Hyd-2 and Hyd-3 of *E. coli*, while HypA and HycI together are involved only in the maturation of Hyd-1 and Hyd-3 (Casalot and Rousset, 2001). HypA functions as metallo-chaperone in the insertion of nickel in Hyd-3 and one of the important proteins in the maturation of [Ni-Fe] (Watanabe *et al.*, 2009). On the other hand, HycI is an important processing protease that has a molecular mass of 17 kDa. It removes a 32 amino-acid peptide from the C-terminus extension of pre-HycE to allow the metal insertion and convert the HycE precursor into its mature form. The hycI

cleavage reaction is not inhibited by any conventional inhibitors or Syrine and metalloproteases (Rossmann *et al.*, 1994)

HypB act as nickel donating system, in which GTP hydrolysis involved in releasing HypB from pre-HycE (Maier *et al.*, 1993). HypC is thought to be specific chaperone-type protein which is able to form a complex with pre-HycE during the maturation process, which keeps the protein in a folded state ready for the nickel insertion (Drapal and Bock, 1998). It has another function, which is preventing the small subunit (HycG) from associating with pre-HycE during the maturation process (Magalon and Bock, 2000). HypC and D together forming a complex the iron ligand, making Fe is available to be transferred into the mature large subunit (Blokesch and Bock, 2002). HypE and F are partners in the maturation pathway. HypF interacts with pre-HycE and helps in the synthesis and/or the insertion of cyanide and carbon monoxide ligands in the active site of the large subunit (Jacobi *et al.*, 1992). Mutation in *hypF* stops any hydrogenase activity, and a defect in *hypA* or *hypB* can be overcome by adding high concentration of nickel (> 0.4 mM). However, mutation in all *hyp* genes will prevent the maturation of all hydrogenases (Jacobi and Bock, 1992).



**Figure 1.18. Genetic organisation of *hyc* and *hyp* operon on the *E. coli* chromosome.** Hyd: hydrogenase. *hydN*: is gene product involve in electron transfer from/ to FdhH. White arrows are other accessory proteins which are dividing Hyd-3 from *hypF*. As shown by (Pinske and Sawers, 2014).

### 1.9. Aims and objectives of this study:

The overall aim of the experimental programme is to consider the biochemical purposes of Hyf and FocB in *E. coli* K-12, by answering the specific questions below:

- ❖ Bioinformatics analysis on *hyf* strains
  - What are the species distribution of *hyf* and is there any link between strains with *hyf* and those without?
- ❖ Does *hyf*-encoded FocB function like its homologue, FocA, as a formate transporter?
  - What is the effect of active expression of *focA* and *focB* on hypophosphite (toxic formate analogue) sensitivity.
  - Does transport activity for FocA and FocB react differently to pH?
  - Does formate addition protect the cell from hypophosphite toxicity?
  - Is formate/hypophosphite transport via FocB affected by the growth conditions?
- ❖ Does Hyf, like Hyc, generate H<sub>2</sub>?
  - Does Hyf produce H<sub>2</sub> if the *hyf* operon is expressed from an inducible promoter?
  - Does pH affect the hydrogen production by Hyf?
  - Does formate addition increase the quantity of hydrogen produced by Hyf?
  - Does Hyf show any hydrogenase activity upon activity staining?
  - Does Hyf allow an energy conserving FHL reaction?
- ❖ Is Hyf a second formate hydrogen lyase (Fhl-2)?
  - Is Hyf-dependent hydrogen production dependent on Fdh-H (*fdhF*)?
  - Does the cloned *hyf* operon, together with cloned *fdhF*, (double transformation) enable production of hydrogen in *hycE<sup>-</sup> fdhF* double

mutant?

- Is there any inducer effect for the double transformation with respect to hydrogen production?
- Does formate addition increase hydrogen production in the double mutant strain?
- ❖ Is there any accessory systems/metals effect for Hyf-dependent hydrogen production?
  - Does nickel addition affect Hyf-dependent hydrogen evolution?
  - What are the effects of the metals iron, nickel, molybdenum and selenium on Hyf-dependent hydrogen production in parallel with the bacterial growth?

## Chapter 2: Materials and Methods

### 2.1. Chemicals

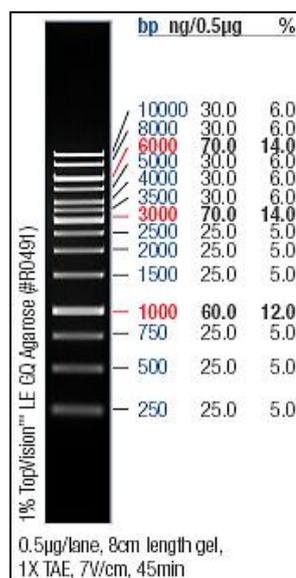
All chemicals were of analytical grade or higher and were purchased from Bio-Rad, Fisher Scientific, Melford or Sigma Aldrich unless otherwise stated.

### 2.2. Enzymes

All restriction enzymes were purchased from Promega, New England Biolabs, or Fermentas (Thermo Fisher Scientific). DNA polymerase and Quick ligase were purchased from Bioline and Fermentas. All the enzymes were used according to the manufacture's instruction with the appropriate buffers.

### 2.3. The size marker

GeneRuler™ 1 kb DNA ladder from Fermentas was used to determine the size and quantity of DNA following gel electrophoresis (Fig 2.1).



**Figure 2.1: DNA marker.** GeneRuler 1kb ladder from Fermentas. DNA marker was used to estimate the size and quantity of DNA. Source: [http://2009.igem.org/wiki/images/3/3f/Generulers\\_1kb\\_marker\\_Fermentas.jpg](http://2009.igem.org/wiki/images/3/3f/Generulers_1kb_marker_Fermentas.jpg).

## 2.4. Antibiotics

Antibiotics were prepared as described (**Table 2.1**) with those dissolved in water being filter sterilised through a sterile 0.22 µm membrane (Millipore) and stored at -20 °C.

**Table 2.1: Antibiotics used in this study.**

Antibiotic	Mode of action	Uses	Working Strength
<b>Ampicillin:</b> (100 mg/mL stock in nano pure H <sub>2</sub> O).	Gram negative bacterial. Inhibits cell wall peptidoglycan synthesis at the transpeptidation step	Selection and maintenance of <i>E. coli</i> strain carrying the β-lactamase gene	100 µg/mL
<b>Chloramphenicol:</b> (50 mg/mL in ethanol).	Bacteriocidal, inhibits 50S ribosomal elongation	Selection and maintenance of <i>E. coli</i> strains carrying the cat gene	35 µg/mL
<b>Kanamycin:</b> (50 mg/mL stock in nano pure H <sub>2</sub> O).	Interacts with 30S subunit of bacterial ribosomes and inhibits translocation during protein synthesis	Selection and maintenance of <i>E. coli</i> strains carrying the Kan <sup>R</sup> gene.	35 µg/mL

## 2.5. Microbiological media

### 2.5.1. Lysogeny broth (LB) and agar

Lysogeny broth (LB), was used for routine bacterial work, contained 10 g/l tryptophan, 5 g/l yeast extract and 5 g/l NaCl. LB was then made up to one litre with qH<sub>2</sub>O (18.2 MΩ, purified through active carbon filter, reverse osmosis, ultrafilter and final passage through

0.22 µm filter using Nanopure™ diamond filtration system from Barnstead laboratories) and autoclaved to ensure sterility before use. Once sterile and cooled, antibiotics or sterile growth additives were added according to experimental requirements. To make LB agar, one litre of L broth was made as above, to this, 15 g of agar were added and well mixed using a magnetic stirrer. Once thoroughly mixed, the medium was dispensed into appropriate aliquots and autoclaved ensuring the lids were slightly loosened. The agar was then cooled to 50 °C before adding any antibiotics or other heat labile additives, the medium was then solidified before use.

### **2.5.2. Werkmann Minimal media (WM)**

This medium was used for hypophosphite inhibition test. WM media contains 50 mM disodium hydrogen phosphate, 100 mM potassium dihydrogen phosphate, 1 mM magnesium sulphate, 0.1 mM calcium chloride, 15 mM ammonium sulphate and 80 mM glucose. This medium was adjusted at pH 6.5, by adding potassium hydroxide. Supplements, when required, were added to the following final concentrations; 5 mM sodium hypophosphite as maximum, 20 mM sodium formate as maximum and 0.02% (w/v) rhamnose. Workmen solid medium contained 1.5% w/v agar. When required, the antibiotic chloramphenicol was added to a final concentration of 50 µg/ml. In case if it is used in measuring hydrogen production, 0.4-0.8% trace elements were used (section 2.5.2.1).

#### **2.5.2.1. Trace elements**

The trace elements (0.2%) were used in WM medium, for hydrogen production testing as well as the growth trend. These consisted of 0.2% ammonium sulphate, 0.1% potassium phosphate, 0.1% sodium dihydrogen phosphate, 0.02% magnesium sulphate

hepta-hydrated, 5  $\mu$ M nickel chloride, 1.6  $\mu$ M ferric citrate, 1  $\mu$ M sodium molybdate and 1  $\mu$ M sodium selenite.

### **2.5.3. Rich media**

This medium was used in measuring the hydrogen production in *hyc* mutant transformants complemented with the inducible *hyf* operon. Rich medium contained 20 g/l peptone, 2 g/l  $K_2HPO_4$ , 5 g/l NaCl and 0.01 M glucose. The medium was adjusted to pH 6.5 by adding potassium hydroxide. Supplements, when required, were added to the following final concentrations; 5 mM sodium hypophosphite as maximum, 20 mM sodium formate as maximum and 0.02% w/v rhamnose. Rich medium contained 1.5% w/v agar. When required, the antibiotic chloramphenicol was added to a final concentration of 50  $\mu$ g/ml or ampicillin to 100  $\mu$ g/ml.

### **2.5.4. TGYEP media**

This medium was used for determining the activity of hydrogenase 4 in *hyc* mutant transformants complemented with the *hyf* operon. This medium contains 10 g/l tryptone, 5 g/l yeast extract, 12 g/l  $K_2HPO_4$ , 3 g/l  $KH_2PO_4$  and 5 g/l glucose. The pH is adjusted to 6.5 by adding potassium hydroxide. Supplements, when required were, added to the following final concentrations of 5 mM as maximum for sodium hypophosphite, 20 mM sodium formate and 0.02% w/v rhamnose. Rich medium contained 1.5% w/v agar. When required, the antibiotic chloramphenicol was added to give 50  $\mu$ g/ml.

## **2.6. Media sterilisation**

Bacterial medium components were prepared as described by Sambrook *et al.* (2001). All media and heat stable solutions were sterilised by autoclaving at 121°C, 20 lb/in<sup>2</sup> for

20 min. Sterilisation of heat labile solutions was achieved by filtration through a sterile 0.22  $\mu\text{m}$  membrane (Whatman). Media were solidified with 1.5% w/v agar before autoclaving. Glassware used in microbiological procedures was sterilised by dry heat (150 °C for 2 to 2.25 h). For all iron-restricted growth, acid-washed glassware was used.

## 2.7. Equipment

A Biofuge bench top Micro Centrifuge (MSE) or Eppendorf centrifuge 5424 were used to centrifuge volumes smaller than 1.5 ml, a Biofuge bench top centrifuge (Centaur) was used for centrifugation of large samples. A benchtop BR401 refrigerated centrifuge (Denley) or refrigerated Eppendorf centrifuge 5804 R were used for centrifugation of large samples at 4 °C.

## 2.8. Strains

Strains used in this study are listed in Table 2.2. Liquid cultures of bacteria were grown at 37 °C and 250 rpm in a Sanyo Gallenkamp orbital shaker in Luria-Bertani broth containing appropriate antibiotics in different concentrations as described in Table 2.1. Bacterial stocks stored in LB liquid medium at -80 °C with 17.5% (v/v) glycerol.

**Table 2.2: Strains used in this study.**

Bacterial Strain	Genotype	Reference/Source
<b>MG1655</b>	Wild type	Lab stock
<b>Top10</b>	<i>E. coli</i> F <sup>-</sup> , <i>mcrA</i> , $\Delta(mrr\text{-}hsdRMS\text{-}mcrBC)$ , $\phi 80lacZ\Delta M15$ , $\Delta lacX74$ , <i>nupG</i> , <i>recA1</i> , <i>araD139</i> , $\Delta(ara\text{-}leu)7697$ , <i>galE15</i> , <i>galK16</i> , <i>rpsL(Str<sup>R</sup>)</i> , <i>endA1</i> , $\lambda^-$	(Sambrook <i>et al.</i> , 2001)
<b>BW25113</b>	F <sup>-</sup> , $\Delta(araD\text{-}araB)567$ , $\Delta lacZ4787(::rrnB\text{-}3)$ , $\lambda^-$	(Datsenko and Wanner, 2000)

	<i>ΔfocB740::kan, rph-I, Δ(rhaD-rhaB)568, hsdR514</i>		
<b>JW0887-1</b>	BW25113 <i>ΔfocA::kan</i>		(Baba <i>et al.</i> , 2006)
<b>JW0887-A-1</b>	BW25113 <i>ΔfocA</i>		This study
<b>JW0887-2</b>	BW25113 <i>ΔfocA::kan</i>		(Baba <i>et al.</i> , 2006)
<b>JW2477-1</b>	BW25113 <i>ΔfocB::kan</i>		(Baba <i>et al.</i> , 2006)
<b>JW2477-A-2</b>	BW25113 <i>ΔfocB</i>		This study
<b>JW2477-2</b>	BW25113 <i>ΔfocB::kan</i>		(Baba <i>et al.</i> , 2006)
<b>BW25113</b>	BW25113 <i>ΔhycE::kan</i>		(Baba <i>et al.</i> , 2006)
<b>BW25113-A-3</b>	BW25113 <i>ΔhycE</i>		This study
<b>BW25113</b>	BW25113 <i>ΔhycA-E::kan</i>		(Baba <i>et al.</i> , 2006)
<b>BW25113-4</b>	BW25113 <i>ΔhycA-E</i>		This study
<b>BW25113</b>	BW25113 <i>ΔhyfG::kan</i>		(Baba <i>et al.</i> , 2006)
<b>BW25113-A-5</b>	BW25113 <i>ΔhyfG</i>		This study
<b>BW25113</b>	BW25113 <i>ΔfdhF::kan</i>		(Baba <i>et al.</i> , 2006)
<b>BW25113-A-6</b>	BW25113 <i>ΔfdhF</i>		This study
<b>BW25113</b>	<i>ΔfdhF::cat Δhyc::kan</i>		This study
<b><i>focA- pfocA</i></b>	<i>ΔfocA E. coli</i> (JW0887) carrying pBAD <sub>rha-focA</sub>	Cm <sup>R</sup>	This study
<b><i>focB- pfocA</i></b>	<i>ΔfocA E. coli</i> (JW0887) carrying pBAD <sub>rha-focB</sub>	Cm <sup>R</sup>	This study
<b>Wt- <i>pfocA</i></b>	Wildtype <i>E. coli</i> (MG1655) carrying pBAD <sub>rha-focA</sub>	Cm <sup>R</sup>	This study
<b><i>focA- pfocB</i></b>	<i>ΔfocA E. coli</i> (JW0887) carrying pBAD <sub>rha-focB</sub>	Cm <sup>R</sup>	This study
<b><i>focB- pfocB</i></b>	<i>ΔfocB E. coli</i> (JW0887) carrying pBAD <sub>rha-focB</sub>	Cm <sup>R</sup>	This study
<b>Wt- <i>pfocB</i></b>	Wildtype <i>E. coli</i> (MG1655) carrying	Cm <sup>R</sup>	This study

pBAD <sub>rha-focB</sub>			
<b><i>hycE- phyc</i></b>	<i>ΔhycE E. coli</i> (BW25113) carrying pBAD <sub>rha-hyc</sub>	Cm <sup>R</sup>	This study
<b><i>hycA-E-phyc</i></b>	<i>ΔhycA-E E. coli</i> (BW25113) carrying pBAD <sub>rha-hyc</sub>	Cm <sup>R</sup>	This study
<b>Wt- <i>phyc</i></b>	Wildtype <i>E. coli</i> (MG1655) carrying pBAD <sub>rha-hyc</sub>	Cm <sup>R</sup>	This study
<b><i>hycE- phyf</i></b>	<i>ΔhycE E. coli</i> (BW25113) carrying pBAD <sub>rha-hyf</sub>	Cm <sup>R</sup>	This study
<b><i>hycE- phyfG</i></b>	<i>ΔhycE E. coli</i> (BW25113) carrying pBAD <sub>rha-hyfG</sub>	Cm <sup>R</sup>	This study
<b><i>hycA-E- phyf</i></b>	<i>ΔhycA-E E. coli</i> (BW25113) carrying pBAD <sub>rha-hyf</sub>	Cm <sup>R</sup>	This study
<b>Wt- <i>phyf</i></b>	Wildtype <i>E. coli</i> (MG1655) carrying pBAD <sub>rha-hyf</sub>	Cm <sup>R</sup>	This study
<b><i>fdhF- pfdhF</i></b>	<i>ΔfdhF E. coli</i> (BW25113) carrying pBAD <sub>ara-fdhF</sub>	Amp <sup>R</sup>	This study
<b><i>hycE- pfdhF</i></b>	<i>ΔhycE E. coli</i> (BW25113) carrying pBAD <sub>ara-fdhF</sub>	Amp <sup>R</sup>	This study

<b>Wt- <i>pdfhF</i></b>	Wildtype <i>E. coli</i> (MG1655) carrying pBAD <sub>ara</sub> - <i>fdhF</i>	Amp <sup>R</sup>	This study
<b><i>hycE+fdhF-phyf</i></b>	( $\Delta fdhF$ )::cat $\Delta hyc$ <i>E. coli</i> (BW25113) carrying pBAD <sub>rha</sub> - <i>hyf</i>	Cm <sup>R</sup>	This study
<b><i>hycE+fdhF-phyc</i></b>	( $\Delta fdhF$ )::cat $\Delta hyc$ <i>E. coli</i> (BW25113) carrying pBAD <sub>rha</sub> - <i>hyc</i>	Cm <sup>R</sup>	This study
<b><i>hycE+fdhF-pdfhF</i></b>	( $\Delta fdhF$ )::cat $\Delta hycE$ <i>E. coli</i> (BW25113) carrying pBAD <sub>ara</sub> - <i>fdhF</i>	Amp <sup>R</sup>	This study
<b><i>hycE+fdhF-phyf+pdfhF</i></b>	( $\Delta fdhF$ )::cat $\Delta hyc$ <i>E. coli</i> (BW25113) carrying pBAD <sub>rha</sub> - <i>hyf</i> and pBAD <sub>ara</sub> - <i>fdhF</i>	Cm <sup>R</sup> and Amp <sup>R</sup>	This study
<b><i>hycE+fdhF-phyc+pdfhF</i></b>	( $\Delta fdhF$ )::cat $\Delta hyc$ <i>E. coli</i> (BW25113) carrying pBAD <sub>rha</sub> - <i>hyc</i> and pBAD <sub>ara</sub> - <i>fdhF</i>	Cm <sup>R</sup> and Amp <sup>R</sup>	This study
<b><i>fdhF-phyf+pdfhF</i></b>	$\Delta fdhF$ <i>E. coli</i> (BW25113) carrying pBAD <sub>rha</sub> - <i>hyf</i> and pBAD <sub>ara</sub> - <i>fdhF</i>	Cm <sup>R</sup> and Amp <sup>R</sup>	This study
<b>Wt- <i>phyf+pdfhF</i></b>	Wildtype <i>E. coli</i> (MG1655) carrying pBAD <sub>rha</sub> - <i>hyf</i> and pBAD <sub>ara</sub> - <i>fdhF</i>	Cm <sup>R</sup> and Amp <sup>R</sup>	This study

## 2.9. Plasmids

All plasmid stocks were maintained at -20 °C in ultra-pure water. Plasmid details are listed in Table 2.3.

**Table 2.3. Plasmids used in this study.**

Plasmid name	Genotype	Antibiotic resistance	Source / Reference
<b>pJET1.2/blunt</b>	Cloning vector (Appendix 1)	Cm <sup>R</sup>	Fermentas
<b>pBAD<sub>rha</sub></b>	Cloning vector with rhamnose inducible promoter (P <sub>rha</sub> ) (Appendix 2)	Cm <sup>R</sup>	Ford <i>et al.</i> , 2014
<b>pBAD<sub>ara</sub></b>	Cloning vector with arabinose inducible promoter (Para)	Amp <sup>R</sup>	Guzman <i>et al.</i> , 1995
<b>pKD3</b>	Derived from pANTS <sub>γ</sub> , containing FRT-flanked <i>cat</i> gene from pSC140	Cm <sup>R</sup>	Datsenko and, Wanner 2000
<b>pKD4</b>	Derived from pINT-ts, containing <i>araC-ParaB</i> and $\gamma$ $\beta$ <i>exo</i> . Also contains the tL3 terminator	Amp <sup>R</sup>	Datsenko and, Wanner 2000
<b>pCP20</b>	Temperature sensitive plasmid (30 °C) encoding a Flp-recombinase	Amp <sup>R</sup> and Cm <sup>R</sup>	Mori <i>et al.</i> , 2000
<b>pBAD<sub>rha</sub>-<i>focA</i></b>	pBAD <sub>rha</sub> plus <i>focA</i> gene from <i>E. coli</i> MG1655	Cm <sup>R</sup>	This study
<b>pBAD<sub>rha</sub>-<i>focB</i></b>	pBAD <sub>rha</sub> plus <i>focB</i> gene from <i>E. coli</i> MG1655	Cm <sup>R</sup>	This study

<b>pBAD<sub>rha</sub>-<i>hyc</i></b>	pBAD <sub>rha</sub> plus <i>hyc</i> operon from <i>E. coli</i> MG1655	Cm <sup>R</sup>	This study
<b>pBAD<sub>rha</sub>-<i>hyf</i></b>	pBAD <sub>rha</sub> plus <i>hyf</i> operon from <i>E. coli</i> MG1655	Cm <sup>R</sup>	This study
<b>pBAD<sub>ara</sub>-<i>fdhF</i></b>	pBAD <sub>ara</sub> plus <i>fdhF</i> gene from <i>E. coli</i> MG1655	Amp <sup>R</sup>	This study
<b>pBAD<sub>rha</sub>-<i>hyfG</i></b>	pBAD <sub>rha</sub> plus <i>hyfG</i> operon from <i>E. coli</i> MG1655	Cm <sup>R</sup>	This study

## 2.10. DNA manipulation and analysis methods

### 2.10.1. Extraction of genomic DNA

Genomic DNA was extracted from the cells of *E. coli* MG1655 using a DNeasy<sup>®</sup> Blood & Tissue kit (Qiagen). One colony was inoculated in 2.5 ml LB and incubated overnight at 37 °C on a shaker (250 rpm). 0.5 ml of overnight cultures were inoculated into 5 ml pre-warmed LB. Cells were harvested in the exponential phase and centrifuged at 7,500 rpm for 10 min. The pellet was re-suspended in 180 µl Animal Tissue Lysis (ATL) buffer. 20 µl of proteinase K were added to the sample and mixed, the sample was incubated at 56 °C in a water bath for 4 h and vortexed every 30 min. 200 µl of Animal Lysis (AL) buffer were added and mixed with the sample by vortex. 200 µl of ethanol (absolute) were added and mixed by vortex. The sample was applied to the DNeasy<sup>®</sup> Mini Column and centrifuged at 8,000 rpm for 1 min, 500 µl of Washing-1 (AW1) buffer were added and the column was centrifuged again at 8,000 rpm for 1 min. Washing-2 (AW2) buffer was then added and the column was centrifuged at 14 rpm for 3 min. 200 µl of Elution (AE) buffer were added and incubated at room temperature for 1 min. Then the column

was centrifuged at 8,000 rpm for 1 min and the elutant was transferred to the freezer at -20 °C for later use.

### **2.10.2. Polymerase chain reaction (PCR)**

A Thermo Scientific Phusion<sup>®</sup> High Fidelity DNA polymerase kit was used for DNA amplification. Each 20 µl PCR reaction contained the following components; 5x Phusion GC buffer or 5x Phusion GC buffer, 200 µM dNTPs, primers at 0.5 µM, 3% DMSO, 1-2 µl DNA (60-100 ng/µl) and up to 20 µl sdH<sub>2</sub>O. Also, CloneAmp HiFi PCR premix/Universe HotStart was used for *hyf* operon amplification. The 20 µl PCR reaction contained 12.5 µl of CloneAmp HiFi PCR premix; 0.2-0.3 µM of each primer as a final concentration; < 100 ng of DNA templet and up to 25 µl sdH<sub>2</sub>O. After pre-heating at 94 °C for 3 min, PCR was performed using 30 thermal PCR cycles (98 °C for 10 secs, 60-64 °C for 15 secs, and 72°C for 1-4 min). This was followed by a final extension step, 72°C for 10 min.

#### **2.10.2.1. Colony PCR**

Colony PCR is a method used for determining the presence or absence of insert DNA in plasmid constructs in multiple colonies. This was achieved by designing primers at the cloning flanking regions of the selected plasmid. Typical reaction volumes used for colony PCR were 25 µl and consisted of 2.5 µl 10x Dream Taq<sup>™</sup> DNA polymerase buffer, 0.5 µl 10 mM dNTP's, 1 µl 10 µM forward primer, 1 µl 10 µM reverse primer, 19.75 µl qH<sub>2</sub>O and 0.25 µl Dream Taq<sup>™</sup> DNA polymerase added on ice. The desired colony was then touched using a sterile fine yellow tip attached to a Gilson pipette set to 20 µl, the tip was gently touched onto an agar plate containing selective antibiotics before being inserted into one of the aliquots containing 25 µl of the reaction mixture, the colony was then mixed with the reaction constituents by gently pipetting up and down taking

care not to introduce too much air to the PCR reaction. Once carefully mixed, the tip was discarded and the procedure repeated using separate aliquots for the desired number of colonies to be screened. Positive controls of colonies known to contain the correct plasmid insert were used to ensure that the PCR reaction was working properly. Once all mixtures were inoculated with a colony, they were placed into an Eppendorf Mastercycler<sup>®</sup> gradient PCR machine to be amplified. The following conditions were used as a standard with the annealing temperature adjusted according to appropriate primer  $T_m$ . The led heated to 105 °C followed by an initial heating step at 95 °C for 5 min to break open bacterial cells and denature the DNA. This was subsequently followed by 35 cycles with 1 min at 95 °C (denaturation), 1.5 min at 45 °C (annealing) and 1 min at 72 °C (extension). Following the 35 cycles there was a final extension step at 72 °C for 5 min to ensure elongation was complete. The PCR reactions were cooled to 4 °C and analysed for the appropriate plasmid insert by agarose gel electrophoresis.

### **2.10.3. Purification of PCR products**

PCR products were purified using the GeneJET<sup>™</sup> PCR Purification Kit (Fermentas), according to manufacturer's instructions. 1:1 (v/v) of binding buffer to PCR product were combined, up to 800 µl of the re-suspension solution was transferred to the GeneGET<sup>™</sup> purification column which was then centrifuged for 45 sec. 700 µl of wash buffer were added and the column was then centrifuged again for 45 sec. The column was centrifuged for 1 min, then 50 µl of elution buffer were added and the column was centrifuged for 1 min. The eluate was then stored at -20 °C. All micro-centrifugation was 12,000 rpm.

#### **2.10.4. Restriction digestion**

DNA digestion of PCR products and vectors was generally performed in 20-30  $\mu$ l volumes containing 60-100 ng/ $\mu$ l DNA, 1X react buffer and appropriate restriction enzymes (1-5 U). The mixture was incubated at 37 °C for 60 min without shaking. The digested DNA was purified using PCR product purification kit (Fermentas).

#### **2.10.5. DNA concentration determination**

Prior to ligation reaction, the concentration of the plasmid DNA was determined using the Nanodrop spectrophotometer. Two  $\mu$ l of plasmid DNA were placed onto the spectrophotometer's pedestal and the absorbance of the sample at 260 nm was used to determine DNA concentration.

#### **2.10.6. Ligation reactions**

The digested and purified PCR products and vector were ligated together in a volume of 10  $\mu$ l containing 70-100 ng of the digested PCR product and 100-200 ng of the digested plasmid, 1X T4 ligase buffer (Fermentas) and 3 U of T4 ligase (Fermentas<sup>®</sup> or New England BioLabs<sup>®inc</sup>). Sticky end ligations were incubated at room temperature for 10 min whereas blunt end ligations were incubated for 1 h. In both cases, 3-5  $\mu$ l was used for transformation into 100  $\mu$ l of chemically competent cells.

#### **2.10.7. Plasmid 'minipreps' isolation**

One millilitre of an overnight cell suspension was centrifuged at 13,000 rpm for 8 min and 100  $\mu$ l of buffer 1 (50 mM Tris-HCl, 10 mM EDTA, 100  $\mu$ g/ml RNase A, pH 8.0) were added to each cell pellet and mixed by vortexing. 200  $\mu$ l of buffer 2 (1% SDS, 0.2 M NaOH) were added to each tube and mixed by inversion. Tubes were stood on ice for

5 min, 150  $\mu$ l of ice-cold buffer 3 (3.0 M potassium acetate, pH 5.5.) were added to each tube and mixed as before and stood on ice for another 5 min. Samples were centrifuged at 13,000 rpm for 5 min, the supernatant was transferred to a new tube. 400  $\mu$ l of isopropanol were then added to each tube which was mixed vigorously and left to stand in room temperature for 2 min. The samples were centrifuged as before, the supernatant was discarded, 200  $\mu$ l of absolute ethanol were added to each tube and mixed by inversion. Samples were centrifuged as before; the supernatant was discarded and the tubes with open lids were left at 37 °C. 20  $\mu$ l of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) were added to each tube followed by mixing. Samples were finally stored at -20 °C (Birnboim and Doly, 1979).

Plasmids, were also isolated by using Thermo Scientific GeneJET™ Plasmid Minipreps (Fermentas). One transformant colony was inoculated in 2.5 ml LB containing a selective antibiotic and incubated overnight at 37 °C on a shaker (250 rpm). The entire overnight culture was centrifuged at 13,000 rpm for 5 min at room temperature, and the supernatant was discarded. The pellet was resuspended in 250  $\mu$ l of Resuspension Solution (containing RNase A) and vortexed until no cell clumps remained. The bacterial suspension was centrifuged as before, 250  $\mu$ l of Lysis Solution was added to the pellet and mixed thoroughly by inverting the tube 4-6 times until the solution became viscous and slightly clear. Neutralization Solution (350  $\mu$ l) was added to the sample and mixed immediately and thoroughly, as described before. The sample was then centrifuged as before for 5 min to pellet the cell debris and chromosomal DNA. The supernatant was carefully transferred to a GeneJET spin column and centrifuged for 1 min, the flow-through was discarded and column was returned to the collection tube. Plasmid DNA in the spin column was washed twice with 500  $\mu$ l of Wash Solution (diluted with ethanol) and then centrifuged as above for an additional 30-60 sec to remove residual Wash

Solution. The final step in plasmid isolation was addition of 50  $\mu$ l of the Elution Buffer to the centre of the GeneJET spin column membrane to elute the plasmid DNA, the spin tube was then incubated at the room temperature for 2 min prior to centrifugation as before. The purified plasmid DNA was then stored at -20 °C for future work.

## **2.10.8. Gel electrophoresis**

### **2.10.8.1. Agarose gel electrophoresis**

For analysing plasmid DNA and PCR products, agarose gel electrophoresis was performed. Gels were generally 0.7% (w/v) agarose in 0.5X TBE buffer (0.4 M Tris, 0.4 M borate, 1 mM EDTA, pH 8.0). Gel staining was performed with Biotium Nucleic Acid Stain GelRed™ (dilution of GelRed™ 10,000X stock reagent into the agarose gel solution i.e. 1  $\mu$ l of the GelRed™ 10,000X stock reagent added to 50 ml of the gel solution). DNA samples (2  $\mu$ l) were loaded with 2  $\mu$ l of DNA loading dye (10 mM Tris-HCl pH 7.6, 0.03% bromophenol blue, 0.03 xylene cyanol FF, 60% glycerol 60 mM EDTA) and 6  $\mu$ l sterile ultrapure water (suH<sub>2</sub>O). GeneRuler™ 1 kb DNA ladder (Fermentas) was used as the DNA size marker (1  $\mu$ l GeneRuler™, 2  $\mu$ l of DNA loading dye, 7  $\mu$ l suH<sub>2</sub>O). Samples were electrophoresed (50-60 volts) for around 90 min in a BioRad horizontal gel tank containing 0.5X TBE buffer. Agarose gels were then visualised under UV illumination provided by a short-wave length UVP GelDoc-it ultraviolet transilluminator Bio Imaging system and a G: Box Chemi-XL1.4 GENESys (Synoptics Ltd).

### **2.10.8.2. Native Polyacrylamide gel electrophoresis**

The hydrogenase activity band was determined using native 7.5% polyacrylamide gel with 0.1% (w/v) Triton X-100. The gel contained resolving and stacking gels. The

resolving gel contained 1.3 ml 62.5 mM Tris-HCl (pH 7.5), 11.3 ml 30 % w/v acrylamide (Bio-Rad), 0.05 ml 10% w/v ammonium persulphate, 0.003 ml TEMED and 2.35 ml distilled water. The gel was cast and, once set, the stacking gel applied to the top.

The stacking gel was made up of 0.13 ml 34 mM Tris-HCl (pH 5.5), 0.17 ml 30% w/v acrylamide, 0.01 ml 10% w/v ammonium persulphate, 0.001 ml TEMED, 0.69 ml distilled H<sub>2</sub>O. The running buffer (Tris/glycine system) was made up of 0.1 M Tris; 0.1 M glycine and 0.1% (v/v) Triton X-100, pH 8.5. Another gel system was used at neutral pH (barbitone gel system) which was made of 82.5 mM Tris; 26.8 mM barbital and 0.1 % v/v Triton X-100, pH 7.0. The loading dye contained 15 g of ficoll-400, 250 mg of xylene cyanol, 250 mg of bromophenol blue and water up to 100 ml.

## 2.11. Cloning into plasmids

### 2.11.1. Primers designing for gene amplification using PCR

Oligonucleotides primers were designed using Vector NTI *Express* and ordered from Eurofin (formally MWG). All primers used in this study are listed in Tables 2.4-2.9.

**Table 2.4. Primers designed for PCR amplification for *focA* and *focB* cloning into pBAD<sub>rha</sub> and their sequence primers used in this study. The underline sequences are the restriction enzyme sequence.**

Primer name	Forward/ Reverse	5' -3'	Restriction enzyme	Primer Length (bp)	TM (°C)	GC (%)
pBAD <sub>rha</sub> - <i>focA</i>	TCAGCAGGATCACATATG AAAGCTGACAACCCTTT		<i>NdeI</i>	35	63.3	43
pBAD <sub>rha</sub> - <i>focA</i>	GACTCTAGAGGATCCTTA ATGGTGGTCGTTTTCAC		<i>BamHI</i>	35	64.4	46
pBAD <sub>rha</sub> - <i>focB</i>	TCAGCAGGATCACATATG AGAAACAAACTCTCTTTC G		<i>NdeI</i>	37	63.3	41
pBAD <sub>rha</sub> - <i>focB</i>	GACTCTAGAGGATCCTCA GGGTTCTGACGTAAATA AA		<i>BamHI</i>	38	65.5	45

<b>T7F</b>	TAATACGACTCACTATAG GG	-	20	47.7	40
<b>pJET1.2 reverse</b>	<u>AAGAACATCGATTTTCCA</u> <u>TGGCAG</u>	-	24	54	42
<b>pBAD<sub>rha</sub> aF</b>	ATGCCATAGCATT TTTATC C	-	20	45.6	35
<b>pBAD<sub>rha</sub> aR</b>	<u>GATTTAATCTGTATCAGG</u>	-	18	41.2	33

**Table 2.5. Primers designed for PCR amplification of *hyf* operon of *E. coli* MG1655 for In-Fusion cloning and its sequence after the cloning into pBAD<sub>rha</sub>. The underline sequences are the restriction enzyme sequence.**

Primer name	Forward/ Reverse 5' -3'	Restriction enzyme	Primer length (bp)	T <sub>m</sub> (°C)	GC (%)
pBAD <sub>rha</sub> - <i>hyfA</i>	TCAGCAGGATCACATATGAACCG CTTTGTGGTGGC	<i>NdeI</i>	35	66.8	54
pBAD <sub>rha</sub> - <i>focB</i>	GACTCTAGAGGATCCTCAGGGTT CCTGACGTAATAAAA	<i>BamHI</i>	38	65.5	45
pBAD <sub>rha</sub> -F1	CTTCCCTGGTTGCCAATG	-	19	51.1	53
pBAD <sub>rha</sub> - <i>hyf</i> F2	CCTTTTGAATTTGCCGAT	-	18	46.8	42
pBAD <sub>rha</sub> - <i>hyf</i> F3	CGGCGCACTGTTTCATCTG	-	19	53.2	58
pBAD <sub>rha</sub> - <i>hyf</i> F4	GGCTAAATAAAGGTCTTGC	-	19	46.8	42
pBAD <sub>rha</sub> - <i>hyf</i> F5	TGTTGCTGAAGCAGAACA	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyf</i> F6	TATTACCAGAGCGATAAAG	-	19	44.6	37
pBAD <sub>rha</sub> - <i>hyf</i> F7	TTCGCCTGGTTTATTCGC	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf</i> F8	CTGCTTTTTTCGCTGCTC	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf</i> F9	GTTGTTGCTCATCTTCGG	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf</i> F10	GGTCACCATTACGGTGAA	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf</i> F11	GGAAAAGTCGATGACGAT	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyf</i> F12	CTGGCTTTCGCGGAAAAC	-	18	50.3	56
pBAD <sub>rha</sub> - <i>hyf</i> F13	GCGGTATTTCCACGATC	-	18	48	50

pBAD <sub>rha</sub> - <i>hyf</i> F14	CGTGAGAAGGATTTCTCA	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyf</i> F15	GAATCTGCGCTCCTTATA	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyf</i> F16	AATGGCGCGCCATATGAA	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf</i> F17	TAATCTGGTTTTCCGGGC	-	18	48	50
pBAD <sub>rha</sub> - R1	AAGCTTGCATGCCTGCAGGT	-	20	53.8	55
pBAD <sub>rha</sub> - <i>hyf</i> R2	AGCAGATTACAACACCTC	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyf</i> R3	TTTGACCAGCGGCTTGTCG	-	19	53.2	58
pBAD <sub>rha</sub> - <i>hyf</i> R4	CCAGTTAAACAGTGTGCA	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyf</i> R5	CACTCTTCAGTCATGTAAC	-	19	56.8	42
pBAD <sub>rha</sub> - <i>hyf</i> R6	GGTTTTGTGGCGCATT	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyf</i> R7	TAAAGCCTTCGGTCAGCA	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf</i> R8	CTCATTGATAAACGGATACG	-	20	47.7	40
pBAD <sub>rha</sub> - <i>hyf</i> R9	CAGCAGGACGATAATCAG	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf</i> R10	CAGAAACGCCGAGCTTAA	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf</i> R11	TATTTTCGTGCAGCGTAAG	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyf</i> R12	GCTGTGGCAAATACATCAG	-	19	48.9	47
pBAD <sub>rha</sub> - <i>hyf</i> R13	AAACGAAACGACCAGTGC	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf</i> R14	ATCAGACCTGAAGATGTC	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyf</i> R15	GTATTTGGCACTTCCTGT	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyf</i> R16	TGAAAGCGTGCGGAAACT	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf</i> R17	GCAGCAAAGCATGTTGTG	-	18	48	50

Table 2.6. Primers designed for PCR amplification of *hyc* operon of *E. coli* MG1655 for In-Fusion cloning and its sequence after the cloning into pBAD<sub>rha</sub>.

Primer name	Forward/ Reverse 5' -3'	Restriction enzyme	Primer length (bp)	Tm (°C)	GC (%)
-------------	----------------------------	--------------------	--------------------	---------	--------

pBAD <sub>rha</sub> - <i>hycA</i>	TCAGCAGGATCACATA TGACTATTTGGGAAAT AAGCGAGA	<i>NdeI</i>	40	53	36
pBAD <sub>rha</sub> - <i>hycI</i>	GGGTACCATGGCATA TACTCTTCTTCCACCGC TAACTGC	-	40	59	52
pBAD <sub>rha</sub> - <i>hyc-s F1</i>	GCAGGATCACAATGAC TA	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyc-s F2</i>	ATGGGGCCGTGCAGTT GA	-	18	52.6	61
pBAD <sub>rha</sub> - <i>hyc-s F3</i>	TCATGGCCCTGTGCGCG G	-	18	57.2	72
pBAD <sub>rha</sub> - <i>hyc-s F4</i>	AATAGTGGCGCGTTTGT T	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyc-s F5</i>	TCCGGCTGGGTGTTCCG C	-	18	57.2	72
pBAD <sub>rha</sub> - <i>hyc-s F6</i>	GCGCTGAATGAGGCAT TT	-	18	48	50
pBAD <sub>rha</sub> - <i>hyc-s F7</i>	ACGCTTGCACTCGCATC T	-	18	50.3	56
pBAD <sub>rha</sub> - <i>hyc-s F8</i>	AGAAGAGCAAAGTGGT GC	-	18	48	50
pBAD <sub>rha</sub> - <i>hyc-s F9</i>	TTTCGCCGCTGTTTGAT G	-	18	48	50
pBAD <sub>rha</sub> - <i>hyc-s F10</i>	GCAAAATGGAAGGCGT GC	-	18	50.3	56
pBAD <sub>rha</sub> - <i>hyc-s R1</i>	CAGGGTTATTGTCTCAT G	-	18	45.8	44
pBAD <sub>rha</sub> - <i>hyc-s R2</i>	GGTTTAGCCCCATATCC G	-	18	50.3	56
pBAD <sub>rha</sub> - <i>hyc-s R3</i>	CGCAGCGGCTGCACCA TA	-	18	54.9	67
pBAD <sub>rha</sub> - <i>hyc-s R4</i>	AAGTCTTCTTTCTTCCA C	-	18	43.5	39

pBAD <sub>rha</sub> - <i>hyc-s</i> R5	ATAGCCGACAAACGGG TG	-	18	50.3	56
pBAD <sub>rha</sub> - <i>hyc-s</i> R6	AAAGTTCATCCGGCCA GT	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf-s</i> R7	GTTTGCCCATTTTCGATA A	-	18	43.5	39
pBAD <sub>rha</sub> - <i>hyf-s</i> R8	CGTGGTCGTAACCGCA CA	-	18	52.6	61
pBAD <sub>rha</sub> - <i>hyf-sf</i> R9	ATCATGCCGAGCACCA GC	-	18	52.6	61
pBAD <sub>rha</sub> - <i>hyf-s</i> R10	GCTCATTTAGCCTCTCC A	-	18	48	50
pBAD <sub>rha</sub> - <i>hyf-s</i> R11	TAGCGGTGTATAAGCT GT	-	18	45.8	44

Table 2.7. Primers designed for PCR amplification of *fdhF* of *E. coli* MG1655 for In-Fusion cloning into pBAD<sub>ara</sub>.

Primer name	Forward/ Reverse 5' -3'	Restriction enzyme	Primer length (bp)	T <sub>m</sub> (°C)	GC (%)
pBAD <sub>ara</sub> - <i>fdhF</i> -F	GAGGAATTAACCATGAAAA AGTCGTCACGGTTTGCC	<i>Nco</i> I	37	64.5	43
pBAD <sub>ara</sub> - <i>fdhF</i> -R	AAAACAGCCAAGCTTTTACG CCAGTGCCGCTTCGCG	<i>Hind</i> III	36	64.4	44

Table 2.8. Primers designed for PCR amplification of *fdhF* and *hycE* in *E. coli* MG1655 for post transduction and *kan* cassette removal confirmation.

Primer name	Forward/ Reverse 5' -3'	Primer length (bp)	T <sub>m</sub> (°C)	GC (%)
<i>fdhF</i> -F	CATTACTGATGCTGGACAGCC	21	54.4	52
<i>fdhF</i> -R	GCGGCTTTCTGTGCCAATA	20	53.8	55
<i>hycE</i> -F	GCGTCTTGATATTACTCCGC	20	51.8	50

<i>hycE</i> -R	TTATCAACCGCAATCGGCTC	20	51.8	50
----------------	----------------------	----	------	----

**Table 2.9. Primers designed for PCR amplification of pKD3 as chloramphenicol template for *fdhF* knockout in the  $\Delta hycE$  strain.**

Primer name	Forward/ Reverse 5' -3'	Primer length (bp)	T <sub>m</sub> (°C)	GC (%)
<i>fdhF</i> -fwd	CGTATGCGTGATTTGATTA <del>ACTGGAGCG</del> AGACCGATGAAAAAAGTCTGTGTAGGCT GGAGCTGCTTC	67	74.4	48
<i>fdhF</i> -rev	<del>CAGCCTCCGAAAGGAGGCTGTAGAAAG</del> GACGGTATTACGCCAGTCATATGAATAT CCTCCTTAGT	65	74.1	48

**Table 2.10. primers designed for PCR amplification of *hyfG* in *E. coli* MG1655 for cloning into pBAD<sub>rha</sub> in this study. The underline sequences are the restriction enzyme sequence.**

Primer name	Forward/ Reverse 5' -3'	Restriction enzyme	Primer length (bp)	T <sub>m</sub> (°C)	GC (%)
pBAD <sub>rha</sub> - <i>hyfG</i> -F	TCAGCAGGATC <u>CACAT</u> GTGAA CGTTAATTCATCGTC	<i>Nde</i> I	35	63.3	43
pBAD <sub>rha</sub> - <i>hyfG</i> -R	<del>GACTCTAGAGGATCCTTATTC</del> AGCGGCGAACGGTT	<i>Bam</i> HI	35	65.6	49

### 2.11.2. Preparation of plasmid DNA

A single *E. coli* colony carrying the desired plasmid was used to inoculate 2.5 ml LB containing the appropriate antibiotic. This was incubated overnight at 37 °C with shaking at 250 rpm. The cells were then harvested and plasmid DNA was isolated using a GeneJET™ plasmid Miniprep kit. All plasmids were eluted in ultra-pure water and stored at -20 °C.

## **2.12. Preparation of competent cells for transformation**

Chemically competent cells were created by treatment with calcium chloride. 0.5 ml of an overnight culture of *E. coli* (different strains) were inoculated into 250 ml conical flask containing pre-warmed 50 ml LB, and grown up to an OD<sub>600</sub> of 0.5 (Lederberg & Cohen, 1974) at 37 °C in a Gallenkamp orbital shaker (250 rpm). Cells were harvested in pre-cooled 50 ml Falcon tubes by centrifugation at 4 °C and 5,000 rpm for 5 min in a Biofuge Stratos bench-top centrifuge. Following the centrifugation, the supernatant was decanted, and the cells were re-suspended in 30 ml of ice-cooled 0.1 M of MgCl<sub>2</sub> and incubated on ice for 10 min. Cells were centrifuged as before, the pellet was re-suspended again in 30 ml ice cold 0.1 M of CaCl<sub>2</sub> and incubated on ice for 30 min. Cells were centrifuged again as before and pellets re-suspended in 8 ml ice cold 0.1 M of CaCl<sub>2</sub> plus 20% glycerol. Cells were then divided into 200 µl aliquots and stored at -80 °C.

### **2.12.1. Transformation into chemically competent cells and plating of the cells**

#### **2.12.1.1. Heat-shock transformation**

*E. coli* chemically competent cells (100 µl) were inoculated with 2 µl plasmid DNA or ligation reaction (Hanahan, 1983; Lederberg & Cohen, 1974; Sambrook, 1989). The cells were incubated on ice for 30 min, heat shocked at 42 °C for 2.5 min, and then immediately returned to ice for a further 5 min. Next, 1 ml of pre-warmed L-broth was added to the cells, and they were incubated at 37 °C for a further 45 min. following incubation, the tubes were micro-centrifuged at 13,000 rpm for 6 min and 900 µl of supernatant were removed. The pellet was re-suspended in the remaining liquid, and then 30 µl aliquots were plated on to separate LB-agar plates containing the selective antibiotic. Plates were allowed to dry were then incubated (inverted) for 18 h at 37 °C.

### **2.12.1.2. Electroporation transformation**

Electroporation was carried out as described previously (Yu *et al.*, 2000). Briefly, pre-cooled *E. coli* competent cells (90  $\mu$ l) were incubated with 5-10  $\mu$ l plasmid containing solution. The mixture was incubated on ice for 30 min and then transferred into a pre-cooled electroporation cuvette (0.1 cm). Electroporation was performed by using a Bio-Rad Gene Pulser set at 1.8 kV, 25  $\mu$ F with a pulse control of 200 ohms. The electroporated cells were immediately diluted with 900  $\mu$ l of Lb. Around 985  $\mu$ l was decanted into Eppendorf tubes and incubated with shaking at 37 °C in a Gallenkamp orbital shaker (250 rpm) for 2 h. The cells were then centrifuged, 400  $\mu$ l of the supernatant was removed and the cells were re-suspended in the remainder. They were then spread on LB agar plates containing the appropriate antibiotic and incubated at 37°C for 18-24 h.

### **2.12.2. Gene cloning into vectors**

DNA digestion of PCR products and vectors was generally performed as described above (2.10.4). The digested PCR products were purified according to manufacturer's instruction using a PCR purification kit (Fermentas) or reaction GeneJET™ cleanup Kit (Fermentas) (section 2.10.3). Ligation was performed at 22 °C with T4 DNA ligase (Fermentas) as described (section, 2.10.6). The ligation reaction was then used to transform chemically competent *E. coli*. The identity of plasmids within transformants obtained was confirmed by plasmid DNA isolation, restriction mapping and DNA sequencing.

#### **2.12.2.1. In-Fusion® HD cloning**

An In-Fusion® HD Cloning Kit from Clontech Laboratories, Inc. was used for Gibson assembly cloning. The In-Fusion cloning kit fuses DNA fragments (e.g. PCR-generated

sequences and linearized vectors), efficiently and precisely by recognizing a 15 bp overlap at the ends of the fragments. This 15 bp overlap can be engineered by designing primers for amplification of the desired sequences. The kit includes exonuclease for generating sticky ends, a DNA polymerase to fill gaps in annealed fragments and a DNA ligase to seal nicks.

### **2.12.2.2 Primers designing for In-Fusion® HD cloning**

The 5' end of every In-Fusion primer was designed to possess 15 nucleotides that are homologous to 15 of those at one end of the DNA fragment to which the amplification product was to be joined. The 3' end of the primer was designed to be between 18-25 nucleotides in length, have a GC-content between 40-60%, a melting temperature between 58-65 °C and contain a sequence that is specific to the other end of the fragment.

### **2.12.2.3 In-Fusion cloning protocol**

Approximately, 3 µl of around 10-200 ng of gel purified PCR product (e.g. insert) were mixed with 1 µl of around 50-200 ng of the linearized vector, 2 µl of 5X In-Fusion HD Enzyme Premix, and the total reaction volume was adjusted to 10 µl using deionised H<sub>2</sub>O, then the reaction components were mixed. Subsequently, the reaction mixture was incubated for 15 min at 50 °C followed by holding on ice and immediately for transformation.

### **2.12.2.4 In-Fusion transformation procedure using Stellar™ competent cells**

Stellar™ competent cells (Clontech) were used to transform the cells with 2.5 µl of the In-Fusion reaction mixture. A ratio of 1/100<sup>th</sup>-1/5<sup>th</sup> of each transformation reaction was

placed into separate tubes and the volume was adjusted to 100  $\mu$ l with a nutrient-rich SOC medium. Each diluted transformation reaction was spread on a separate LB plate containing a desired concentration of an appropriate antibiotic for the cloning vector. Subsequently, each transformation reaction was centrifuged at 6,000 rpm for 5 min. The supernatant was discarded and each pellet was resuspended in 100  $\mu$ l fresh SOC medium. Each sample was spread on a separate LB plate containing the appropriate antibiotic and all plates were then incubated overnight at 37 °C. Next day, individual isolated colonies were picked up from each transformation plate. Later, plasmid DNA was isolated using GeneJET™ Plasmid Miniprep Kit (Fermentas). To determine the presence of insert, the DNA was analysed by restriction digestion.

### **2.13. Elimination of the Kanamycin resistance cassette**

The kanamycin resistance cassette was removed as part of the strain construction process, and the Flippase recognition target (*flr*) sites were used in order to do so. The method used to delete the antibiotic resistance gene is described by Cherepanov and Wackernagel (1995). Strains from which kanamycin resistance genes needed to be removed were transformed with pCP20 plasmid (Table 2.3). This is an ampicillin and chloramphenicol resistant plasmid that has temperature sensitive replication and thermal induction of FLP synthesis (Cherepanov and Wackernagel, 1995). The transformed cells were plated onto LB solid medium containing ampicillin and incubated overnight at 30 °C. A few colonies were selected, plated on LB agar and incubated overnight at 44 °C in order to delete the kanamycin resistance cassette from the bacterial chromosome. Single colonies were picked and streaked onto LB agar, LB agar plus ampicillin and LB agar plus kanamycin and grown overnight at 30 °C. The mutants that grew only on LB agar (without any additional antibiotics) were those that had the kanamycin cassette removed and had also

lost the plasmid. The deletion of the kanamycin resistance genes was confirmed by colony PCR (section 2.10.2.1).

## **2.14. Gene inactivation procedure “The knockout”**

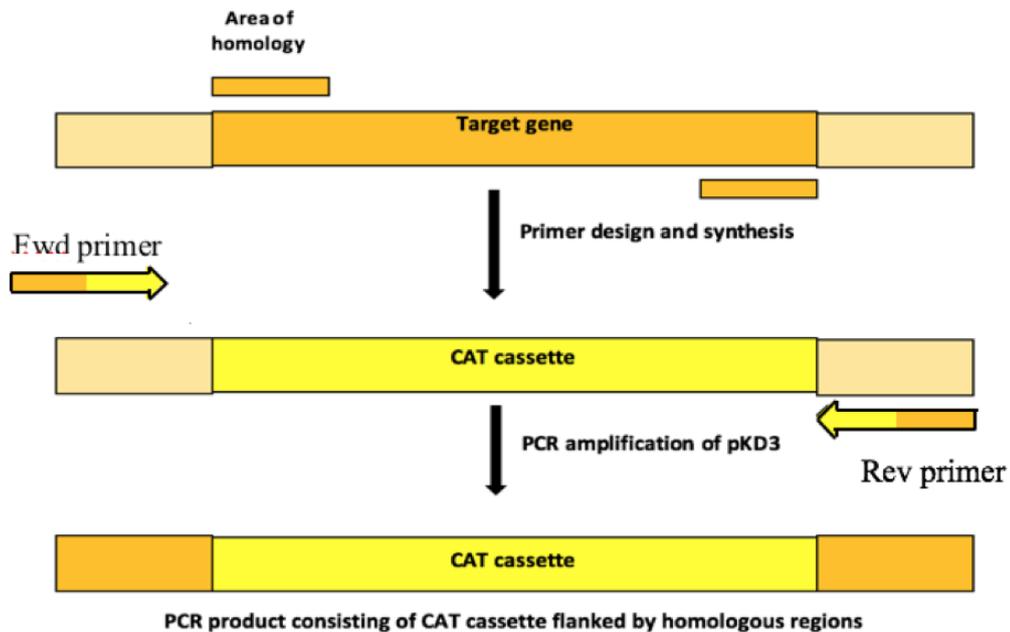
### **2.14.1. Lambda ( $\lambda$ ) Red disruption system**

Gene knockout of *E. coli*  $\Delta hycE$  strain was generated using the Wanner method (Datsenko and Wanner, 2000). This method relies upon the presence of a low copy, temperature sensitive “helper” plasmid encoding components of the homologues recombination system found in bacteriophage  $\lambda$ . These components are called Exo (a 5'-3' exonuclease, which processes along double-stranded DNA), Bet (a single-stranded DNA-binding protein, which is able to anneal complementary single strands) and Gam (an inhibitor of host RecBCD exonucleases). Expression of these genes is under the control of an arabinose-inducible promoter ( $P_{bad}$ ). When cells carrying the plasmid are grown in the presence of arabinose, exogenously applied linear DNA is able to undergo homologues recombination with the bacterial chromosome. In this way, it is possible to generate an in-frame gene deletion using a PCR product.

### **2.14.2. Primer design**

Primers were designed to anneal at the 4<sup>th</sup> codon and the penultimate codon of the target gene, allowing generation of an in-frame deletion with minimal downstream effects. The 5' end of each primer (between 45-48 nucleotides) was homologous to the target gene, whereas the 3' end of each primer was designed to amplify the chloramphenicol resistance cassette encoded by pKD3. This plasmid was selected for use due to the fact that it bears very little similarity to the *E. coli* genome and so potential generation of

unwanted side product is reduced. This process of primer design and the resulting of PCR product is summarized in Fig. 2.2.



**Figure 2.2. Diagram describing the process of primers design and PCR product generation for use in generating single gene knockout strain.**

### 2.14.3. PCR amplification of CAT cassette

The plasmid pKD3 was used as a template for PCR so that linear DNA encoding the  $Cm^R$  cassette could be generated. PCR was carried out as described in section 2.10.2 and the product was purified as described in section 2.10.3.

### 2.14.4. Induction and preparation of host cells

Cells carrying the pKD46 plasmid were grown in LB (containing antibiotics as appropriate) at 30 °C, 250 rpm for 4 h. At this point, arabinose was added to a final concentration of 10 mM in order to induce expression of the homologous recombination system. The cells were incubated under the same conditions for 1 h and then harvested

by centrifugation at 4,000 rpm for 20 min at 4 °C. The cell pellet was then aspirated and re-suspended in 1 ml ice cold water. The cells were then centrifuged at 13,000 rpm for 1 min, the supernatant was removed and the pellet was re-suspended in the same volume of ice cold water. This washing process was repeated five times in total, after which the cells were re-suspended in a volume of ice cold water approximately double that of the pellet. The cells were then aliquoted into pre-chilled electroporation cuvettes and incubated on ice for 15 minutes prior to use.

#### **2.14.5. Electroporation with linear DNA**

About 2 µg of the linear DNA was added to each electroporation cuvette and mixed by pipetting. The cell/DNA mixture was then electroporated. The cells were incubated at 30 °C for 1-3 h and subsequently spread on solid medium containing chloramphenicol (8 µg/ ml). The plates were then incubated at 37 °C overnight. Next day, single colonies were selected for further work and propagated on L-agar plates containing (34 µg/ ml) chloramphenicol.

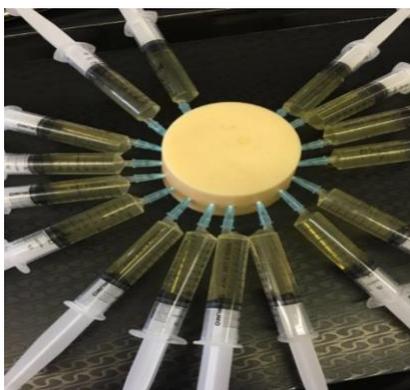
#### **2.15. Formate –hypophosphite inhibition test**

This test was done according to Suppmann and Sawers (1994). The strains were selected to test their growth ability against hypophosphite and formate under aerobic and anaerobic conditions. One single colony was inoculated in 5 ml WM-medium with/without selected antibiotic. Bacteria were grown overnight at 250 rpm, 37 °C in a Sanyo Gallenkamp orbital shaker. After 16-18 h of incubation, optical density (OD) was measured at 600 nm using a portable spectrophotometer. This OD was used to calculate the volume of culture to be added to fresh WM-medium to give a starting OD of 0.01 in Bijoux bottles. These bottles have rubber inserts within their metal screw lids, providing

an air-tight seal. The bottles were incubated at 37 °C for 48 h.

### 2.16. Hydrogen gas production assay

Hydrogen evolution activity was assayed using syringes. The bacterial strains and the transformants to be assayed were grown overnight aerobically at 37 °C in a Sanyo Gallenkamp orbital shaker (250 rpm). Next day, the optical density was measured at 600 nm using a portable spectrophotometer. Then this OD was used to calculate the volume of culture to be added to fresh rich medium (section 2.5.3) to give a starting OD of 0.01 in the 10/20 ml syringes. Formate was added in different concentrations to measure its effect on hydrogen gas evolution and 0.02% rhamnose was added as an inducer for the transformants. Also, an appropriate antibiotic was added to select for the transformants. The whole mixture was then trapped in a syringe without any bubble and the needle inserted into a bung to prevent any oxygen from interrupting the bacterial culture; it was incubated anaerobically at 37 °C for 24-48 h. The results were measured after the incubation period by visualizing the distance that the plunger had moved due to the evolution of gas (Fig 2.3).



**Figure 2.3. Hydrogen assay using syringes.** The syringe is filled fully with medium and the inoculum and the needle is inserted into a bung to provide anaerobic environment.

### 2.17. Formate assay

Formate assay is done by the use of Abcam's Formate Assay kit, where formate is oxidized in the presence of  $\text{NAD}^+$  by the enzyme formate dehydrogenase (FDH) to generate NADH in color formation. The amount of NADH is stoichiometric with the amount of formate. Thus, NADH is the one which was measured at the absorbance ( $\lambda = 450 \text{ nm}$ ).

The strains were grown aerobically overnight at  $37^\circ\text{C}$  and 250 rpm. Next day, 0.01 OD cultures in WM-medium were used to inoculate in 20 ml syringes. The cultures were incubated anaerobically at  $37^\circ\text{C}$  and the formate assay test was performed each 2-3 h during growth. These samples were tested for formate concentration by adding into 96 well plate, 4  $\mu\text{l}$  of the sample; 21  $\mu\text{l}$  formate assay buffer and 25  $\mu\text{l}$  of the reaction mix. Each well was adjusted to a total of 50  $\mu\text{l}$ . The standards were prepared by taking 10  $\mu\text{l}$  of 100 mM formate standard solution and diluted into 990  $\mu\text{l}$  pure distilled water to make 1 mM concentration. Then, 1 to 5  $\mu\text{l}$  of the diluted formate standards were added into 96 well plate with the addition of 25  $\mu\text{l}$  of formate assay buffer and 25  $\mu\text{l}$  of the reaction mix (23  $\mu\text{l}$  formate assay buffer + 1  $\mu\text{l}$  of formate enzyme mix + 1  $\mu\text{l}$  formate substrate mix) to each diluent to generate 1 to 5 nmol/ well of the formate standard. The blank was contained 24  $\mu\text{l}$  formate assay buffer and 1  $\mu\text{l}$  formate substrate mix. The reaction was mixed well and incubated for 60 minutes at  $37^\circ\text{C}$ . After incubation period, the 96 well plate (reactions) were measured at  $\text{OD}_{450\text{nm}}$  in a microplate reader (SpectraMax). The data were analyzed by plotting a standard curve of nmol/ well vs  $\text{OD}_{450\text{nm}}$ . The formate concentration in the test samples was calculated as follow:

Concentration (nmol/  $\mu\text{l}$ ) = the amount of formate (nmol) of each sample from the standard curve)/ the sample volume ( $\mu\text{l}$ ) added to the reaction well.

## **2.18. Protein cell extract and hydrogenase activity staining**

### **2.18.1. Protein sample preparation**

This work was performed according to Pinske *et al.* (2012) by growing bacterial strains overnight in YGEP media (section 2.5.4) at pH 6.5 and 37 °C. Fresh cultures of OD 0.01 were used to inoculate 20 ml syringes which were incubated at 37 °C anaerobically until the mid-exponential phase. The cells were then harvested by centrifugation (10,000 x g for 15 min at 4 °C). The pellets were washed in the same volume of 50 mM MOPS buffer (pH 7.0) and was then re-suspended in one tenth of the sample volume with the same buffer. The cells were kept on ice and ready for broken by sonication (Sonics Vibra-Cells) at 80% amplitude with a 6 mm diameter tip for 10 minutes in short bursts of 20 seconds on and 20 seconds off to lyse the cells and any remaining intact cells were removed by centrifuge at 15,000 rpm for 30 min. The crude extract was then re-suspended at a protein concentration of 10 mg/ml in 50 mM MOPS (pH 7.0). Protein concentration was measured by Bradford assay and nanodrop spectrophotometer using Nanodrop ND-1000 spectrophotometer (section 2.18.2.1; section 2.18.2.2). The sample was adjusted to a final concentration of 5% (w/v) Triton X-100 prior to electrophoresis.

### **2.18.2. Determination of protein concentration**

Protein concentration was determined according to the dye-binding method of Bradford (1976) using a prefabricated assay from Bio-Rad Laboratories (Hercules, CA) and bovine serum albumin (BSA) as protein standard. Standards were prepared by diluting bovine serum albumin (BSA) in a series ranging from 0.025 to 5 mg/ml. Volumes of 10 µl for each duplicate BSA dilution (standards) and 10 µl of each protein sample (unknowns) were added into 96-wells microtiter plate, followed by adding 190 µl of 1x diluted (Bio-Rad) protein assay dye reagent. The absorbance was measured at 595 nm using a

microtiter-plate reader. Using measurement obtained from the standards, a standard curve was plotted, from which the protein concentration was calculated.

The protein concentration was also assessed using a Nandrop ND-1000 (Nanodrop Technologies) which measures the absorbance of 2  $\mu$ l of protein sample at 280 nm. Both data from Nanodrop and Bradford assay were compared to ensure consistency and accuracy.

### **2.18.3. Hydrogenase activity staining**

Activity staining was performed to detect hydrogenase 4 activity by showing a clear band following native PAGE and treatment with a continuous hydrogen flush. Native 7.5% PAGE was as in section 2.10.8.3). After loading in the gel, 30 mA was applied for 3-4 h, the gel was then soaked in 0.2 mM benzylviologen (BV) and 1 mM triphenyl tetrazolium chloride (TTC) with continuous flushing with pure hydrogen gas for ~10 min, until the activity bands appeared.

## **2.19. Phenotypic studies**

### **2.19.1 Phenotypic analysis of $\Delta$ *focA*, $\Delta$ *focB*, $\Delta$ *hyc* and $\Delta$ *hyf***

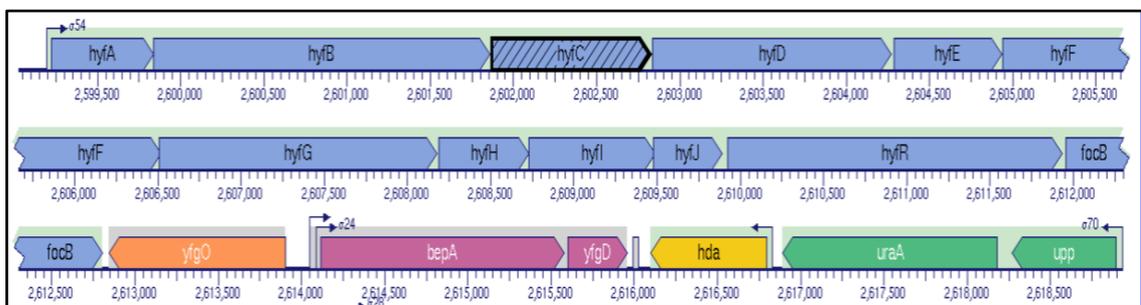
Bacterial strains used were grown in the appropriate medium unless otherwise specified. To prepare an overnight culture, 5 ml of medium was inoculated with a single bacterial colony from a freshly streaked plate or scraped directly from seed stocks. Overnight cultures were grown in a shaking incubator for 16 h (stationary phase). The pH levels were adjusted using three different Good buffers; 0.2 M MES (acidic), 0.2 M MOPS (neutral) and 0.2 M HEPES (basic). After 16-18 h of growth, OD was measured to calculate the volume of the culture needed for addition to fresh medium to give an OD of

0.01. To measure the effect of certain chemicals on the bacterial growth, such as the effect of hypophosphite and different formate concentrations under different pH conditions, the cultures were grown in Bijoux bottles with lids and rubber seals to achieve anaerobiosis.

## Chapter 3: Bioinformatics analysis

### 3.1 Bioinformatic data collections

*E. coli* contains an operon called *hyf* that encodes a type of [Ni-Fe] hydrogenase called hydrogenase-4 (Hyd-4). The *hyf* operon of *E. coli* K-12 starts at 2,599,500 bp and ends at 2,612,500 bp (Andrews *et al.*, 1997) (Fig 3.1). Seven genes of the *hyf* (*hyfABCDEFGHIJ*) operon are homologues to corresponding *hyc* genes (encoding Hyd-3). However, *hyfD*, *hyfE*, *hyfF* encode integral membrane subunits which have no direct equivalent in Hyd-3 (Skibinski *et al.*, 2002).



**Figure 3.1. Physical map and gene organisation of *hyf* operon in the *E. coli* K-12 chromosome (BioCyc Database Collection).**

To identify species which contain *hyf* and *hyc*, the *hyfD*, *hyfE* and *hyfF* sequences of *E. coli* K-12 were obtained from the NCBI database ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) in FASTA format. These were then used to search for homologues using the Basic Local Alignment Search Tool (BLAST; [www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)). In this way, the number of *hyfD*, *hyfE* & *hyfF* homologues in the sequence database, and their degree of similarity and taxonomic distribution could be discerned.

### 3.2. Phylogenetic tree

Since the gene encoding ‘the RNA-polymerase’  $\beta$  (RpoB) subunit is conserved, this was used as a phylogenetic indicator for *E. coli* strains and close relatives. The corresponding amino acid sequence from a K-12 strain was obtained from Uniprot and used for BLAST search against the microbial genomic database in NCBI. This provided a list of all *E. coli* strains and close relatives (*Salmonella* and *Shigella*) within the database, 400 in total (Table 3.1). Each strain was checked for the presence of the three genes *hyfD*, *hyfE* and *hyfF* that distinguish *hyf* from *hyc*. No *Salmonella* species was found that contained *hyf* genes. Furthermore, *E. coli* (SE15, NA114, CFT073, APEC\_O1, IHE3034, ED1a, LF82 and O127\_H6, E2348/69) were also found to lack *hyf* genes. A phylogenetic tree was constructed using RpoB amino acid sequences (Fig 3.2) and the phylogeny.fr program ([www.phylogeny.fr](http://www.phylogeny.fr)). Each strain found to carry *hyfD*, *hyfE* and *hyfF* was checked for the presence of the *hyc* operon by searching for *hycE* (large subunit of the *hyc* operon) using ‘BLAST’, and it was found that all *Escherichia*, *Salmonella* and *Shigella* strains (Table 3.1) contain *hyc*.

The phylogenetic tree showed that all of the *E. coli* strains that lacked *hyf* are clustered into a single clad and that all members of this clad are *hyf*-free (Fig 3.2). Although these strains were isolated from different sources (Fig 3.2), most of them cause urinary tract infection (UTI) (Krause *et al.*, 2011, Andersen *et al.*, 2013, Avasthi *et al.*, 2011, Lu *et al.*, 2011, Matsumoto, 2001, Iguchi *et al.*, 2009, McCusker *et al.*, 2014 and Zhou *et al.*, 2014). For example, strains D\_i14 and D\_i2 were isolated from stool samples and were found to cause urinary tract infection (Fig 3.3) (Peris-Bondia *et al.*, 2013). Most of the *hyf* carrying strains (*E. coli* and *Shigella*), where information is available, were isolated from faecal samples (Fig 3.2), for instance *E. coli* K-12 (Oshima *et al.*, 2008).

**Table 3.1. The presence of *hyfD*, *hyfE* and *hyfF* in *E. coli* strains and close relatives with percentage match indicated.**

Species (strain)	<i>hyfD</i>	<i>hyfE</i>	<i>hyfF</i>
<i>Salmonella agona</i> (24249)	X	X	X
<i>Salmonella agona</i> (SL483)	X	X	X
<i>Salmonella javiana</i> (CFSAN001992)	X	X	X
<i>Salmonella heidelberg</i> (CFSAN002069)	X	X	X
<i>Salmonella heidelberg</i> (41578)	X	X	X
<i>Salmonella heidelberg</i> (B182)	X	X	X
<i>Salmonella heidelberg</i> (SL476)	X	X	X
<i>Salmonella paratyphi-C</i> (RKS4594)	X	X	X
<i>Salmonella cubana</i> (CFSAN002050)	X	X	X
<i>Salmonella typhi</i> (Ty21a)	X	X	X
<i>Salmonella typhi</i> (P-stx-12)	X	X	X
<i>Salmonella paratyphi-A</i> (AKU12601)	X	X	X
<i>Salmonella paratyphi-B</i> (SPB7)	X	X	X
<i>Salmonella enteritidis</i> (P125109)	X	X	X
<i>Salmonella bareilly</i> (CFSAN000189)	X	X	X
<i>Salmonella thompson</i> (RM6836)	X	X	X
<i>Salmonella typhimurium</i> (CFSAN002069)	X	X	X
<i>Salmonella schwarzengrund</i> (CVM19633)	X	X	X
<i>Salmonella dublin</i> (CT02021853)	X	X	X
<i>Salmonella newport</i> (USMARC S3124.1)	X	X	X
<i>Salmonella newport</i> (SL254)	X	X	X
<i>Salmonella arizonae</i> (RSK2980)	X	X	X
<i>E. albertii</i> (KF1)	X	X	X
<i>E. fergusonii</i> (ATCC_35469)	√ (75%)	√ (79%)	√ (77%)
<i>E. coli</i> (SMS-3-5)	√ (98%)	√ (99%)	√ (98%)
<i>E. coli</i> o7_K1 (CE10)	√ (98%)	√ (99%)	√ (98%)
<i>E. coli</i> (IAI39)	√ (98%)	√ (99%)	√ (98%)
<i>E. coli</i> (JJ1886)	X	X	X
<i>E. coli</i> (SE15)	X	X	X
<i>E. coli</i> (NA114)	X	X	X
<i>E. coli</i> (ABU 83972)	X	X	X
<i>E. coli</i> (clone Di14)	X	X	X
<i>E. coli</i> (clone Di2)	X	X	X
<i>E. coli</i> (CFT073)	X	X	X
<i>E. coli</i> (536)	X	X	X
<i>E. coli</i> (APEC-O1)	X	X	X
<i>E. coli</i> (S88)	X	X	X
<i>E. coli</i> (IHE3034)	X	X	X
<i>E. coli</i> (PMV-1)	X	X	X
<i>E. coli</i> (UTI89)	X	X	X
<i>E. coli</i> (ED1-a)	X	X	X
<i>E. coli-O83_H1</i> (NRG_857C)	X	X	X
<i>E. coli</i> (LF82)	X	X	X
<i>E. coli</i> (042)	√ (99%)	√ (98%)	√ (98%)
<i>E. coli</i> (UMN026)	√ (99%)	√ (99%)	√ (98%)
<i>E. coli-O145_H28</i> (RM13514)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli-O145_H28</i> (RM13516)	√ (99%)	√ (98%)	√ (99%)
<i>E. coli-O157_H7</i> (TW14359)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli- O157_H7</i> (EC4115)	√ (99%)	√ (99%)	√ (99%)

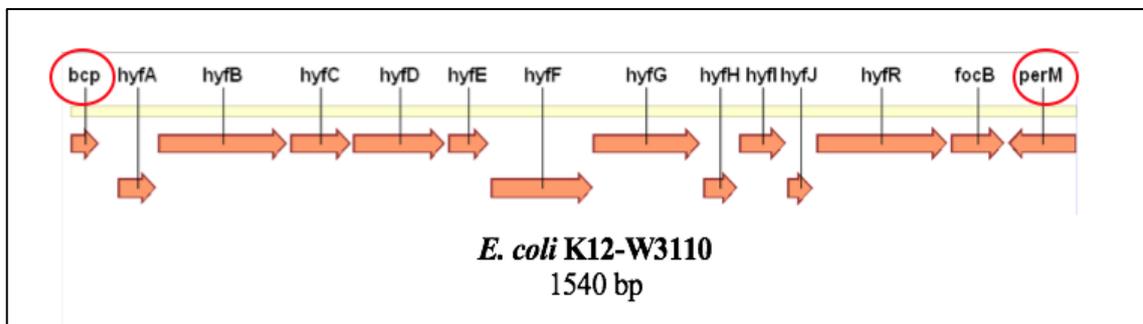
<i>E. coli- O157_H7</i> (EDL933)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli</i> (Xuzhou21)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli- O55_H7</i> (RM12579)	√ (98%)	√ (98%)	√ (99%)
<i>E. coli- O55_H7</i> (CB9615)	√ (98%)	√ (98%)	√ (99%)
<i>E. coli- O157_H7</i> (Sakai)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli-BL21</i> (DE3)	√ (99%)	√ (99%)	√ (100%)
<i>E. coli</i> (P12b)	√ (100%)	√ (100%)	√ (99%)
<i>E. coli-B</i> (REL606)	√ (99%)	√ (99%)	√ (100%)
<i>E. coli_BL21-Gold_DE3</i>	√ (99%)	√ (99%)	√ (100%)
<i>E. coli</i> (ATCC_8739)	√ (99%)	√ (99%)	√ (99%)
<i>Shigella dysenteriae</i> (Sd197)	√ (99%)	√ (98%)	√ (99%)
<i>Shigella sonnei</i> (53G)	√ (99%)	√ (99%)	√ (99%)
<i>Shigella sonnei</i> (Ss046)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli O111_H</i> (11128)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli O26_H11</i> (11368)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli</i> (E24377A)	√ (98%)	√ (99%)	√ (98%)
<i>Shigella flexneri-5</i> (8401)	√ (99%)	√ (99%)	√ (98%)
<i>Shigella flexneri</i> (2002017)	√ (99%)	√ (99%)	√ (99%)
<i>Shigella flexneri-2a</i> (301)	√ (99%)	√ (99%)	√ (99%)
<i>Shigella flexneri-2a</i> (2457T)	√ (99%)	√ (99%)	√ (99%)
<i>Shigella boydii</i> (3083-94)	√ (98%)	√ (99%)	√ (99%)
<i>Shigella boydii</i> (Sb227)	√ (98%)	√ (99%)	√ (99%)
<i>E. coli-O104-H4</i> (2009EL-2050)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli-O104-H4</i> (2011C-3493)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli-O104-H4</i> (2009EL-2071)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli</i> (55989)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli -O103-H2</i> (12009)	√ (99%)	√ (99%)	√ (98%)
<i>E. coli</i> (SE11)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli</i> (LY180)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli APEC</i> (O78)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli</i> (W)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli</i> (KO11FL)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli</i> (HS)	√ (100%)	√ (100%)	√ (100%)
<i>E. coli</i> (C321.deltaA)	√ (100%)	√ (100%)	√ (100%)
<i>E. coli-K12</i> (MG1655)	√ (100%)	√ (100%)	√ (100%)
<i>E. coli-K12</i> (MDS42)	√ (100%)	√ (100%)	√ (100%)
<i>E. coli-K12</i> (W3110)	√ (100%)	√ (100%)	√ (100%)
<i>E. coli</i> (DH1)	√ (100%)	√ (100%)	√ (100%)
<i>E. coli</i> (ETEC-H10407)	√ (99%)	√ (99%)	√ (99%)
<i>E. coli</i> (UMNK88)	√ (100%)	√ (100%)	√ (100%)
<i>E. coli</i> (BW2952)	√ (100%)	√ (100%)	√ (100%)
<i>E. coli-K12</i> (DH10B)	√ (100%)	√ (100%)	√ (100%)
<i>E. coli-K12</i> (MC4100)	√ (100%)	√ (100%)	√ (100%)



**Figure 3.2. Phylogenetic tree of all *E. coli* strains and close relatives within the complete microbial chromosome database of NCBI.** The tree was derived from an amino acid sequence alignment of the ‘RNA Polymerase  $\beta$ ’ subunit of ~500 different strains in the database selected on the basis of their sequence similarities of 80%-100%. The boxes indicate the strains with no *hyf*. The circles indicate the source of the isolate strain: Urine, ●; Faeces, ●; Spinal fluid, ●; Blood, ●; Food industry contamination, ●; Intestine, ●; and Environmental contamination, ●. A subset of strains were selected randomly to assess their niches.

### 3.3. Physical map

A number of strains were selected, from representative points in the phylogenetic tree, from each group (strains with *hyf* and without *hyf* region) to view the upstream and downstream regions of the *hyf* operon in order to determine the differences between *hyf*<sup>+</sup> and *hyf*<sup>-</sup> strains at the *hyf* locus. This was done by using various programs (Xbase, Ecobase and NCBI). The strains with no *hyf* used were *E. albertii* (KF1) and *E. coli* (O127:H6 - E2348/69), (SE15), (NA114), (CFT073), (APEC-O1), (IHE3034), (ED1a) and (LF82); while the strains with *hyf* were *E. fergusonii* (ATCC35469) and *E. coli* (SMS\_3\_5), (042) and (O157:H7), (E24377A), (55989) and (MG1655), Sakai, *Shigella sonnei* (53G), and *E. coli* K-12 (W3110). For *E. coli* K-12, upstream of *hyfA* is *bcp* (thioredoxin dependent thiol peroxidase), and downstream of *focB* is *perM* (permease/UP0118 family) (Fig 3.3).



**Figure 3.3. Physical map of the *hyf* region of *E. coli* (W3110) and adjacent genes (2,599,134 – 2,614,537). Red circles show the upstream and downstream of *hyf* operon.**

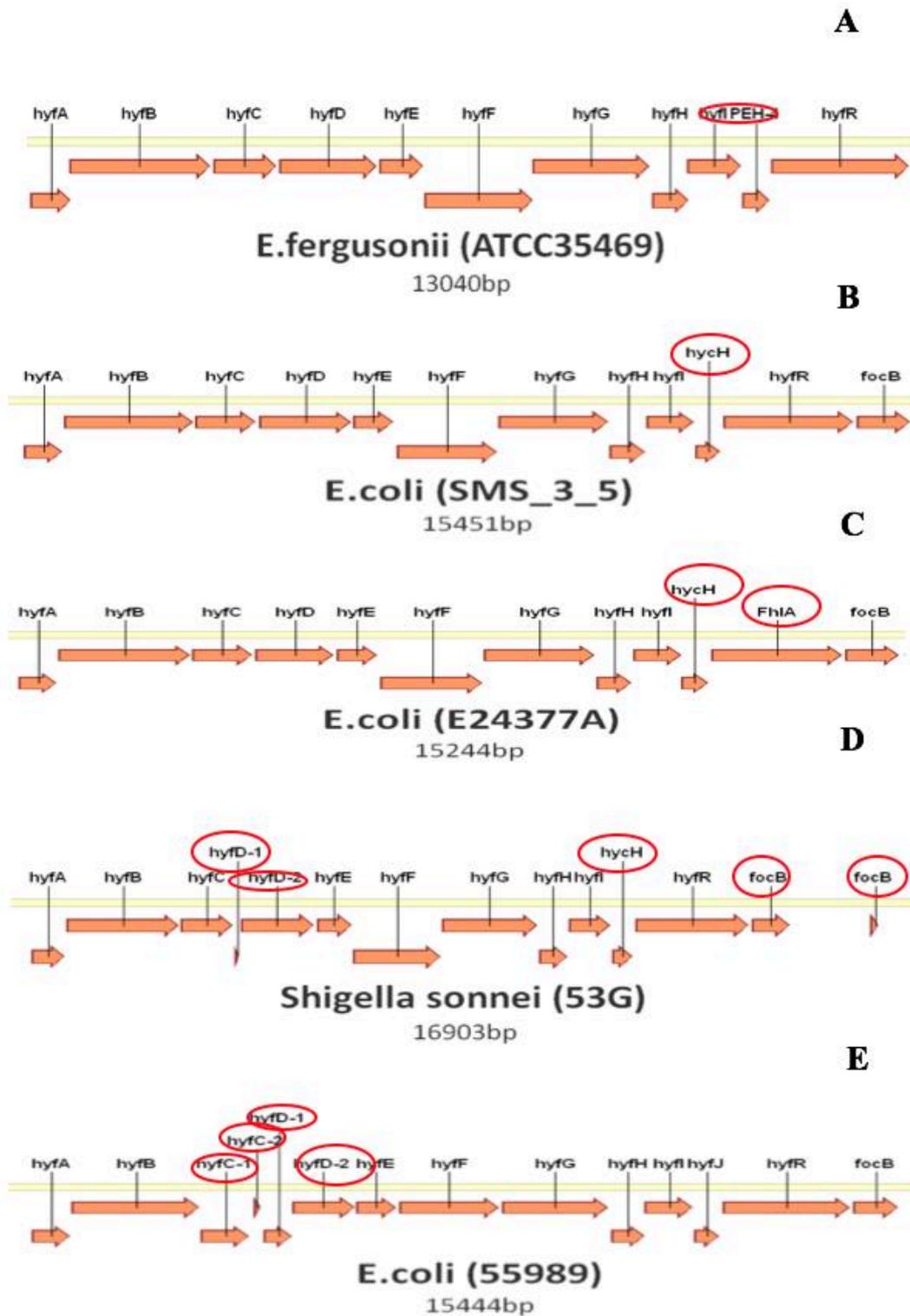
Upon comparison with the *hyf* operons of the other strains listed above, three different organisations were apparent: some selected strains contain a full *hyf* operon whereas others do not. For example, *E. fergusonii* ATCC35469 has a full *hyf* region except that *focB* is not present. This could be due to the presence of *focA* elsewhere on the chromosome (adjacent to *pflB*) which might render *focB* redundant. *E. coli* SMS\_3\_5, as

annotated, contains a full operon except for *hyfJ*, where a *hycH* homologue is reported to be present instead. However, since BLAST showed that the identity between *hyfJ* and the SMS\_3\_5 *hycH* gene is 99%, it indicates that *hycH* is mis-annotated and is indeed *hyfJ*. On the other hand, the *hyf* operon of *E. coli* (E24377A) lacks two *hyf* genes (*hyfJ* and *hyfR*) but includes two *hyc*-related genes instead (*hycH* and *fhIA*) and because the identity level, according to BLAST, between *hyfJ* and *hycH*, and *hyfR* and *fhIA*, is 99%, this shows that these genes should be annotated *hyfJ* and *hyfR* instead of *hycH* and *fhIA*. In *Shigella sonnei* (53G), *hyfD* and *focB* are each split into two fragments. When BLAST analysis was performed for the two gene fragments with the *E. coli* K-12 equivalents, both *hyfD* and *focB* were shown to have 98% identity to each of the split genes in K-12. However, although *hyfD* gave 100% query cover, *focB* gave just 89% suggesting loss of part of the open reading frame (orf). This division of these two genes into two orfs suggests that the genes may not be functional. In addition, *hyfJ* was also reported as absent, with *hycH* present instead, but according to BLAST, *hycH* is mis-annotated and is indeed, *hyfJ*. *hyfC* and *hyfD* were also each split in two fragments in *E. coli* (55989), and BLAST showed that they were 99% identical with the corresponding W3110 protein with 88-89% query cover (Fig 3.4). For the two strains in which *hyf* genes (*hyfD*, *hyfC* and *focB*) are divided into two open reading frames (orfs), the divided genes were more closely examined using Vector NTI. The two orfs of *focB* in *S. sonnei* are separated by ~1 kb indicating a major disruption of *focB* (Fig 3.4) which is thus unlikely to be functional. Additionally, in *S. sonnei*, the two orfs for *hyfD* (37 and 415 codon) (Fig 3.5) overlap sufficiently to suggest the possibility that this gene might retain function if the two predicted translation products are able to assemble together within the Hyf complex. For the two orfs of *hyfC* in *E. coli* 55989, the separation between the up and downstream orfs is 19 codons, and the two predicted coding products lack 19 amino acids

(QELSLACLLTSLVVTTLLKV) with respect to the intact *hyfC* product which suggests that they would not be functional (Fig 3.6). For *hyfD* of strain 55989, the two *hyfD* orfs (152 and 329 codons; Fig 3.7) overlap in a fashion consistent with translational coupling which suggests they could each be expressed to generate functional products that assemble together within the proposed Hyf multi-enzyme complex.

The second group of *E. coli* strains considered here are those that fully lack the *hyf* operon region: *E. coli* (SE15, NA114, CFT073, APEC-O1, ED1a, LF82, O127: H6, E2348/69) and *E. albertii* (KF1) (Fig. 3.8). It can be seen that no additional genes are found at the *hyf* locus, between *bcp* and *perM*. Thus, the inclusion or exclusion of the *hyf* operon within *E. coli* appears to represent a genetic insertion/deletion process, rather than major genetic substitution.

The third group have the entire *hyf* operon without any exceptions or mis-annotations: e.g. *E. coli* (042), (MG1644) and (O157: H7-Sakai) (Fig. 3.9).



**Figure 3.4.** *E. coli* strains with some missing *hyf* genes. (a) *hyfJ* and *focB* are missing, instead there is processing element hydrogenase-4; which is '*hyfJ*'. (b) *hyfJ* is absent and *hycH* in its place. (c) *hyfJ* and *hyfR* are missing and there are *hycH* and *fhIA* respectively instead. (d) Two fragments of *hyfD* and *focB* appeared, and *hyfJ* is absent and *hycH* is instead. (e) Two split fragments of *hyfC* and *hyfD* appeared.

```

+2 MET GLU ASN LEU ALA LEU THR THR LEU MET LEU PRO PHE ILE GLY ALA LEU VAL VAL SER PHE SER PRO
4301 GGTAAACCGG TCTGTAAGGA GCACTGACGG AATATGGAAA ATCTTGCTCT GACGACGTTA ATGCTGCCTT TTATCGCGCG ACTGGTCGTT TCGTTTTCGG
+2 PROGLN ARG ARG ALA ALA GLU TRP GLY PHE CYS SER PRO HIS TER
+1 LEU PHE ALA ALA LEU THR THR LEU LEU THR THR LEU CYS MET LEU SER LEU ILE SER ALA PHE TYR GLN ALA ASP LYS VAL ALA
4401 CACAACGTCG GCGCCGCCGA TGGGGGTTTT GTTCGCCGCA CTGACCAACGC TGTGCATGTT GTCGCTGATC TCCGGGTTTT ATCAGGCCGA TAAAGTTGGC
+1 VAL THR LEU THR LEU VAL ASN VAL GLY ASP VAL ALA LEU PHE GLY LEU VAL ILE ASP ARG MET SER THR LEU ILE LEU PHE VAL VAL VAL PHE LEU GLY LEU
4501 GTCACGTTGA CGTGGTCAA CGTGGGGGAT GTGGCATTGT TTGGCTTGGT CATTGATCGC ATGAGTACGC TGATTCTGTT TGTGGTGGTG TTTCTCGGTT
+1 LEU LEU VAL THR ILE TYR SER THR GLY TYR LEU THR ASP LYS ASN ARG GLU HIS PRO HIS ASN GLY THR ASN ARG TYR TYR ALA PHE LEU LEU VAL PHE ILE
4601 TGCTGGTCAC GATCTACTCC ACGGGTTATC TGACGGATAA AAATCGCGAA CACCCGCATA ACGGCACGAA TCGTTATTAC GCATTTTTGC TGGTGTATTAC
+1 ILE ILE ALA MET ALA ILE LEU VAL LEU SER SER THR LEU LEU LEU THR THR LEU LEU PHE PHE GLU ILE THR GLY GLY CYS SER TRP ALA LEU ILE SER TYR TYR
4701 CCGCGCGATG GCGGGACTGG TACTCTCCTC GACGCTGCTC GGTCAGTTGT TGTTTTTGA AATTACGGGC GGCTGCCTCT GGGCGTTGAT CAGTTATTAC
+1 GLN SER ASP LYS ALA GLN LEU SER ALA LEU LYS ALA LEU LEU ILE THR HIS ILE GLY SER LEU GLY LEU TYR LEU ALA ALA ALA THR LEU PHE LEU GLN THR
4801 CAGAGCGATA AAGCACAAGT TTAGCACTA AAAGCGTTAC TTATCACTCA TATCGGCTCG TTGGGGTGT ATCTTGGCGC CGCCACGCTG TTTTTCGAGA
+1 TRHGLY THR PHE ALA LEU SER ALA MET SER GLU LEU HIS GLY ASP ALA ARG TYR LEU VAL TYR GLY GLY ILE LEU PHE ALA ALA TRP GLY LYS SER ALA GLN
4901 CCGAARCTT TGGCCTTAGC GCGATGAGCG AGTTACACCG CGACGCACGT TATCTGGTTT ATGGCGCAT CCTGTTTGGC GCGTGGGGGA AATCGGCCA
+1 GLN LEU PRO MET GLN ALA TRP LEU PRO ASP ALA MET GLU ALA PRO THR PRO ILE SER ALA TYR LEU HIS ALA ALA SER MET VAL LYS VAL GLY VAL TYR ILE
5001 GCTACCGATG CAAGCTGGC TACCGGATGC AATGGAAAGC CCAACACCGA TCAGTGCTTA CTCCACGCC GCATCGATGG TGAAGTGGG CGTTTACATT
+1 PHE ALA ARG ALA ILE ILE ASP GLY GLY ASN ILE PRO HIS VAL ILE GLY GLY VAL GLY MET VAL MET ALA LEU VAL THR ILE LEU TYR GLY PHE LEU MET TYR
5101 TTTGCCCGCG CCATTATCGA CCGCGGCAAT ATCCCGCATG TGATTGGCGG CGTTGGCATG GTCATGGCGC TGCTCACCAT TCTTATGTC TTTCTGATGT
+1 TYR LEU PRO GLN GLN ASP MET LYS ARG LEU LEU ALA TRP SER THR ILE THR GLN LEU GLY TRP MET PHE PHE GLY LEU SER LEU SER ILE PHE GLY SER ARG
5201 ATTTGCCACA GCAGGATATG AAGCGTTGC TGGCCTGGTC GACCATCACT CAACTTGGCT GGATGTTCTT CCGCTTGTGC CTCTCCATCT TCGGCTCGCG
+1 ARG LEU ALA LEU GLU GLY SER ILE ALA TYR ILE VAL ASN HIS ALA PHE ALA LYS SER LEU PHE PHE LEU VAL ALA GLY ALA LEU SER TYR SER CYS GLY THR
5301 GCTGGCGCTG GAGGATAGCA TCGCTACAT CGTCAACACC CGCTGCGCTA AAAGCCTGTT TTTCTTGTGA GCAGTGCCG TGAGTTACAG CTGGCCGACG
+1 ARG LEU LEU PRO ARG LEU ARG GLY VAL LEU HIS THR LEU PRO LEU PRO SER VAL GLY PHE CYS VAL ALA ALA LEU ALA ILE THR GLY VAL PRO PRO PHR SM
5401 CGCTTGTGTC CCGCTGTGCG TGGCGTATTG CACACCTGCG CGTTGCCAAG CGTGGGTTTC TGGTGGCAG CGCTGGCGAT TACCGGTGTG CCGCCGTTCA
+1 ASN GLY PHE PHE SER LYS PHE PRO LEU PHE ALA VAL GLY PHE ALA LEU SER VAL GLU TYR TRP ILE LEU LEU PRO ALA MET ILE LEU LEU MET ILE GLUSER
5501 ACGGCTTCTT CAGTAAATTC CCGCTGTTG CCGCTGGTTT TGCCTGTCA GTGGAGTACT GGATCCTGCT GCCCGCATC ATTCTGCTGA TGATTGAATC
+1 SER VAL ALA SER PHE ALA TRP PHE ILE ARG TRP PHE GLY ARG VAL VAL PRO GLY LYS PRO SER GLU ALA VAL ALA ASP ALA ALA PRO LEU PRO GLY SER MET
5601 GGTGCCAGT TGGCCTGGT TTATTGCTG GTTTGGTGC GTCGTGCGG GCAAACCGAG CGAGGCCGTC GCCGATGCC CACCGCTGCC AGGGTCAATG
+1 ARG LEU VAL LEU ILE VAL LEU ILE VAL MET SER LEU ILE SER SER VAL ILE ALA ALA THR TRP LEU GLN TER
5701 CGCTGGTGT TGATTGTA CTGTTGATG TCGCTGATT CCAGCGTAAAT CGCCGCGACC TGGTTGCAGT AAGGAGATGA TGAATGACCG GTTCTATGAT

```

**Figure 3.5. Translated sequences of the two fragments of *hyfD* in *Shigella sonnei* (53G).** The green highlight indicates the first orf of *hyfD* where the start codon is predicted to be (ATG), while the yellow highlight shows the second orf of *hyfD* and the start codon could be TTG or an ATG (9 codons further downstream). Azure highlight is where the two orfs overlap.

```

+1 LEU PRO SER LYS GLU ILE THR MET ARG GLN THR LEU CYS ASP GLY TYR LEU VAL
3301 TCTATTGCC TATGTTGGT CCGCACTGG TTGTGCTGCT AATCGCTATT GCCGCTAAG GAAATCACCA TGAGACAAAC TCTTTGGCAG GGATATCTGG
+1 VAL ILE PHE ALA LEU ALA GLN ALA VAL ILE LEU LEU MET LEU THR PRO LEU PHE THR GLY ILE SER ARG GLN ILE ARG ALA ARG MET HIS SER ARG ARG GLY
3401 TTATTTTTG GTTAGCACAG GCCGTGATTC TGCTGATGCT AACCCCACTT TTTACGGGTA TTTCCCGCA GATACGCGCG CGTATGCATC CCGCCGCGCG
+1 GLY PRO GLY ILE TRP GLN ASP TYR ARG ASP ILE HIS LYS MET PHE LYS ARG GLN GLU VAL ALA PRO THR SER SER GLY LEU MET SER ARG LEU MET PRO TRP
3501 GCGCGGGATC TGGCAGGATT ATCGCGATAT CCACAAAATG TTTAAACGCC AGGAAGTTGC GCGACATCT TCAGGTCTGA TGTTCGCGCT GATGCGCGTG
+1 VAL LEU ILE SER SER MET LEU VAL LEU ALA MET ALA LEU PRO LEU PHE ILE THR VAL SER PRO PHE ALA GLY GLY GLY ASP LEU ILE THR LEU ILE TYR LEU
3601 GTATTAATCA GCAGCATGCT GGTGCTGGC ATGGCCTTAC CACTGTTAT TACCCTTTC CCTTTTGGC GCGCGCGCA TCTGATCACC CTTATCTATC
+1 LEU LEU ALA LEU PHE ARG PHE PHE PHE ALA LEU SER GLY LEU ASP THR GLY SER PRO PHE ALA GLY VAL GLY ALA SER ARG GLU LEU THR LEU GLY ILE LEU
3701 TCTTGGCCT GTTTCGTTTT TTCTTGCTC TTCCGGGCT GATACCCGGA AGTCCGTTTG CCGGAGTCCG TGCCAGTCCG GAGTTGACGC TCGGCATTCT
+1 LEU ALA LEU LEU SER LEU LEU VAL LEU ALA LEU ILE ALA LEU ILE THR THR HIS ILE GLU MET ILE SER ASN THR LEU ALA MET GLY TRP ASN
3801 GGTGGAACCA ATGCTTATTC TCTCACTGCT GGTATTGGC CTGATAGCAG GTTCCACGCA TATCGAGATG ATCAGCAATA CGCTGGCGAT GGGCTGGAAC
+1 SER PRO LEU THR THR VAL LEU ALA LEU LEU ALA CYS GLY PHE ALA CYS PHE ILE GLU MET GLY LYS ILE PRO PHE ASP VAL ALA GLU ALA GLU GLN GLU LEU
3901 TCGCCGCTAA CCACCGTACT GGCCTTACT GCCTGTGGTT TTGCGTCTT CATTGAGATG GGAATAATC CTTTGTATGT TGCTGAAGCA GAACGAGAA
+1 LEU GLN GLU GLY PRO LEU THR GLU TYR SER GLY ALA GLY LEU ALA LEU ALA LYS TRP GLY LEU GLY LEU LYS GLN VAL VAL MET ALA SER LEU PHE VAL ALA
4001 TACAGGAAG CCCGCTAAC GAATATTCCG GTGCCGGCT GCGCTAGCG AAATGGGGC TGGGGCTGAA ACAGGCTGCT ATGGCATCAC TGTTTGTAGC
+1 ALA LEU PHE LEU SER PHE GLY ARG ALA TER LEU LEU ILE PHE VAL LEU
4101 CCGTITCTG TCCTTTGGC GCGGTAAGA ACTTCTCTC GCCTGCCTGC TGACTTCACT TGTGTTACG CTGCTCAAGG TTTTGTGAT TTTTGTACTG
+1 ALA SER ILE ALA GLU ASN THR LEU ALA ARG GLY ARG PHE LEU LEU ILE HIS HIS VAL THR TRP LEU GLY PHE SER LEU ALA ALA LEU ALA TRP VAL PHR P
4201 GCCTCAATCG CAGAAAACAC GCTGGCACGC GGGGTTTTT TACTCATTCA CCATGTGACC TGGCTTGGCT TCAGCCTTGC TGGCTTGA TGGGCTTCT
+1 GLY SER TRP VAL CYS LYS
4301 GGTAAACCGG TCTGTAAGGA G

```

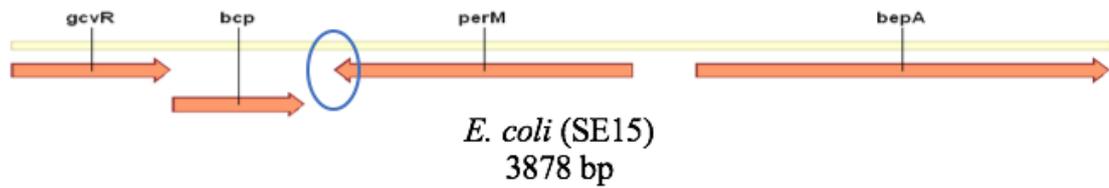
**Figure 3.6. Translated sequences of the two fragments of *hyfC* in *E. coli* (55989).** The yellow highlight indicates the first orf of *hyfC* where the start codon is predicted to be TTG (or more likely, the ATG 7 codons downstream), while the green highlight is for the second orf of *hyfC*. It is clear there is no overlapping between the two open reading frames which suggests that this gene is non-functional.

```

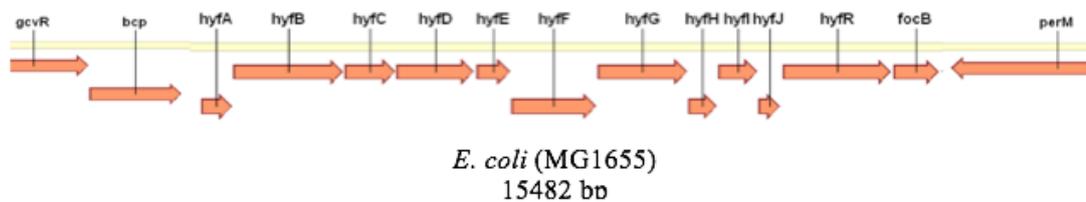
+2 MET GLU ASN LEU ALA LEU THR THR LEU LEU LEU PRO PHE ILE GLY ALA LEU VAL VAL SER PHE SERRO
4301 GGTAAACCGG TCTGTAAGGA GCACTGACGG AATATGGAAA ATCTTGCTCT GACGACGTTA TTGCTGCCTT TTATCGCCGC ACTGGTGGTT TCGTTTTCCG
+2 PROGLN ARG ARG ALA ALA GLU TRP GLY VAL LEU PHE ALA ALA LEU THR THR LEU CYS MET LEU SER LEU ILE SER ALA PHE TYR GLN ALA ASP LYS VALALA
4401 CACAACGTCG GGGCCCGGAA TGGGGGGTTT TGTTCCGCCG GCTGACCACG CTGTGCATGT TGCACTGAT CTCCGCGTTT TATCAGGCCG ATAAAAGTTGC
+2 ALAAL THR LEU THR LEU VAL ASN VAL GLY ASP VAL ALA LEU PHE GLY LEU VAL ILE ASP ARG VAL SER THR LEU ILE LEU PHE VAL VAL VAL PHE LEU GLY
4501 CGTCACGTTG ACATTGGTCA ACGTGGGGGA TGTGGCATTG TTTGGCCTGG TCATTGATCG CGTGAGTACG CTGATTCTGT TTGTGGTGGT GTTTCTCGGT
+2 LEU LEU VAL THR ILE TYR SER THR GLY TYR LEU THR ASP LYS ASN ARG GLU HIS PRO HIS ASN GLY THR ASN ARG TYR TYR ALA PHE LEU LEU VAL PHE ILE
4601 TTGCTGGTCA CGATCTACTC CACGGGTTAT CTGACGGATA AAAATCGCGA ACACCCGCAT AACGGCACGA ATCGTTATTA CGCATTTTTG CTGGTGTTTA
+2 ILEGLY ALA MET ALA GLY LEU VAL LEU SER SER THR LEU LEU GLY GLN LEU LEU PHE PHE GLU ILE ARG ALA ALA ALA PRO GLY ARG TER
+1 LEU ILE SER TYR TYR
4701 TCGGCGCGAT GGGGGGACTG GACTCTCCT CGACGCTGCT CGGTCAGTTG TTGTTTTTGG AAATTAGGGC GGCTGCTCCT GGGCGTTGAT CAGTTATTAC
+1 GLN ASN ASP LYS ALA GLN ARG SER ALA LEU LYS ALA LEU LEU ILE THR HIS ILE GLY SER LEU GLY LEU TYR LEU ALA ALA ALA THR LEU PHE LEU GLNHR
4801 CAGAACGATA AAGCGCAGCG TTCAGCACTA AAAGCGTTGC TTATCACCCA TATCGGTTCC TTGGGGTTGT ATCTTGCCGC CGCCACGCTG TTTTTCAGAA
+1 THRGLY THR PHE ALA LEU SER ALA MET SER GLU LEU HIS GLY ASP ALA ARG TYR LEU VAL TYR GLY GLY ILE LEU PHE ALA ALA TRP GLY LYS SER ALA GLN
4901 CCGGAACGTT TCGCCTTAGC GCAATGAGCG AGTTACACGG CGACGCACGT TATCTGGTTT ATGGCGGGAT CCTGTTTGCC GCGTGGGGGA AATCGGCCCA
+1 GLN LEU PRO MET GLN ALA TRP LEU PRO ASP ALA MET GLU ALA PRO THR PRO ILE SER ALA TYR LEU HIS ALA ALA SER MET VAL LYS VAL GLY VAL TYR ILE
5001 GCTACCGATG CAAGCGTGGC TACCGGATGC AATGGAAGCG CCAACACCGA TCAGCGCCTA TCTCCACGCC GCATCGATGG TGAAGTGGG CGTTTACATT
+1 PHE ALA ARG ALA ILE ILE ASP GLY GLY ASN ILE PRO HIS VAL ILE GLY GLY VAL GLY MET VAL MET ALA LEU VAL THR ILE LEU TYR GLY PHE LEU METYR
5101 TTTGCCCGCG CAATTATCGA CCGCGGCAAT ATCCCGCATG TGATTGGCGG CGTTGGCATG GTCATGGCAC TGGTCACCAT TCTTTATGGC TTTCTGATGT
+1 TYR LEU PRO GLN GLN ASP MET LYS ARG LEU LEU ALA TRP SER THR ILE THR GLN LEU GLY TRP MET PHE PHE GLY LEU SER LEU SER ILE PHE GLY SERARG
5201 ATTTGCCACA GCAGGATATG AAGCGGTTGC TGGCCTGGTC GACCATCACT CAACTTGGCT GGATGTTCTT CGCCTTGTGC CTCTCCATCT TCGGCTCGCG
+1 ARG LEU ALA LEU GLU GLY SER ILE ALA TYR ILE VAL ASN HIS ALA PHE ALA LYS SER LEU PHE LEU VAL ALA GLY ALA LEU SER TYR SER CYS GLY THR
5301 GCTGGCGCTG GAGGGTAGCA TCGCCTACAT CGTCAACCAC GCGTTGCTA AAAGCCTGTT TTTCTTGTGA GCAGGTGCGC TGAGTTACAG CTGCGGCACG
+1 ARG LEU LEU PRO ARG LEU ARG GLY VAL LEU HIS THR LEU PRO LEU PRO GLY VAL GLY PHE CYS VAL ALA ALA LEU ALA ILE THR GLY VAL PRO PRO PHEASN
5401 CGCTTGTTC CGCGTCTGCG TGGCGTATTG CACACCCTGC CCGTTGCCAGG CGTGGGTTTC TGCGTGGCAG CGCTGGCGAT TACCGGCGTG CCGCCGTTCA
+1 ASNGLY PHE PHE SER LYS PHE PRO LEU PHE ALA ALA GLY PHE ALA LEU SER VAL GLU TYR TRP ILE LEU LEU PRO ALA MET ILE LEU LEU MET ILE GLUSER
5501 ACGGCTTCTT CAGTAAATTC CCGCTGTTTG CTGCCGTTT TGCGTTGTCA GTGGAGTACT GGATCCTGCT GCCCGCCATG ATTCTGCTGA TGATTGAATC
+1 SERVAL ALA SER PHE ALA TRP PHE ILE ARG TRP PHE GLY ARG VAL VAL PRO GLY LYS PRO SER GLU ALA VAL ALA ASP ALA ALA PRO LEU PRO GLY SER MET
5601 GGTCCGCACT TTCGCTGGT TTATTCGCTG GTTTGGTCCG GTTGTGCTCG GCAAACCGAG CGAGGCCGTC GCCGATGCCG CACCGCTGCC AGGGTCAATG
+1 ARG LEU VAL LEU ILE VAL LEU ILE VAL MET SER LEU ILE SER SER VAL ILE ALA ALA THR TRP LEU GLN TER
5701 CGCCTGGTGT TGATTGTA CTGTTGTGATG TGCGTGATTT CCAGCGTAAT CGCCGCGACC TGTTGCACT AAGGAGATGA TGAATGACCG GTTCTATGAT

```

**Figure 3.7. Translated sequences of the two orfs of *hyfD* in *E. coli* (55989).** The yellow highlight is the sequence of the first orf of *hyfD* where the start codon is could be (ATG) and the green highlight is for the second orf of *hyfD* and the start codon is predicted to be (TTG) not (ATG). The azure highlight shows where the overlapping occurs.



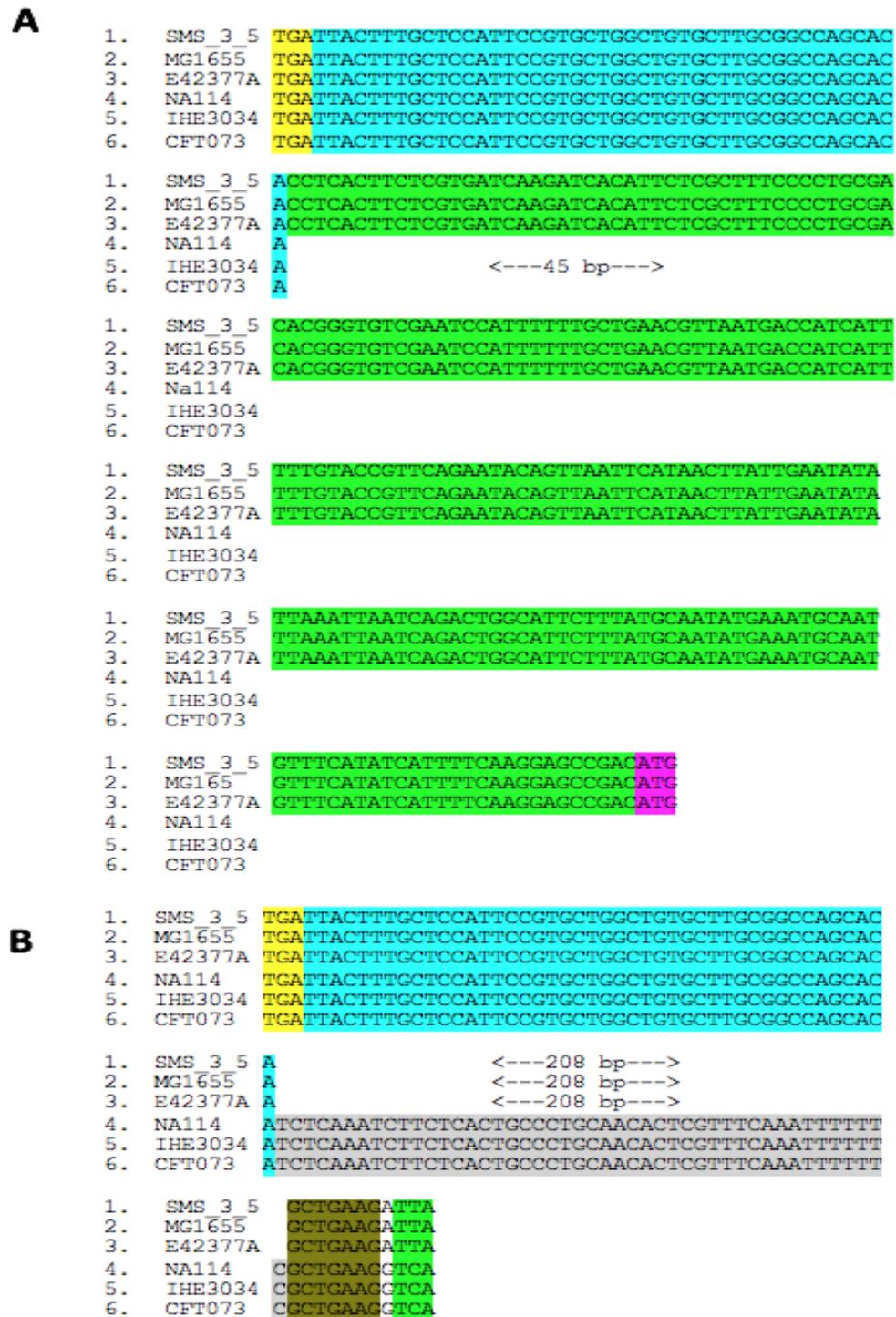
**Figure 3.8. Different *E. coli* (SE15) strain with no *hyf* operon.** Blue circle indicates the ‘missing’ region of the *hyf* operon. The orange arrows showing the direction of the genes and the position of these genes on both sides of the missed *hyf* operon. This map is identical to some other *E. coli* strains, NA114; CFT073; APEC-01 and ED1a.



**Figure 3.9. *E. coli* (MG1655) with full *hyf* region.** The orange arrows indicate the gene direction and its position among the rest within the operon. The upstream gene (*bcp*) and the downstream gene (*perM*) are shown clearly. bp: base pair. This map look similar to some other *E. coli* strains, 042 and 0157:H7 Sakai.

### 3.4. Detailed comparison of the *hyf* locus in *hyf*<sup>+</sup> and *hyf*<sup>-</sup> *E. coli* strains

To determine the precise differences in the genomes of the *hyf*<sup>+</sup> and *hyf*<sup>-</sup> strains, six *E. coli* strains were selected for genome comparison: three strains with *hyf* (SMS\_3\_5, E24377A and MG1655) and three strains without (NA114, CFT073 and IHE3034). These strains were used to produce alignments of the upstream (*bcp* and *hyfA*) and downstream (*focB-perM*) regions of the *hyf* operon. The alignments show that the *hyf*<sup>+</sup> strains have 208 bp of non-encoding DNA upstream of *hyf*, as well as 29 bp downstream of *focB*, that are absent in the *hyf*<sup>-</sup> strains. Also, there is a 44 bp non-encoding segment shared by all *E. coli* strains located downstream of *bcp* (Fig 3.10-A) as well as 7 bp of non-encoding DNA downstream of *perM* shared by all *E. coli* strains. Interestingly, there is a 45 bp region of DNA found only in the *hyf*<sup>-</sup> strains (Fig 3.10-B & Fig 3.11). BLAST analysis indicates that this sequence is unique among the 18 *E. coli* strains that are *hyf*<sup>-</sup>. On the other hand, the 208 bp segment is found (with 97-100% identity) in the 45 *E. coli* strains that possess *hyf* and are included in Fig 3.2. In addition, the 29 bp segment was also found in the same 45 *E. coli* strains which contain the *hyf* region. Figure 3.12 shows a schematic map of the *hyf* region in *hyf*<sup>+</sup> and *hyf*<sup>-</sup> strains which highlights the differences between these *E. coli* strains.



**Figure 3.10. Nucleotide sequence alignments of the upstream *hyf* region of three *hyf*<sup>+</sup> *E. coli* strains. (1. SMS\_3\_5, 2. MG1655, 3. E42377A) and three *hyf* (4. NA114, 5. IHE3034 and 6. CFT073) strains. **A)** Yellow highlights indicates stop codon of *bcp*, while azure highlighting shows 44 bp shared sequence within all strains. Green highlighting indicates 208 bp found only in *hyf*<sup>+</sup> strains and purple highlights show start codon of *hyfA*. **B)** Yellow highlights indicate stop codon of *bcp*, while azure highlighting shows 44 bp shared sequence within all strains. Gray highlight shows 45 bp of the unique sequence, found only in *hyf* strains, while khaki highlighting indicates a 7 bp shared sequence within all strains. The green highlight shows the stop codon of *perM*.**

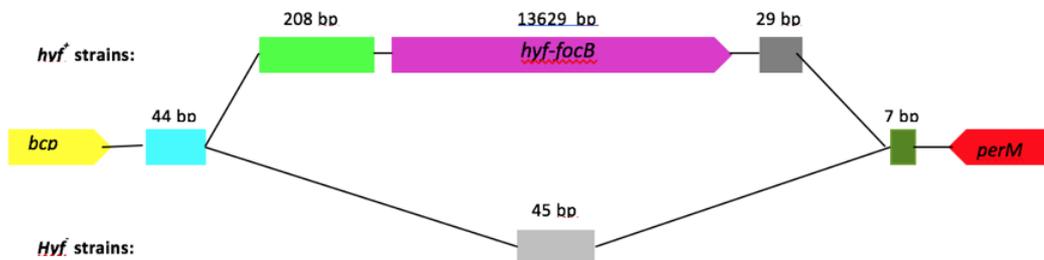
```

1. SMS_3_5 TCAGGAACCCTAAAAAATTCAGCCCGGCGAAACAGTTCGTCCG
2. MG1655 TCAGGAACCCTAAAAAATTCAGCCCGGCGAAACAGTTCGTCCG
3. E42377A TCAGGAACCCTAAAAAATTCAGCCCGGCGAAACAGTTCGTCCG
4. NA114
5. IHE3034 <---45 bp--->
6. CFT073

1. SMS_3_5 GCTGAAGATTATCA
2. MG1655 GCTGAAGATTATCA
3. E42377A GCTGAAGATTATCA
4. NA114 GCTGAAGTCA
5. IHE3034 GCTGAAGTCA
6. CFT073 GCTGAAGTCA

```

**Figure 3.11. Nucleotide sequence alignments of the downstream *hyf* region of three *hyf*<sup>+</sup> *E. coli* strains. (1. SMS\_3\_5, 2. MG1655, 3. E24377A) and three *hyf*<sup>-</sup> (4. NA114, 5. IHE3034 and 6. CFT073) strains. Yellow highlighting indicates the stop codon of *focB* gene, while dark gray indicates 29 bp downstream of *focB* that is absent in the *hyf*<sup>-</sup> strains. Azure highlight shows shared sequence within all strains and the red highlight indicates the stop codon of *perM*.**



**Figure 3.12. Schematic map showing the organisation of the *hyf* region in *hyf*<sup>+</sup> and *hyf*<sup>-</sup> *E. coli* strains. Unique regions for *hyf*<sup>+</sup> and *hyf*<sup>-</sup> strains are indicated in different colours corresponding to those used in the previous figure.**

### 3.5. Discussion and Conclusion

This study described in this chapter was conducted in order to provide further understanding of the purpose of the *hyf* operon of *E. coli*. *hyf* encodes a putative hydrogenase (designated hydrogenase 4, Hyf or HYD-4 from a 12 gene operon (*hyfABCDEFGHIIR-focB*). *focB* encodes a potential formate transporter that may supply Hyf with formate. Hyf may function like the homologous Hyc system in the breakdown of formate into hydrogen, but unlike Hyc may be energy conserving (Andrews *et al.*, 1997)

A bioinformatics comparison was performed to distinguish those *E. coli* strains that are *hyf* containing from those that are not. It was found that of the 94 *E. coli*; *Salmonella* and *Shigella* strains considered there were, 45 *E. coli* and 9 *Shigella* strains containing *hyf*, and 18 *E. coli* strains *hyf* free. Also, all 22 *Salmonella* species are *hyf* free, as is *E. albertii*, although *E. fergusonii* is *hyf*<sup>+</sup>. *E. albertii* is more distantly related to *E. coli* than is *E. fergusonii* which would be consistent with the conservation of the *hyf* system within *E. fergusonii* and most *E. coli* and all *Shigella* strains. However, all 94 *E. coli*, *Shigella* and *Salmonella* strains contain *hyc* indicating its universal importance for fermentative growth in the intestine.

A phylogenetic tree was constructed using the RpoB protein amino acid sequences of *E. coli*, *Shigella* and *Salmonella* strains (Case *et al.*, 2007; Mollet *et al.*, 1997). This tree clearly supports the view that the *Salmonella* genus is phylogenically distinct from that of the *Escherichia/Shigella* genus, as expected from previous analyses e.g. using *gyrB* genes (Fukushima *et al.*, 2002 ). Such previous results indicate that *S. sonnei* is more close to certain *E. coli* strains than to *S. boydii*, which matches the results provided in Fig. 3.2. The tree also suggests that *E. albertii* is more distantly related to *E. coli*, than is *E. fergusonii*, again as previously reported (Maheux *et al.*, 2014; Ooka *et al.*, 2015).

Thus, the phylogenetic tree shown in Fig 3.2 is generally consistent with previous reports. Importantly, the tree shows that all the *E. coli* strains without the *hyf* operon are clustered into one major clad. This clad corresponds to the B2 group, which in other phylogenetic analyses is the most deeply branched (the first phylogroup to diverge from the rest of the *E. coli* group) (Sims and Kim, 2011; Chakraborty *et al.*, 2015). Thus, the *hyf* minus phylogroup (B2) is considered as the most basal of the *E. coli* phylogroups (Sims & Kim, 2011). Interestingly, a high proportion of these *hyf*-free *E. coli* strains are associated with urinary tract infection whereas most of the *hyf*-carrying strains are faecal isolates. Indeed, the B2 group is recognised to be dominated by UPEC and APEC strains, with the UPEC strains being considered as opportunistic pathogens, and the group as a whole thought to be derived from a progenitor that was a facultative/opportunistic pathogen in nature (i.e. commensal or harmless, depending on the environmental niche occupied). This would suggest that the presence of *hyf* might be important for a more specialised lifestyle associated with colonisation of the intestine, indicating that *hyf* may not be important for *E. coli* survival in sites outside of the intestine, such as the urinary tract. Furthermore, in the intestine the ability to respire aerobically and anaerobically provides fitness and competition between *E. coli* strains (Jones *et al.*, 1997) which indicates that diverse respiratory (energy generating) capacity is important for gut colonisation of *E. coli* strains because of the highly competitive nature of the intestinal niche. Nucleotide-sequence alignment of the *bcp-perM* region in *E. coli* (corresponding to the locus where *hyf* is normally found) of *hyf*-free strains with the equivalent region of *hyf*-carrying strains showed that in addition to the *hyf-focB* coding region (~13,630 bp), the *hyf*<sup>+</sup> *E. coli* strains possess a unique 208 bp segment of DNA upstream of *hyf* as well as a unique 29 bp segment downstream of *focB*. Inspection of the sequences at the flanks of the *hyf*-specific region (the *hyf* genome island, GI) indicates that there are not any significant inverted

repeats associated with these regions that would be suggestive of integration by transposition. Also, there is 44 bp of non-encoding DNA downstream of *bcp* that is shared by all *E. coli* strains, as well as 7 bp of non-encoding DNA downstream of *perM*. In addition, there is a 45 bp segment of DNA found only in *hyf*-free strains in place of the *hyf* operon. This sequence was found to be unique among the 18 *E. coli* strains with no *hyf* region. On the other hand, the 208 and 29 bp regions are found only in *E. coli* strains containing the *hyf* region and can thus be considered to be *hyf* associated. The above suggests that the *hyf* region was gained during the evolution of the *E. coli/fergusonii* common ancestor, just prior to the split with the *E. albertii* group.

It should be noted that for *E. albertii*, the *perM-bcp* locus resembles that of the *hyf*-free *E. coli* strains. It is thus possible that the *hyf* operon arrived within the common ancestor of the *hyf*<sup>+</sup> strains by horizontal gene transfer, and that the *hyf-focB* region thus represents a GI that provides *E. coli* with an advantage during intestinal colonisation, but possibly not during UTI. It is likely that the clad in the *E. coli* cluster lacking *hyf-focB* suffered a subsequent deletion of this region in order to gain fitness for survival during UTI (or in another niche, possible non-intestinal, that is distinct from that occupied by the *hyf*<sup>+</sup> strains). However, as *hyc* is present universally in both the *Salmonella* and the *Escherichia* genus, it appears that the ability to dispose of formate is a common requirement for these organisms. However, the *hyf*-encoded Hyf system is hypothesised to be capable of participating in an energy-conserving formate hydrogen lyase reaction, which would be expected to provide an advantage under conditions where formate is available but respiratory electron acceptors are absent, i.e. during fermentation in the anaerobic conditions of the intestine. Thus, Hyf may offer the advantage over Hyc of serving to enable energy production from the disposal of formate. An important clue to the rationale for the presence or absence of Hyf in *E. coli* strains might be gleaned through

further understanding the unique features (e.g. niche preference, metabolic capacity) of the *hyf* group.

## Chapter 4: Analysis of *focB* function

### 4.1. Introduction

FocA belongs to the Formate-Nitrite Transporter (FNT) family. Members of this family are channels responsible for the transport of nitrite and formate, or structurally related compounds (Suppmann and Sawers, 1994). Under fermentative conditions, formate builds up outside of the *E. coli* cell, thus the concentration of formate increases and the pH level begins to drop as fermentative growth progresses. This triggers the FocA channel to act as an importer, instead of an exporter, allowing consumption of toxic formate by formate hydrogenlyase (FHL). Thus, FocA has the ability to allow selective movement of a formate in both directions across a cytoplasmic membrane depending on the environmental requirements (Wang *et al.*, 2009; Falke *et al.*, 2010). Thus, FocA can change its transport mode from a passive exporter channel at alkaline pH to a secondary active formate/H<sup>+</sup> symporter at lower pH (Rossmann *et al.*, 1991). The structural basis for this was shown in the crystal structure of the FocA of *Salmonella enterica* serovar Typhimurium, where a change in pH induced a major rearrangement of the residues in the C-terminal region of the individual promoters in a region that is well conserved (Lü *et al.*, 2011). The FocA structure of *E. coli*, reported by Wang *et al.* (2009) and the FocA structure of *V. cholerae* by Waight *et al.* (2010) showed similar homopentameric structures for FocA and aquaporin and glycerol porin, supporting its function as a channel, and showed that the first three helices are structurally superimposable on the second three helices, despite poor amino acid sequence similarity. Thus each protomer displays partial two fold symmetry. They also found that there is a short, conserved amino acid sequence (Y-L-R/K) located immediately after the sixth transmembrane helix projecting into the cytoplasm. Hunger *et al.* (2017) found that removal of R280 (in the YLR/K motif) prevented formate uptake via FocA in a strain with no PflB (pyruvate

formate lyase) but it did not prevent formate export via FocA. On the other hand, mutation of L279 strictly prevented formate export. It was concluded that the C-terminal YLR motif is extremely important for bidirectional transport of formate via FocA, as indicated above by the pH-triggered structural change in this region.

*focA* is located with the *pflB* genes (pyruvate formate lyase and Pfl-activating enzyme) forming the *focA-pflB* operon in *E. coli*. On the other hand, FocB is a putative formate channel homologous with FocA (50% amino acid sequence identity), encoded by the last gene in *hyf* operon (Saier, 1999). Supmann and Sawers (1994) found that in a *focA* mutant strain, the levels of formate accumulated under fermentative conditions are only ~50% of those seen for the wildtype so it was concluded that FocA is the major active formate channel in *E. coli*, although there might be another channel responsible for the residual formate accumulation observed in the *focA* mutant. It was further shown that FocA is specific for formate (and not other acidic fermentation end products) *in vivo* at least, and that formate is the only fermentation product that is later reimported by *E. coli* (Beyer *et al.*, 2013). Trchounian and Trchounian (2014) studied the effect of *focA* and *focB* mutation on the hydrogen production rate and on formate transport. They found that, if glucose is available as a carbon source, formate is produced and translocated via FocB at pH >7. They also found that deletion of both channels increases hydrogen production, presumably through intracellular retention of formate, with loss of *focB* resulting in a twofold reduction in hydrogen production. Furthermore, they discovered that during glycerol fermentation at pH 7.5, FocB is active in formate import, while FocA is effective in formate export. The studies by Trchounian and Trchounian (2014) thus indicate a role for FocB in formate translocation, although they contradict other studies indicating that the *hyf* operon is transcriptionally silent (Skibinski *et al.*, 2002; Self *et al.*, 2004).

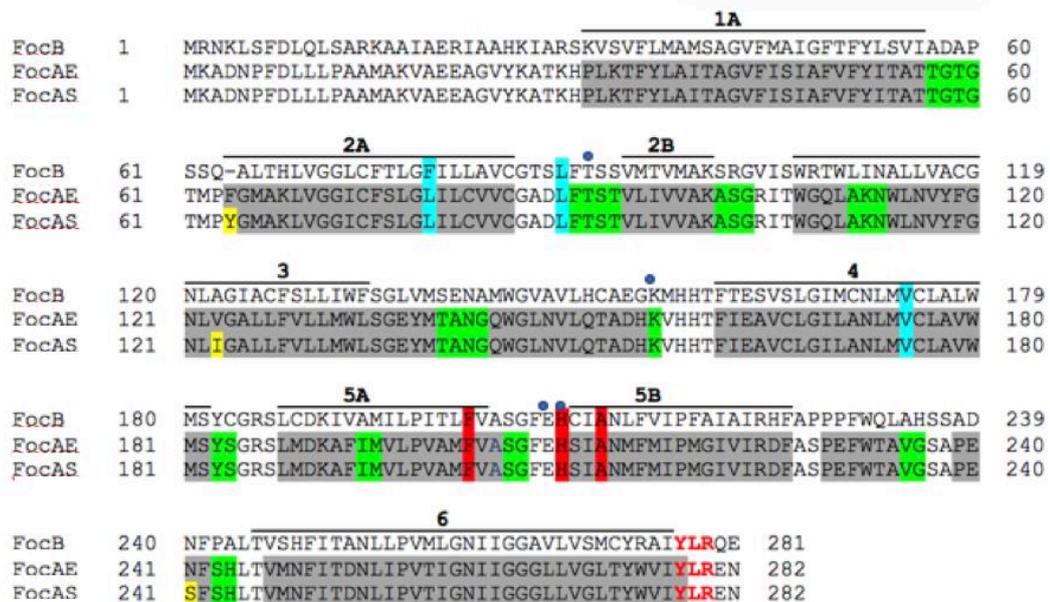
In this chapter, the potential for FocB to act as a formate channel in place of FocA is investigated. In addition, the impact of pH on FocB activity is considered, and compared to the effect that pH has on FocA activity.

## 4.2. Comparison between FocA and FocB amino acid sequences

Initially, a simple bioinformatic analysis was performed to compare the topologies of FocA and FocB, and to determine the degree to which residues shown to be important for FocA function are conserved in FocB. The amino acid sequence alignment shows that FocA and FocB can be aligned optimally with just one padding character required, and thus they have almost fully congruous sequences (Fig. 4.2). The padding character, in the FocB sequence, is inserted directly at the beginning of helix 2A adjacent to a loop region of random coil secondary structure; a deletion at this point is unlikely to influence the fold of the protein. The alignment also indicates that the secondary structure elements observed in the *E. coli* (and *Salmonella*) FocA are fully conserved in FocB, and that FocB is thus likely to assume a similar structure to that of FocA (Andrews *et al.*, 1997; Beyer *et al.*, 2013). Furthermore, the alignment shows that the pore amino acid residues at the periplasmic (F75, A212, F202) and cytoplasmic constriction (L79, L89 and V175) points, as found in the *E. coli* (and *Salmonella*) FocA, are conserved also in FocB (except L79F), a conservative change. This similarity therefore predicts a similar role for these residues in FocB (Lü *et al.*, 2012).

The 'YLR' motif at the C-terminus of FocA (*E. coli*/ *Salmonella*) is absolutely conserved in FocB, suggesting, that like FocA, FocB formate transport activity is regulated in response to pH. Importantly, the residues of FocA identified as formate ligands (Fig. 4.1) are absolutely conserved in FocB, which is consistent a similar role for FocB in formate transport as FocA. In addition, residues that constrict the channel are highly conserved

also. Thus, the comparison of the FocA and FocB sequences supports a role for FocB as a pH responsive channel. However, these proteins only exhibit 50% sequence identity indicating that they may also display some differences in their biochemical activities.



**Figure 4.1. Sequence alignment of FocA from *E. coli* (FocAE) and *Salmonella* (FocAS) with FocB from *E. coli*.** Gray highlight indicate the six membrane helices. Green highlight indicates turns in the transmembrane helices. Yellow indicates the amino acids found only with *Salmonella* FocA. Four blue small dots indicate conserved ligands of formate binding site (E, H, K and T). Red bold amino acids indicate the conserved sequence (Y-L-R) at the C-terminal. Red and cyan highlights indicate the conserved pore constriction residues at the cytoplasmic and periplasmic surfaces, respectively.

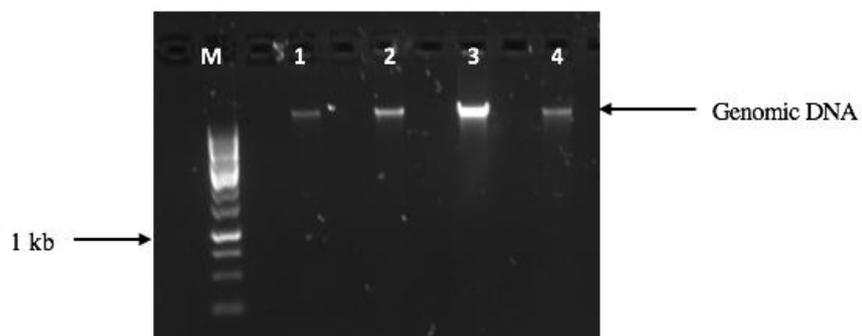
### 4.3. Cloning of *focA* and *focB* into inducible vectors

In order to enable controlled expression of *focA* and *focB*, the open-reading frames for each gene were cloned into the pBAD<sub>rha</sub> vector, which allows controlled expression in response to rhamnose. In this way, both *focA* and *focB* function could be assessed under any desired environmental condition without influence of the regulatory factors that would normally control their expression. In particular, it would be possible to ensure that

*focB* is expressed since previous studies have indicated that the *hyf-focB* operon is transcriptionally silent in the wildtype (Skibinski *et al.*, 2002; Self *et al.*, 2004).

### 4.3.1 Genomic DNA extraction of wildtype (MG1655)

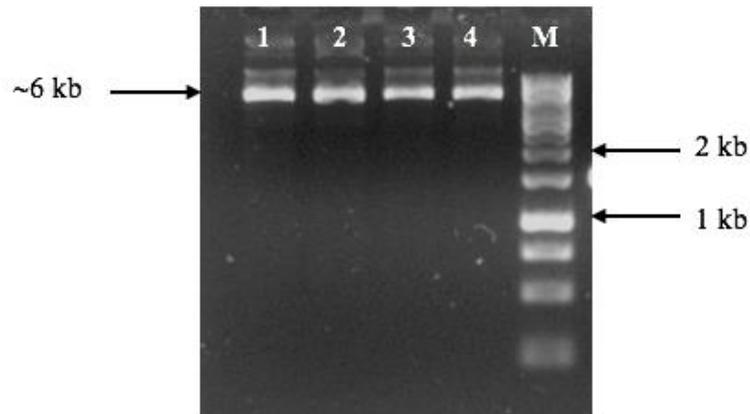
In order to obtain *focA* and *focB* genes, genomic DNA was extracted from the wildtype *E. coli* K-12 strain, MG1655, using a GeneJET™ genomic DNA purification kit (Fermentas) (section 2.10.1). The presence of isolated DNA was confirmed by agarose gel electrophoresis (section 2.10.8.1) (Fig 4.2).



**Figure 4.2.** Agarose gel (0.7%) electrophoretic analysis of genomic DNA from *E. coli* MG1655. Lane 1-4, genomic DNA samples. M, 1 kb ladder (Fermentas). Loadings were 1  $\mu$ l for ladder and 2  $\mu$ l for DNA sample.

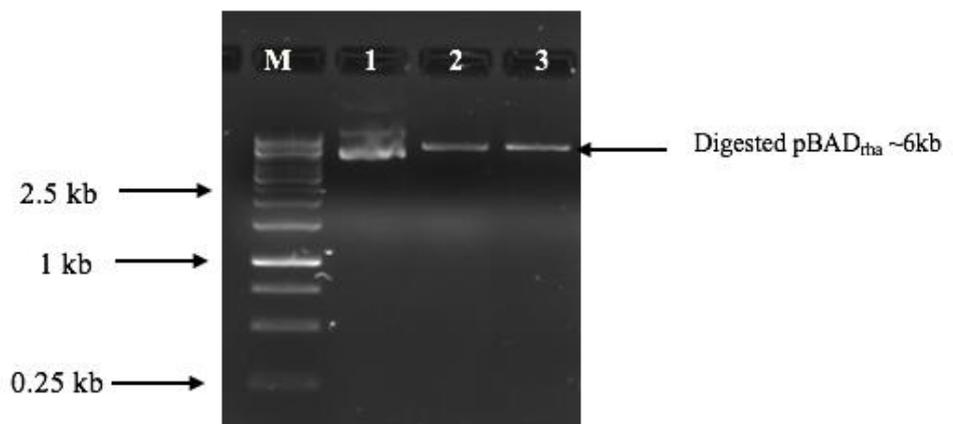
### 4.3.2 Isolation and digestion of pBAD<sub>rha</sub>

The plasmid pBAD<sub>rha</sub> was obtained from laboratory stocks. pBAD<sub>rha</sub> transformants were then generated and plasmid DNA was isolated from four samples using a GeneJET™ plasmid Miniprep Kit (Fermentas) (section 2.10.7). The isolated plasmids were screened by electrophoretic analysis and the presence of plasmid DNA of high mass corresponding to the expected size (6.1 kb; Fig. 4.3) was confirmed.



**Figure 4.3 Agarose gel (0.7%) electrophoretic analysis of undigested pBAD<sub>rha</sub>.** M, Marker 1 kb Generuler (Fermentas); lane 1-4, 2  $\mu$ l pBAD<sub>rha</sub> Plasmid. Arrow indicates position of plasmid DNA (~ 6 kb).

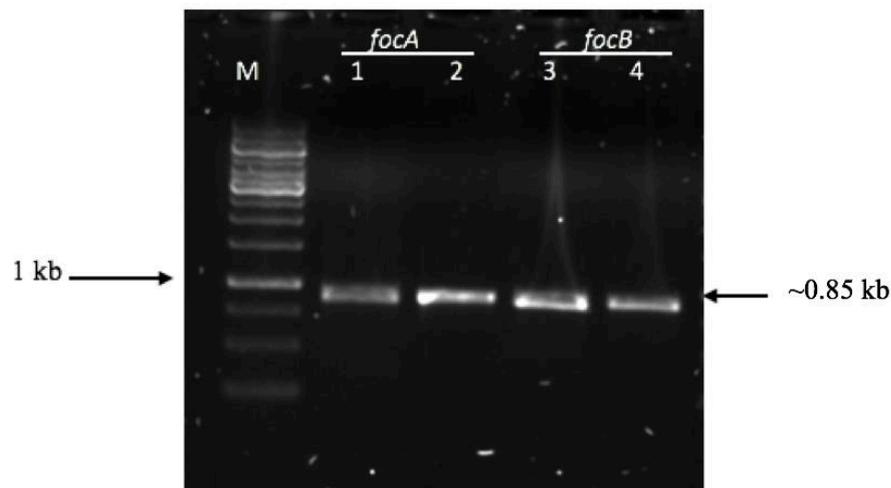
The identity of the plasmid DNA thus isolated was confirmed by digestion with *Nde*I and *Bam*HI followed by electrophoresis. The double digestion converted the plasmid from the supercoiled form to a single linear form of mobility matching that expected for pBAD<sub>rha</sub>. (Figure 4.4).



**Figure 4.4. Agarose gel (0.7%) electrophoretic analysis of pBAD<sub>rha</sub> following restriction digestion with *Nde*I and *Bam*HI.** M, 1 kb ladder (Fermentas). Lane 1, pBAD<sub>rha</sub> undigested; lane 2 & 3, pBAD<sub>rha</sub> double digested.

### 4.3.3. PCR amplification of *focA* and *focB*

Genomic DNA of *focA* and *focB* was amplified by PCR (section 2.10.2; Fig. 4.6) using specific primers pBAD<sub>rha</sub>-*focA* and pBAD<sub>rha</sub>-*focB* (forward and reverse) (Table 2.4). These primers were designed for use in In-Fusion cloning (Gibson methodology) and as described in sections 2.12.2.1 and 2.12.2.2. The Amp HiFi PCR premix DNA polymerase (Clontech) was used as it provides accurate and efficient DNA amplification and is recommended for In-fusion PCR cloning. The amplified fragments were analysed by agarose gel electrophoresis and the sizes expected were 857 and 848 bp for *focA* and *focB*, respectively, which are a close match to those observed (Fig 4.5).

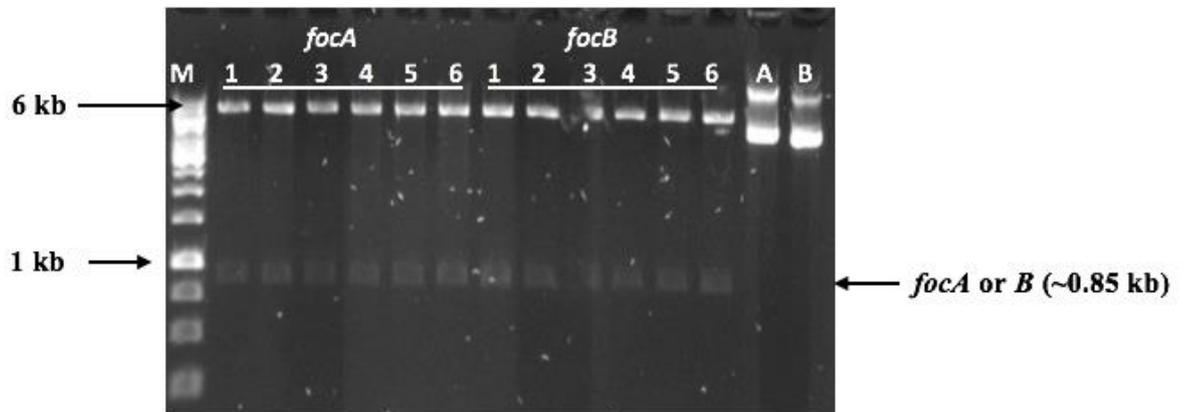


**Figure 4.5. Gel electrophoretic analysis of the *focA* and *focB* PCR products.** Electrophoresis was performed using a 0.7% agarose gel. Lane M, the GeneRuler 1 kb ladder (Fermentas); lane 1-2, *focA* PCR product; and lane 3-4, the *focB* PCR product.

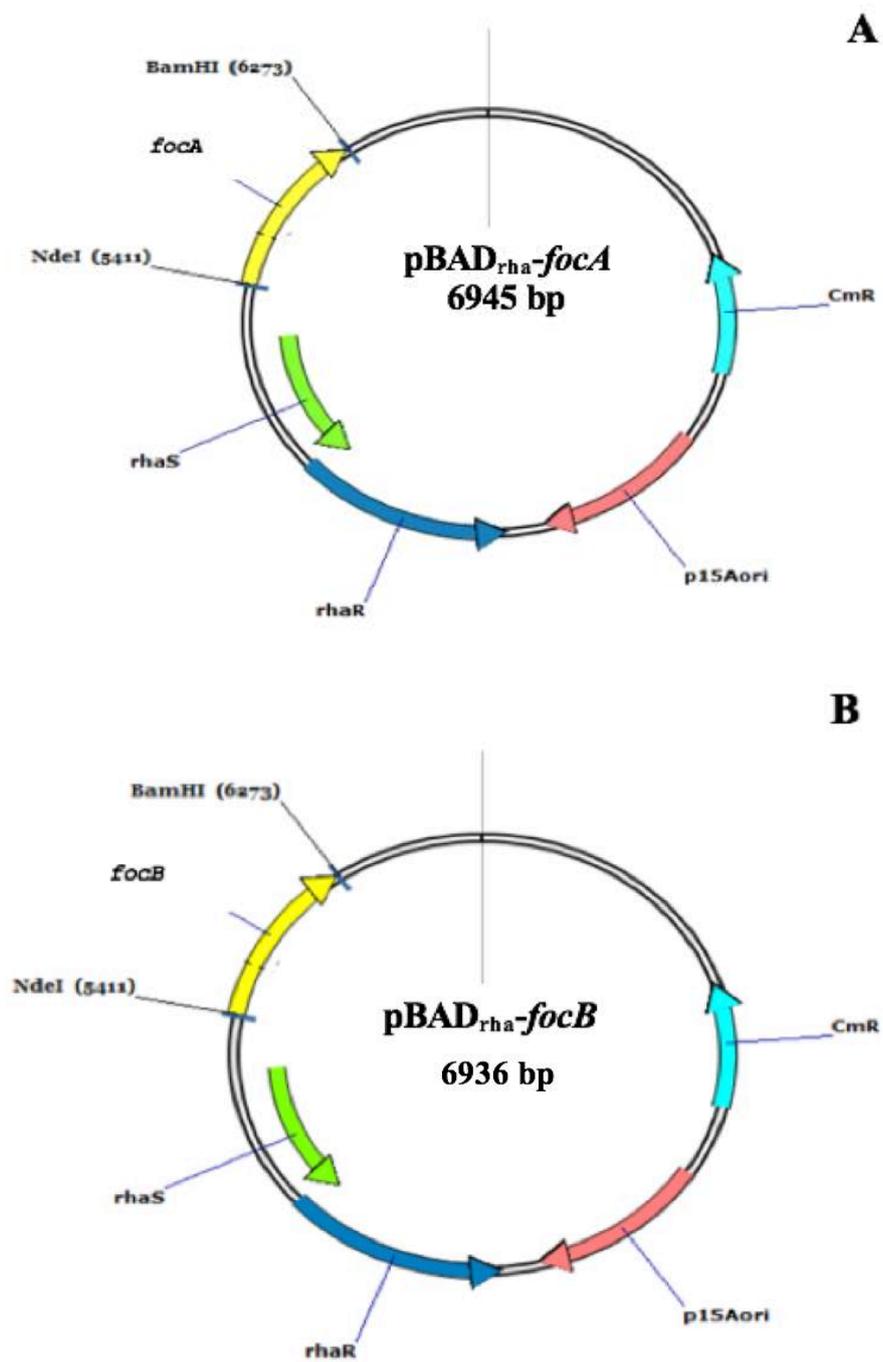
### 4.3.4. Cloning of *focA* and *focB* into the vector, pBAD<sub>rha</sub>

To enable controlled induction of *focA* and *focB*, the *focA* and *focB* PCR products were cloned into pBAD<sub>rha</sub>, at the *NdeI* and *BamHI* cloning sites of the vector, using Gibson cloning methodology (Gibson *et al.*, 2009; section 2.12.2.3.). The resulting reaction products were used to transform chemically competent *E. coli* Stellar cells (section

2.12.2.4). Twelve resulting Cm<sup>R</sup> colonies were selected for plasmid ‘miniprep’ isolation using GeneJET™ Plasmid Miniprep Kit (Fermentas) (section 2.10.7), six of each type. These plasmids were then further analysed by restriction digestion (section 2.10.4; Fig. 4.6) with *NdeI* and *BamHI* to release the insert from the vector. Following analysis by agarose gel electrophoresis, all 12 plasmids were shown to carry an insert of the expected size (~850 bp) and were subsequently designated pBAD<sub>rha</sub>-*focA* and pBAD<sub>rha</sub>-*focB* (Fig. 4.7).



**Figure 4.6. Gel electrophoric analysis of pBAD<sub>rha</sub>-*focA* or *focB*.** Electrophoresis was performed using a 0.7% agarose gel. Constructed plasmid was digested with *NdeI* and *BamHI*. Expected sizes; A is undigested pBAD<sub>rha</sub>-*focA* and B is undigested pBAD<sub>rha</sub>-*focB*; 1-6 digested plasmid DNA (~6 kb and 0.85 kb). M, 1 kb DNA ladder (Fermentas).



**Figure 4.7. Restriction map of A: pBAD-*focA* and B: pBAD-*focB*.** The plasmid contains *focA* or *focB* cloned into *NdeI* and *BamHI* sites of pBAD<sub>rha</sub>. The origin (p15Aori); Cm<sup>R</sup> gene, *rhaS* and *rhaR* are genes (Egan and Schleif, 1993) are shown.

Two of each of the pBAD<sub>rha</sub>-*foc* plasmids (sample 2 and 5 in each case) were subjected to nucleotide sequencing by Source Bioscience using pBAD<sub>rha</sub>-F and pBAD<sub>rha</sub>-R primers (Table 2.4). The sequences obtained were compared with the sequence database using BLAST which confirmed that the inserts have the correct sequence correctly located at the desired cloning sites (Appendix 4 &5). These plasmids were employed in future studies, as described below.

#### **4.4. Complementation of the *focA* mutant phenotype by *focB***

##### **4.4.1. Hypophosphite anaerobic-growth inhibition is FocA dependent**

The availability of *focA* and *focB* in an inducible vector allowed progression of experiments aimed at complementation of a *focA* mutant. A well-established phenotype for *focA* mutants of *E. coli* is resistance to the anaerobic growth inhibition caused by the formate analogue, hypophosphite (HP) (Suppmann and Sawers, 1994). Thus, the ability of the pBAD<sub>rha</sub> constructs to reverse this resistance was tested. Initially, various HP concentrations were tested on the *E. coli* parental strain (BW25113) and the isogenic mutant strains (note, both the *focA* and *focB* mutants used throughout this thesis were deletions with the *kan* cassette removed, thus limiting any polarity effect on downstream expression; Table 2.2). Suppmann and Sawers (1994) used 5-80 mM HP under anaerobic condition and found that as the concentration of HP increased, growth of the wildtype was decreased by more than 50%. However, in *focA* mutant strains there was very little effect of HP on bacterial growth, except at higher concentration (80 mM). In addition, HP did not inhibit aerobic growth, presumably because the gene encoding the PFL enzyme that is inhibited by HP (and which is the direct cause of HP growth inhibition) is not induced or required (like *focA*) aerobically. These findings thus show that FocA mediates HP uptake that results in growth inhibition and demonstrate that inhibition by

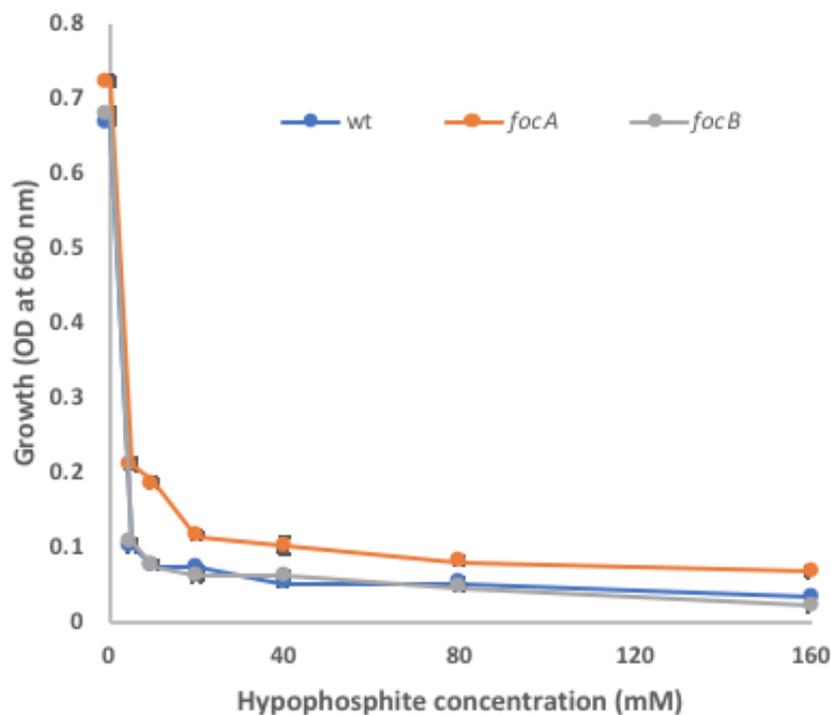
HP can be used to monitor FocA transport activity, suggesting that such an approach may be applicable for FocB.

To investigate the effect of HP on FocA and FocB in comparison to the wildtype under anaerobic conditions, overnight cultures of the *focA* and *focB* mutant strains, and a wildtype strain were cultivated in WM-medium (section 2.5.2) with 0.8% trace elements and 120 mM MOPS. Next day, the strains were tested with a range of HP concentrations (0, 5, 10, 20, 40, 80 and 160 mM) following inoculation to a final OD<sub>600</sub> of 0.01 in fresh WM-medium (80 mM glucose) at starting pH of 6.5. The strains were grown in a series of 7 ml bijoux tubes (for 48 hours at 37 °C) sealed with lids carrying a rubber insert to ensure no oxygen can enter the tubes and that an anaerobic environment is maintained (once established during the early stages of growth). All growths were performed in triplicate, and were repeated twice with similar results obtained. The results shown are representative of one repeat.

The results showed that all three strains are inhibited at all concentrations of HP (HP) employed. At just 5 mM HP, the growth of the wildtype was reduced from 0.7 to 0.1 OD units, and growth continued to decline further as HP levels increased to the maximum employed (160 mM). However, although the *focA* mutant was also HP sensitive, it was less affected by HP than the wildtype (Fig 4.10). At 5 mM HP, its growth declined from 0.7 to 0.2 OD units, and so its growth was 2 fold greater than that of the wildtype at this HP level. The greatest difference between the wildtype and *focA* mutant was seen at 10 mM HP (2.5 fold) ( $P = 0.04$ ). This increased resistance to HP for the *focA* strain was observed at all concentrations employed, with a 2-2.5 fold higher growth seen with respect to the wildtype at 5-160 mM HP (Fig. 4.8). As expected, the *focB* mutant displayed the same level of growth inhibition in response to HP as the wildtype ( $P = 0.91$ ). Thus, absence of *focB* did not increase HP resistance. This effect is presumed to

be caused by the absence of *hyf-focB* expression, as previously reported (Skibinski *et al.*, 2002; Self *et al.*, 2004.) However, this contradicts the work of Trchounian and Trchounian, (2013), which indicated that FocB is biologically active at pH 7.5 with glucose.

Thus, the above experiment confirms the role of FocA in HP toxicity and its likely role as a HP transporter (Suppmann and Sawers, 1994). However, no such evidence was obtained for FocB, which is assumed to relate to its lack of expression.

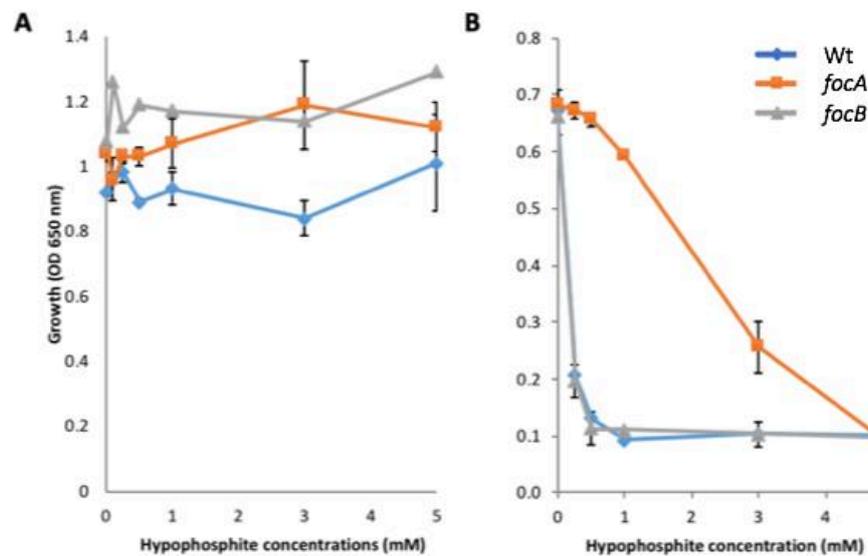


**Figure 4.8. Effect of hypophosphite on anaerobic growth and the impact of *focA/focB* status.** The wildtype (BW25113), and the *focA* (JW0887-1) and *focB* mutant strains (JW2477-1) were grown anaerobically in WM-medium for 48 h at 37 °C in triplicate. Data is the average of three technical repeats with error bars indicated standard deviation; the experiment was repeated twice with similar results obtained

The HP inhibition test was repeated as above but with a lower HP concentration range (0-5 mM) to test the effect of lower HP levels. The wildtype and *focB* mutant displayed high sensitivity to low HP levels with 70% growth inhibition observed at just 0.25 mM HP, and 85% inhibition at just 0.5 mM HP (Fig. 4.11B). However, as before, the *focA*

mutant strain was much less sensitive to HP, showing only a 1.4 and 3% reduction in growth at 0.25 and 0.5 mM HP, respectively. This correlates to a significant 3-fold increase in resistance for the *focA* strain with respect to the wildtype, at 0.25 mM HP ( $P = 0.02$ ) (Fig. 4.11). However, as the concentration of HP increased the difference in growth between the *focA*<sup>+</sup> and *focA*<sup>-</sup> strains diminished due to the increasing growth inhibition of the *focA* strain. Note that the results shown in Figs 4.8 and 4.11 are mostly consistent, but the data in Fig. 4.9 indicates no growth difference between the *focA*<sup>+</sup> and *focA*<sup>-</sup> strains at 5 mM HP, whereas in Fig. 4.8 a clear increase in growth is shown for the *focA* mutant at this concentration. The *focB* mutant showed no notable difference in HP sensitivity with respect to the wildtype at any concentration employed (Fig. 4.9).

The inhibition test was also performed under aerobic conditions. The results showed no notable impact of HP on the growth of any of the three strains tested; this is as expected because the *focA-pflB* operon is expressed (and required) only under anaerobic conditions (Sawers and Böck, 1989) and therefore not only would Pfl inhibition have little impact aerobically, the enzyme would be expected to be absent with pyruvate dehydrogenase acting in its place (Fig 4.9A).



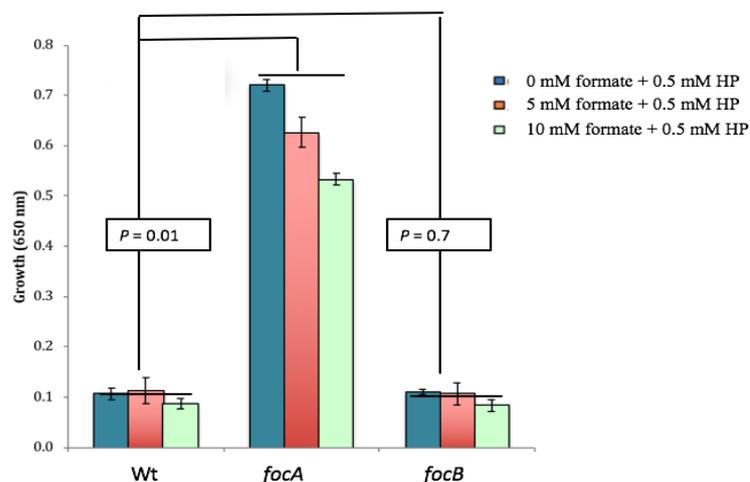
**Figure 4.9. Effect of hypophosphite under aerobic and anaerobic growth. A)** Aerobic conditions, growth in 6 inch test tubes with 3 ml medium at 250 rpm, for 24 h. **B)** Anaerobic growth, 48 h. Other details are as for Fig. 4.10.

#### 4.4.2. Can formate protect against hypophosphite toxicity?

There is a possibility that formate might reverse the inhibition of growth caused by HP, by acting as a competitor for transport of HP into the cell via FocA or by displacing HP from the Pfl enzyme. To test this possibility, formate was added to WM-medium (section 2.5.2) with 0.8% trace elements and 120 mM MOPS under anaerobic conditions with 0.5 mM of HP, at pH 6.5. Syringes of 10 ml with needles inserted into rubber stoppers were used to provide the anaerobic condition for 48 h. Fig 4.10 shows that addition of formate to the wildtype and *focB* mutant does not reverse sensitivity to HP and thus does not protect the cell from HP toxicity ( $P = 0.7$ ). The *focA* mutant remained highly resistant to HP compared to the wildtype when formate was added, with a 5.6-6.1 fold higher growth exhibited ( $P = 0.01$ ); pH was measured before and after growth with little difference observed (Table 4.1).

The above observation raises the question of the impact of pH on FocA-dependent sensitivity to HP, since at pH around neutral, or above, FocA acts as in passive facilitator-diffusion mode for formate export, thus FocA may not be very effective in promoting substrate uptake at higher pH. However, under acid conditions FocA switches to a proton motive force (PMF) driven importer of formate (Lv *et al.*, 2013), and so might thus be expected to promote higher HP sensitivity at low pH, and weaker sensitivity at high pH. Since the pKa of hypophosphite is approx. 1, which is much lower than the pKa of formate (3.7), it is suggested that hypophosphite will compete with formate for import into the cell where it causes toxicity. Therefore, the HP inhibition experiment was repeated at three different pH values (pH 5, 6 and 8) to test the influence of external pH on HP sensitivity (Fig. 4.11). 120 mM MOPS was added to WM-medium to maintain the desired pH, and pH was monitored before and after growth with little difference observed (Table 4.2). The results show that at pH 5 growth was severely inhibited in both the wildtype and *focA* mutant, so the effect of HP could not be determined. At pH 6, HP caused a major inhibition (2.5 and 3.5 fold;  $p = 0.15$ ) of growth for the wildtype at both 0.5 and 3 mM, respectively; formate failed to reverse this effect when added at up to 3 mM. It should be noted that the effect of formate on anaerobic growth in WM medium was consistently negative, with increasing growth reduction seen as formate levels were raised (Fig. 4.10 & 4.11). Also, under all conditions investigated, formate failed to provide any major protection against HP toxicity. However, unlike at pH 6, at pH 8 the wildtype showed only a modest reduction in growth (17%), from OD 1.15 to 0.95 ( $P = 0.05$ ), with addition of 0.5 mM HP (Fig 4.11A). Thus, FocA-dependent HP toxicity appears far less severe at high pH than at low pH with 0.5 mM HP. At higher HP levels (3 mM), the wildtype was fully growth inhibited at pH 8, as was the case at pH 6 (Fig. 4.10A). Thus, HP remains toxic at higher pH, but higher HP concentrations are

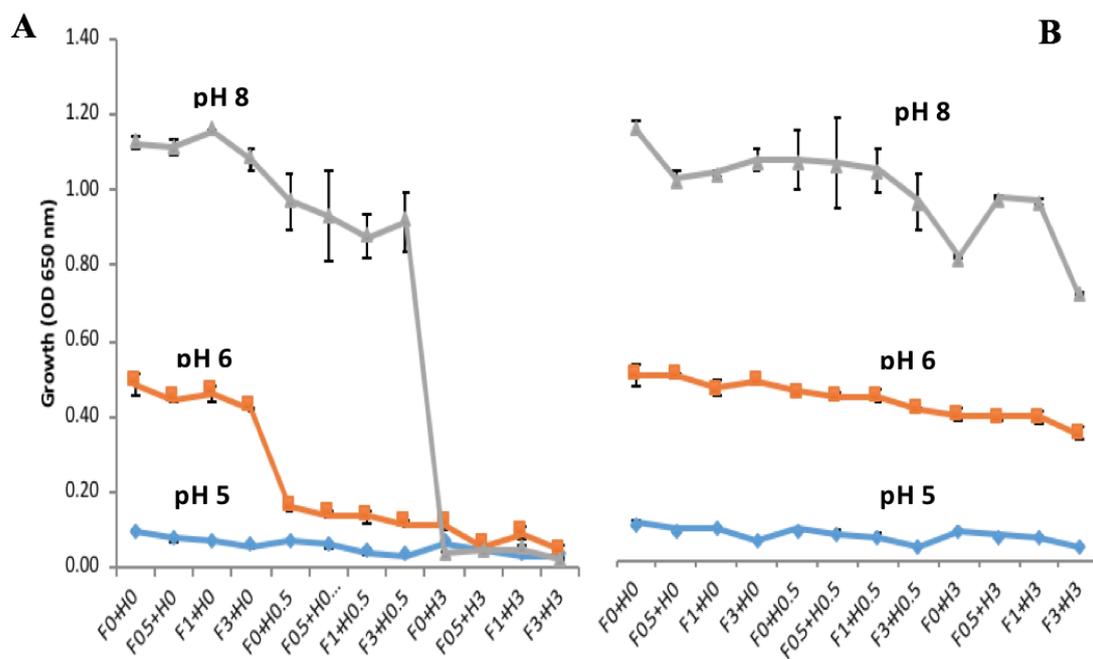
required to achieve this toxicity. On the other hand, the *focA* mutant remained highly resistance to HP at both pH 6 and 8 (Fig 4.11B). Thus, the results indicate that FocA acts as a channel for HP entry into the cell at both low and high pH, but that higher concentrations of HP are required at higher pH in order for toxicity to be exhibited. This suggests that FocA HP-transport activity is, at best, only weakly pH regulated. It should be noted that it is possible that pH might influence *pflB* and/or *focA* expression which could in turn impact HP toxicity in response to pH independently of FocA transport status. Further, the weaker growth of *E. coli* at pH 6 cf. pH 8 might also influence the capacity of the cell to cope with HP toxicity. Thus in summary, the relatively impact of pH on the concentration-dependence of HP toxicity, as mediated by FocA, provides little support for a pH dependent switch for FocA from an importer at low pH, to an exporter at high pH.



**Figure 4.10. The effect of formate on inhibition by 0.5 mM hypophosphite at pH 6.5.** Strains were grown anaerobically in WM medium as above, with 0.5 mM HP and 0-10 mM formate, in triplicate. The experiment was performed three times, with similar results obtained. Data is the average and standard deviation of three replicates.

Table 4.1. pH values before and after growth related to Fig. 4.12

strain	Starting pH	pH after 48 h
Wildtype	6.5	6.2
<i>focA</i>		6.3
<i>focB</i>		6.25



**Figure 4.11. The effect of pH on HP growth inhibition under anaerobic fermentative conditions.** The wildtype (A) and *focA* mutant (B) were propagated for 48 h anaerobically in WM medium, pH 6.5 as starting pH and OD was recorded at the end of growth. F, formate (0-3 mM); H, hypophosphite (0-3 mM). Data is the average of three replicates, and the experiment was performed three times with similar results obtained.

**Table 4.2. pH values difference before and after growth related to fig. 4.13.**

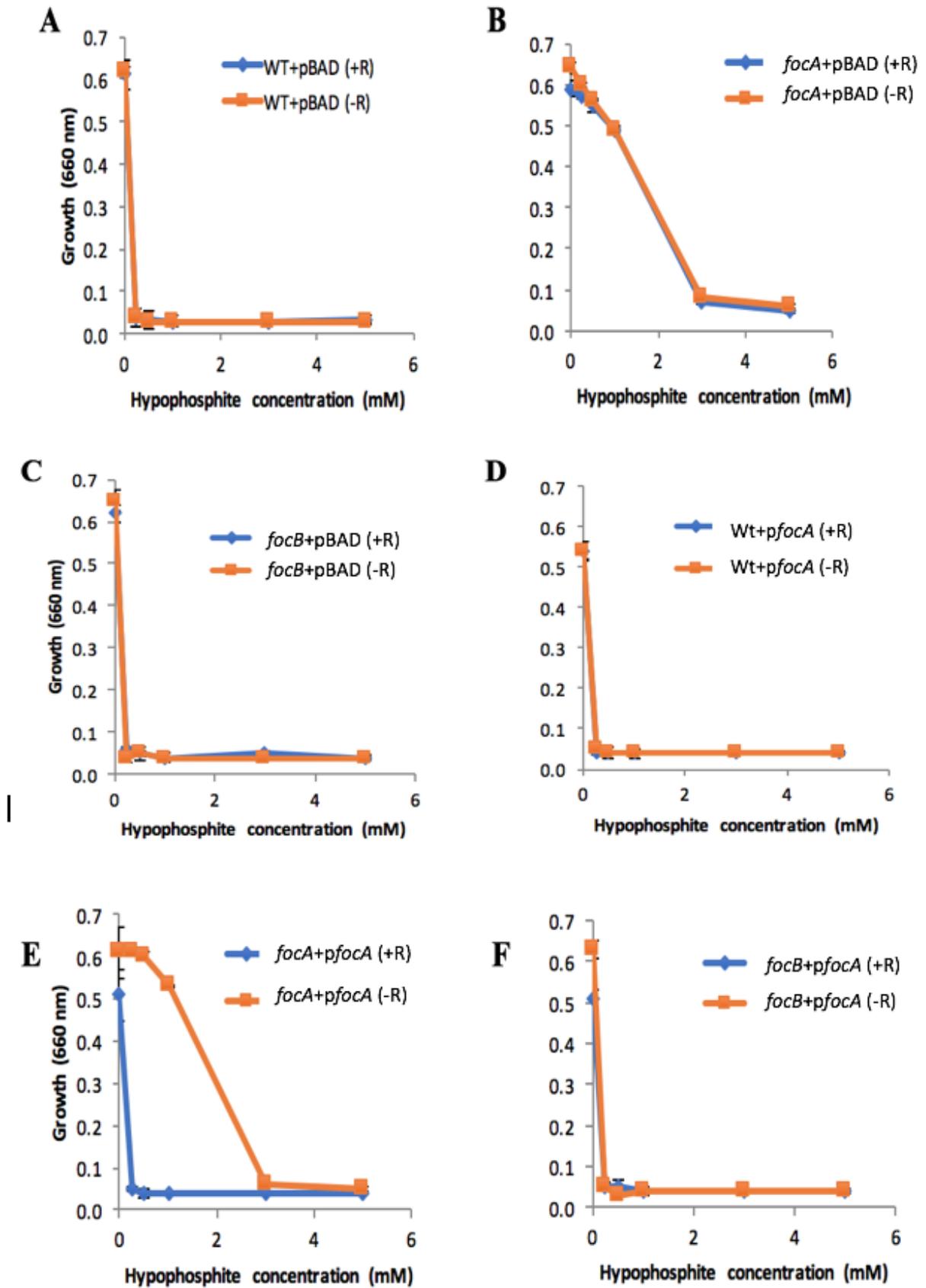
strain	Starting pH	pH after 48 h
Wildtype	5	4.8
<i>focA</i>		4.7
Wildtype	6	6.3
<i>focA</i>		6.3
Wildtype	8	7.8
<i>focA</i>		7.7

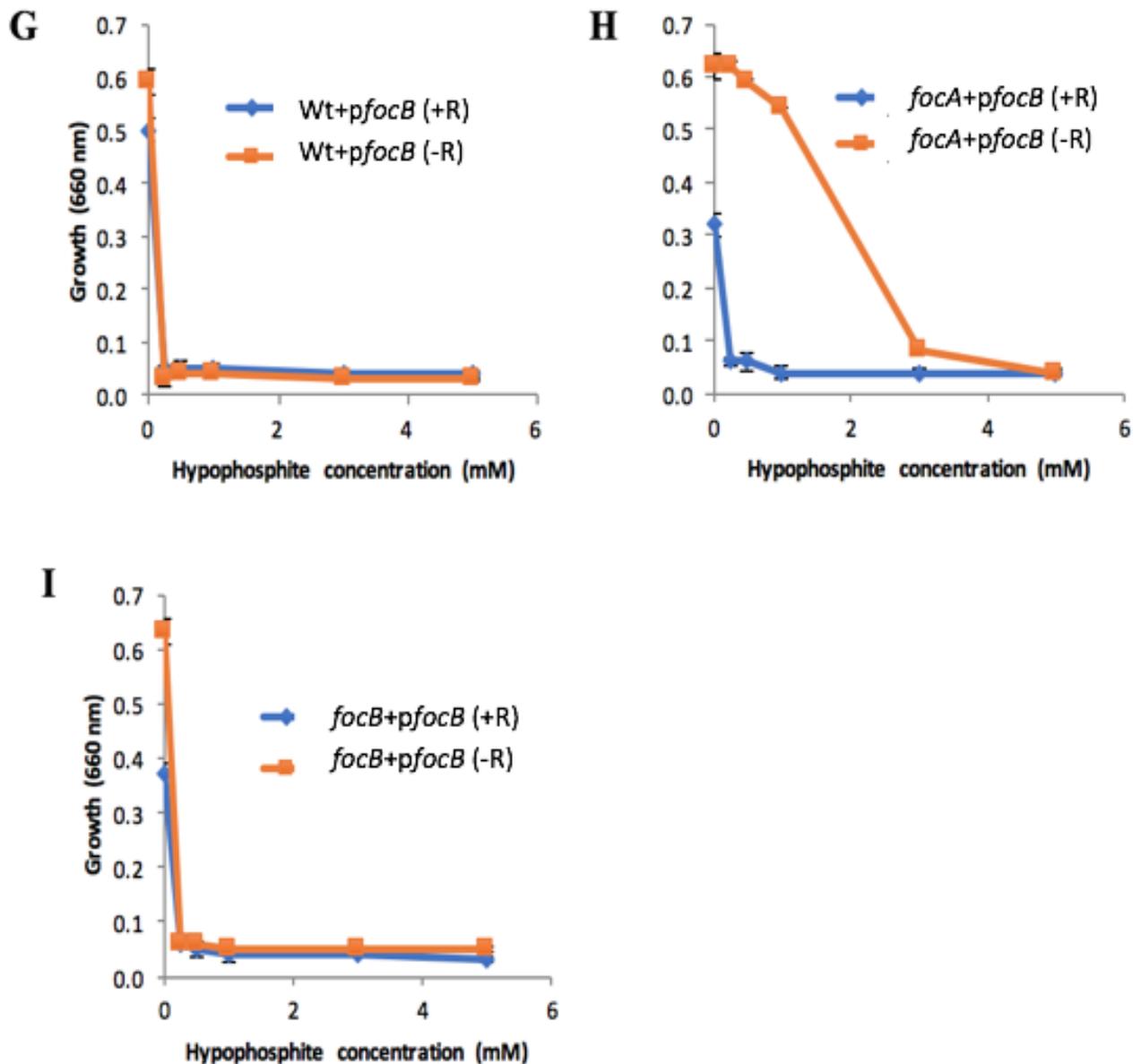
#### 4.4.3. Can the resistance of the *focA* mutant to hypophosphite be reversed by complementation with *focA* and/or *focB*?

To explore the role of FocB in formate transport and to compare it with that of FocA, the pBAD<sub>rha</sub> vectors (Appendix-2) expressing *focB* and *focA* (along with the vector control) were used to transform the wildtype, and the *focA* and *focB* mutants. The resulting pBAD<sub>rha</sub>-*focA*, pBAD<sub>rha</sub>-*focB* and pBAD<sub>rha</sub> transformants were grown anaerobically, as above, in WM-medium (80 mM glucose; 0.8% trace elements and 120 mM MOPS) with a range of HP levels and with chloramphenicol (50 µg/ml), with and without 0.02% w/v rhamnose as inducer.

Results confirmed (Fig 4.12A-C) that the *focA* mutant is highly resistance to HP compared to the wildtype and *focB* mutant. They show that the vector alone (pBAD<sub>rha</sub>) has little impact on HP sensitivity, either with or without rhamnose. However, the pBAD<sub>rha</sub>-*focA* plasmid restored sensitivity of the *focA* mutant to HP when rhamnose was included, but not in its absence (Fig 4.12D-F). There was a significant 12.6 fold difference ( $P = 0.02$ ) between the growth sensitivity of the *focA*-complemented *focA* mutant to 0.5 mM HP with and without rhamnose. The pBAD<sub>rha</sub>-*focA* plasmid had no notable effect in the other strains since they are already sensitive to HP due to their

chromosomal *focA*<sup>+</sup> status. Importantly, the pBAD<sub>rha</sub>-*focB* plasmid also restored HP sensitivity to the *focA* mutant when rhamnose was provided, with a significant 10.5 fold increase in growth observed at 0.5 mM HP without rhamnose ( $P = 0.03$ ) (Fig 4.12 G-I). Thus, the results indicate that FocB acts as a HP transporter and is therefore likely to function as a second formate transporter in *E. coli*. This provides support for the role of the Hyf system in formate metabolism as part of a second FHL complex.





**Figure 4.12. The effect of *focA* and *focB* complementation of hypophosphite sensitivity.** Growths were as described in Fig. 4.10 in WM medium (80 mM glucose; 8% trace elements and 120 mM MOPS) (pH 6.5) containing Cm and the indicated levels of HP, for 48 h. Strains were BW25113 and the *focA* and *focB* mutants, carrying either pBAD<sub>rha</sub>, pBAD<sub>rha-focB</sub> or pBAD<sub>rha-focA</sub>, as indicated. **A)** Wildtype with pBAD<sub>rha</sub>. **B)** *focA* mutant with pBAD<sub>rha</sub>. **C)** *focB* mutant with pBAD<sub>rha</sub>. **D)** Wildtype with pBAD<sub>rha-focA</sub>. **E)** *focA* mutant with pBAD<sub>rha-focA</sub> in. **F)** *focB* mutant with pBAD<sub>rha-focA</sub>. **G)** Wildtype with pBAD<sub>rha-focB</sub>. **H)** *focA* mutant with pBAD<sub>rha-focB</sub>. **I)** *focB* mutant with pBAD<sub>rha-focB</sub>. ‘+R’, with rhamnose; ‘-R’ without rhamnose. Error bars are included but are too small to see, and it corresponds to standard deviation of three technical replicates. The experiment was performed three times and similar results obtained.

#### 4.5. The effect of pH and formate on hypophosphite sensitivity upon induction of *focA* or *focB*

The above study clearly shows that FocB mediates HP toxicity under anaerobic conditions when induced from pBAD<sub>rha</sub>, and that a similar effect is observed for FocA, which is consistent with previous work (Suppmann and Sawers, 1994). Since FocA-mediated HP toxicity was found to require higher HP levels at high pH cf. low pH (Fig. 4.11), the possibility that the HP toxicity mediated by FocB is also pH dependent was investigated. The wildtype and *focA* strains, complemented with the *focA* or *focB* encoding plasmids, were grown anaerobically in WM medium at pH 6, 7 and 8, with HP at 0-0.5 mM, and formate at 0-3 mM (Fig. 4.13).

Induction of *focA* in the wildtype at pH 6 had little notable impact on the final OD achieved at pH 6 (Fig 4.13A), although *focA* induction did result in a modest reduced growth (16.6%) in the absence of HP, particularly with 3 mM formate. However, for induction of *focB* at pH 6 there was a considerable (up to 7.7- 12.4 fold) reduction in growth observed when formate (1-3 mM, respectively) was added; such an effect was not seen in the absence of formate (Fig 4.13A). This suggests that induction of FocB at pH 6, but less so for FocA, enables formate toxicity. In the *focA* mutant at pH 6, induction of *focA* restored sensitivity to HP (as expected and as seen above) but had little impact in the absence of HP. *focB* induction in the *focA* mutant, at the same pH, also resulted in restoration of HP toxicity (Fig. 4.13B), although not to the same degree as seen for *focA* induction. Interestingly, *focB* induction in the *focA* mutant resulted in a marked growth reduction in the absence of HP (2.5 fold), which was enhanced by the presence of formate (1.2 and 2.3 fold for 1 and 3 mM, respectively). This supports the suggestion above that, at pH 6, FocB causes an increase in formate toxicity. As before, formate did not reverse HP toxicity (at pH 6) under any condition tested with FocA.

Results at pH 7 were similar to those achieved at pH 6, in most cases, although the overall growth was approx. three times greater (Fig. 4.13C & D). However, unlike at pH 6, induction of the *focB* gene in either the wildtype or *focA* mutant resulted in no growth under all six growth conditions. This suggests that anaerobic production of FocB at pH 7 results in toxicity that is independent of added HP or formate. This effect is similar to that seen at pH 6, where *focB* expression reduced growth in the presence of formate (although not in its absence). Thus, it appears that *focB* induction is deleterious for *E. coli* growth, more so at pH 7 than pH 6. Interestingly, under non-inducing conditions at pH 7, the *focB* plasmid enhanced survival of the wildtype in the presence of HP (Fig. 4.13C). This indicates that low level *focB* expression counteracts HP toxicity, possibly by interfering with FocA activity. Such an effect was not seen in the *focA* strain nor at pH 6 (Fig. 4.13A & D).

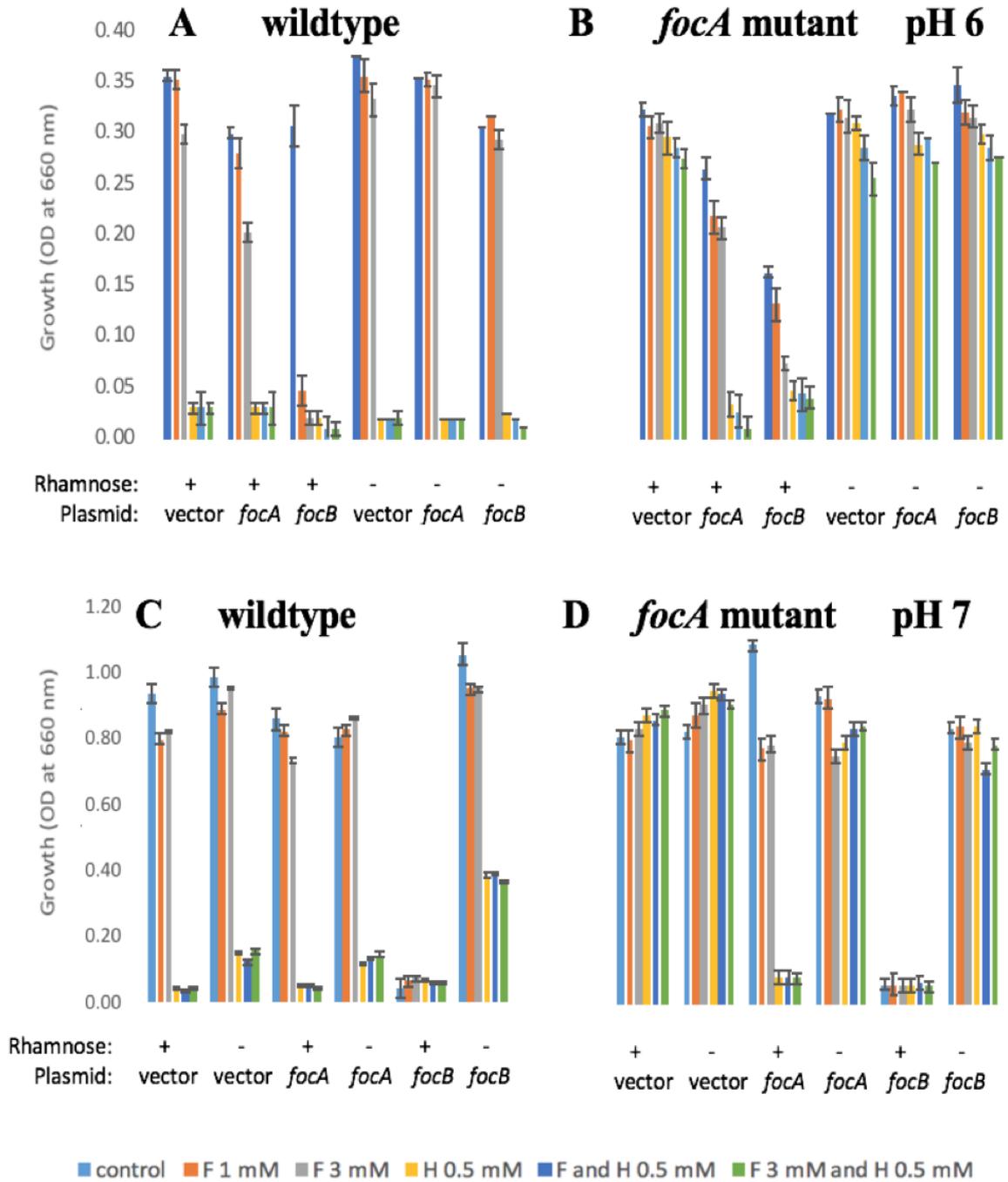
At pH 8, HP toxicity was relatively weak for the wildtype carrying the vector only (Fig. 4.15E). HP only caused a ~2.6-fold growth reduction whereas at lower pH (6 and 7) the HP growth inhibition was 12-20-fold ( $P = 0.034$ ;  $P = 0.01$ ), respectively (Fig. 4.13A & C). This effect is presumed to reflect poor expression of the chromosomal *focA* gene since, when *focA* was induced from pBAD<sub>rha</sub>, there was strong (12.5 fold) growth inhibition by HP at pH 8 (Fig. 4.13E & F). *focA* is known to be induced anaerobically by FNR and ArcBA, however, it is not reported to be subject to pH control (Levanon *et al.*, 2005) as far as I am aware. It should be noted that an increase in the relative expression of *focA* was previously reported to enhance HP toxicity (Beyer *et al.*, 2013). Induction of either *focA* or *focB* in the wildtype enhanced the otherwise modest HP toxicity at pH 8, and for *focB* induction, growth of the wildtype was also inhibited in the presence of formate, as seen at pH 6. The weak toxicity of HP in the wildtype at pH 8 was reversed by the presence of the *focA* plasmid in the absence of inducer. This suggests that, at pH

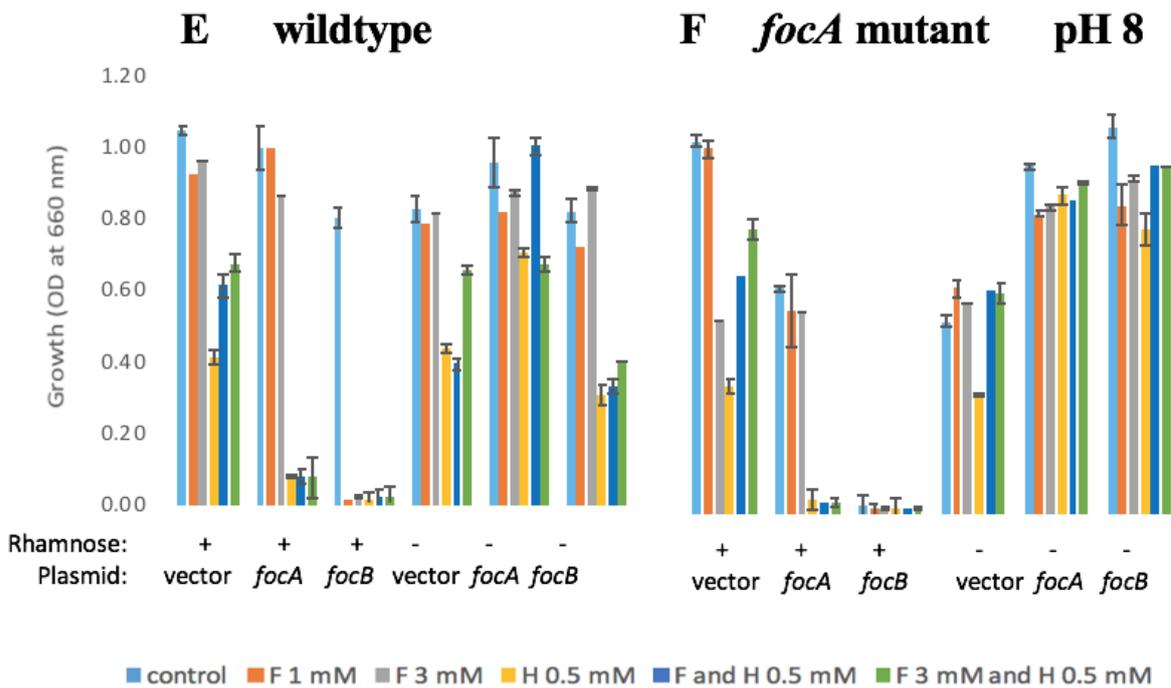
8, weak expression levels of *focA* from pBAD<sub>rha</sub> might lower HP toxicity. This would be consistent with the export role of FocA at high pH. Indeed, the *focA* mutant showed increased (~2.5 fold) growth with HP when the *focA* plasmid was present without inducer (with respect to the vector control). A similar effect was observed with the *focB* plasmid in the *focA* mutant (Fig. 4.13F). It is interesting to note that, at pH 8, the wildtype and *focA* mutant vector controls gave very similar growth patterns with HP, which further indicates that the chromosomal *focA* gene is strongly down regulated at this pH. At pH 8, the wildtype and *focA* mutant both displayed increased growth with HP when formate was added, but this was seen only for the vector control (Fig. 4.13E& F). This indicates that formate can counter HP toxicity, but only at higher pH. It also indicates that the effect is unrelated to FocA since the *focA* mutant exhibits this behaviour (Fig. 4.13F) in a manner similar to that of the wildtype.

Experiments were performed as above, but under aerobic conditions, to confirm that the HP and *focA/focB* growth effects observed are dependent on anaerobiosis (Fig. 4.14). The wildtype and *focA* strains showed little response to HP and formate at pH 6 with the vector control or in the absence of inducer, aerobically. However, both strains showed a 2-3 fold reduced growth upon induction of either *focA* or *focB* when HP and formate were present together at their highest concentrations. This suggests that aerobically, induction of either *focA* or *focB* increase toxicity towards a combination of HP and formate – this effect may be caused by FocA/FocB mediating the ‘inappropriate’ uptake of HP and formate resulting in toxicity.

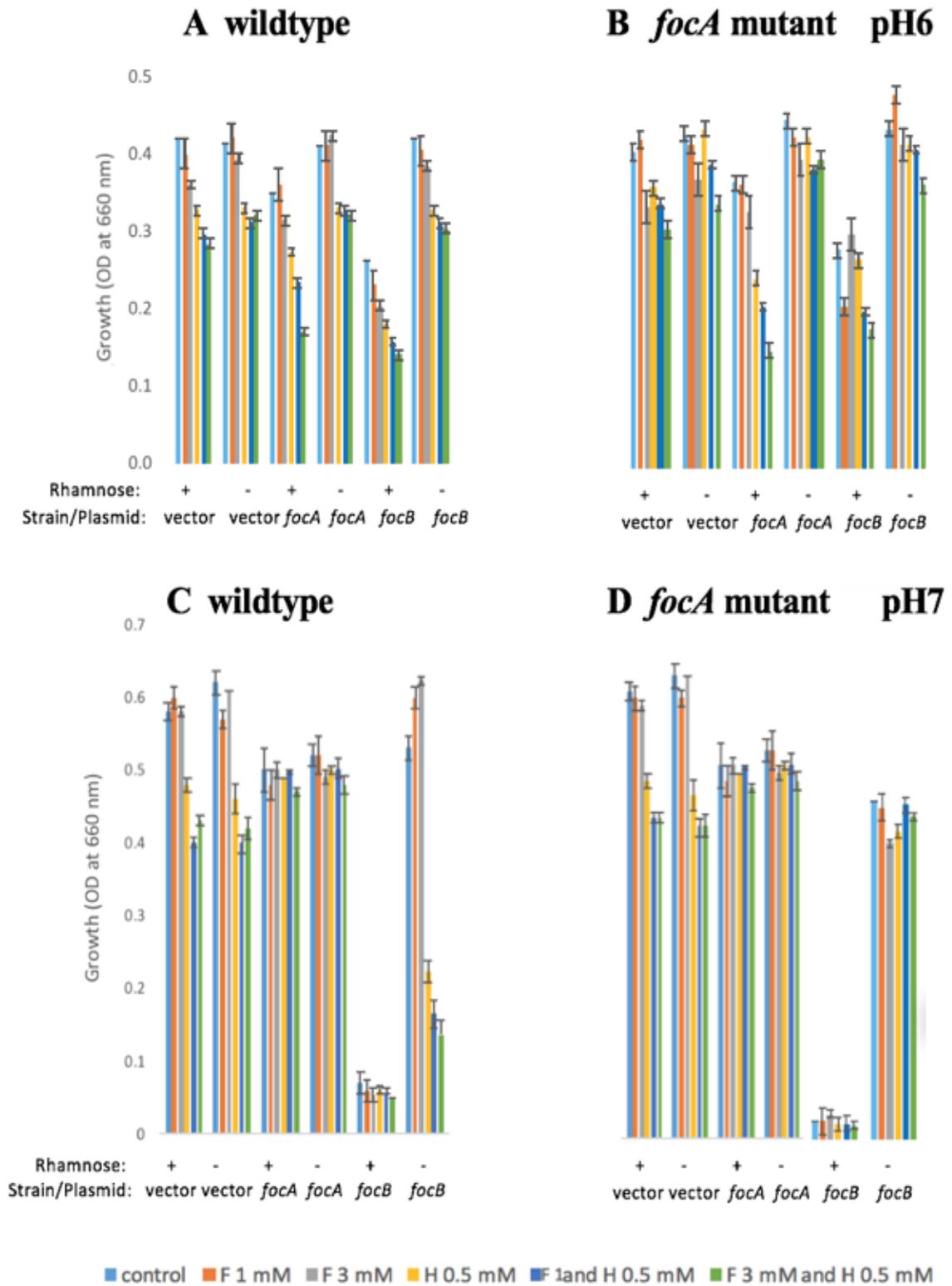
Aerobically at pH 7 and 8, induction of either *focA* or *focB* resulted in lack of growth for both the wildtype and *focA* mutant (Fig. 4.14C & D), with a greater effect seen with *focB* induction. The reason for this effect is unclear but could be related either to a membrane disruption effect caused by excessive levels of FocA/FocB, or by inappropriate transport

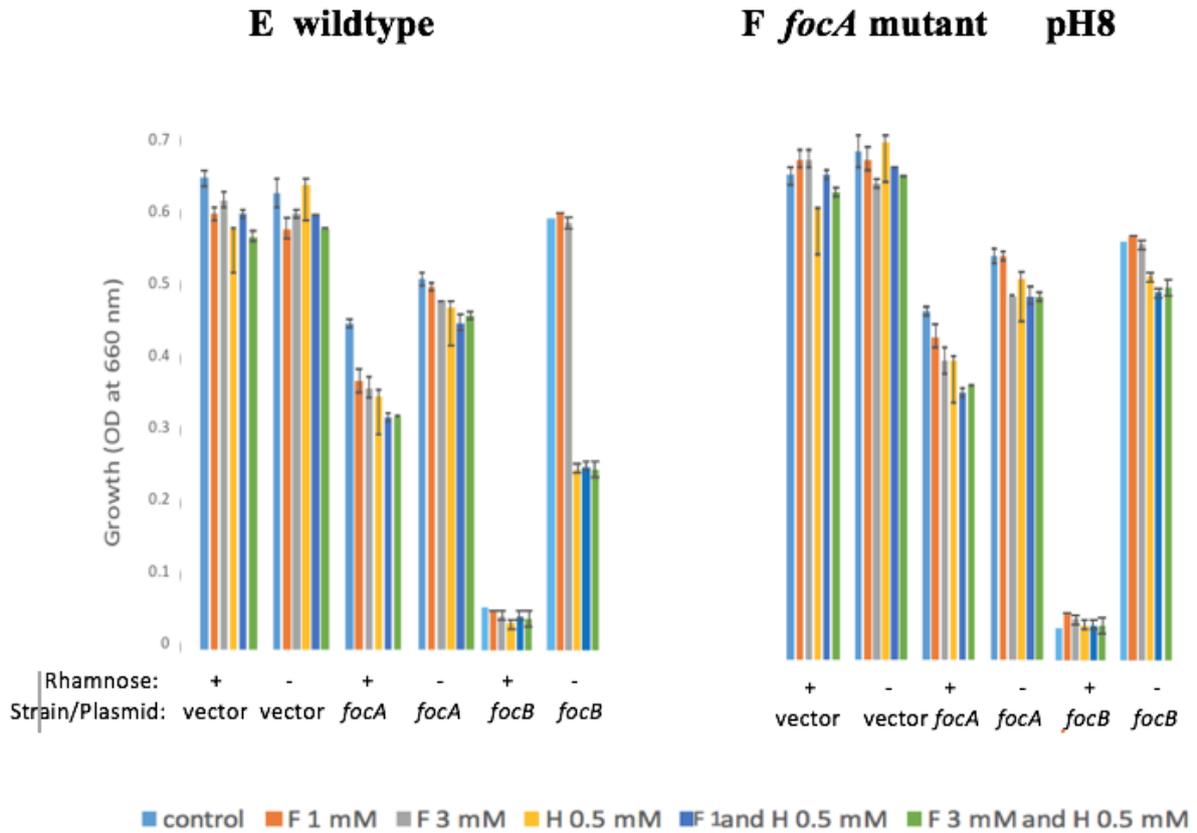
activity of these proteins which is stimulated at high pH but not at lower pH aerobically. Such an effect would be consistent with the previously reported pH-dependent switch in FocA transport activity (Lü *et al.*, 2011). Furthermore, HP did not cause a notable reduction in growth as *focA* induced in wildtype with/ without an inducer at pH 7 and 8. On the other hand, HP caused a 2.5 fold reduction in growth at similar pH when the *focB* plasmid was present in wildtype in the absence of inducer. This indicates that low levels of FocB drive HP toxicity aerobically. It should be noted that this effect was not seen at pH 6 which suggests that the transport behaviour of FocB towards HP is distinct at pH 6 versus pH 7-8.





**Figure 4.13. Effect of pH on the toxicity of hypophosphite upon induction of *focA* or *focB* in a wildtype or *focA* background during anaerobic growth.** Conditions were as described in Fig. 4.12, with or without 0.02% w/v rhamnose. Strains and transformed plasmid were as in Fig. 4.14, except no *focB* mutant was included. **A**, wildtype with vector control, pBAD<sub>rha</sub>-*focA* or pBAD<sub>rha</sub>-*focB* at pH 6. **B**, as A except for the presence of the *focA* mutant in place of the wildtype. **C** and **D**, as for A and B, respectively, except for pH 7. **E** and **F**, as for A and B, except for pH 8. F, formate; H, hypophosphite (concentrations are indicated). Growths were performed in triplicate for 48 h and averages and standard deviations are indicated. The experiment was repeated twice, with similar results obtained.





**Figure 4.14. Effect of pH on the toxicity of hypophosphite upon induction of *focA* or *focB* in a wildtype or *focA* background during aerobic growth.** Conditions are as for Fig. 4.15 except for growth under aerobiosis. **A**, wildtype plus *focA*, *focB* or vector only at pH 6. **B**, as A except for use of *focA* mutant in place of wildtype. **C** and **D**, as for A and B, except at pH 6 and 8, respectively. Growths were performed in triplicate and standard deviations are indicated. The experiment was repeated twice, with similar results obtained.

Thus, to summarise the above work, a number of pH dependent FocA and FocB effects on growth were observed.

- FocB increases formate toxicity at pH 6 and 8, anaerobically
- FocB prevents growth at pH 7, anaerobically
- HP imposes very little toxicity at pH 8 anaerobically in the wildtype, apparently due to weak *focA* expression at this pH
- Aerobically, induction of *focA* or *focB* enhances toxicity of HP at pH 7 and 8, but has little impact at pH 6.

- Aerobically, at pH 8, weak *focB* expression results in HP toxicity.

Such effects are suggestive of a pH dependent transport behaviour for FocB, which would be consistent with the pH dependence of formate transport as mediated by the homologous FocA protein (Lü *et al.*, 2011).

## **4.6. Effect of *focA/focB* status on formate production and consumption, and H<sub>2</sub> production**

### **4.6.1. The *focA* and *focB* mutants**

To confirm a role for FocA in mediating the release of formate during anaerobic growth, the levels of formate were measured during the growth of the *focA* mutant and the wildtype (section 2.15). Growth was in WM medium (section 2.5.2) plus 80 mM glucose with 0.8% trace elements (section 2.5.2.1), using 20 ml plastic syringes to provide an anaerobic environment. The two strains grew similarly and formate concentration in the medium increased during the first 5 h of growth for both the wildtype and the  $\Delta$ *focA* mutant (deletion strain with *kan* cassette removed), which represents the early stages of growth. Levels then decreased rapidly to reach a stable level at 10 h where the growth density was approx. 45 or 42% of the final maximum achieved, respectively (Fig 4.15). The amount of formate released into the medium for the *focA* mutant was 1.9 fold less than that for the wildtype at 3-5 h, and at 10-30 h (when levels had declined to a steady low concentration) formate was 2.5-10-fold less in the mutant (Fig. 4.14). A similar pattern of formate release and consumption was reported previously, along with a 50% reduction in formate release for a strain lacking FocA (Beyer *et al.*, 2013). These results clearly indicate that *E. coli* uses FocA to secrete formate, but can still secrete formate, although less effectively, in the absence of FocA which indicates that there must be a second formate export system, as proposed previously (Beyer *et al.*, 2013). Previous

work has also showed that the loss of formate from the medium, following its initial accumulation, is due to degradation by seleno-Cys dependent, and FdhD/E dependent, formate dehydrogenases (Beyer *et al.*, 2013). Thus, under the fermentative conditions employed here (lacking respiratory acceptors), the only likely option for the observed formate degradation would be the FdhH-dependent FHL which would require import of formate into the cytoplasm. This further indicates that a second formate importer can act in the absence of FocA. Degradation via FHL would require re-import of formate since, unlike Fdo and Fdn, its active site is on the internal side of the cytoplasmic membrane. The higher formate levels at 10-30 h in the wildtype cf. the *focA* mutant are presumed to reflect the greater initial formate levels at 5 h resulting in the observed higher residual formate levels at  $\geq 10$  h.

The *focA* mutation also resulted in a 1.85 fold ( $P= 0.015$ ) reduction in H<sub>2</sub> production (Fig. 4.16). This reduced H<sub>2</sub> production correlates well with the observed 11.6 fold reduction in formate export (Fig. 4.15) which indicates that H<sub>2</sub> production is still possible (although reduced) in the absence of the FocA channel and also indicates that higher external levels of formate drive a greater degree of H<sub>2</sub> production (Sawers, 2005; Yoshida *et al.*, 2005).

A caveat associated with the above is the impact that the  $\Delta$ *focA* mutation might have on the expression of the downstream *pflB* gene. PflB is responsible for the anaerobic production of formate from pyruvate and so any polarity effect caused by the *focA* mutation could diminish PflB levels. Indeed, previous work suggests a strong dependence of *pflB* expression on upstream expression of *focA*, with expression reduced by  $\sim 10$  fold in absence of *focA* expression (Beyer *et al.*, 2013). Thus, the lower level of formate shown here for the *focA* mutant could be, at least partly, a consequence of poor *pflB* expression. However, results below (section 4.6.2) indicate that complementation of the *focA* mutant allows an increase in formate export and H<sub>2</sub> production to give levels

similar to those of wildtype, which suggests that the *focA* mutation does not impact *pflB* expression to a degree that greatly affects formate (or hydrogen) production.

#### 4.6.2 Complementation with pBAD<sub>rha</sub> expression *focA* and *focB*

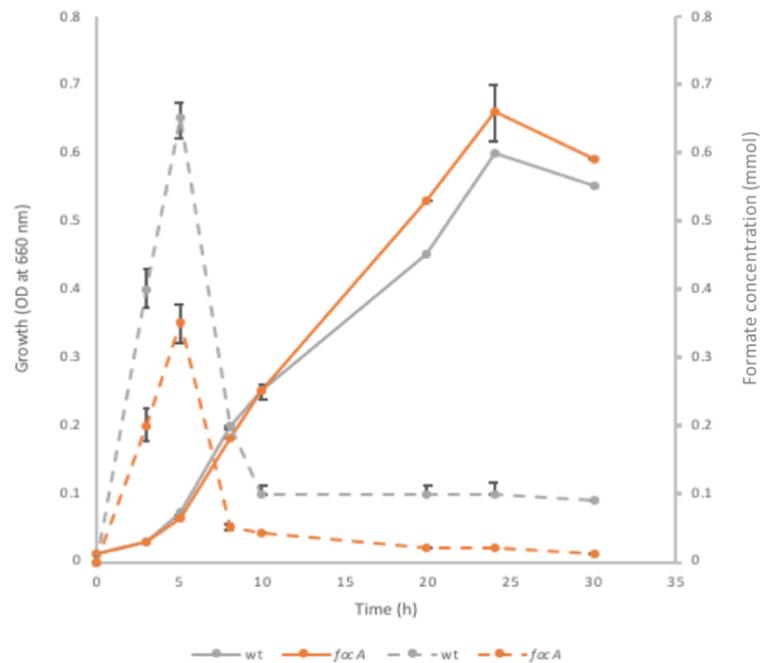
Experiments were performed as above, but at pH 6.5 and 7.5, and with the *focA* and *focB* plasmids, to determine whether the inducible *focA* and *focB* genes can restore formate production in the *focA* mutant strain, and also to test the impact of *focA/B* status on H<sub>2</sub> production. A similar overall pattern of formate production was seen as above, with a peak level at 5 h in all cases (Figs. 4.15 & 4.17). In the wildtype, induction of *focA* and *focB* at pH 6.5 resulted in lower formate production in the first 5 h, an effect that was particularly marked at 3 h with a 3.4 fold lower level cf. the vector control (Fig. 4.17A). This indicates that induction of *focA* or *focB* reduces formate release at pH 6.5, possibly through enabling enhanced formate uptake or interference with Pfl activity. The uninduced controls and vector controls gave similar formate levels at pH 6.5 in the wildtype at 5 h. However, after 10 h, the vector controls had lower levels of formate in comparison to the induced and non-induced *focA/focB* plasmid bearing strains, which suggests a decreased formate uptake (or enhanced release) in the presence of these plasmids – the reason for this effect is unclear. As expected from previous results (Fig. 4.13), induction of *focB* lowered growth of the wildtype by 2.5 fold at H 6. Interestingly, at pH 6.5, H<sub>2</sub> production in the wildtype was 1.8-fold reduced by *focB* induction of the vector control, and was also slightly reduced (1.2 fold) by *focA* induction (Fig. 4.17). This appears to be correlated with the correspondingly reduced formate release. In the *focA* mutant, at the same pH, production of formate was 2-fold lower than in the wildtype, for the vector controls, at 5 h (Fig. 4.17C). This effect was reversed by induction of the *focA* plasmid. Induction of *focB* also gave a modest increase in

extracellular formate levels in the *focA* mutant strain at pH 6.5, which was particularly noticeable at 3 h (1.7 fold higher levels cf. the vector control. Induction of *focA* or *focB* increased the growth of the *focA* mutant with respect to the induced-vector control by 2.4 and 1.5 -fold at 10 h and 2 and 1.4-fold at 20 h, respectively;  $P= 0.01$  at both time points respectively). This indicates that lack of formate import/export inhibits growth under fermentative conditions, and that either *focA* or *focB* in *trans* can counter absence of the chromosomally encoded FocA channel. H<sub>2</sub> production was 5.6 fold lower in the *focA* mutant cf. the wildtype (uninduced vector controls; Fig 4.17). Just as formate release was restored by induction of *focA* in the *focA* mutant, so was H<sub>2</sub> production reinstated to levels similar to those of the wildtype. Interestingly, this was also the case without inducer, indicating that even low level *focA* expression enables sufficient internalisation of formate to return H<sub>2</sub> production to wildtype levels. Also of note is the observation that at pH 6.5, induction of *focB* did not re-enable H<sub>2</sub> production to any degree, whereas in the non-inducing state the *focB* plasmid returned H<sub>2</sub> levels to those approaching wildtype (increase of 3.2 fold cf. the vector control; Fig, 4.19). These findings would suggest that FocA is the formate importer largely responsible for delivering exogenous formate to FHL-1 for H<sub>2</sub> production, but that FocB can function in a similar fashion when produced at low levels, but not when at high levels.

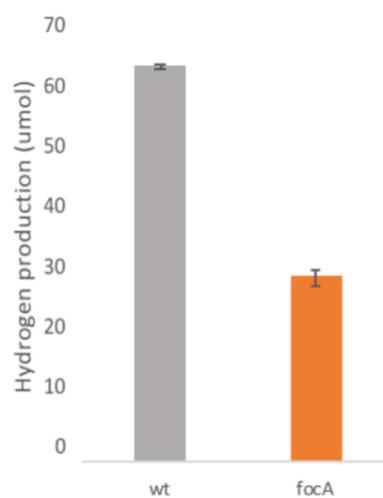
At pH 7.5, similar to previous results (Fig. 4.18B & D), *focB* induction led to growth inhibition and the induction of *focA* in the wildtype led to decrease in formate release (by 1.5 fold cf. the vector control) – as seen at pH 6.5 (Fig. 4.18A & 4.17A). *focA* induction in the wildtype also led to a slight (20% lower compared to the vector control) reduction in H<sub>2</sub> production, as seen at pH 6.5. The uninduced *focB* plasmid gave a twofold reduced H<sub>2</sub> production, indicating that low level FocA production reduces formate delivery to FHL-1. In the *focA* mutant at pH 7.5, the level of formate produced was only slightly

lower (13% lower in the vector controls) than that of the wildtype, which indicates no major role for FocA in formate production at this pH – which is possibly related to lack of expression at higher pH, as suggested above. Induction of *focA* lowered formate levels at 5 h, as it was the case in the wildtype. In addition, *focA* and *focB* in the uninduced states also reduced formate levels at 5 h in the *focA* mutant at pH 7.5 (unlike in the wildtype). Further, induction of *focA* in the *focA* mutant at pH 7.5 resulted in a growth enhancement (1.2 fold greater growth cf. the induced vector control at 20 h;  $P= 0.02$ ). The reason for this effect is unclear but might be caused by the higher extracellular formate levels seen in the *focA* mutant when *focA* induced (compared to vector control) at 10 h (23 versus 7.5 nmol), since the growth difference indicated above was largely achieved between the 10 and 20 h time points. The *focA* mutant strain exhibited a 2.2 fold lower H<sub>2</sub> level than that of the wildtype, for the uninduced vector controls, and the presence of the induced or uninduced *focA* or *focB* plasmids did not return production to wildtype levels (Fig. 4.19B). Indeed, the presence of the *foc* plasmids had, in general, an inhibitory effect of hydrogen production. Whether this effect is caused by a reduced *pflB* expression in the *focA* mutant is unclear.

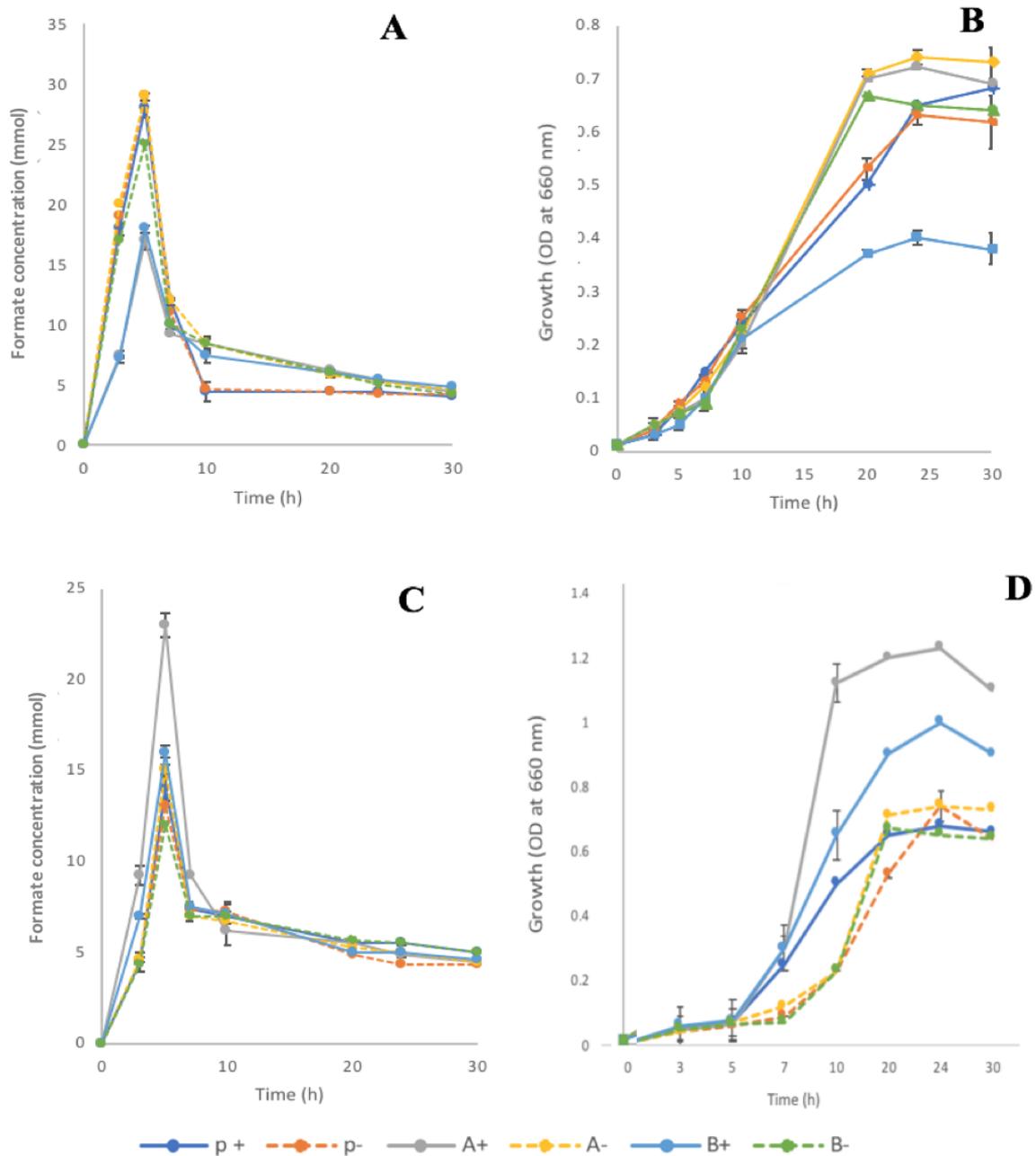
In summary, the results show that FocA supports H<sub>2</sub> production at low pH, but appears to have little impact at high pH. The results also indicate that low levels of FocB can compensate for absence of FocA. Plasmid induced *focA* restores the levels of formate released in a *focA* mutant to those seen in the wildtype, whereas plasmid induced *focB* increases such levels more modestly. Thus, the results indicate that, like FocA, FocB also acts as a formate exporter and can deliver exogenous formate to FHL-1 indicating a role in formate import also. These roles for FocA and FocB seem to be largely restricted to acidic pH conditions.



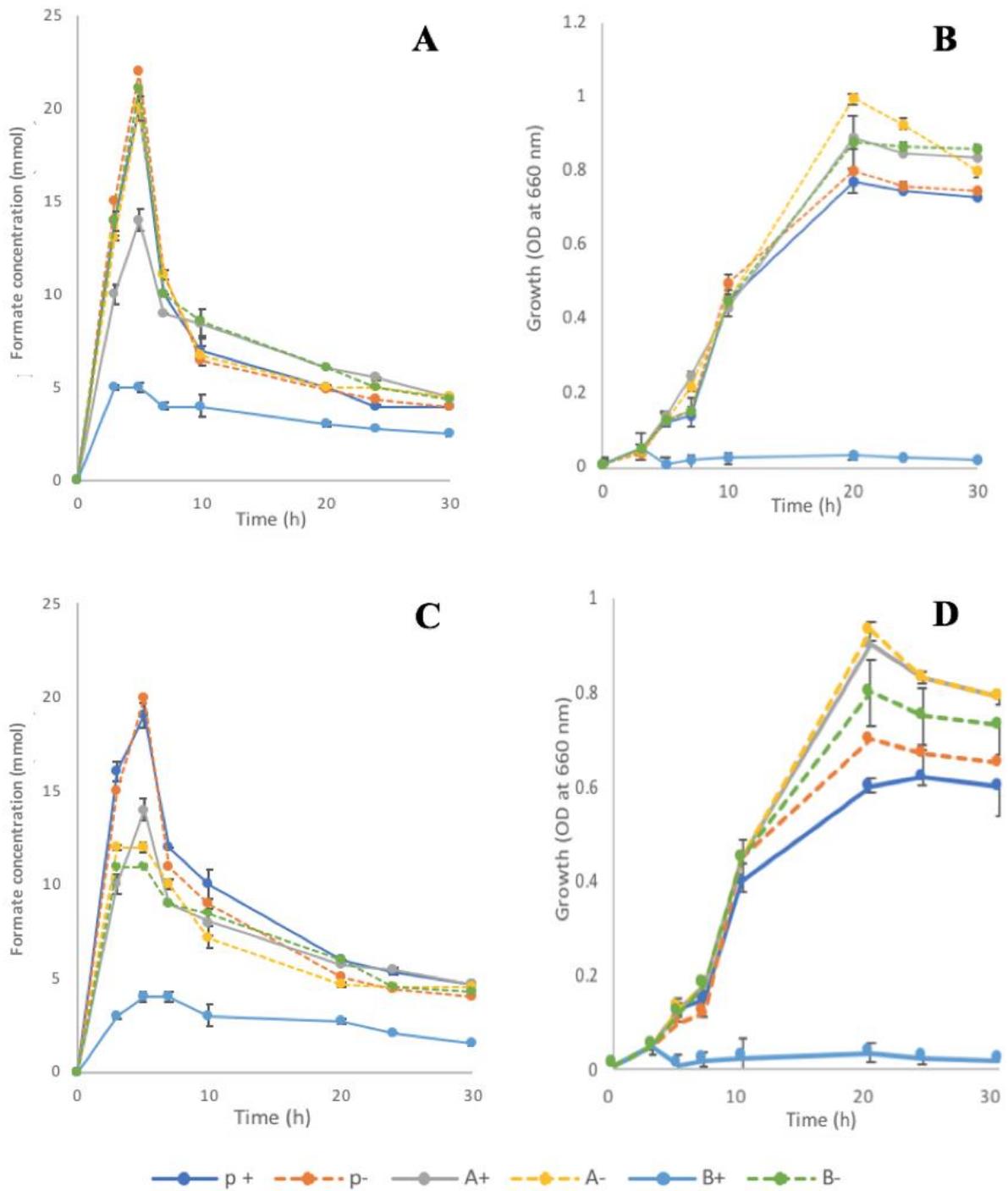
**Figure 4.15.** The effect of *focA* mutation on formate concentration during fermentative growth compared to wt in WM-medium at pH 6.5. Dotted line indicates the formate concentration while the solid line indicates bacterial growth (660 nm). The growth was in WM-medium at 37 °C, 80 mM glucose, with 0.8% trace elements and 120 mM MOPS. Wt, and BW25113  $\Delta$ *focA* was generated) Growths were performed in triplicate and standard deviations are indicated. The experiment was repeated twice, with similar results obtained.



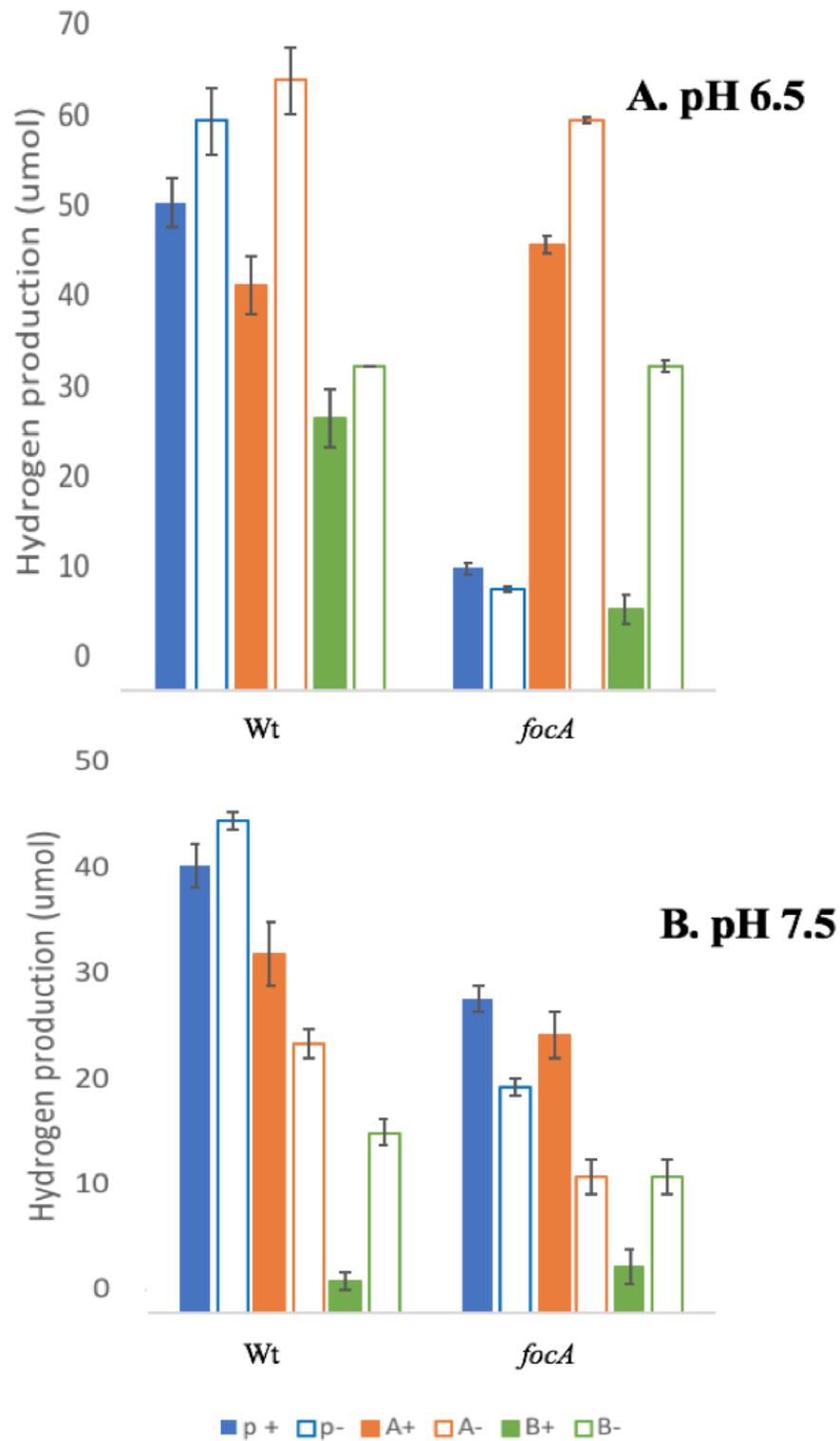
**Figure 4.16.** Hydrogen production in *focA* mutant compared to wildtype at pH 6.5. Measurements were made from the experiment presented in Fig. 4.15.



**Figure 4.17. Effect of *focA* or *focB* induction on formate production and consumption in a wildtype or *focA* mutant at pH 6.5.** Conditions are as for Fig. 4.17. **A** and **B**, wildtype plus *focA* (A) or *focB* (B) inducible plasmids, or the vector only (p), with (+) or without inducer (-). **C** and **D**, as A/B, except for the use of the  $\Delta$ *focA* mutant in place of the wildtype. A and C, growth; B and D, corresponding formate levels. Growths were performed in triplicate and standard deviations are indicated. The experiment was repeated twice, with similar results obtained.



**Figure 4.18.** Effect of *focA* or *focB* induction on formate production and consumption in a wildtype or *focA* mutant at pH 7.5. Conditions are as for Fig. 4.18, except for the pH employed.

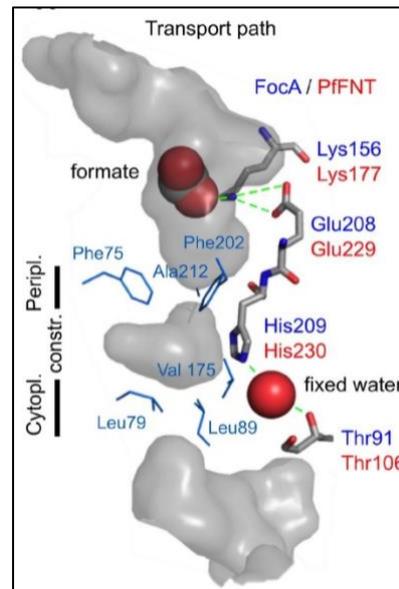


**Figure 4.19. Hydrogen production upon induction of *focA* or *focB* in a wildtype or  $\Delta focA$  background.** Details are as for Figs. 4.16 & 4.17. **A**, pH 6.5 (data from experiment in Fig. 4.16). **B**, pH 7.5 (data from experiment in Fig. 4.17). ‘P’, vector only; ‘A’, *focA* plasmid; ‘B’, *focB* plasmid; ‘+’, with rhamnose; ‘-’, minus rhamnose.

## 4.7. Discussion

In this chapter, the role of FocB was analyzed and compared with that of the primary formate channel FocA. Both the *focB* and *focA* genes were cloned in an inducible vector to test their ability to mediate HP (a formate analogue) toxicity as well as to mediate the release and uptake of formate at different pH, and the production of hydrogen gas.

Firstly, a simple bioinformatics analysis was performed to compare the features of FocA that play important roles in FocA structure-function with such features that may be conserved in FocB. The results indicated that FocB of *E. coli* shares all the key structural and functional elements of FocA (Fig. 4.1). This suggests that it will form a cytoplasmic membrane channel that adopts a pentameric structure very similar to that of FocA, including all six transmembrane helices (Lü *et al.*, 2011). In addition, there are conserved formate binding ligands (Lys156, Glu208, His209 and Thr91), which forms a hydrogen bond between a formate substrate molecule and the FocA channel (*E. coli* and *Salmonella*) (Fig 4.20) (Waight *et al.*, 2010; Wiechert and Beitz, 2017), these ligands were found fully in FocB, which is consistent with the suggestion that FocB is a fully functional formate channel. Moreover, the *E. coli* and *Salmonella* FocA pore constriction residues are also conserved in FocB. Furthermore, the conserved C-terminal residues that are important in formate transport and pH-regulated transport mode in FocA (Hunger *et al.*, 2017) are completely found in FocB (Fig 4.1). This suggests that FocB is a formate transporter and a pH dependent channel.



**Figure 4.20.** A view of FocA channel in *Salmonella enterica* serovar Typhimurium. The formate binding ligands indicates as blue sticks and the constriction residues indicates in light blue. Green dashed line indicate the hydrogen bond between the formate substrate and the ligands in the channel. Red circles indicates formate substrate and grey colour indicates the internal shade of the transport path (Wiechert and Beitz, 2017).

In order to test the role of FocB experimentally as a formate transporter, three isogenic strains were used; the wildtype, and a  $\Delta focA$  and  $\Delta focB$  mutant. The result showed that the toxicity effect of hypophosphite (HP) appears at low concentrations under anaerobic conditions (Fig 4.8) in the wildtype and *focB* mutant, which both possess FocA such that the formate analogue (HP) can be imported into the cell causing a reduced growth through inhibition of Pfl (Sawers and Böck, 1989). This effect was not seen in the *focA* mutant strain unless higher HP levels were applied, as reported previously (Suppmann and Sawers, 1994). The reason for the lack of any impact of the *focB* mutation is in contrast to previous work (Trachounian and Trachounian, 2014), but matches that of others indicating that the *hyf-focB* operon is cryptic (Skibinski *et al.* 2002; Self *et al.*, 2004.). The work performed here supports that of Suppmann and Sawers (1994) demonstrating that FocA is required for HP toxicity and additionally show that FocB is not involved in

HP toxicity under the conditions tested, presumably because of poor *focB* expression. Suppmann and Sawers (1994) showed that 75 mM HP in WM medium caused complete inhibition of the anaerobic growth of *E. coli* (MC4100), but at lower concentration (20 mM) the toxicity effect was much clearer (1.5 fold compared to the initial optical density). The results here showed that an even lower concentration (0.5 mM) of HP gave the best growth differential between the *focA*<sup>+</sup> and *focA*<sup>-</sup> strains (Fig. 4.9B).

Direct experimental analysis of FocB and its ability in formate transport was achieved by cloning *focA* and *focB* into pBAD<sub>rha</sub> to test their induction and complementation in a *focA* strain. An effect of induction of *focB* in the *focA* mutant was observed under low concentration (0.5 mM) of HP, as a major growth HP-mediated inhibition effect under fermentation conditions (Fig. 4.12). A very similar effect was seen for *focA* induction under the same conditions. This results thus suggest that the genomic *focB* is indeed weakly or non-expressed compared the genomic *focA* gene, and to its induction from the pBAD<sub>rha</sub> vector.

Under aerobic conditions, no effect of HP toxicity was seen in WM medium (Fig. 4.9A), which is in agreement with Suppmann and Sawers (1994). Normally, HP would only be expected to inhibit growth anaerobically as the *focA-pflB* operon is not expressed nor required under aerobic conditions; FocA is produced under anaerobic condition (Sawers, 1994). Furthermore, HP is thought to act as an inhibitor of pyruvate formate-lyase (Pfl). Since Pfl is only expressed and required under fermentative/anaerobic conditions (Sawers, 1995), then any aerobic uptake of HP enabled by induction of *focA* or *focB* might not be expected to have an influence, which matches the results obtained here.

Lü *et al.* (2011) performed electrophysiological experiments showing that the FocA channel is a passive facilitator for formate at pH  $\geq 6.8$ . Thus, as formate is produced by Pfl under fermentative conditions, the excess formate is able to diffuse out of the

cytoplasm and accumulate in the medium as a secreted waste product. However, under low pH (< 6.8), the FocA channel is shifted to an active proton symporter driving uptake of formate for consumption by FHL-1. Since the WM-medium used here has a starting pH of 6.5, which would be expected to drop further during fermentative growth (Fig. 4.10), it is anticipated that FocA would act as a proton/formate-symporter which would help to explain why HP was toxic in WM medium.

The results show that induction of *focB* increases formate toxicity at pH 6 and 8, and prevents growth at pH 7 under anaerobic conditions (Fig. 4.11). On the other hand, HP showed little toxicity at pH 8, which appears to be due to weak *focA* expression since induction of *focA* from pBAD<sub>rha</sub> restored HP toxicity at this pH. Thus, it would be very interesting to test the effect of pH on *focA* expression. The observation that chromosomally encoded *focA* has little effect at higher pH (due to poor expression) would suggest that the proposed high-pH-dependent switch in activity to a formate exporter through facilitated diffusion (Lv et al., 2013; Lü et al., 2011) is not physiologically relevant. Indeed, it could be argued that FocA could function effectively solely as a passive diffusion facilitator (without any need for active uptake) such that formate uptake and export would be entirely driven by formate production (by Pfl) and consumption (by FHL).

In order to determine how *focA* mutation affects formate production during fermentative growth, formate build up in the medium was monitored during growth. The *focA* mutant generated twofold less formate than that for the wildtype at 3-5 h, and at 10-30 h the formate in the medium was up to 10-fold less in the mutant (Fig. 4.15). A similar pattern of formate release and consumption was reported previously, along with a 50% reduction in formate release for a strain lacking FocA (Suppmann and Sawers, 1994; Beyer et al., 2013). These results clearly suggest that *E. coli* secretes formate through FocA, but in

the absence of FocA is still able to release formate but in less quantity, thus suggesting that there must be a second formate export channel (Beyer *et al.*, 2013; Christiansen & Pedersen, 1981). In addition, the absence of FocA resulted in lower H<sub>2</sub> production by ~twofold, indicating that the reduced formate release in the *focA* mutant led to a similar reduction of H<sub>2</sub> production.

Induction of *focA* and *focB* in the wildtype at pH 6.5 resulted in a reduction in formate release compared to the vector control by 1.7 fold at 3-5 h, although the amount of hydrogen produced was close to that of the vector control. However, complementation of the *focA* mutant by induction of plasmid-encoded *focA* led to a restoration of formate levels in the medium and thus hydrogen was produced at levels that were similar to that in wildtype. On the other hand, complementation by induction of plasmid-encoded *focB* had only a slight effect on formate release and no notable positive effect on hydrogen production, although provision of *focB* in the non-induced state did increase H<sub>2</sub> production significantly without affecting formate release. This might be because of the lower growth of the *focB*-induced strain at this pH level as shown before at pH 6 (Fig. 4.13)

Induction of *focA* in the wildtype resulted in lower levels of peak formate accumulation (at 5 h) in the medium at pH 7.5, and a similar result was observed at pH 6.5 for both *focA* and *focB* induction. However, the total hydrogen produced was in similar quantities for the *focA*-induced and vector control wildtype, which reflects the fact that the final formate levels were similar at 30 h. However, the induction of *focB* in the wildtype at pH 6.5 did result in a lower H<sub>2</sub> production which would appear to reflect the weaker growth achieved.

In summary, FocB was found to be capable of compensating for lack of FocA and thus can function as a second formate channel in the absence of *focA*. Further, evidence was

obtained to suggest that the chromosomal version of *focA* is weakly expressed at pH 7.5 cf. pH 6.5, although FocA is functional at this pH if expressed from a surrogate promoter. Since FocB showed activity in HP toxicity and formate release, this provides evidence to support a role for Hyf in metabolism of formate as part of a second H<sub>2</sub>-producing hydrogenase forming a second FHL complex involved in formate disposal and possibly energy conservation. This possibility is further addressed in the following chapters.

## Chapter 5: Phenotypic analysis of the effect of *hyf* operon induction in *E. coli* K-12

### 5.1. Introduction

Four [NiFe] hydrogenases have been identified in *E. coli* (Sawers, 1995). Hydrogenase 1 (Hyd-1) and hydrogenase 2 (Hyd-2) oxidize hydrogen, while, hydrogenase 3 (Hyd-3) is a part of the formate hydrogen lyase (FHL) complex, which produces hydrogen from formate. Under anaerobic conditions, synthesis of Hyd-1 and Hyd-2 occurs, and during fermentation these enzymes are generally associated with the appearance of the third hydrogenase, Hyd-3, which forms a part of the hydrogen evolving activity of FHL (Soboh *et al.*, 2011). However, Hyd-2 switches between hydrogen consuming and hydrogen production depending on the redox status of the quinone pool and the carbon source provided (Brøndsted and Atlung, 1994; Pinske *et al.* 2015).

Bock and Sawers (1996), demonstrated the role of Hyd-3 as a reversible hydrogenase. By combining deletion mutations they were able to show that there is no hydrogen uptake in a cell that lacks of Hyd-1 and 2, but there is a 30 fold increase in hydrogen production in  $\Delta hya \Delta hyb$  strain compared to a strain lacking all three hydrogenases. Noguchi *et al.* (2010) revealed that there was no hydrogen produced under anaerobic conditions in a  $\Delta hycE$  strain at pH 5.5, 6 and 6.5. However, under similar conditions in the wildtype, hydrogen was produced at all pH values tests, but was highest at pH 5.5. Hyd-1 and Hyd-2 possess six conserved metal-binding motifs (L0-L5) in the large hydrogenase subunit. However, Hyd-3 differs because there is no L0 motif and the order of the other five motifs is different (Meeda *et al.*, 2007).

Andrews *et al.* (1997) sequenced the 55.8 min region of *E. coli* genome revealing a putative 12-gene *hyfABCDEFGHIJRfocB* operon. The sequence similarities with Hyd-3 strongly suggested that *hyfG* and *hyfI* genes encode the large and small subunits

respectively of a fourth [NiFe] hydrogenase and therefore the operon was denoted as *hyf* (hydrogenase four) encoding a putative fourth hydrogenase (Hyd-4, or Hyf). Moreover, the *hyf* operon was found to be similar to *hyc* but it includes three additional genes, two of them (*hyfB* and *hyfD*) are homologues of the antiporter-like subunits from Complex-I and the third (*hyfE*) encodes a protein homologous to NuoL (Efremov and Sazanov, 2012).

It was shown by Yoshida *et al.* (2005) that as formate concentration increases, hydrogen production increases in *E. coli*. The maximum formate concentration was tested 100 mM, and hydrogen was still increasing. Trchounian and Trchounian (2013) reported that Hyd-3 is not the only hydrogen producer in *E. coli* as Hyd-4 was found to generate H<sub>2</sub> fermentatively in the presence of glucose. This effect was shown by use of a *hyc* mutant strain and testing for hydrogen production using Ti-Si electrode which is sensitive to hydrogen as well as oxygen. Trchounian *et al.* (2015) reported that Hyd-4 produces hydrogen at pH 7.5 with glycerol fermentation and formate supplementation. The most recent finding by Trchounian *et al.* (2017) indicated that hydrogen production via Hyd-4 mainly depends on formate and low pH levels (pH 5.5), and for hydrogen production at higher pH (pH 7.5) Hyd-1 and Hyd-2 need to be removed. However, all this work was performed using an electrode to sense H<sub>2</sub> and no second method of detection was applied. In addition, the findings appear to contradict those reported by others indicating no hydrogen production by Hyd-4 and lack of *hyf* expression unless *hyfR* is induced synthetically (Skibinski *et al.*, 2002; Self *et al.*, 2004).

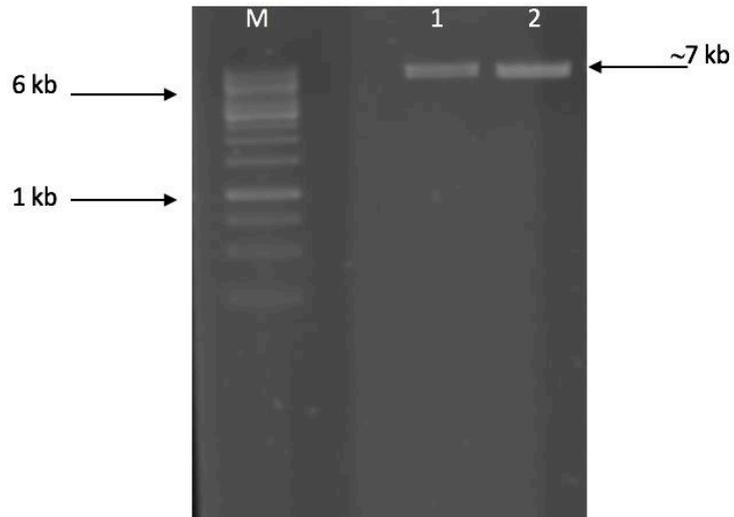
In this chapter, Hyd-4 function as a hydrogen producer was examined by placing the *hyf* operon under the control of an inducible promoter. Furthermore, the effect of pH on the hydrogen yield compared to the wildtype and the effect that pH has on Hyd-3 activity and Hyd-4 were studied.

## 5.2. Cloning of *hyc* and *hyf* into inducible vectors

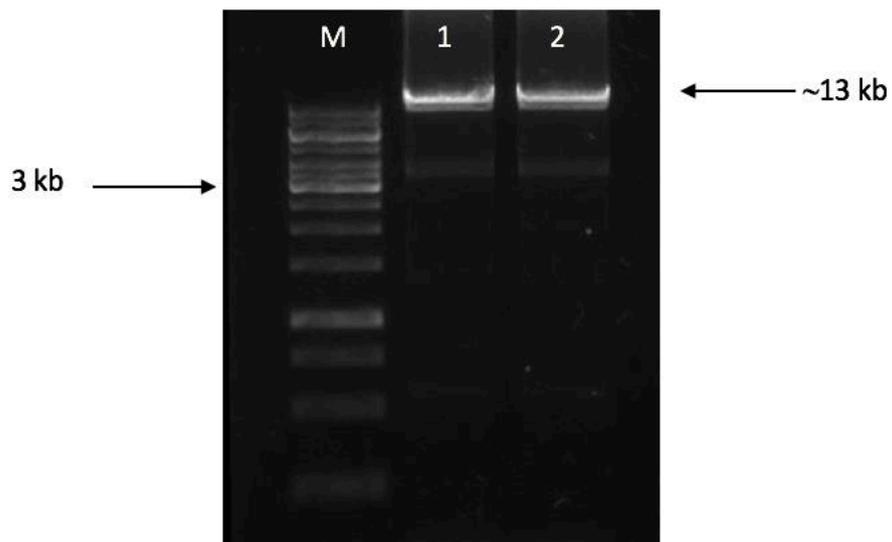
In order to enable the controlled expression of *hyc* and *hyf*, each operon was cloned into the pBAD<sub>rha</sub> vector; this allows controlled expression by rhamnose. In this way, both *hyc* and *hyf* function can be measured under any desired environmental condition without any impact of the regulatory factors (e.g. formate, FhlA, HyfR, anaerobiosis,  $\sigma^{54}$ ) which normally control their expression. In particular, it would be possible to determine the effect of *hyf* expression since while studies indicated that the *hyf-focB* operon is silent in the wildtype (Skibinski *et al.*, 2002; Self *et al.*, 2004).

### 5.2.1 Extraction and amplification of *hyc* and *hyf* from the wildtype (MG1655)

In order to obtain the *hyc* and *hyf* operons for cloning purposes, genomic DNA was extracted from the wildtype *E. coli* K-12 strain, MG1655, using a GeneJET™ genomic DNA purification kit (Fermentas) (section 2.10.1). PCR amplification of the ~7 kb *hyc* and ~13 kb *hyf* operons (section 2.10.1) was achieved using primers in Tables 2.5 and 2.6 (Appendix 6 & 9), and Amp HiFi PCR premix DNA polymerase (Clontech)/ Universe HotStart polymerase. Amplified fragments were analysed by agarose gel electrophoresis to estimate their size (Figs. 5.1 and 5.2). The fragments generated had the desired size and were thus used for cloning into pBAD<sub>rha</sub>.



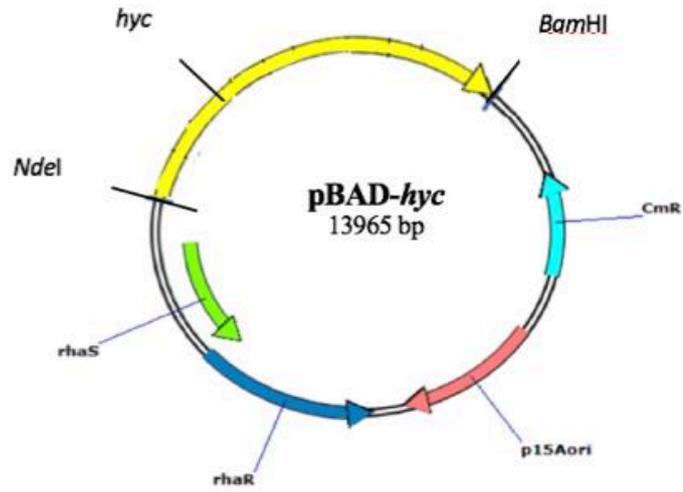
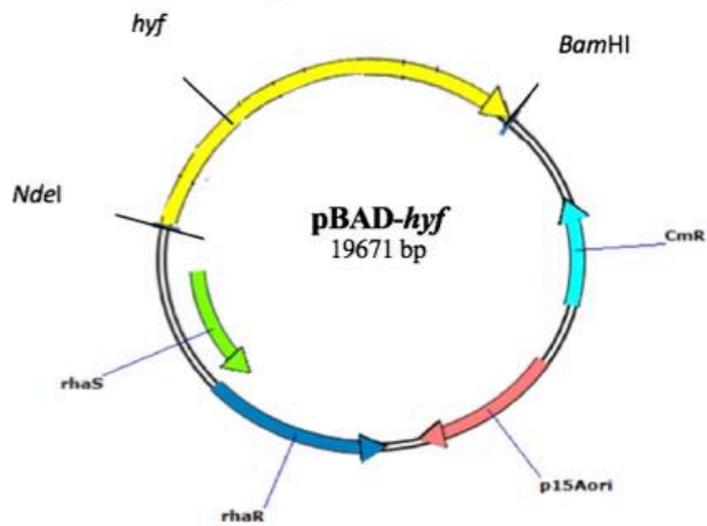
**Figure 5.1 Agarose gel (0.7%) electrophoretic analysis of purified *hyc* operon DNA.** M, Marker 1 kb Generuler (Fermentas); lanes 1-2, 2  $\mu$ l of *hyc* DNA PCR product using Amp HiFi PCR premix DNA polymerase. Arrows indicate position of DNA bands.



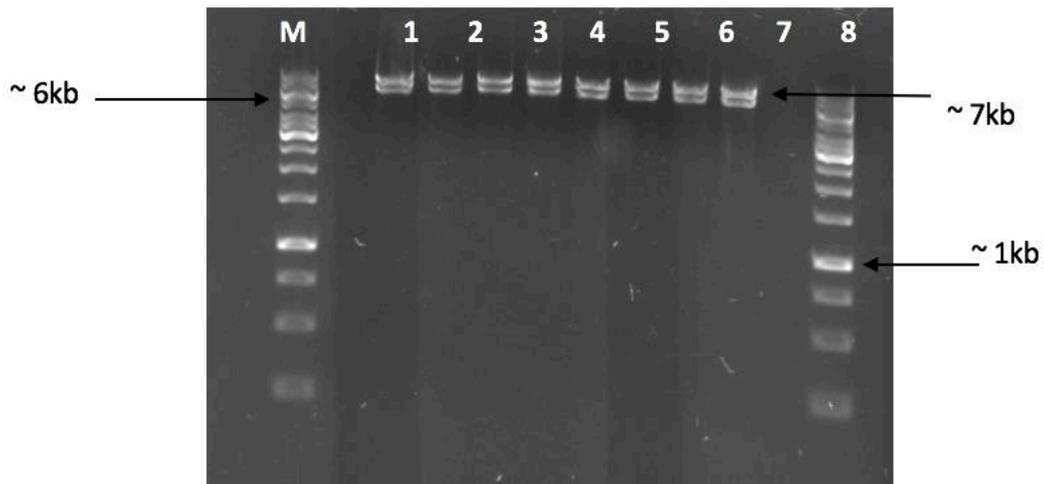
**Figure 5.2 Agarose gel (0.7%) electrophoretic analysis of purified *hyf* operon DNA.** M, Marker 1 kb Generuler (Fermentas); lanes 1-2, 2  $\mu$ l purified *hyf* as PCR product. For other details, see above.

### 5.2.2. Cloning of *hyc* and *hyf* into the pBAD<sub>rha</sub> vector

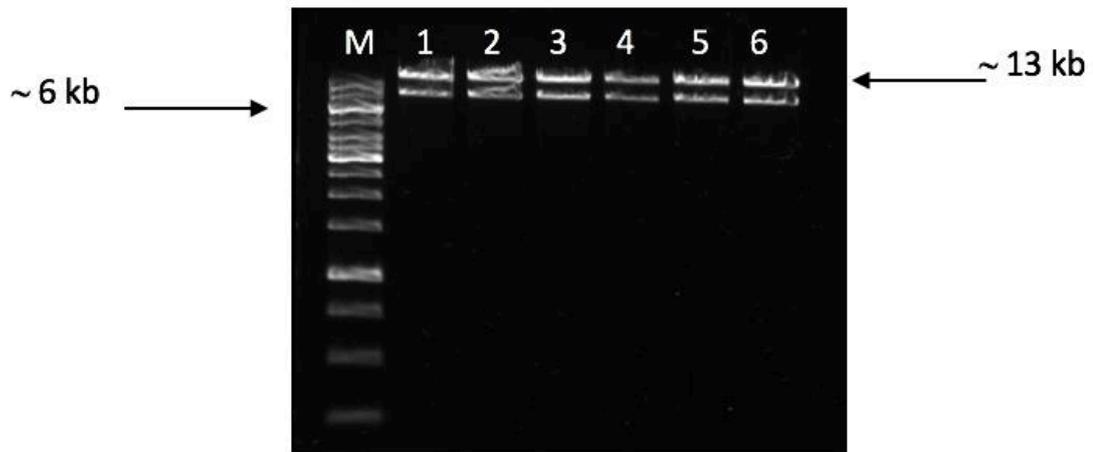
To allow controlled induction of *hyc* and *hyf*, the PCR products of *hyc* and *hyf* were cloned into pBAD<sub>rha</sub>, such that expression would be dependent on the vector-borne *Prha* promoter and ribosome-binding site. Gibson cloning methodology (Gibson *et al.*, 2009; section 2.12.2.3) was used for the cloning process. The resulting reaction products were used to transform chemically competent *E. coli* Stellar cells (section 2.12.2.4). Twelve resulting Cm<sup>R</sup> colonies were selected for plasmid ‘miniprep’ isolation using GeneJET<sup>TM</sup> Plasmid Miniprep Kit (Fermentas) (section 2.10.7), six of each type. These plasmids were then further analysed by restriction digestion (section 2.10.4) with *EcoRI* and *NdeI* for cloned *hyc* (Fig 5.4) and *NdeI* and *BamHI* for cloned *hyf* (Fig 5.5) to release the insert from the vector. Following analysis by agarose gel electrophoresis, all 12 plasmids were shown to carry an insert of the expected size (~7 kb for *hyc* and ~13 kb for *hyf*) and were subsequently designated pBAD<sub>rha</sub>-*hyc* and pBAD<sub>rha</sub>-*hyf* (Fig. 5.3).

**A****B**

**Figure 5.3. Map of the pBAD<sub>rha-hyc</sub> and pBAD<sub>rha-hyf</sub> plasmids. Relevant restriction sites are shown. The origin (*p15AoriV*), *Cm*<sup>R</sup> locus, and *rhaS* and *rhaR* genes (Egan and Schleif, 1993) are shown.**



**Figure 5.4 Agarose gel (0.7%) electrophoretic analysis of the pBAD-*hyc* constructs.** M, Marker 1 kb Generuler (Fermentas); lanes 1-6, plasmid DNA double digested with *EcoRI* and *NdeI* isolated from six distinct transformants.



**Figure 5.5. Gel electrophoretic analysis of pBAD<sub>rha-hyf</sub> constructs.** Electrophoresis was performed using a 0.7% agarose gel. M, 1 kb DNA ladder (Fermentas). Lanes 1-6, pBAD<sub>rha-hyf</sub> plasmids digested with *NdeI* and *BamHI*, isolated from six distinct transformants. Expected fragments sizes were ~6 and ~13 kb).

Two of each of the pBAD<sub>rha-hyc</sub> and pBAD<sub>rha-hyf</sub> plasmids (sample 1 and 3 in each case) were subjected to nucleotide sequencing by Source Bioscience using pBAD<sub>rha-F</sub> and

pBAD<sub>rha</sub> - R primers (Table 2.4) (Appendix 7 & 10) plus 17 specific forward and 17 specific reverse primers for the *hyf* operon (Table 2.5), and 10 specific forward and 11 specific reverse primers for the *hyc* operon (Table 2.6) (see Appendix for summary of the sequencing data- Appendix 8 & 11). The sequences obtained were compared with the sequence database using DNASTAR which confirmed that the inserts have the correct sequence and are correctly located at the desired cloning sites. These plasmids were employed in future studies, as described below.

### **5.3. Can pBAD<sub>rha</sub>-encoded *hyf* produce hydrogen in a *hyc* mutant background?**

#### **5.3.1 The effect of pH and formate on hydrogen production in the wildtype, $\Delta hyf$ and $\Delta hyc$ mutants.**

The availability of *hyc* and *hyf* in an inducible vector allowed experiments aimed at complementation of a *hyc* mutant to proceed. A well-established phenotype for *hyc* mutants of *E. coli* is lack of hydrogen production under fermentative conditions at acidic pH (Meeda *et al.*, 2007; Noguchi *et al.*, 2010). Thus, the ability of the pBAD<sub>rha</sub> constructs to reverse this activity was tested. Initially, the effect of various pH levels (6.5, 7 and 7.5) on hydrogen production was tested on the *E. coli* parental strain (BW25113) and the isogenic mutant strains (note, both the *hycE* and *hyfG* mutants used throughout this thesis are deletions with the *kan* cassette removed, thus limiting any polarity effect on downstream expression; Table 2.2).

Noguchi *et al.* (2010) used three different pH levels, 5.5, 6 and 6.5, to check the hydrogen yield in the wildtype and a  $\Delta hycE$ , and found that as the pH level reduces from 6.5 to 5.5 in the wildtype, the production of hydrogen increased 70 fold, but in a *hycE* mutant strain hydrogen production was hardly detected compared to the wildtype. However, Mnatsakanyan *et al.* (2004) found less than a twofold increase in hydrogen production

from pH 7.5 to pH 5.5. The difference between the two studies might be due to the media employed: Noguchi *et al.* (2010) used peptone medium with no glucose, whereas Mnatsakanyan *et al.* (2004) used the same medium with 0.2% glucose, which might cause a repression in *hyc*. These findings thus show that *hyc* produces more hydrogen at pH below 7. Thus, it is possible that Hyf may also be subject to the pH dependent effects on its activity. In addition, the previous work on the effect of pH on hydrogen production did not separate *hyc* expression effects from any influence of pH on Hyd-3 enzyme activity. This is something that can be explored here using the pBAD<sub>rha</sub>-*hyc* plasmid.

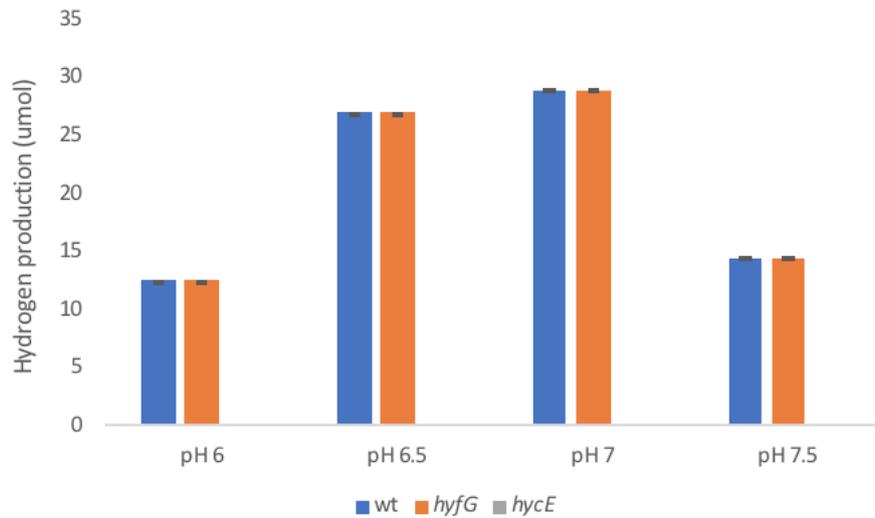
To test the effect of different pH levels on hydrogen production by *E. coli*, H<sub>2</sub> production was compared in the wildtype and the  $\Delta hycE$  and  $\Delta hyfG$  mutants under anaerobic conditions in rich medium at a range of pH values (6, 6.5, 7 and 7.5) with 10 mM glucose (section 2.5.3) and 120 mM MOPS. Overnight cultures were inoculated into fresh medium to give a final OD<sub>600</sub> of 0.01 at pH 6, 6.5, 7 and 7.5 in 10 ml syringes to provide an anaerobic environment (section 2.16); growth was for 24 h at 37 °C. All growths were performed in triplicate, and were repeated twice with similar results obtained. The results shown are representative of one of three repeats.

The results showed that there was no hydrogen production in the *hycE* mutant strain (Fig 5.6), although hydrogen was produced in the wildtype and the *hyfG* mutant at each of the different pH levels tested. Hydrogen was produced in equivalent amounts for these two strains at all pH values tested (Fig. 5.6). However, the levels of hydrogen generated were approx. twice as high at pH 6.5 and 7, than at pH 6 and 7.5. Yoshida *et al.* (2005) also reported a maximal H<sub>2</sub> production at pH 6.5-7 with ~twofold lower levels at pH 6 and 7.5 - as observed here. The results thus indicate that in the absence of *hycE*, which encodes the large subunit of Hyd-3, there is no hydrogen production at any of the pH values tested, as expected and as reported before (Noguchi *et al.* 2010; Mnatsakanyan *et*

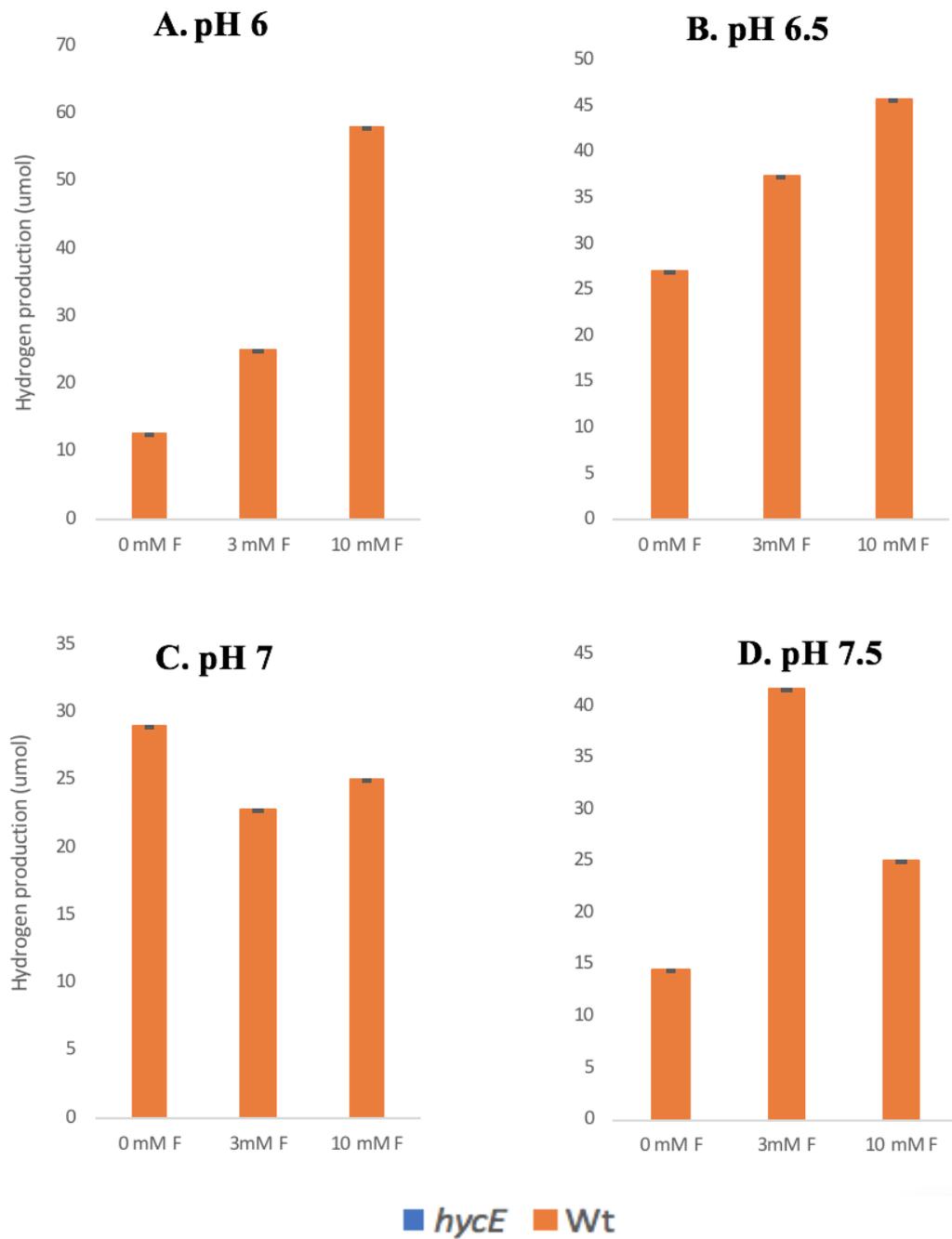
*al.* 2004). On the other hand, the *hyfG* mutation had no apparent effect on total hydrogen production which indicates, as suggested in previous work, that the Hyf system does not contribute to H<sub>2</sub> production (under the conditions tested) due to the silent nature of the *hyf-focB* operon (Skibinski *et al.*, 2002; Self *et al.*, 2004). It should be noted that this finding is not in agreement with the report of Trchounian and Trchounian (2013), which indicated that Hyf is biologically active in hydrogen production at pH > 7 in the presence of glucose.

Previous studies indicate that as the formate concentration increases to a maximum of 100 mM, the production of hydrogen by the wildtype also increases (Sawers, 2005; Yoshida *et al.*, 2005). Thus, the experiment was repeated as above but with addition of different concentrations of formate (3 and 10 mM) but without the *hyfG* mutant since this strain gave a similar result to that of the wildtype in the previous experiment (Fig 5.6). The *hycE* mutant strain showed no production of hydrogen with formate, as was the case in the absence of formate (Fig. 5.7). However, the wildtype showed increased hydrogen production as the levels of exogenous formate increased at all pH values used (except pH 7). At pH 6, 3 and 10 mM formate increased hydrogen production by 3.3 and 3.7 fold ( $P = 0.02$ ), respectively (Fig 5.7). At pH 6.5, 3 mM formate had little effect on hydrogen production but 10 mM formate increased hydrogen production by 2.3 fold ( $P = 0.036$ ) (Fig 5.7). At pH 7, formate addition up to 10 mM did not affect the amount of hydrogen released (Fig 5.7). However, at pH 7.5, formate at 3 mM increased hydrogen production by 2.8 fold ( $P = 0.02$ ), although at 10 mM the increase was only 1.7 fold ( $P = 0.03$ ) (Fig 5.7). Thus, the above results generally match those of previous work indicating that a roughly neutral pH favours maximal total hydrogen production and that increased formate levels stimulate greater hydrogen production (Yoshida *et al.*, 2005). The one notable exception was the lack of any major effect of formate at up to 10 mM at pH 7.

The reason for this lack of formate stimulation is unclear but it is possible that H<sub>2</sub> production would be stimulated at higher formate concentrations since Yoshida *et al.* (2005) reported increases in hydrogen production with increasing formate at up to 100 mM.



**Figure 5.6. Effect of  $\Delta hycE$  or  $\Delta hyfG$  mutation on hydrogen production at different pH levels (pH 6, 6.5, 7 and 7.5).** The wildtype (BW25113), and the  $\Delta hycE$  and  $\Delta hyfG$  mutants (BW25113-A-3 and BW25113-A-4, respectively) were grown in rich medium containing (section 2.5.3) 10 mM glucose and 120 mM MOPS for 24 h at 37 °C under anaerobic conditions, in triplicate. Data are the average of three technical repeats with error bars indicating standard deviation; the experiment was repeated twice with similar results obtained (not shown).



**Figure 5.7. Effect of formate on hydrogen production in *E. coli* wildtype and  $\Delta hycE$  mutant at pH 6-7.5.** Details are as for Fig. 5.6 except for the use of 3 or 10 mM formate.

### 5.3.2 The effect of *hyc* or *hyf* induction on H<sub>2</sub> production in the wildtype and $\Delta hycE$ mutant.

To explore the role of *hyf* in hydrogen production and to compare it with that of *hyc*, the pBAD<sub>rha</sub> vectors expressing *hyc* and *hyf*, along with the vector control, were used to transform the wildtype and the  $\Delta hycE$  mutant. The resulting pBAD<sub>rha</sub>-*hyc*, pBAD<sub>rha</sub>-*hyf* and pBAD<sub>rha</sub> transformants were grown anaerobically, as above, but with chloramphenicol (50 µg/ml) and with or without 0.02% w/v rhamnose as an inducer (Fig. 5.8).

Results showed that induction of *hyf* or *hyc* in the  $\Delta hycE$  strain results in the restoration of hydrogen production under most conditions tested (Fig 5.8). At pH 6, induction of *hyf* in the  $\Delta hycE$  mutant resulted in hydrogen levels similar to those of the wildtype vector control whereas induction of *hyc* gave hydrogen levels at just 50% of those of the control. Nevertheless, induction of both *hyf* and *hyc* successfully complemented the  $\Delta hycE$  mutant, albeit that *hyf* did so more effectively. At pH 6 without inducer, the non-induced *hyf* operon gave ~twofold higher H<sub>2</sub> levels than the corresponding wildtype vector control (37.4 µmol versus 20.8 µmol, respectively) – levels that were 50% higher than those achieved with inducer; the non-induced *hyc* operon also allowed hydrogen production in the  $\Delta hycE$  strain, but to a far lower level than seen in the wildtype vector control (0.1 versus 0.5 ml, respectively).  $P = 0.013$  values. Interestingly, non-induced *hyf* also raised H<sub>2</sub> production in the wildtype (from 0.5 ml for the vector control, to 0.9 ml), which contrasts with the effect seen with non-induced *hyc* and for the induced forms of *hyf* and *hyc* in the wildtype where reductions in hydrogen production were observed. In summary, the results at pH 6 show that *hyf* encodes an active H<sub>2</sub> producing hydrogenase that can compensate for absence of Hyd-3 and can raise the overall levels of hydrogen production in the wildtype. However, this effect requires synthetic expression of *hyf*

operon since it appears that the chromosomal version is not well expressed under the conditions tested.

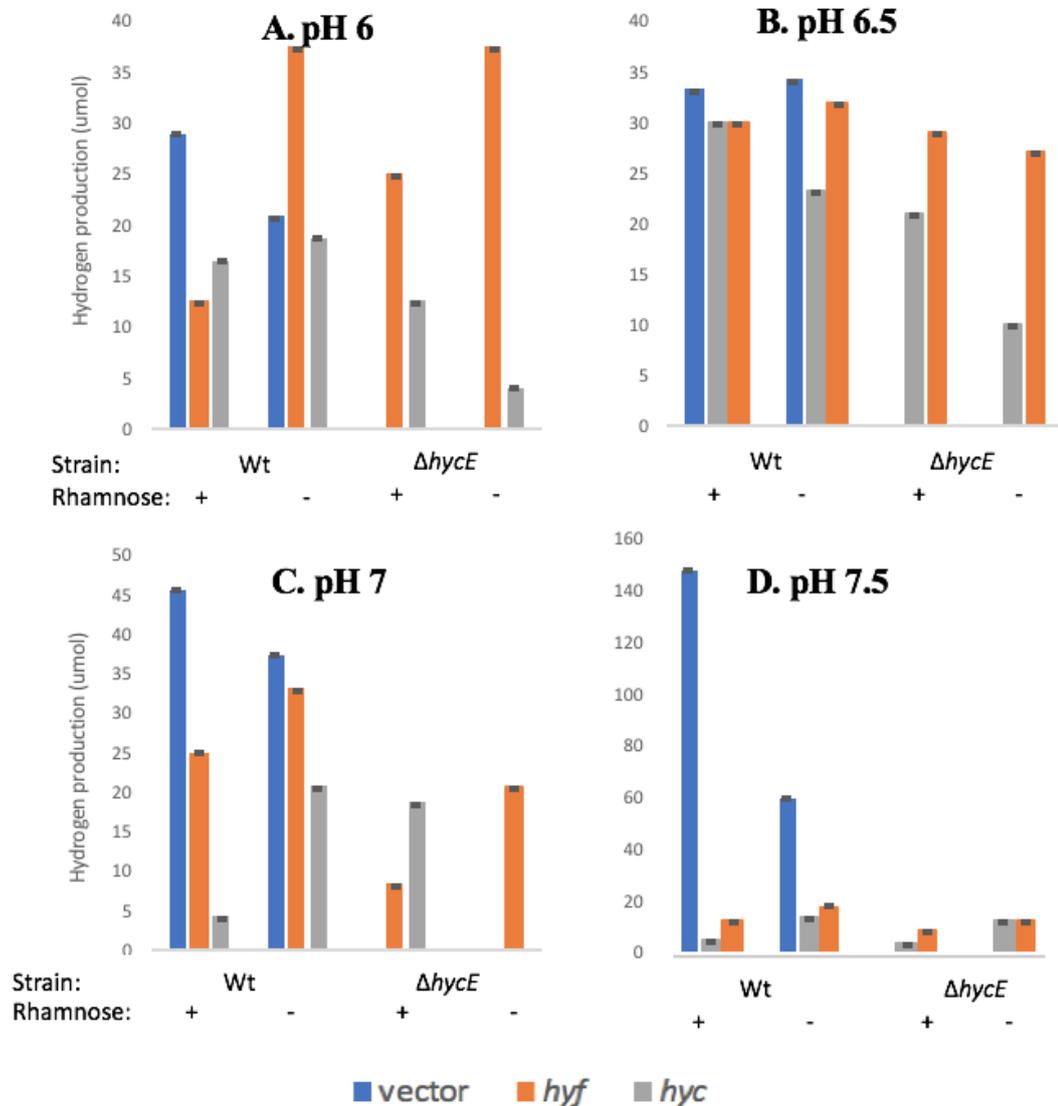
At pH 6.5, similar results were obtained to those at pH 6 except that the overall level of H<sub>2</sub> production by the wildtype was lower (26.2 versus 29 μmol for the wildtype vector control without inducer at pH 6.5 and 6, respectively); a similar pH dependent effect on hydrogen production was seen above (Fig. 5.7). *hyf* complementation of the *hycE* mutant gave higher H<sub>2</sub> production than *hyc* complementation, with or without inducer, and the effect of *hyf* complementation was little affected by inducer at pH 6.5. However, the ability of the *hyc* plasmid to complement the *hycE* mutant was enhanced by the presence of the inducer, by ~twofold. Levels of hydrogen produced with inducer were similar for the *hyf* and *hyc* complemented strains as seen in the vector control. Thus again, *hyf* was able to effectively complement the absence of Hyd-3 in terms of hydrogen production, and did so more effectively than *hyc* in wildtype. The induced *hyf* and *hyc* plasmids either reduced hydrogen production in the wildtype or had little effect, i.e. in no case did the *hyf* or *hyc* plasmids provide increased H<sub>2</sub> production in comparison to the corresponding wildtype vector control at pH 6.5.

At pH 7, hydrogen production was again relatively high (as observed above; Fig. 5.7) in comparison to that seen at pH 6 and 7.5 in the wildtype vector controls. Induction of *hyf* or *hyc* resulted in restoration of hydrogen production in the *hycE* mutant, as seen at pH 6 and 6.5. However, in this case the level of H<sub>2</sub> production was greater for *hyc* induction than for *hyf* induction, by ~twofold. The difference between H<sub>2</sub> production in the wildtype vector control and the complemented mutant (with inducer) was greater at this pH than any other explored, with the wildtype vector control generating 45.7 μmol H<sub>2</sub> gas with inducer compared to just 8.3 and 18.7 μmol for the *hycE* mutant complemented with *hyc* or *hyf* (with inducer). The uninduced *hyf* plasmid restored H<sub>2</sub> production to the

*hycE* strain to a greater degree than the induced *hyf* plasmid; however, the uninduced *hyc* plasmid failed to enable H<sub>2</sub> production in the *hycE* strain. Of particular note is the major (11-fold) reduction in hydrogen production caused by induction of the *hyc* plasmid in the wildtype, whereas; induction of *hyf* caused just a twofold reduction in H<sub>2</sub> production. This suggests that excess levels of Hyc components disrupt hydrogen production, might effect the Hyp apparatus, which prevent hydrogenase maturation, and consequently affect H<sub>2</sub> production, although it is unclear why such an effect would be stronger at pH 7 cf. pH 6, 6.5 and 7.5.

At pH 7.5, the overall H<sub>2</sub> production level was higher than that at pH 7 by 2.9 and 1.2 fold consistent with the results in Fig 5.7 above. Induction of both the *hyf* and *hyc* plasmids again restored H<sub>2</sub> production in the *hycE* mutant, to similar degrees, but did not restore production to the same level as in the wildtype vector control (5 and 12.4 μmol H<sub>2</sub> cf. 148 μmol, respectively). Unlike the *hyc* plasmid, the uninduced *hyf* plasmid also restored hydrogen production in the *hycE* mutant (12 μmol H<sub>2</sub>), to a similar degree than that of the induced *hyf* plasmid.

In summary, from these results, is clear that the plasmid-borne *hyf* complements the  $\Delta hycE$  mutant strain in allowing the production of hydrogen. This effect was seen at all pH values tested (pH 6-7.5). Induced *hyf* (in the *hycE* strain) gave greater H<sub>2</sub> levels than induced *hyc* at acid pH (6-6.5) whereas induced *hyc* gave the greater H<sub>2</sub> production at pH 6.5-7. This may indicate a more important physiological role for Hyf at lower pH with an opposite function for Hyc; although it should be noted that chromosomally encoded Hyd-3 enables greater H<sub>2</sub> production at pH 6.5-7 than pH 6 or 7.5, which matches the results reported in Fig 5. 8 showing that *hyc* induction in the *hycE* strain results in nearly twofold more H<sub>2</sub> at pH 6.5-7 than at pH 6 or 7.5.



**Figure 5.8. Plasmid induced *hyc* or *hyf* enables production of hydrogen gas in a  $\Delta hycE$  mutant.** Conditions were as for Fig. 5.6 (Rich medium with 10 mM glucose and 120 mM MOPS). The strains were grown anaerobically for 24 h at 37 °C in triplicate. Data is the average of three technical repeats with error bars indicating standard deviation; the experiment was repeated twice with similar results obtained. **A:** pH 6.0; **B:** pH 6.5; **C:** pH 7.0 and **D:** pH 7.5. ‘+’ indicates with rhamnose; ‘-’ indicates without rhamnose (inducer). Strains were BW25113 and the  $\Delta hycE$  mutant carrying pBAD<sub>rha</sub> (vector), pBAD<sub>rha</sub>-*hyc* (*hyc*) or pBAD<sub>rha</sub>-*hyf* (*hyf*), as indicated.

**Table 5.1. pH values difference before and after growth related to Fig. 5.8. ‘+’ indicates with rhamnose induction. ‘-’ indicates without rhamnose induction.**

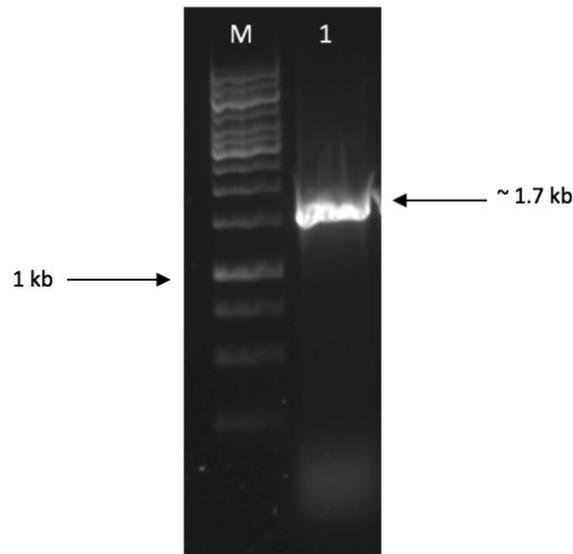
strain		Strating pH	pH after 24 h			
Wildtype	Vectot+	Vector -	6	5.8	5.8	
	<i>hyf</i> <sup>+</sup>	<i>hyf</i> <sup>-</sup>		5.8	5.92	
	<i>hyc</i> <sup>+</sup>	<i>hyc</i> <sup>-</sup>		5.83	5.9	
$\Delta$ <i>hycE</i>	Vectot+	Vector -		5.1	5.3	
	<i>hyf</i> <sup>+</sup>	<i>hyf</i> <sup>-</sup>		5.8	5.8	
	<i>hyc</i> <sup>+</sup>	<i>hyc</i> <sup>-</sup>		5.76	5.79	
Wildtype	Vectot+	Vector -		6.5	6.3	6.31
	<i>hyf</i> <sup>+</sup>	<i>hyf</i> <sup>-</sup>			6.4	6.43
	<i>hyc</i> <sup>+</sup>	<i>hyc</i> <sup>-</sup>			5.48	6.47
$\Delta$ <i>hycE</i>	Vectot+	Vector -	5.7		5.5	
	<i>hyf</i> <sup>+</sup>	<i>hyf</i> <sup>-</sup>	6.4		6.43	
	<i>hyc</i> <sup>+</sup>	<i>hyc</i> <sup>-</sup>	6.3		6.37	
Wildtype	Vectot+	Vector -	7		6.8	6.83
	<i>hyf</i> <sup>+</sup>	<i>hyf</i> <sup>-</sup>			6.82	6.82
	<i>hyc</i> <sup>+</sup>	<i>hyc</i> <sup>-</sup>			6.81	6.8
$\Delta$ <i>hycE</i>	Vectot+	Vector -		6.3	6.2	
	<i>hyf</i> <sup>+</sup>	<i>hyf</i> <sup>-</sup>		6.88	6.9	
	<i>hyc</i> <sup>+</sup>	<i>hyc</i> <sup>-</sup>		6.85	6.87	
Wildtype	Vectot+	Vector -		7.5	7.3	7.4
	<i>hyf</i> <sup>+</sup>	<i>hyf</i> <sup>-</sup>			7.4	7.41
	<i>hyc</i> <sup>+</sup>	<i>hyc</i> <sup>-</sup>			7.38	7.4
$\Delta$ <i>hycE</i>	Vectot+	Vector -	6.5		6.3	
	<i>hyf</i> <sup>+</sup>	<i>hyf</i> <sup>-</sup>	7.42		7.4	
	<i>hyc</i> <sup>+</sup>	<i>hyc</i> <sup>-</sup>	7.4		7.4	

### 5.3.2.1. Is *hyf* complementation of the $\Delta$ *hycE* mutant achieved by *hyfG* replacement of *hycE* function?

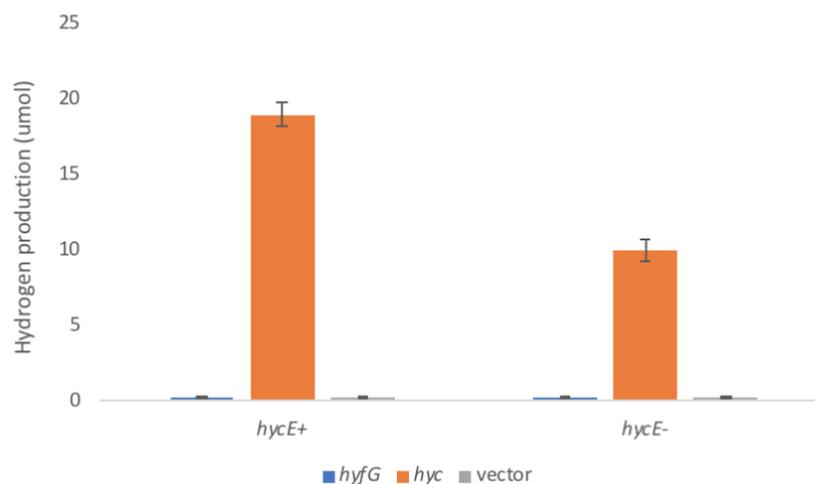
The above study shows that the *hyf* operon enables production of H<sub>2</sub> under anaerobic conditions in a  $\Delta$ *hycE* mutant background when induced from pBAD<sub>rha</sub>. Since previous studies showed that there is 73% identity between *hycE* and *hyfG* (Andrews *et al.*, 1997),

this raises the possibility that the *hyfG* gene of the *hyf* operon may replace *hycE* in the *hycE* mutant strain upon *hyf* induction such that the observed hydrogen production might be largely *hyc* dependent, not *hyf*. The possibility was thus tested.

PCR amplification of the ~1.7 kb *hyfG* (section 2.10.1) gene was achieved using primers listed in Table 2.10 using Amp HiFi PCR premix DNA polymerase (Clontech). The amplified fragment was analysed by agarose gel electrophoresis (Fig. 5.9) and found to have the desired size; this was thus used for cloning into pBAD<sub>rha</sub>. The resulting plasmid was subjected to nucleotide sequencing by Source Bioscience using pBAD<sub>rha</sub>-F and pBAD<sub>rha</sub>-R (Table 2.4). The sequence obtained was compared with BLAST, which confirmed that the insert has the correct sequence and is correctly located at the cloning sites. This plasmid was transformed into the  $\Delta hycE$  mutant strain. The  $\Delta hycE$  was also complemented with *hyc* and the vector. The transformants were grown in rich medium (section 2.5.3) with 120 mM MOPS at pH 6.5 with and without inducer (0.02% w/v rhamnose) and 50 ug/ ml chloramphenicol, anaerobically for 24 hours. Results showed (Fig 5.10) that complementation with *hyfG* did not result in production of H<sub>2</sub>, as was the case for the vector control. However, the *hycE* mutant complemented with plasmid encoded *hyc* did produce H<sub>2</sub> with inducer (more than 2.1 fold greater than that achieved without inducer). This result clearly suggests that *hyfG* does not replace *hycE* in the *hycE* mutant, and the H<sub>2</sub> which was produced in section 5.3.2. upon *hyf* induction in the  $\Delta hycE$  strain is entirely Hyf, not Hyc, dependent.



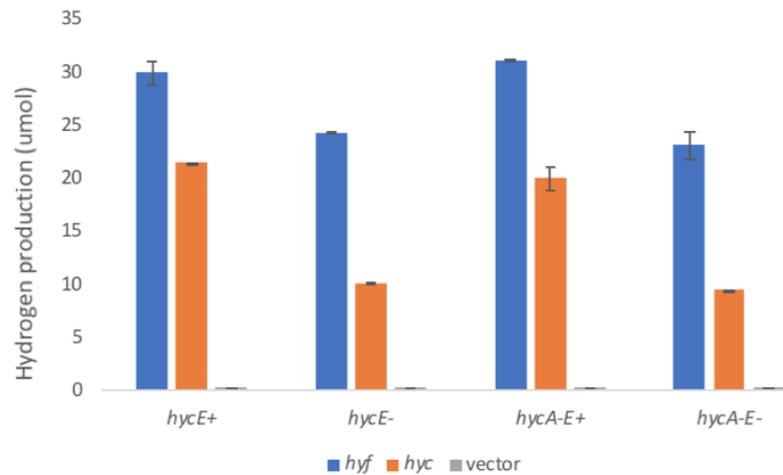
**Figure 5.9. Agarose gel (0.7%) electrophoretic analysis of *hyfG* DNA.** M, Marker 1 kb Generuler (Fermentas); lane 1, 2  $\mu$ l of *hyfG* DNA PCR product using Amp HiFi PCR premix DNA polymerase. Arrows indicate position of DNA bands.



**Figure 5.10. Hydrogen production upon *hyc* and *hyfG* induction along with vector control in  $\Delta$ *hycE* mutant.** Conditions were as for Fig. 5.6. Data are average of three technical replicates with error bars indicating standard deviation; the experiment was repeated twice with similar results obtained. '+' indicates with rhamnose; '-' indicated without rhamnose (inducer). Strains were in a  $\Delta$ *hycE* mutant background carrying pBAD<sub>rha</sub> (vector) as control, pBAD<sub>rha</sub>-*hyfG* and pBAD<sub>rha</sub>-*hyc*, as indicated.

Another experiment was also conducted to prove that Hyf is the source of H<sub>2</sub> production seen in the *hyf*-complemented *hycE* mutant, and is not Hyc/hyfG dependent. The pBAD<sub>rha</sub> plasmids expressing *hyc* and *hyf*, along with the vector control, were used to transform the  $\Delta hycE$  and  $\Delta hycA-E$  mutants. The resulting transformants were grown under anaerobic environment as above. The results (Fig 5.11) show that induction of *hyf* in the *hycE* or *hycA-E* mutant backgrounds restored H<sub>2</sub> production. It is clear that H<sub>2</sub> was produced in similar concentrations (~30  $\mu$ mol), which is higher than the uninduced, strains by 1.2 fold in  $\Delta hycE$  mutant and 1.3 fold in  $\Delta hycA-E$  mutant backgrounds. This shows also that H<sub>2</sub> production, upon *hyf* induction, is more than that seen for *hyc* induction in the  $\Delta hycE$  and  $\Delta hycA-E$  strains by 1.4 and 1.5 fold, respectively. No detectable H<sub>2</sub> was generated for the vector control, with or without inducer, as expected. This result demonstrates that the *hyf* operon allows H<sub>2</sub> production in a strain where most of Hyc subunits are absent. This clearly indicates that Hyf is able to produce H<sub>2</sub> even in the absence of the Hyc system. However, the *hycI* encoded HycI protease might play a role in Hyd-4 maturation, since it is available in the background strain; this possibility should be checked in the future.

Since H<sub>2</sub> levels upon induction of *hyf* in the  $\Delta hycE$  or  $\Delta hycA-E$  strains are similar, and *hyfG* does not replace *hycE*, only  $\Delta hycE$  will be used as the *hyc* mutant strain in all further experiments.



**Figure 5.11. Hydrogen production upon induction of *hyc* or *hyf* in the  $\Delta hycE$  and  $\Delta hycA-E$  mutants.** Conditions were as in Fig 5.10. ‘+’ indicates with rhamnose; ‘-’ indicates without rhamnose (inducer)

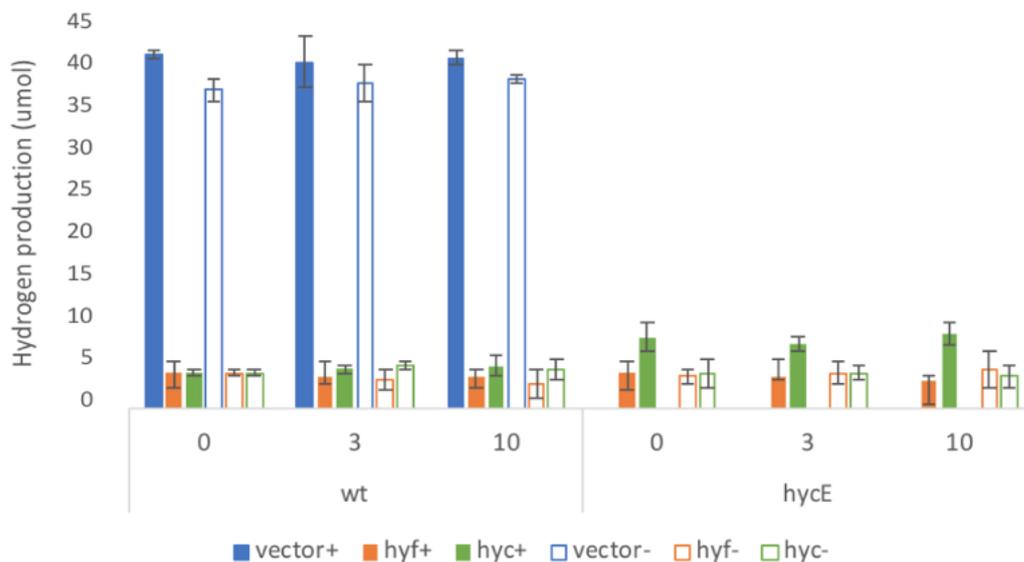
#### 5.4. Does formate stimulate H<sub>2</sub> production by Hyf?

It has been shown previously and above (Fig. 5.7) that formate addition increases the amount of Hyd-3-dependent hydrogen production in the wildtype (Sawers, 2005; Yoshida *et al.*, 2005). To test this possibility for Hyf-dependent hydrogen production, and for Hyd-3 under conditions where *hyc* expression is independent of FhlA (the formate-dependent transcriptional regulator), the *hyc* and *hyf* plasmids were induced in wildtype and *hycE* mutant backgrounds at different added formate concentrations (3 and 10 mM) and pH levels (5.5, 6.5 and 7.5), using conditions as described above.

At pH 5.5, hydrogen was produced by the wildtype, but not the  $\Delta hycE$  mutant. However, formate had little impact on hydrogen production for the wildtype. The presence of the *hyc* or *hyf* plasmids, with or without inducer, resulted in a major reduction in H<sub>2</sub> production (by approx. tenfold; Fig. 5.12). The reason for this is unclear, but such an effect was not seen at pH 6.5. Both the *hyc* and *hyf* plasmids weakly reversed the lack of H<sub>2</sub> production in the *hycE* mutant strain at this pH (with or without inducer), but as indicated above, the levels of production achieved were far lower (5-10 fold) than in the

wildtype vector control, and indeed were similar to those obtained in the wildtype carrying the *hyf* and *hyc* plasmids. Thus, this result suggest that at pH 5.5 the vector-borne *hyc* and *hyf* operons mediate production of Hyc and Hyf components that results in a limited Hyd-3 and Hyd-4 activity, and indeed such production considerably interferes with endogenous Hyd-3 function. Induced *hyc* gave an approx. twofold higher H<sub>2</sub> production in the *hycE* mutant than did induced *hyf*, indicating that at pH 5.5 (unlike pH 6 and 6.5), induction of *hyc* results in a greater hydrogenase activity than induction of *hyf*.

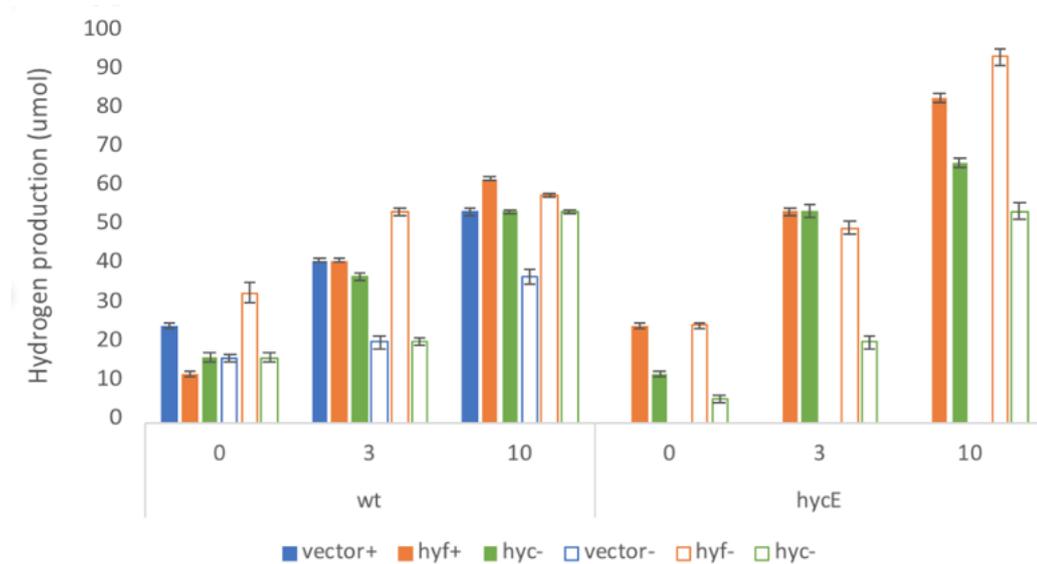
In summary, of particular note is the lack of formate-dependent induction of hydrogenase activity at pH 5.5 in all cases, which was somewhat unexpected given previous results reported (Noguchi *et al.* 2010; Yoshida *et al.*, 2005).



**Figure 5.12. Hydrogen production upon induction of *hyc* or *hyf* in a wildtype or  $\Delta hycE$  background, at pH 5.5.** The conditions are as in Fig 5.8, but growth was for 30 h. ‘+’, with rhamnose and ‘-’, minus rhamnose, and with 0, 3 or 10 mM formate (as indicated by 0, 3, 10). The experiment was performed three times and similar results obtained. Final OD values were similar in all cases 0.21 and 0.37 OD<sub>600</sub>.

At pH 6.5 (Fig. 5.13), the level of hydrogen production was lower than at pH 5.5, in the absence of formate, by ~twofold. As before, the uncomplemented *hycE* mutant failed to produce hydrogen. At this pH, as before (Fig. 5.7), formate clearly enhanced hydrogen production in the wildtype (vector controls), with a ~twofold induction seen with 10 mM formate. Provision of the *hyc* or *hyf* plasmid had little impact on H<sub>2</sub> production in the wildtype, nor the response to formate. In the *hycE* mutant, complementation with either *hyc* or *hyf* allowed formate-induced hydrogen production, with levels of hydrogen generated with formate exceeding those seen for the wildtype vectors controls by up to 40%. For the induced *hyf* plasmid in the *hycE* mutant, H<sub>2</sub> production increased from 25 to 83.2 umol with 10 mM formate (a 3.3 fold increase), and for the induction of the *hyc* plasmid in the *hycE* mutant, 10 mM formate increased H<sub>2</sub> production from 12.5 to 66.6 umol (a 5.3 fold increase). Thus, at pH 6.5, Hyf-dependent H<sub>2</sub> production, like that of Hyd-3, is clearly stimulated by formate. Further, the *hyf* plasmid in the *hycE* mutant gave a greater level of H<sub>2</sub> production (on average, by ~twofold) than did the *hyc* plasmid, in nearly all cases (Fig. 5.13).

In summary, these results show that the formate dependence of Hyd-3 activity is not solely related to its transcriptional control by FhlA. The mechanism by which formate levels influence the Hyd-3/Hyd-4 activities at pH 6.5, as measured here, could simply be related to availability of formate as a substrate for its enzymatic conversion into hydrogen and CO<sub>2</sub>. But in addition, formate availability might affect expression of other genes required for Hyd-3/4 activity e.g. the FhlA dependent *fdhF* gene. Further work would be required to test these alternative explanations.



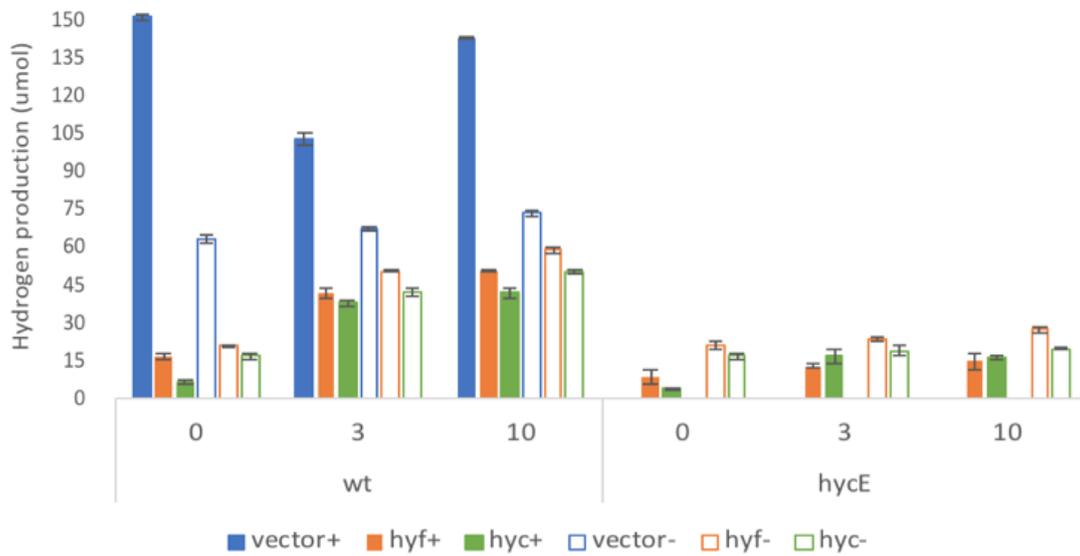
**Figure 5.13. Hydrogen production upon induction of *hyc* or *hyf* in a wildtype or  $\Delta hycE$  background at pH 6.5.** The conditions are same as Fig 5.9, except for the pH employed. Final OD values were similar in all cases 0.5 and 0.37 OD<sub>600</sub>.

At pH 7.5, the wildtype vector control produced 3.6-6 fold more hydrogen than at pH 5.5 or 6.5 (Fig. 5.14), and addition of formate had little impact on the amount of hydrogen produced by the wildtype vector control. This is in contrast to the results obtained with the wildtype (without vector) in Fig. 5.7 where formate induced H<sub>2</sub> production – this difference may relate to the presence of the plasmid and/or antibiotic. The *hycE* mutant failed to generate any hydrogen, either with or without formate (as seen above). Upon complementation with the *hyc* or *hyf* plasmids, hydrogen production was restored, both with and without formate, but at levels much lower about (30-10 fold) than seen for the wildtype vector controls. Indeed, the *hyc* and *hyf* plasmids appeared to lower the amount of hydrogen generated in the wildtype by 4.1-3.8 fold, which resembles the effect seen at pH 5.5. Slightly better H<sub>2</sub> production was seen in the *hyc/hyf*-complemented *hycE* strain without inducer, suggesting that weak *hyc* and *hyf* expression is more beneficial in terms

of hydrogen production than strong expression at this pH. Although the H<sub>2</sub> production of the wildtype vector control showed no response to formate, most of the other strain/plasmid combinations displayed an enhanced H<sub>2</sub> production with formate, with the *hycE* mutant carrying induced *hyf* or *hyc* giving a 2- or 4-fold increase (respectively) in response to formate (Fig. 5.14). Thus, this effect resembles the formate-induction of H<sub>2</sub> production seen above for the wildtype at pH 7.5 (Fig. 5.7). This induction effect was not seen in the absence of rhamnose. Interestingly, the response to formate was particularly marked for the wildtype carrying the *hyc* or *hyf* plasmids, with increases in H<sub>2</sub> production of between 2.6 and 9.7 fold observed with formate. Whether this effect reflects a simple substrate-dependent increase in H<sub>2</sub> production or a formate gene induction effect (e.g. for *fdhF*), as discussed above, is unclear.

In summary, both Hyf and Hyd-3 activity respond to formate levels at pH 7.5, as seen at pH 6.5, when the plasmid-borne *hyc* and *hyf* operons are provided in the *hycE* mutant in the presence of inducer.

Thus, the above results demonstrate a formate induction of both Hyd-3 and Hyf (Hyd-4) dependent H<sub>2</sub> production at pH 6.5 and 7.5, but not at pH 5.5.



**Figure 5.14. Hydrogen production upon induction of *hyc* or *hyf* in a wildtype or  $\Delta hycE$  background at pH 7.5.** The conditions is same as Fig 5.8, except for the pH employed. Final OD values were similar in all cases of 0.2 OD<sub>600</sub>.

## 5.5. Effect of *hyc/hyf* status on formate production and consumption.

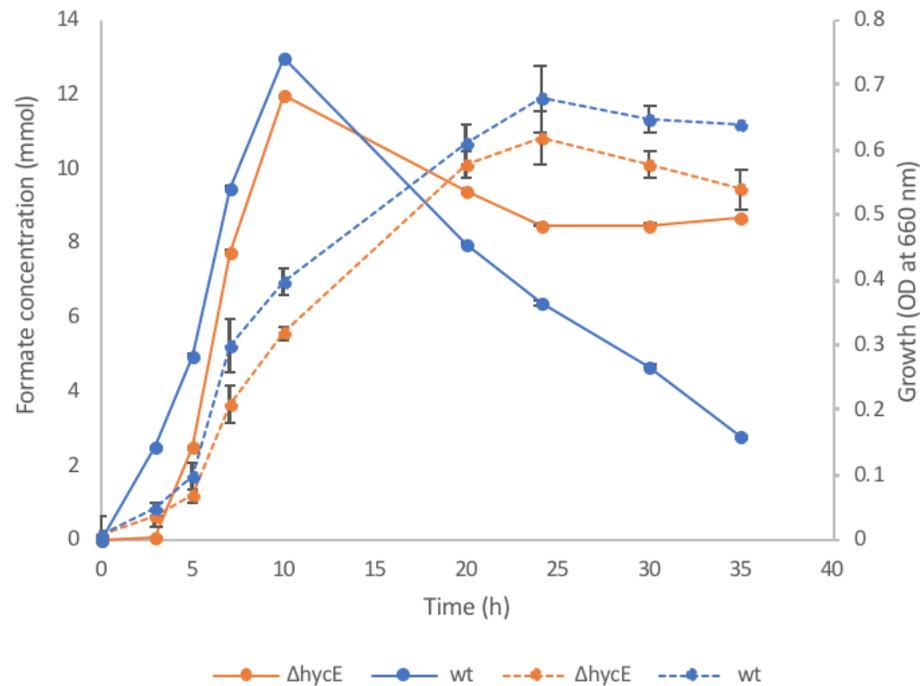
### 5.5.1. The *hycE* mutant.

To confirm a role for Hyd-3 in mediating the consumption of formate during anaerobic growth and its correlation to hydrogen production, the levels of formate were measured during the growth of the *hycE* mutant and the wildtype (section 2.15). Growth was in WM medium (section 2.5.2) plus 80 mM glucose with 0.8% trace elements and 120 mM MOPS (section 2.5.2.1), in 20 ml plastic syringes to provide an anaerobic condition (section 2.16).

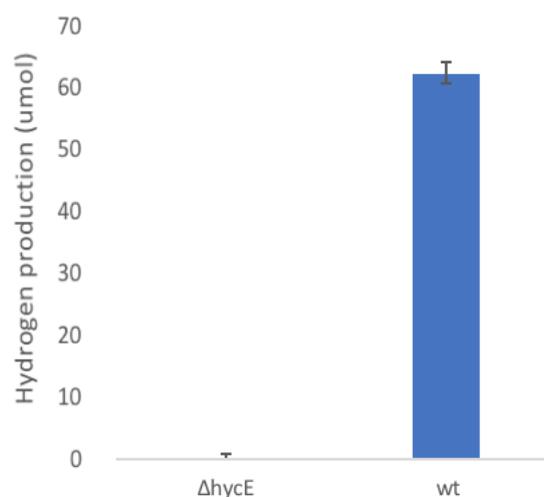
The two strains grew similarly and formate concentration in the medium increased during the first 10 h of growth for both the wildtype and the  $\Delta hycE$  mutant, to reach a peak level (that was similar for both strains) at the mid-point of the growth curve. However, growth

was slightly weaker for the *hycE* mutant and the formate release was also slightly less than that seen for the wildtype (possibly reflecting the lower growth). After 10 h, formate levels for the wildtype steadily decreased over the next 25 h to reach a level that was ~fourfold lower than that measured at 10 h (Fig. 5.15). However for the *hycE* mutant, the decrease in formate after 10 h was far less marked than that seen for the wildtype, with a reduction of just 1.1 fold over the following 25 h, giving a final formate level that was 1.3-fold higher than that obtained for the wildtype (Fig. 5.15). Interestingly, at 3 h, the amount of formate released into the medium for the *hycE* mutant was 25 fold less than that for the wildtype, although at 10 h the formate levels were only 1.08-fold less (Fig. 5.15). As expected, the *hycE* mutant failed to generate hydrogen (Fig. 5.16).

The results thus show a major reduction in formate consumption for the *hycE* mutant, which resembles results previously showing reduction in formate consumption by a *hyc* mutant as *hyc* is a main part of FHL, which consume formate to produce H<sub>2</sub> (Sawers, 1994).



**Figure 5.15.** The effect of *hycE* mutation on formate levels during fermentative growth in WM-medium at pH 6.5. Solid lines indicates the formate concentration (mmol) while the dotted line indicates bacterial growth (660 nm). The growth was anaerobic in WM-medium at 37 °C, 80 mM glucose, with 0.8% trace elements and 120 mM MOPS pH 6.5. BW25113 (wt), and BW25113  $\Delta hycE$  were the strains employed. Growths were performed in triplicate and standard deviations are indicated. The experiment was repeated twice, with similar results obtained.



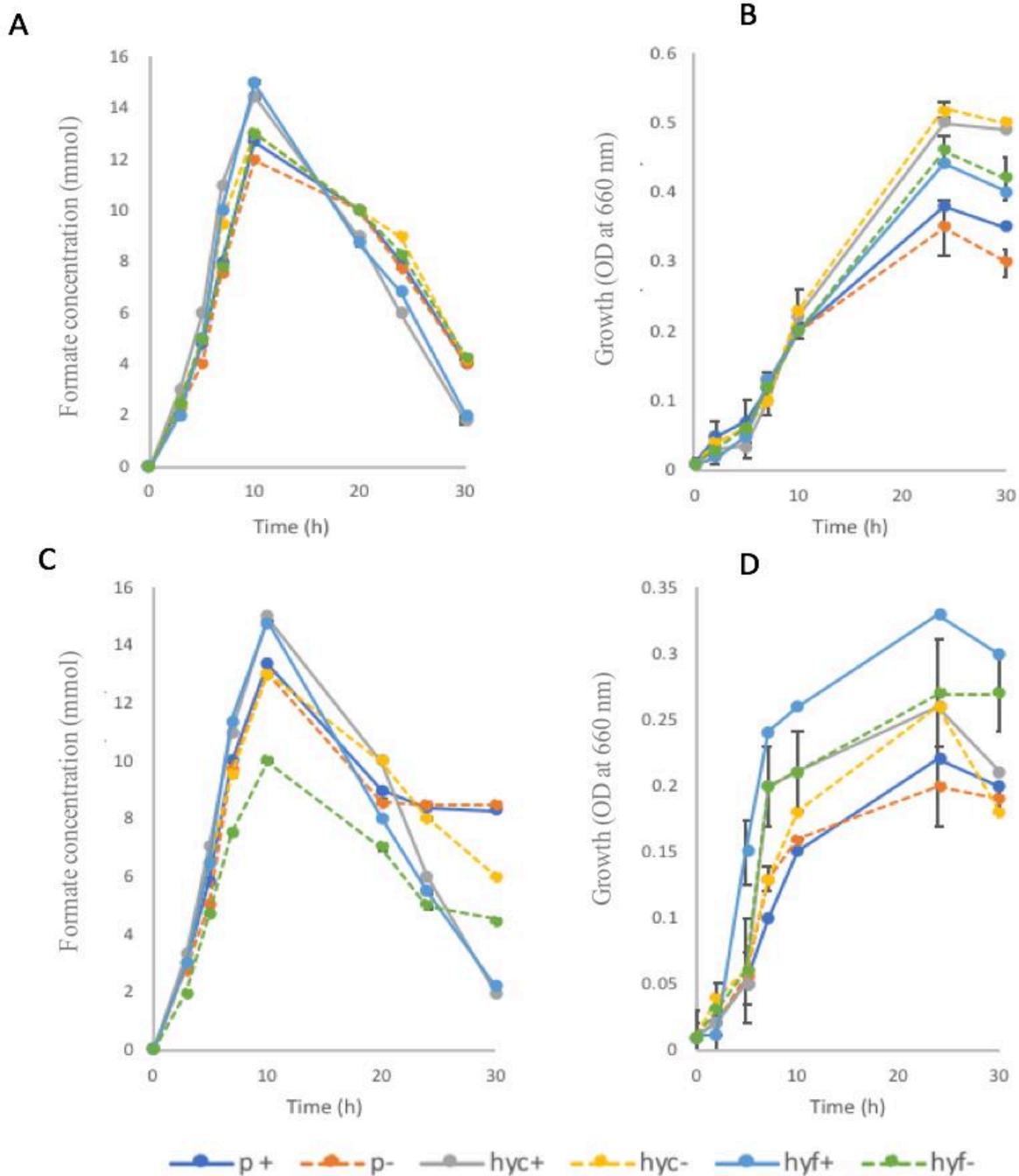
**Figure 5.16.** Hydrogen production in *hycE* mutant compared to wildtype at pH 6.5. Measurements were made from the experiment presented in Fig. 5.12.

### 5.5.2 Effect on formate consumption of complementation of the *hyc* mutant with pBAD<sub>rha</sub>-*hyc* or -*hyf*.

Experiments were performed as above, but at pH 6.5 and 7.5, and with the *hyc* and *hyf* plasmids, to determine whether the inducible *hyc* and *hyf* genes can restore formate consumption in the *hycE* mutant strain. Results showed that a similar overall pattern of formate production occurs to that seen, with a peak level at 10 h in all cases (Figs. 5.17 & 5.18). In the wildtype, induction of *hyc* and *hyf* at pH 6.5 resulted in higher formate production at 7 and 10 h, with 27 and 15% higher levels respect to the vector control (Fig. 5.17). This indicates that induction of *hyc* or *hyf* increases formate release at pH 6.5 in the wildtype, an effect that could be related to growth since the induction of *hyc* and *hyf* also resulted in a higher growth, which was particularly marked at 24 h (30 and 16% increased, respectively, cf. the wildtype vector control). However, the uninduced strains showed similarly increased growth without any notable effect on formate consumption, suggesting that the increased formate levels of the induced *hyc* and *hyf* strains is not caused by their increase in growth. Despite the higher formate levels at 10 h for the induced *hyc* and *hyf* wildtype, at 30 h the levels of formate were lower than those seen in the vector control (and uninduced wildtype) by ~twofold. These results thus indicate that Hyd-3 and Hyf contribute to formate consumption, and that raised expression of either *hyc* or *hyf* in the wildtype increases formate consumption capacity.

In the  $\Delta hycE$  mutant, induction of *hyc* or *hyf* resulted in a slightly raised formate level at 10 h, as seen in the wildtype. Importantly, *hyc* or *hyf* induction also resulted in a major increase in formate consumption with levels dropping to 2 mM at 30 h compared to 8 mM for the vector control (Fig. 5.17C), a fourfold decrease. Thus, induction of either *hyc* or *hyf* resulted in restoration of formate consumption and indeed enhanced consumption of formate with respect to that seen in the wildtype vector control (Fig.

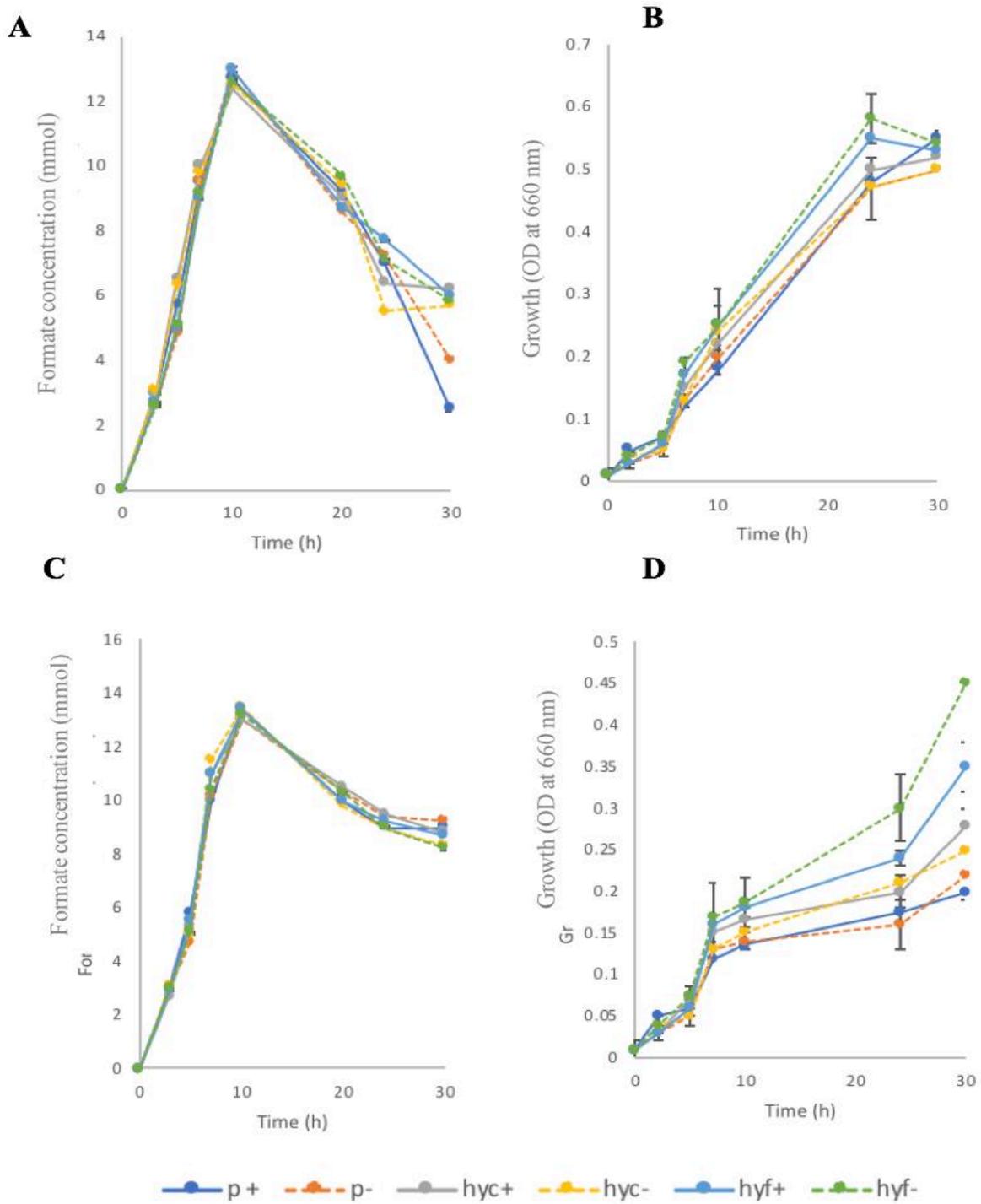
5.17A) by ~twofold. From this result, it is clear that Hyf, like Hyc, contributes to formate consumption and thus likely forms a second FHL complex (Fhl-2) with Fdh-H. Induction of *hyc* or *hyf* in the  $\Delta hycE$  strain resulted in increased growth with respect to the induced-vector control by 1.3 and 1.8 -fold at 10 h and 1.2 and 1.6-fold at 24 h, respectively ( $P = 0.015$  at 24 h  $P = 0.01$  at 10 h and  $P = 0.035$  and  $P = 0.04$  at 24 h, respectively). This indicates that increased capacity to degrade formate, as afforded by Hyd-3 and Hyd-4, enhances growth under fermentative conditions presumably by protecting the cell from formate-induced toxicity. The amount of formate consumed upon *hyc* and *hyf* induction was similar, yet the level of hydrogen generated was 50% greater for *hyf* than for *hyc* induction (Fig. 5.13) which indicates that there is not always a direct correlation between apparent levels of formate consumed and H<sub>2</sub> generated. The higher growth of the *hyc* and *hyf* induced *hycE* mutant strains could also indicate a role for Hyd-3 and Hyd-4 in energy generation. Indeed, the greater growth (24% at 24 h) achieved with *hyf* than with *hyc* induction is consistent with the view that Hyd-4, unlike Hyd-3, enables an energy conserving FHL-1 reaction.



**Figure 5.17. Effect of complementation with *hyc* or *hyf* on formate production and consumption in a wildtype or *hycE* mutant at pH 6.5.** Conditions are as for Fig. 5.12. **A** and **B**, wildtype plus *hyc* (**A**) or *hyf* (**B**) inducible plasmids, or the vector only (*p*), with (+) or without inducer (-). **C** and **D**, as **A/B**, except for the use of the  $\Delta hycE$  mutant in place of the wildtype. **A** and **C**, growth; **B** and **D**, corresponding formate levels. Growths were performed in triplicate and standard deviations are indicated. The experiment was repeated twice, with similar results obtained.

At pH 7.5, the presence of the *hyc* and *hyf* plasmids in the wildtype (induced or uninduced) resulted in a reduced consumption of formate (~2.2 and 2.8 fold,  $P = 0.01$ , more formate at 30 h for the *hyc* and *hyf* induced wildtype cf. the induced vector control; Fig. 5.18A) which reflects the reduced levels of hydrogen generated (Fig. 5.14). This supports the suggestion that at pH 7.5 the induction of *hyc* or *hyf* causes a major perturbation in FHL function. In the *hycE* mutant at pH 7.5, the *hyc* and *hyf* plasmids failed to compensate for the reduced formate consumption of the mutant – levels of formate at 30 h were very similar in all cases at ~8 mM (Fig. 5.18C). This lack of apparent formate consumption via FHL reflects the 15-30 fold lower H<sub>2</sub> production observed for the *hyc* and *hyf* induced *hycE* mutant cf. the wildtype vector control (Fig. 5.14). Interestingly, induction of *hyf* or *hyc* in the *hycE* mutant at pH 7.5 (Fig. 5.18D) resulted in a growth enhancement (a 1.4 and 1.2 fold greater growth, respectively, cf. the induced vector control at 24 h;  $P = 0.02$ ). The reason for this effect is unclear.

In summary, the results show that the *hyc* and *hyf* operons encode hydrogenases that support formate consumption, as well as hydrogen production. At pH 7.5, induction of *hyc* or *hyf* results in a weakened FHL activity indicating that the Hyc and Hyf components thus produced interfere with endogenous FHL activity. This effect could be caused by an excess of hydrogenase complex subunits that swamp Hyp pathways and/or block proper assembly of the FHL complex. The results also support the suggestion that Hyd-4 forms a second FHL complex (FHL-2) catalysing the conversion of formate to H<sub>2</sub> and CO<sub>2</sub>. In addition, the enhanced growth observed with Hyf cf. Hyc is consistent with the notion that FHL-2 is more energy conserving than FHL-1 at pH 6.5 as illustrated.



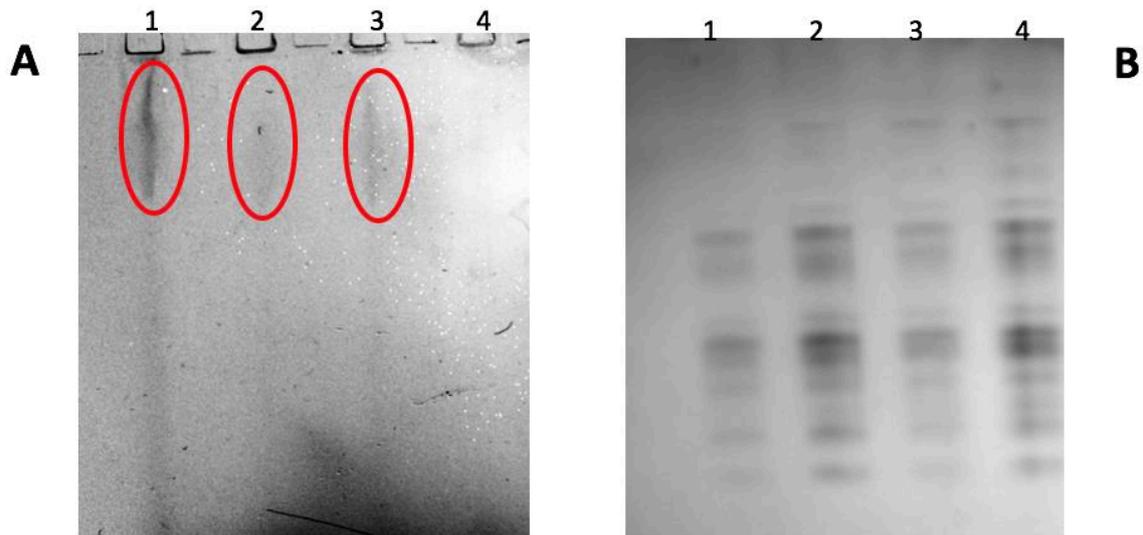
**Figure 5.18.** Effect of *hyc* or *hyf* induction on formate production and consumption in a wildtype or *hycE* mutant at pH 7.5. Details are as for Fig. 5.14 except for the pH employed

## 5.6. Identification of Hyd-4 activity by gel activity staining

The above study shows that Hyf is able to mediate the consumption of formate and the production of hydrogen under anaerobic fermentative conditions when induced from pBAD<sub>rha</sub>, and a similar effect was observed for Hyc, with best results seen at pH 6.5. Since previous work showed Hya (Hyd-1) and Hyb (Hyd-2) activity as hydrogen consuming hydrogenases in *E. coli* using a gel activity staining method (Sauter *et al.*, 1992), and Hyc activity as a hydrogen producing hydrogenase was detected by activity staining by Pinske *et al.* (2012), the possibility that such an activity could be shown for Hyf was tested. The wildtype and  $\Delta hycE$  mutant strains complemented with the *hyc* and *hyf* encoding plasmids were thus grown anaerobically in TGYEP medium with 120 mM MOPS at pH 6.5 (section 2.18) 5h with inducer, and about 25  $\mu$ g of protein was subject to non-denaturing PAGE with Triton X-100 (7.5% w/v polyacrylamide). The gel was stained in a mixture of benzyl viologen (BV) and triphenyl tetrazolium chloride (TTC) under 100% hydrogen atmosphere for ~10 min until the bands appeared. The hydrogenase assay was repeated three times to confirm the results (section 2.18.3; Fig 5.19A).

Fig. 5.19A shows clearly that a major hydrogenase activity band was produced at the top of the gel for the wildtype and for both the *hyf*- and *hyc*-plasmid bearing  $\Delta hycE$  mutant strains. However, nothing appeared for the  $\Delta hycE$  strain at this position, which acts as the negative control. The intensity of the band (reduced benzyl viologen) is higher in the *hyf*- and *hyc*-complemented strains than for the wildtype, indicating a higher production of Hyf and Hyc for the pBAD<sub>rha</sub>-*hyc* and -*hyf* transformants cf. the wildtype presumably due to the multicopy nature of the plasmids. Thus, the above results supports the view that Hyf is an active (H<sub>2</sub>-evolving) hydrogenase, and the fourth hydrogenase of *E. coli*, but its activity is only exhibited when *hyf* is induced from pBAD<sub>rha</sub> with no activity

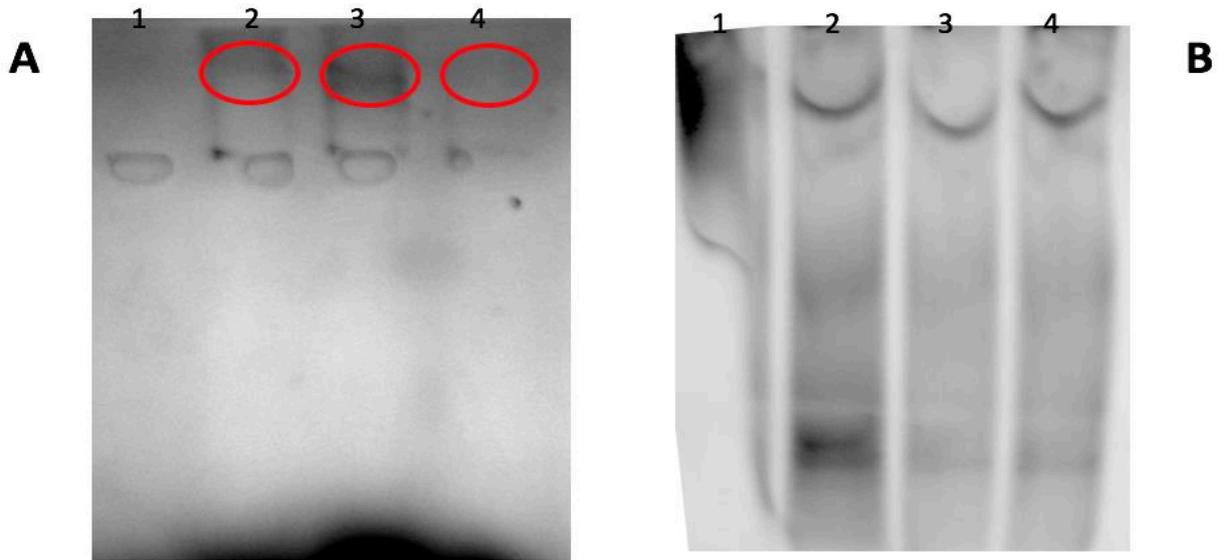
derived from the chromosomally-encoded *hyf* operon under the conditions tested.



**Figure 5.19. Analysis of Hyd-4 hydrogenase activity by activity-staining and native PAGE in Tris-glycine buffer.** **A.** Hydrogenase activity stained gel. Total cell extracts (25  $\mu$ g of protein) from each of the strains: 1, pBAD<sub>rha</sub>-*hyf* in the  $\Delta$ *hycE* strain; 2, wildtype; 3, pBAD<sub>rha</sub>-*hyc* in the  $\Delta$ *hycE* strain; and 4,  $\Delta$ *hycE* strain contain pBAD<sub>rha</sub>. The growth conditions and staining protocol were as indicated above. Red circle indicates activity stain observed. **B.** Coomassie blue stained gel confirming quality and separation of proteins by native PAGE.

The hydrogenase activity was repeated using a Tris-barbitone gel system, as described in Methods (section 2.10.8.3) and by Pinske *et al.* (2012) who showed Hyd-3 activity in both gel systems. However, the resolution of the gel (Fig. 5.20B) was much poorer than that achieved above. Nevertheless, the activity staining result achieved was similar to that obtained in Fig. 5.19A, which supports the proposal that *hyf* encodes a hydrogenase capable of oxidising H<sub>2</sub> and deploying the electrons released for the reduction of BV.

It should be noted that Pinske *et al.* (2012) were able to detect Hyd-1 activity in the wildtype under the growth conditions employed here. But this was a minor band that may not be visible here due to poor resolution of the gel.



**Figure 5.20. Analysis of Hyd-4 hydrogenase activity by activity-staining and native PAGE in Tris-barbitone buffer.** Details are as for Fig. 5.16 except for the use of a barbital buffer. Samples were: 1,  $\Delta hycE$  strain contain pBAD<sub>rha</sub>; 2, wildtype; 3, pBAD<sub>rha</sub>-*hyf* in the  $\Delta hycE$  strain; and 4, pBAD<sub>rha</sub>-*hyc* in the  $\Delta hycE$  strain.

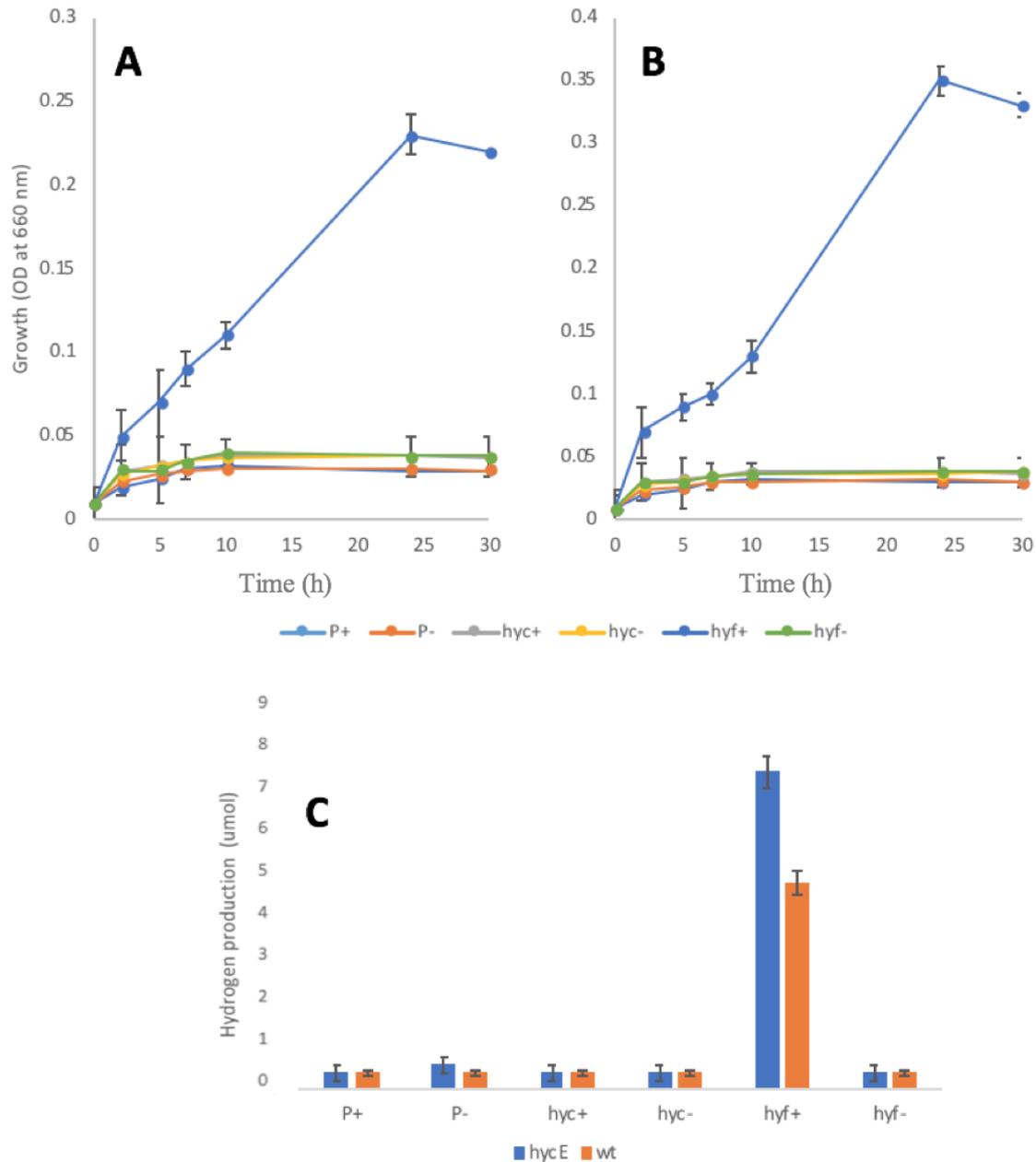
### 5.7. Can *hyf*-expressing strains use formate as an energy source?

The results in section 5.5 suggest that not only can Hyf mediate the consumption of formate and production of hydrogen, it can also enhance growth under fermentative conditions, and does so more effectively than Hyc. In addition, the greater complexity of Hyf with respect to Hyc, and its closer similarity to Complex I, is consistent with a role in FHL activity that is energy conserving through contributing to the pmf across the cytoplasmic membrane (Andrews *et al.*, 1997). Thus, experiments were performed to test whether Hyf can act as an energy conserving hydrogenase through determining its ability to enable fermentative growth with formate as an energy source. The growths were under anaerobic conditions using syringes at 37 °C for 30 h. The *hyf* and *hyc* complemented  $\Delta hycE$  strains were grown in WM-medium, without glucose, but with 10

mM formate as sole energy source (with 0.8% trace elements and 120 mM MOPS at pH 6.5).

For both the wildtype and  $\Delta hycE$  mutant, there was no growth with any of the three plasmids tested except for the condition with pBAD<sub>rha</sub>-*hyf* with inducer. This indicates that expression of the *hyf* operon enables *E. coli* to utilise formate to support growth. Hydrogen was also generated upon induction of *hyf*. Formate consumption by formate dehydrogenases would usually generate energy through formation of NADH for utilisation within the respiratory chain or by direct delivery of electrons to the respiratory chain (Unden & Bongaerts, 1997). Under the conditions used here, there is no respiratory acceptor so such the aforementioned modes of energy generation cannot be deployed. Thus, the best remaining option would be that Hyf is able to support a pmf through an energy-conserving FHL pathway, as suggested previously (Andrews *et al.*, 1997). However, formate is unlikely to serve as a carbon source for anabolic processes since it is converted to CO<sub>2</sub>. However, under induction conditions, 0.02% (1.2 mM) rhamnose is present and this sugar can be utilised by *E. coli* as sole energy and carbon source (Badia *et al.*, 1989). Thus, it is possible that the growth observed in Fig. 5.21 is driven by the presence of rhamnose as well as formate.

Thus, in summary, the results here provide preliminary evidence for a role for Hyf in utilising formate to yield energy for growth, but further work is required to confirm this proposal.



**Figure 5.21. Effect of *hyf* and *hyc* induction on growth with 10 mM formate.** Strains were grown anaerobically in WM- medium without glucose and with 0.8% trace elements and 120 mM MOPS at pH 6.5. Strains used were the wildtype (BW25113) and the isogenic  $\Delta hycE$  mutant. Strains carried either pBAD<sub>rha</sub>, pBAD<sub>rha-hyf</sub> and pBAD<sub>rha-hyc</sub>. **A.** The growth in wild type. The data in Y-axis are linear. **B.** The growth in *hycE* mutant. The data in Y-axis are linear. **C.** Hydrogen production at 30 h. Growths were performed in triplicate and standard deviations are indicated. The experiment was repeated twice, with similar results obtained. Rhamnose was used at 0.02% for induction.

## 5.8. Discussion

In this chapter, the ability of Hyd-4 to generate hydrogen and its formate dependency were explored. The multi-copy-number plasmid, pBAD<sub>rha</sub>, was employed for expression of the 13 kb *hyf* operon; this plasmid allowed induction of *hyf* using rhamnose as an inducer. The resulting pBAD<sub>rha</sub>-*hyf* plasmid was then used to complement a  $\Delta hycE$  strain lacking Hyd-3 activity and the transformant was tested for hydrogen production at different pH levels. It was clearly found that hydrogen production was restored when *hyf* was induced in the  $\Delta hycE$  strain at all pH levels tested. The highest hydrogen production was at pH 6.5 for both *hyc* and *hyf* induction. However, hydrogen was not produced at all in the  $\Delta hycE$  strain carrying the vector control. This suggests that Hyd-4 is a hydrogen producing hydrogenase, as is Hyd-3 at acidic pH (Maupin and Shanmugam, 1990; Bohm *et al.*, 1990), and thus is likely to act as part of a second FHL enzyme along with Fdh-H. Previously, research reported that H<sub>2</sub> production by Hyd-3 is promoted by addition of formate (Yoshida *et al.*, 2005) up to a concentration of 100 mM. Mnatsakanyan *et al.* (2002) also found that hydrogen production is increased (by 3.5 fold) by provision of formate (30 mM). Similar effects were also reported by Pinske and Sargent (2016). Thus, the effect of formate addition on H<sub>2</sub> production was tested for the Hyf system at a range of pH levels since pH has also been shown to affect H<sub>2</sub> production in *E. coli* (Yoshida *et al.*, 2005; ). Induction of *hyc* and *hyf* at pH 5.5 and 7.5 in the  $\Delta hycE$  mutant resulted in relatively little H<sub>2</sub> production (Fig. 5.12; Fig 5.14). However, at pH 6.5, H<sub>2</sub> production was 2.5-fold higher than pH 5.5 and as formate increased to 10 mM, hydrogen production also increased, by 6-fold for Hyf and by 5-fold for Hyc compared to the control (no formate added). No formate induction of H<sub>2</sub>-evolving hydrogenase activity was observed at pH 5.5. This increase in H<sub>2</sub> with formate addition was unrelated to FhlA-dependent (or HyfR-dependent) induction of *hyc* or *hyf* expression, it seems likely that the effect

observed was caused by the raised availability of substrate for FHL activity. This formate dependence provides further evidence for a role for Hyf as a second FHL complex. The work in this chapter therefore provides the first convincing evidence that Hyf is capable of formate-dependent H<sub>2</sub> production and is consistent with the notion that it combines with Fdh-H to form a second FHL. However, it remains unclear under what physiological condition *hyf* would be expressed in a wildtype strain, as previous work indicates that it is not active under fermentative conditions that stimulate *hyc* expression. Also, it remains unclear why two such similar H<sub>2</sub>-evolving hydrogenases would be required by *E. coli*.

To determine whether Hyf can mediate formate consumption, the levels of formate were monitored in anaerobic fermentative cultures of the  $\Delta hycE$  strain in the presence of the induced *hyf* or *hyc* systems. After an initial build-up of formate over the first 10 h at pH 6.5 to give a peak level of ~15 mM, formate declined steadily in the presence of either Hyc or Hyf to ~2 mM the following 20 h (Fig. 5.17C). For the vector controls, formate levels remained relatively high (at ~13 mM) over this period with little consumption observed. Thus, like Hyc, Hyf mediates the consumption of formate produced during fermentative growth on glucose, and can substitute for Hyc when synthetically expressed. This observation provides further evidence that Hyf functions as part of a second FHL pathway in the conversion of formate to H<sub>2</sub> and CO<sub>2</sub>.

Induction of *hyf* gave greater H<sub>2</sub> production levels than did induction of *hyc* at acid pH (6-6.5), whereas *hyc* induction produced more H<sub>2</sub> than for *hyf* at higher pH (7-7.5) (Fig. 5.14), suggesting that Hyf activity is more favoured by lower pH than is Hyc activity.

Preliminary growth experiments with formate as sole energy source under fermentative growth conditions surprisingly revealed that provision of the *hyf* operon in the induced state resulted in substantial growth for both the *hycE* mutant and wildtype (Fig. 5.21). It is suggested that this might be enabled by the presence of the inducer, rhamnose, with

could offer an anabolic carbon source for growth purposes whilst the Hyf system generates energy from formate consumption. However, it would be expected that any energy generation from the FHL reaction would require consumption of the H<sub>2</sub> generated in order to maintain a low H<sub>2</sub> partial pressure, as suggested previously (Andrews *et al.*, 1997). Thus, this result requires further investigation.

Pinske *et al.* (2012), successfully managed to detect Hyd-3 activity by staining native polyacrylamide gels containing *E. coli* extracts grown fermentatively. The activity staining method used BV as an artificial electron acceptor (with TTC as mediator) and H<sub>2</sub> gas as the electron donor for the detection all three hydrogenases. With 100% H<sub>2</sub> gas, only Hyd-1 and -3 were detected in the wildtype – with Hyd-1 a relatively minor band at the bottom of the gel and Hyd-3 a major band at the top. In this chapter, the activity staining method of Pinske *et al.* (2012) was used to detect Hyc and Hyf activity (Fig. 5.19). Although the quality of the resulting gel was poor such that Hyd-1 could not be detected, clear bands were seen at the top of the gel in the tracks with extracts from the Hyf<sup>+</sup> and Hyc<sup>+</sup> strains indicating that Hyf, like Hyc, has the capacity to utilise H<sub>2</sub> as an electron source for reduction of BV, which is consistent with its function as a hydrogenase.

In summary, results provided in this chapter indicate that *hyf* encodes a hydrogen-producing hydrogenase that is part of a formate consuming pathway and that it forms a second FHL system (FHL-2) in *E. coli*. Further work is required to confirm that Hyf activity is dependent on FdhF, as is suggested; this will be explored in the following chapter.

## Chapter 6: Role of FdhH in Hyf activity

### 6.1. Introduction

#### 6.1.1. Hyd-3 as a part of the FHL complex

In *E. coli*, Hyd-3 comprises of six proteins as discussed (section 1.5.2). HycC and HycD are arranged as a membrane domain, while HycBEF and HycG as a cytoplasmic domain (Fig 6.1). Addition of formate dehydrogenase-H (encoded by *fdhF*, which is not a part of *hyc* operon) creates the final formate hydrogenlyase complex-1 (FHL-1). The five subunits FdhF-HycBEFG operate as a closed electron transport system connecting formate oxidation to proton reduction without any evidence for proton translocation (just proton consumption). However, Sauter *et al.* (1992) revealed that deletion of either *hycC* or *hycD* encoding the membrane component result in inactive FHL-1 suggesting that membrane association is important for FHL-1 function. The whole *hyc* operon was engineered to encode affinity-tagged of all the subunits by McDowell *et al.* (2014). This method was used as a natural expression level to detect biosynthesis pathways of FHL-1. It was clear from their findings that all seven subunits are important for the enzyme (FHL-1) activity, including HycC and HycD. Sargent (2016) suggested that the final attachment of Hyd-3 to the membrane is essential for the activation of the enzyme complex (Fig 6.1). However, it remains unclear why FHL-1 is membrane associated.

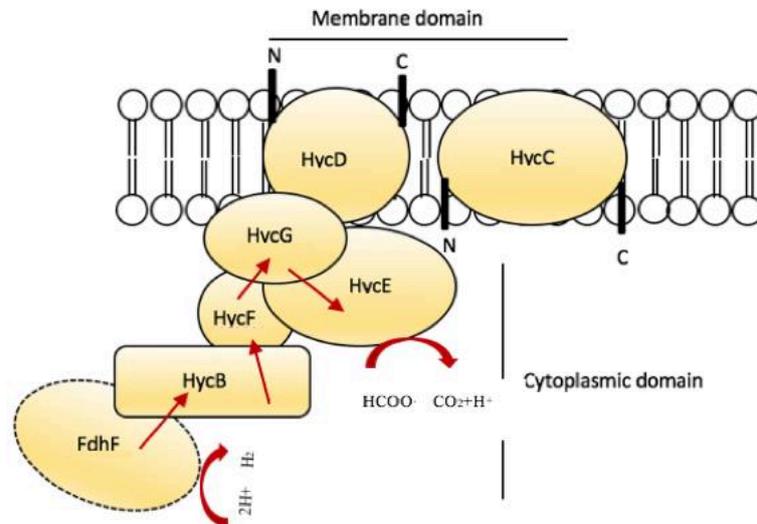
#### 6.1.2. Formate dehydrogenase-H

Formate dehydrogenase H, which is a component of FHL, is hydrogen-linked and was historically named as Fdh-H (Peck *et al.*, 1957). This protein is encoded by *fdhF*, which is located at 93 min on the *E. coli* chromosome. The first cloning of the *fdhF* gene was by Zinoni *et al.*, 1986, which led to a whole new research field in bacterial physiology.

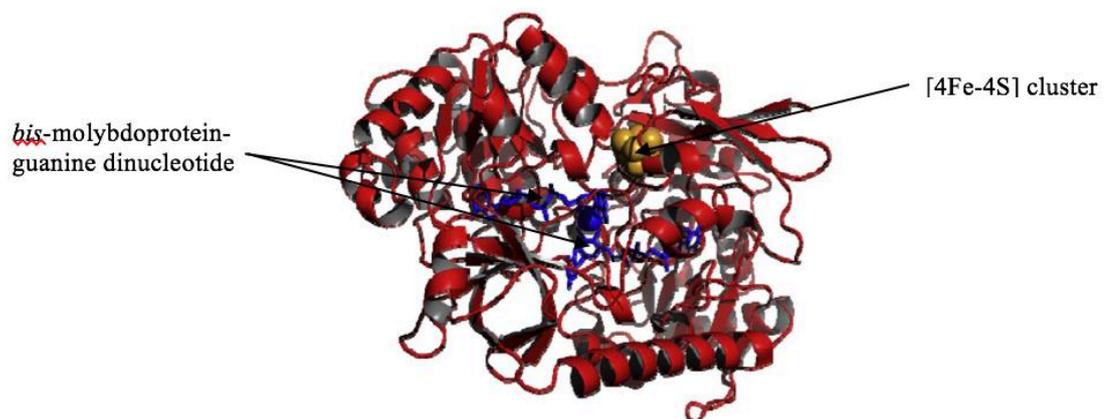
The transcribed *fdhF* gene sequence was found to contain a UGA codon, which is a nonsense codon and is perfectly located at position 140 in the *fdhF* mRNA, near a UGU or UGC cysteine codon. The absence of selenium results in the blockage of translation at the above position and the UGA prevents downstream translation. However, in the presence of selenium, UGA was found to code for selenocysteine required for full activity of formate dehydrogenase in *E. coli*. This incorporation depends on the presence of a unique tRNA species (*selC* product), whose anticodon is UCA which is complimentary to the nonsense codon (Zinoni *et al.*, 1987).

The 80 kDa Fdh-H protein was found to contain a single [4Fe-4S] cluster and a bis-molybdopterin -guanine cofactor, which is a complex between a redox-active molybdenum atom and two cyclic pyranopterins moieties (Mendel and Leimkuhler, 2015). The crystal structure of FdhF in *E. coli* is illustrated in Fig 6.2, which confirmed that selenocysteine was a direct ligand to the molybdenum atom at the active site (Boyington *et al.*, 1997).

The FdhF is found to be loosely attached to the FHL-complex (Fig 6.1) (Sauter *et al.*, 1992). It is also unusual when it is compared to other formate dehydrogenases in *E. coli*, where it is the only dehydrogenase to form a complex with Hyd-3 which can react with benzyl viologen (BV) as an artificial electron acceptor with hydrogen as a donor. Furthermore, FHL catabolizes formate to CO<sub>2</sub> and H<sub>2</sub> under anaerobic condition and the absence of exogenous electron acceptor (Bock and Sawers, 1992).



**Figure 6.1. The structure of hydrogenase-3 and FdhF to form the FHL-1 complex.** It comprises of two domains; membrane and cytoplasmic as illustrated. FdhF is formate dehydrogenase-H, which is a selenoprotein that contains a molybdenum cofactor and [4Fe-4S] cluster. It shows that FdhF is only attached loosely to the cytoplasmic domain through its association with the Hyd-3. Red arrow describe the electron pathway (Sargent, 2016).

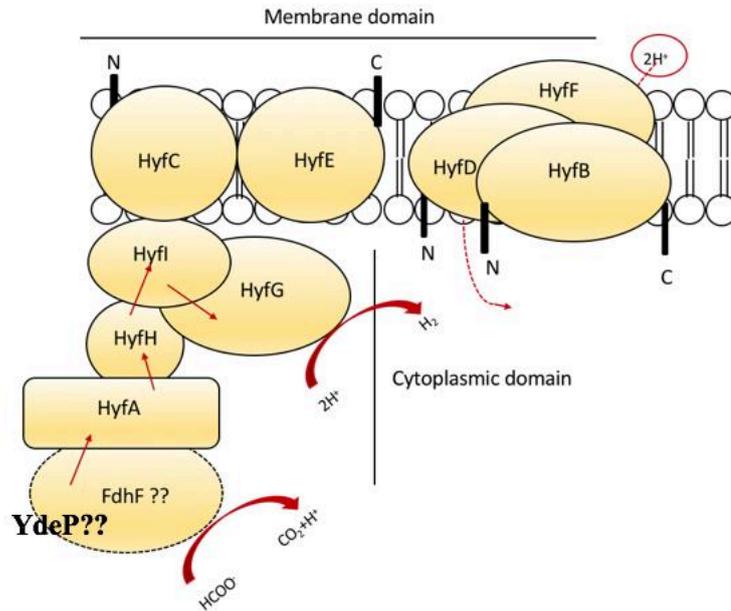


**Figure 6.2. Crystal structure of formate dehydrogenase-H (*fdhF*) of *E. coli*.** The blue colour as indicated shows the *bis*-molybdopterin guanine cofactor and the orange colour illustrates the [4Fe-4S] cluster (Boyington *et al.*, 1997).

### 6.1.3. Formate dehydrogenase-H and hydrogenase-4

The discovery of the fourth hydrogenase in *E. coli* was a surprise to the field (Andrews *et al.*, 1997). Until then, all the biochemical and genetic analysis were able to detect only three active [NiFe]-hydrogenases in *E. coli* (Sawers, 1994). Andrews *et al.* (1997) found that Hyd-4 is closely related to Hyd-3 and it is located at a cytoplasmic and inner membrane, and is related to Complex I. There are several factors that link Hyd-4 to a role as part of a second formate hydrogenlyase complex (FHL-2) (Fig 6.3). The *hyf* operon encodes a gene expressing a putative formate channel (FocB) (Andrews *et al.*, 1997). Indeed, the work in chapter 4 now shows that FocB can function to mediate formate transport. Also, the *hyf* operon got a gene (*hyfR*) that encodes a  $\sigma^{54}$ -dependent regulator (HyR) that drives formate-responsive regulation of the *hyf* operon and is homologous with FhlA (Skibinski *et al.*, 2002). The *hyf* operon is thus considered to specify a fourth hydrogenase (Hyd-4) which is H<sub>2</sub> evolving (like Hyd-3) and like Hyd-3 forms a complex with Fdh-H to generate a FHL enzyme, designated FHL-2. The finding in chapter 5 that Hyf can replace Hyd-3 in formate consumption and hydrogen production provides further evidence for the existence of the proposed FHL-2 complex.

Furthermore, *E. coli* encodes a gene called *ydeP*, which is homologous to *fdhF*. It is predicted to encode a molybdopterin-containing enzyme but this enzyme is not predicted to possess selenocysteine. YdeP functions in acid resistance environment (Masuda and Church, 2003). Vivijs *et al.* (2015) proposed that since YdeP is similar to FdhF, it could act as a component of the FHL complex in place of Fdh-H. So, whether FHL-2 incorporates FdhF with Hyd-4 or whether it uses YdeP instead remains to be proven since an active version of FHL-2 has not been isolated or characterized (Sargent, 2016), and its dependence on FdhF has not yet been explored.



**Figure 6.3. The structure of hydrogenase-4 and FdhF to form the FHL-2 complex.** An illustration of how FHL-2 might be formed. It comprises of two domains; a large membrane domain and cytoplasmic domain. HyfA which is similar to HycB, so it could be the partner subunit to FdhF. All nine Hyf subunits in form FHL-2 either with FdhF or an alternative enzyme (e.g. YdeP?) (Sargent, 2016).

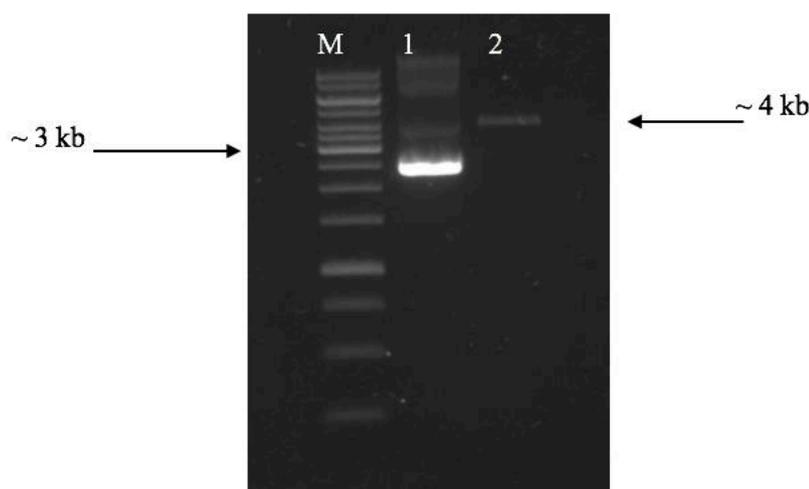
In this chapter, the role of *fdhF* in enabling Hyf-dependent FHL-2 activity was examined.

## 6.2. Cloning of *fdhF* into an inducible vector

In order to enable controlled expression of *fdhF*, the gene is cloned into the pBAD<sub>ara</sub> vector, which is fully compatible with the pBAD<sub>rha</sub> plasmid (distinct *oriV* and antibiotic resistances) (Appendix-3). The pBAD<sub>ara</sub> plasmid allows controlled expression in response to arabinose. In this way, *fdhF* could be tested under any environmental conditions without any impact of regulatory factors (e.g. FHLA) that would usually control its expression according to environmental conditions. Thus, it would be possible to confirm that *fdhF* and *hyf* are able to combine to generate a second FHL.

### 6.2.1. Isolation and digestion of pBAD<sub>ara</sub>

The plasmid pBAD<sub>ara</sub> was obtained from laboratory stocks. pBAD<sub>ara</sub> transformants were then generated and plasmid DNA was isolated from three samples using a GeneJET™ plasmid Miniprep Kit (Fermentas) (section 2.10.7). The isolated plasmids were screened by electrophoretic analysis and the presence of plasmid DNA of high mass corresponding to the expected size (4.1 kb; Fig. 6.4) was confirmed. The identity of the plasmid DNA therefore was confirmed by digestion with *Nco*I and *Hind*III followed by electrophoresis. The double digestion changed the plasmid from the supercoiled form to a single linear of mobility matching that expected for pBAD<sub>ara</sub> (Fig. 6.4).

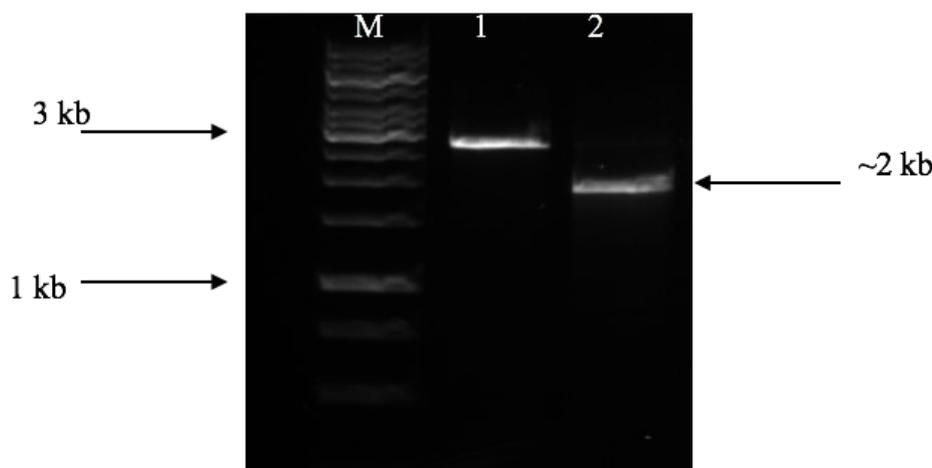


**Figure 6.4.** Agarose gel (0.7%) electrophoretic analysis of pBAD<sub>ara</sub> following restriction digestion with *Nco*I and *Hind*III. M, 1 kb ladder (Fermentas). Lane 1, pBAD<sub>ara</sub> undigested; lane 2, pBAD<sub>ara</sub> double digested.

### 6.2.2. Extraction, amplification and cloning of *fdhF*

Chromosomal DNA was extracted from a wildtype MG1655 (section 2.10.1), and amplified by CloneAmp™ HiFi PCR Premix polymerase from Clontech (section 2.10.2), as it provides accurate and efficient DNA amplification and is recommended for In-fusion PCR cloning (Fig 6.5), using two different primer sets. The first primer set was to confirm

the availability of the gene by choosing 300 bp upstream *fdhF* (pBAD<sub>ara</sub>-*fdhF*-forward) and 300 bp downstream (pBAD<sub>ara</sub>-*fdhF* reverse) (Table 2.8), while the second set were special primers (pBAD<sub>ara</sub>-*fdhF*) forward and reverse designed specifically for use In-Fusion cloning (Gibson methodology) and as described in sections 2.12.2.1 and 2.12.2.2 (Appendix 12). The amplified fragments were analysed by agarose gel electrophoresis and the sizes expected were 2.7 kb (with first set of primers) and 2.1 kb (with the second set), which are a close match to those observed (Fig 6.5). Next, the PCR product was purified (section 2.10.3) and 2 µl of pure PCR product was loaded in an agarose gel electrophoresis (section 2.10.8.1) to confirm its size (Fig 6.6).

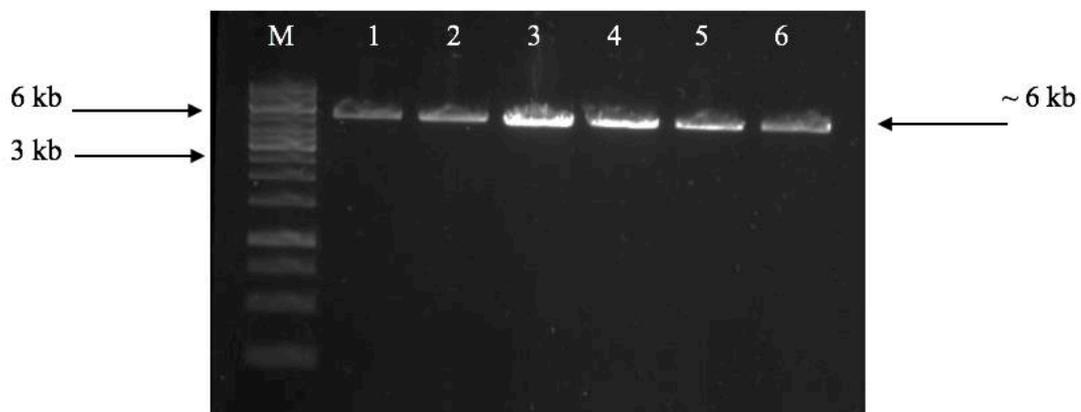


**Figure 6.5. Gel electrophoretic analysis of *fdhF* amplification from MG1655 with two different sets of primers.** Electrophoresis was performed using a 0.7% agarose gel. M is GeneRuler<sup>®</sup> 1 kb Marker. Lane 1, PCR amplified sample of *fdhF* using primers annealing 300 bp upstream and downstream; lane 2, specific primers designed for In-Fusion cloning into pBAD<sub>ara</sub>. 2 µl of samples were loaded and 1 µl of ladder.

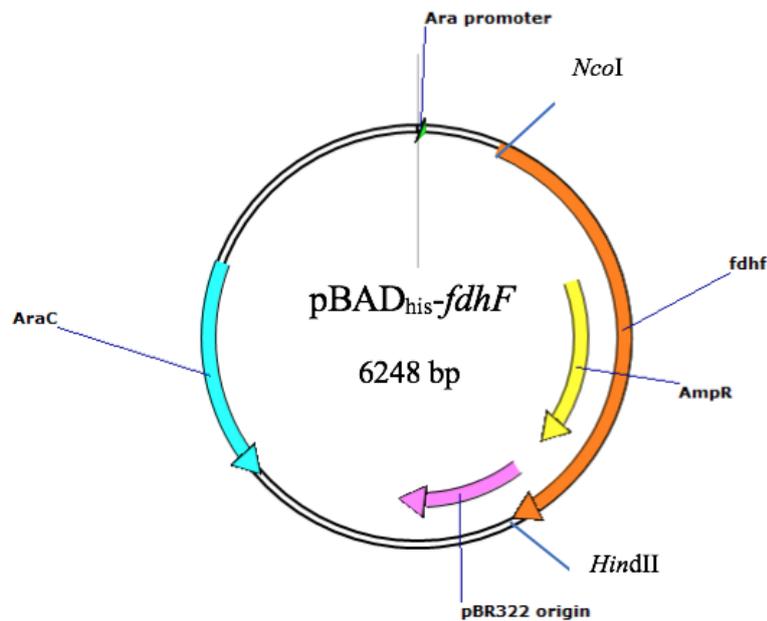
### 6.2.3. Cloning of *fdhF* into the vector, pBAD<sub>ara</sub>

To enable controlled induction of *fdhF*, the *fdhF* PCR product were cloned into pBAD<sub>ara</sub> at the *Nco*I and *Hind*III cloning sites of the vector, using Gibson cloning methodology (Gibson *et al.*, 2009; section 2.12.2.3) (Fig 6.7). The resulting reaction were used to

transform chemically competent *E. coli* Stellar cells (section 2.12.2.4). Six resulting Amp<sup>R</sup> colonies were selected for plasmid ‘miniprep’ isolation using GeneJET™ Plasmid Miniprep Kit (Fermentas) (section 2.10.7). These plasmids were then further analysed by single restriction digestion (section 2.10.4; Fig. 6.6) with *Nco*I to check the size of the plasmid. Following analysis by agarose gel electrophoresis, all plasmids were shown to carry an insert of the expected size (~2.1 kb, not shown) and were then designated pBAD<sub>ara</sub>-*fdhF* (Fig 6.7). Two out of the six successful clones were further confirmed by nucleotide sequencing by Source Bioscience with specific sequencing primers (Table 2.8). The sequences obtained were compared with the sequence database using BLAST which confirmed that the inserts have the correct sequence correctly located at the desired cloning sites (Appendex-13). These plasmids were employed in future studies, as described below.



**Figure 6.6: Gel electrophoretic analysis of cloned *fdhF* in pBAD<sub>ara</sub>.** Electrophoresis was performed using a 0.7% agarose gel. M is GeneRuler® 1 kb ladder (Fermentas). Lane 1-6, cloned samples of pBAD<sub>ara</sub>-*fdhF*. 2 µl of undigested samples were loaded and 1 µl of ladder.



**Figure 6.7. Map of pBAD<sub>ara</sub>-*fdhF*.** The plasmid contains *fdhF* cloned into *Nco*I and *Hind*III sites of pBAD<sub>ara</sub>. The origin of replication (pBR322); Amp<sup>R</sup> locus, *araBAD* promoter and AraC encoding gene are indicated (Reed and Schleif, 1999). The transcription initiation occurs with high concentration level of arabinose (Guzman *et al.*, 1995).

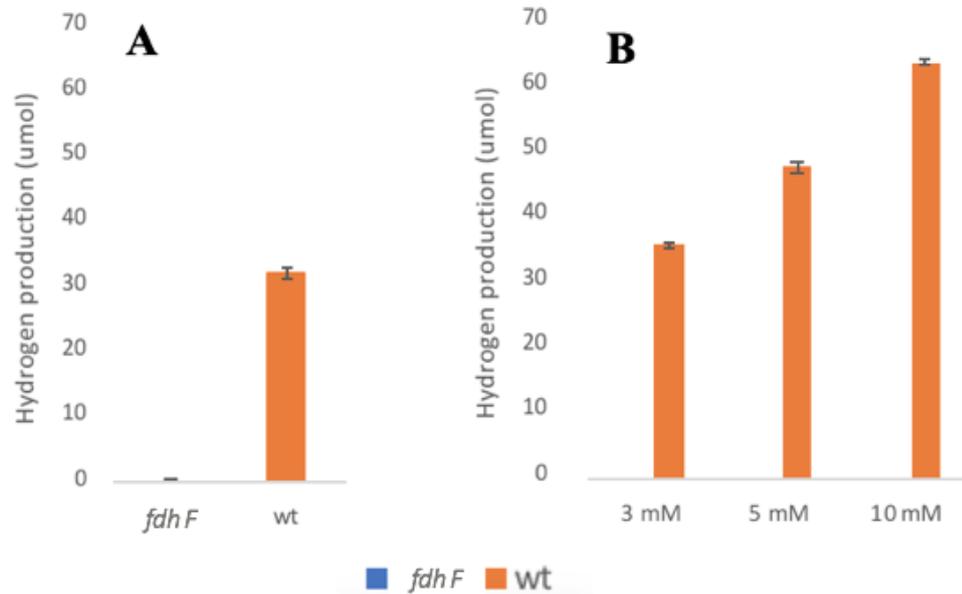
### 6.3. Complementation of *fdhF* and $\Delta$ *hycE* mutant phenotypes by *hyf* and *fdhF*

#### 6.3.1. Confirmation that *fdhF* is required for production of hydrogen

The availability of *fdhF* in an inducible vector allowed progression of the experiments aimed at complementation of  $\Delta$ *hycE* $\Delta$ *fdhF* double mutants. A well-established phenotype of *fdhF* and *hyc* in *E. coli* and their inability to form an FHL complex was discovered earlier (Peck *et al.*, 1957). Thus, the ability of the pBAD<sub>ara</sub>-*fdhF* construct together with pBAD<sub>rha</sub>-*hyf* to allow formation of a second FHL to produce hydrogen was tested using a *hycE fdhF* double (section 6.4; section 2.14.1) mutant strain (note that the

*fdhF* mutant used throughout this thesis was a  $\Delta fdhF$  mutant derived from the Keio collection - i.e. with the *kan* cassette removed; section 2.13). The reason in removing the *kan* cassette is that its presence might have unpredictable affect on downstream gene expression.

To investigate the effect of the  $\Delta fdhF$  mutation in hydrogen production in comparison to the wildtype under anaerobic conditions, overnight cultures of the *fdhF* mutant strain and a wild type strain were cultivated in rich medium (section 2.5.3) with 0.01 M glucose and 120 mM MOPS at pH 6.5. Next day, the strains were tested for hydrogen production following inoculation to a final OD600 of 0.01 in fresh medium, as above. The strains were grown in syringes of 10 ml with needles inserted into rubber stoppers which were used to provide the anaerobic condition for 30 h. The results showed that there was no production of hydrogen for the  $\Delta fdhF$  mutant. That was not surprising because *fdhF* is a required component for FHL and any mutation in this complex would prevent hydrogen production, as reported previously (Soboh *et al.*, 2011). However, there was hydrogen produced in the wild type (Fig 6.8A). Then the same experiment was repeated with the same conditions except with the addition of different concentrations of formate (3, 5 and 10 mM) (Fig 6.8B). The finding was similar in that there was no hydrogen produced by the  $\Delta fdhF$  mutant, but the amount of hydrogen increased as formate concentration also increased. This finding indicates that although *hyc* is functional in the  $\Delta fdhF$  strain and external formate provided, without Fdh-H, no FHL complex will be formed and thus no hydrogen will be produced.

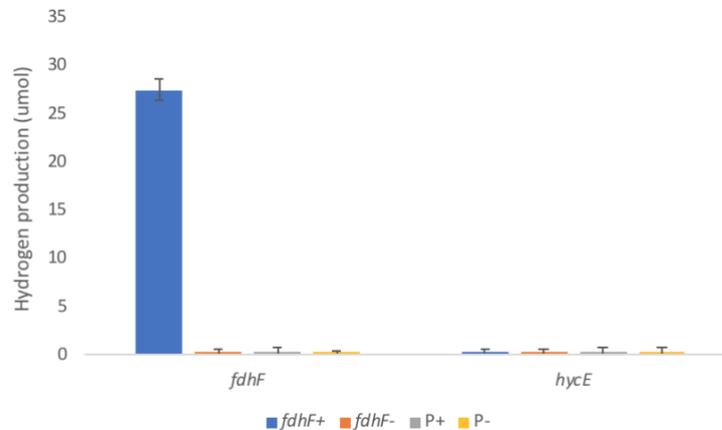


**Figure 6.8. Effect of *fdhF* mutation of H<sub>2</sub> production with and without formate. A.** Without formate. **B.** With formate (mM). The strains (BW25113 and BW25113  $\Delta$ *fdhF*) were grown anaerobically in rich medium (0.01 M glucose and 120 mM MOPS) anaerobically at 37 °C in triplicate for 24 h. Data is the average of three technical repeats with error bars indicated standard deviation; the experiment was repeated twice with similar results obtained. Final OD values were similar in all cases 0.3 and 0.37 OD<sub>600</sub>.

### 6.3.2. Complementation of the $\Delta$ *fdhF* mutant with pBAD<sub>ara</sub>-*fdhF*

To further explore the role of *fdhF* in hydrogen production by Hyc and Hyf, the pBAD<sub>ara</sub> vector expressing *fdhF* (along with the vector control) was used to transform the wildtype and the  $\Delta$ *fdhF* and  $\Delta$ *hycE* mutants. The resulting pBAD<sub>ara</sub>-*fdhF* and pBAD<sub>ara</sub> transformants were grown anaerobically at 37 °C for 24-30 h, as above in rich medium (0.01 M glucose and 120 mM MOPS) with chloramphenicol (50 µg/ml), with and without 0.02% w/v arabinose as inducer. Induction of *fdhF* in the  $\Delta$ *hycE* mutant strain did not enable hydrogen production as was the case with the vector control, which is as expected. However, induction of *fdhF* in the  $\Delta$ *fdhF* mutant resulted in substantial hydrogen production, but only when inducer was included (Fig. 6.9). Thus, this experiment

provides further proof that *fdhF* alone (without Hyc or Hyf) has no ability in hydrogen production. It also shows that the pBAD<sub>ara</sub> encoded *fdhF* is able to complement a  $\Delta fdhF$  mutant and thus is functionally expressed from this vector.



**Figure 6.9. Hydrogen production upon induction of *fdhF* in the  $\Delta fdhF$  and  $\Delta hycE$  mutants.** Growth conditions were similar to Fig 6.8. ‘P’ vector control; ‘+’, with arabinose and ‘-’ without arabinose. Error bars are included but are too small to see, and these correspond to standard deviation of three technical replicates. The experiment was performed three times and similar results were obtained. Growth conditions were as in Fig. 6.8. Strains were BW25113  $\Delta fdhF$  and BW25113  $\Delta hycE$  carrying plasmids pBAD<sub>ara</sub>-*fdhF* or pBAD<sub>ara</sub>.

### 6.3.3. Effect of introduction of pBAD<sub>rha</sub>-*hyc* and -*hyf* into the $\Delta fdhF$ mutant on H<sub>2</sub> production

In order to further assess the role of Fdh-H in enabling Hyc and Hyf-dependent hydrogenase activity and to establish the need for Fdh-H in FHL-2 function, the above experiment was repeated but with cloned pBAD<sub>rha</sub>-*hyf* and pBAD<sub>rha</sub>-*hyc* were transformed into the  $\Delta fdhF$  mutant strain (section 2.12.1). The transformants were cultured anaerobically in rich medium (0.01 M glucose and 120 mM MOPS, pH 6.5) with chloramphenicol (50  $\mu$ g/ml), with and without 0.02% w/v rhamnose as inducer at 37 °C for 24- 30 hours, in syringes to allow measurement the hydrogen production. The results showed (not provided here) that there was no hydrogen produced for the vector control or with the *hyf* and *hyc* plasmids, either with or without inducer. This suggest that neither

*hyf* nor *hyc* alone can produce hydrogen without *fdhF*, and thus indicates that H<sub>2</sub> production by Hyf depends upon Fdh-H (like Hyc). The experiment was repeated as above, but with the addition of formate (at 3, 5 and 10 mM). However, still no hydrogen gas was detected. From this finding, it is clear that formate does not enable hydrogen production by Hyc or Hyf when Fdh-H is absent. Thus, the result provide strong evidence for the requirement for Fdh-H for Hyf-dependent H<sub>2</sub> production, and thus for the presence of a second FHL in *E. coli* (FHL-2) composed of Fdh-H and Hyf. In order to further prove this conclusion, it was considered desirable to create a deletion ( $\lambda$  Red-knockout) of *fdhF* in the  $\Delta hycE$  strain, to generate a double mutant, and then double transformation (section 2.12.2) the resulting  $\Delta fdhF \Delta hycE$  mutant with *hyf* or *hyc* plasmids together with the *fdhF* plasmid for measurement of hydrogen production ability.

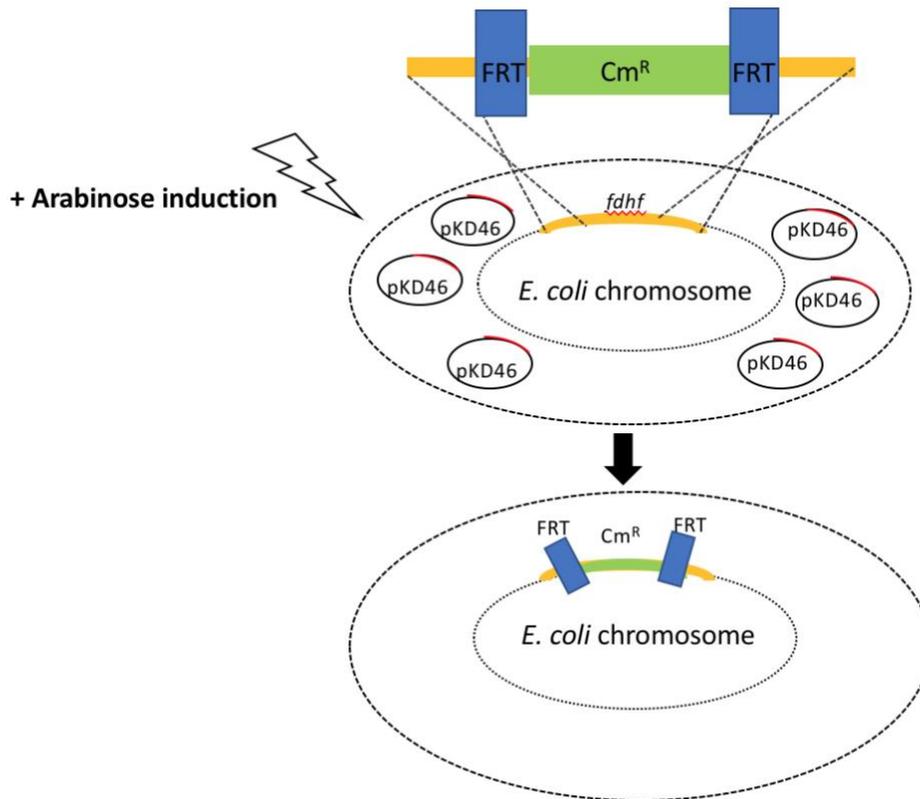
#### **6.4. Inactivation of *fdhF* in the $\Delta hycE$ mutant by knockout using the $\lambda$ red procedure**

A gene knockout is a mutation that inactivates a gene function. Furthermore, recombination is mediated by bacteriophage-based recombination system such as  $\lambda$  Red or RecET of the  $\lambda$  Rac prophage, or similar systems (Thomason *et al.*, 2016). Like genetic engineering, recombineering can be used to make knockout mutations as well as deletions, point mutations, duplications, inversions, fusions and tags (Sawitzke *et al.*, 2011). The gene targeted for knockout is replaced by an antibiotic resistance gene, usually kanamycin or chloramphenicol (Datsenko and Wanner, 2000).

The strategy used in the  $\lambda$  Red system in general involves inserting an antibiotic cassette, which is in this case is the chloramphenicol cassette from pKD3, in place of the target gene (Fig 6.10). The resistance cassette is flanked by FRT sites (Flippase Recognition Target sites), which allow the removal of the cassette once inserted with a FLP helper plasmid (pKD46). The helper plasmid (pKD46) encodes the  $\lambda$  Red recombinase, which

required three phage genes ( $\gamma$ ,  $\beta$  and *exo*) that contribute directly in homologous recombination (Datsenko and Wanner, 2000).

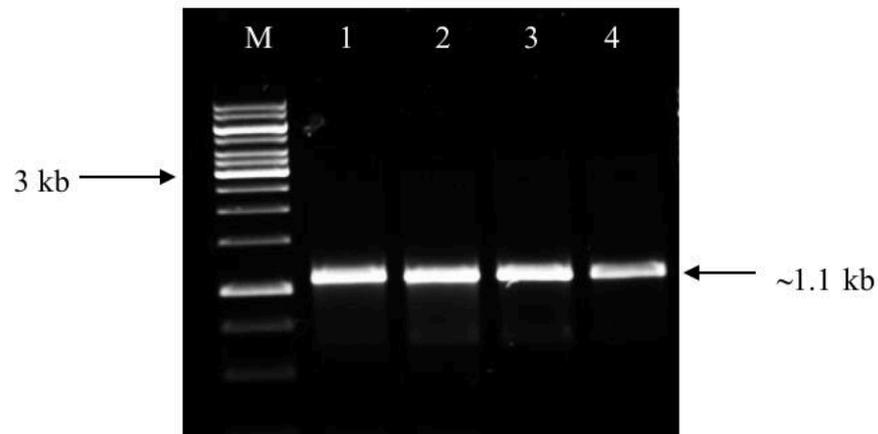
The pKD3 Cm<sup>R</sup> locus was PCR amplified to provide the *cat* antibiotic resistance gene, using specific primers (Table 2.9; section 2.14.2). The amplified resistance cassette was later transformed into the pKD46-carrying  $\Delta hycE$  transformant (see below). pKD46 was transformed into the  $\Delta hycE$  mutant by electroporation (section 2.12.1.2); because pKD46 is a temperature sensitive plasmid, propagation of transformants was at 30 °C. The transformed cells were grown in LB (section 2.5.1) containing ampicillin, at 30 °C, 250 rpm for 4 h. After that, 10 mM L-arabinose was added to induce expression of the  $\lambda$ Red system from the arabinose inducible promoter ( $P_{ara}$ ). The cells next were incubated under the same conditions for 1 h. After that, cells were harvested by centrifugation and prepared for electroporation (section 2.14.4). Then, 2  $\mu$ g of purified PCR fragment (*cat* cassette) were used in the electroporation of the pKD46-carrying  $\Delta hycE$  strain (section 2.12.1.2). Then, the cells were incubated at 30 °C for 2 h and subsequently spread on L-agar (section 2.5) containing chloramphenicol (8  $\mu$ g/ml) and incubated at 37 °C overnight. Next day, single colonies were selected for further testing by growth on L-agar containing 34  $\mu$ g/ml of chloramphenicol and screening for Ap<sup>S</sup>. status. Finally, the identity of the Cm<sup>R</sup> colonies was confirmed by colony PCR (section 2.10.2.1) to check for the replacement of the *fdhF* gene and the presence of the  $\Delta hycE$  mutation.



**Figure 6.10.** A schematic diagram of the knockout strategy using  $\lambda$  Red system. Cm<sup>R</sup>, chloromphenicol resistance gene. FRT, Flippase Recognition Target site. Arabinose is the inducer for the  $\lambda$ Red system carried by pKD46.

#### 6.4.1. Amplification of pKD3 to get the Cm<sup>R</sup> cassette

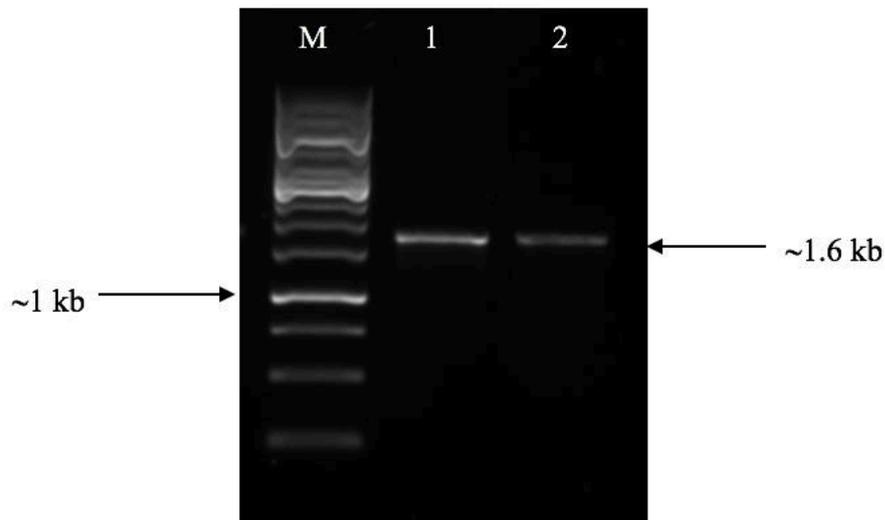
To allow deletion of *fdhF* from the  $\Delta$ *hycE* mutant strain using the  $\lambda$ Red system, pKD3 was required as PCR template for the Cm<sup>R</sup> cassette (section 2.14.2; Table 2.9). The primers used were designed to amplify an ~1100 bp Cm<sup>R</sup> fragment and included the flanking regions of the *fdhF* gene (Fig 6.11).



**Figure 6.11. Agarose gel (0.7%) electrophoretic analysis of the PCR amplified Cm<sup>R</sup> cassette from pKD3.** M, 1 kb Generuler (Fermentas); lanes 1-4, 2  $\mu$ l amplified Cm<sup>R</sup> cassette with *fdhF* flanking regions. Arrow indicates position of the amplified DNA fragment (~ 1.1 kb).

#### 6.4.2. Confirmation of the ( $\Delta$ *fdhF*)::*cat* substitution mutation and $\Delta$ *hycE* mutation in the $\Delta$ *hycE* mutant strain

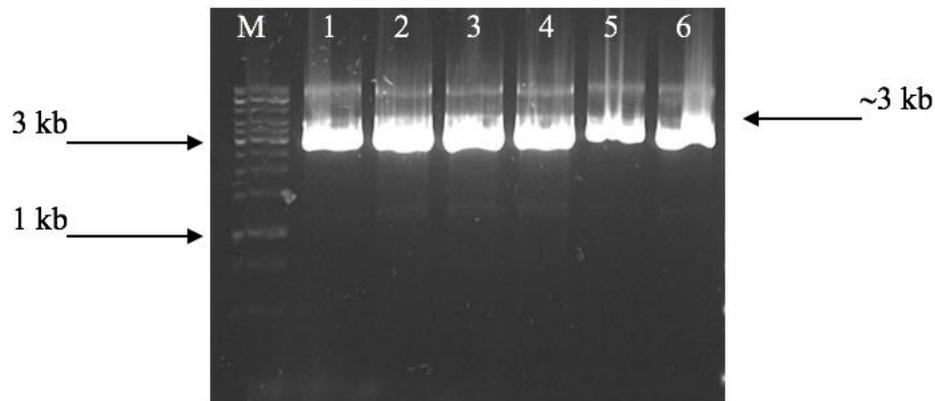
The colonies obtained, on L-agar with 8  $\mu$ g/ml chloramphenicol, following the introduction of the *cat* cassette into the  $\Delta$ *hycE* mutant were streaked onto L-agar plates with 34  $\mu$ g/ml chloramphenicol. Next, colony PCR (using the primers that anneal 300 bp up- and down-stream of *fdhF*) was used to confirm that the deletion had occurred (Fig 6.12). A clear band of the expected size (~1.6 kb) was obtained for two colonies (Fig. 6.12; ~1000 bp from the *cat* gene, and 300 bp upstream and 300 bp downstream of *fdhF*). Note that the wildtype and original  $\Delta$ *hycE* mutant gave a PCR product of ~2.7 kb, as expected (not shown).



**Figure 6.12.** Agarose gel (0.7%) electrophoretic analysis of colony PCR amplification of the *fdhF* locus in two suspected ( $\Delta fdhF$ )::cat mutants. M, 1 kb Generuler (Fermentas); lane 1-2, 2  $\mu$ l of PCR reactions for colonies 1 and 2, respectively. Arrow indicates position of the target band DNA (~ 1.1 kb).

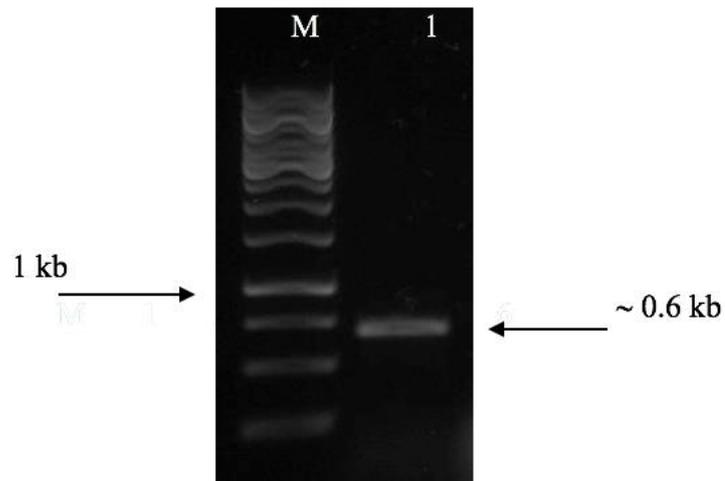
To further confirm the mutation, the PCR fragment generated for one colony was cloned into pJET1.2 to enable nucleotide sequence analysis. Thus, the colony PCR product was purified (section 2.10.3) and cloning into pJET plasmid (Appendix-1) was achieved using the recommended protocol (section 2.12.2). Six of the resulting Cm<sup>R</sup> transformants were subject to plasmid miniprep isolation (section 2.10.7) and the cloning confirmed by electrophoresis (Fig 6.13). Since pJET1.2 is ~3 kb and the PCR product is ~1.6 kb, a plasmid of ~4.6 kb was expected. However, the plasmid is undigested, so the size will appear is 2/3<sup>rd</sup> its real size, i.e. ~3 kb as is observed. One of these pJET1.2-( $\Delta fdhF$ )::cat plasmids was submitted for nucleotide sequencing by Source Bioscience using pBAD<sub>rha</sub>-F and pBAD<sub>rha</sub>-R primers (Table 2.4). The sequences obtained were compared with the sequence database using BLAST which confirmed that the Cm<sup>R</sup> cassette is correctly

located at the desired site and the sequence was 100% identical with the  $Cm^R$  gene of pKD3.

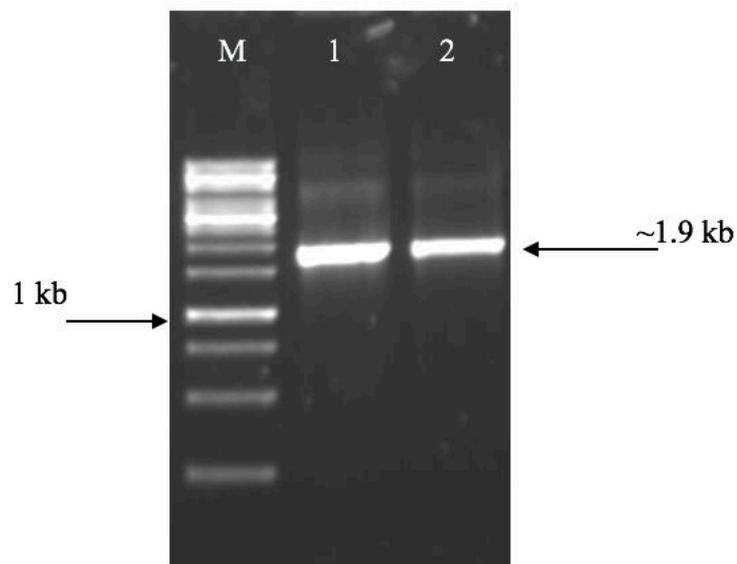


**Figure 6.13. Agarose gel (0.7%) electrophoretic analysis of pJET1.2-( $\Delta fdhF$ )::cat plasmids.** M, 1 kb Generuler (Fermentas); lanes 1-6, 2  $\mu$ l of plasmid DNA isolated from six distinct  $Cm^R$  transformants. All plasmids are undigested.

The above results indicate that the desired *fdhF* mutant has been created. Since the strain which used to knockout the *fdhF* gene was a  $\Delta hycE$  mutant, a further PCR was performed to confirm the *hycE* mutation. Therefore, colony PCR (section 2.10.2.1) was applied with primers specific for the *hycE* locus (Table 2.8; annealing 300 bp before and after the gene). Electrophoresis showed (Fig 6.14) that the resulting PCR product generated was as expected, ~600 bp. Control reactions using wildtype DNA as template gave a PCR product (using Table 2.8, *hycE*-F and *hycE*-R with annealing 100 bp before and 100 bp after) of approx. 2 kb, as expected (Fig 6.15). Thus, from this molecular-genetic confirmation, it can be concluded that the desired double mutant ( $\Delta fdhF$ )::cat  $\Delta hycE$ ) has been successfully generated.



**Figure 6.14.** Agarose gel (0.7%) electrophoretic analysis of the *hycE* PCR product of the putative  $\Delta hycE$  ( $\Delta fdhF$ )::*cat* strain. M, 1 kb Generuler (Fermentas); lane 1,  $\Delta hycE$  PCR product.



**Figure 6.15.** Agarose gel (0.7%) electrophoretic analysis of the *hycE* PCR product of the wildtype strain (contain *hycE*). M, 1 kb Generuler (Fermentas); lane 1-2, *hycE* PCR product.

## 6.5. Effect of *fdhF* and *hyf* status on hydrogen production in the $\Delta hycE \Delta fdhF$ mutant

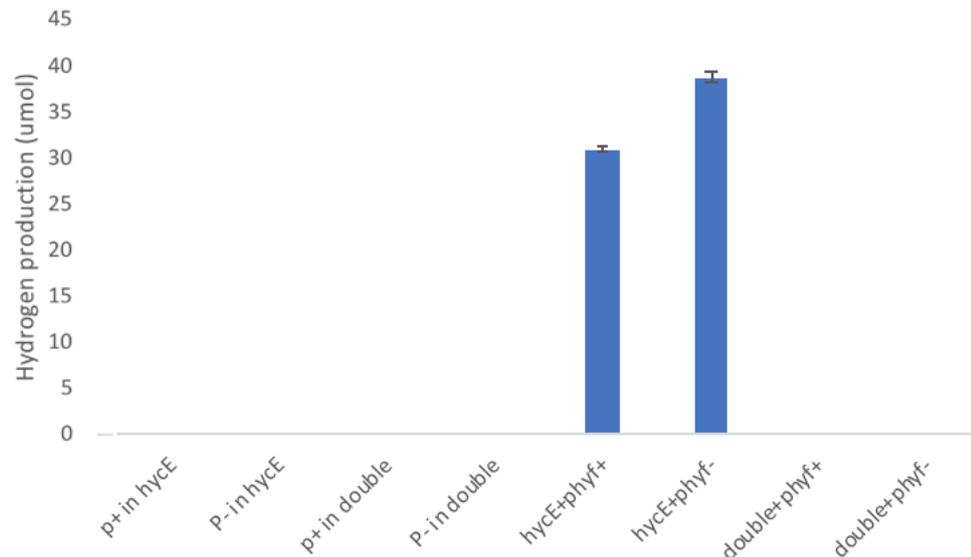
### 6.5.1. Complementation of the double mutant with *fdhF* or *hyf*

In order to enable introduction of pBAD<sub>rha</sub> into the double mutant, the Kan cassette was removed as described in Methods (section 2.13) to generate BW25113  $\Delta fdhF \Delta hycE$ , which was confirmed by PCR (as above) and used in the studies described below.

To provide further evidence that Hyf activity is Fdh-H dependent, as is the case for Hyc, plasmid pBAD<sub>rha-hyf</sub> was transformed into the double mutant and H<sub>2</sub> production was tested by growth in rich medium (with 0.01 M glucose and 120 mM MOPS, pH 6.5, plus Cm) under fermentative conditions using 10 ml syringes at 37 °C for 24-30 h. No hydrogen gas was detected, either with or without inducer (Fig. 6.16). This finding further indicates that Hyf is unable to function without Fdh-H. The experiment was repeated but with formate (at 3-20 mM), but as expected, still no hydrogen gas was produced (data not shown). In addition, pBAD<sub>rha-hyc</sub> was also transformed into the double mutant and propagated under the same conditions as above, and the results were the same - no hydrogen was produced with or without inducer, or with or without formate (data not shown). Thus, as anticipated, both Hyf and Hyc dependent H<sub>2</sub> evolving activity requires Fdh-H.

Addition experiments were performed in parallel with those above, using the  $\Delta hycE$  mutant carrying pBAD<sub>rha-hyf</sub> (as a positive control) or pBAD<sub>rha</sub> (negative control) and the  $\Delta hycE \Delta fdhF$  double mutant with pBAD<sub>rha</sub>, under the same growth conditions as above. Hydrogen was produced for the induced and uninduced *hyf*-complemented *hycE* mutant (as in chapter 5) only, at 0.8 and 1 ml, respectively, but not in the vector controls (Fig 6.16). This confirmed that the *hyf* plasmid enables H<sub>2</sub> production when *fdhF* is functional. In addition, since Fdh-H is needed for both Hyc and Hyf activity, it would

appear that YdeP (Vivijs *et al.* 2015) does not interact with Hyf (or Hyc) to form a FHL complex, at least not under the conditions used here.



**Figure 6.16. Expression of *hyf* fails to restore H<sub>2</sub> production in the  $\Delta fdhF \Delta hycE$  double mutant.** Growth conditions were as in Fig 6.8. ‘P’, vector control; ‘+’, with rhamnose; ‘-’ without rhamnose; ‘double’,  $\Delta hycE \Delta fdhF$  mutant; ‘hycE’,  $\Delta hycE$  mutant; ‘phyf’, pBAD<sub>rha</sub>-*hyf*. Error bars correspond to standard deviation of three technical replicates. The experiment was performed three times and similar results obtained.

### 6.5.2. Complementation of the $\Delta hycE \Delta fdhF$ mutant with pBAD<sub>ara/rha</sub>-expressed *fdhF* and *hyf*

Double transformation involves transforming two plasmids which do not contain the same *oriV* and are thus considered as compatible plasmids. However, plasmids which contain the same *oriV* are considered as incompatible, and they cannot stably co-exist in the same cell (Novick, 1987; Nordstrom and Austin, 1989 and Sambrook *et al.*, 1989). This incompatibility occurs because they will compete for the same replication-initiation machinery, creating an unstable environment. In the research reported here, double transformation was performed for two plasmids with distinct origins of replication and with different antibiotic resistance genes which enables their compatibility and independent selection.

The following transformants were thus generated:

pBAD <sub>rha</sub> - <i>hyf</i> plus pBAD <sub>ara</sub> - <i>fdhF</i>	in the $\Delta hycE \Delta fdhF$ mutant
pBAD <sub>rha</sub> - <i>hyc</i> plus pBAD <sub>ara</sub> - <i>fdhF</i>	in the $\Delta hycE \Delta fdhF$ mutant
pBAD <sub>rha</sub> - <i>hyf</i>	in the $\Delta hycE$ mutant.

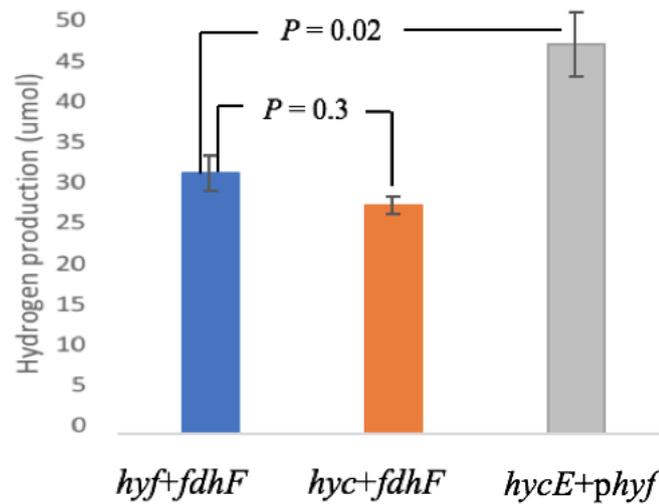
The transformants were grown anaerobically in rich medium (section 2.5.3), as above.

The results show that hydrogen was produced at relatively high levels in the  $\Delta hycE \Delta fdhF$  mutant when complemented with both *hyf* and *fdhF* or with both *hyc* and *fdhF* (Fig 6.17).

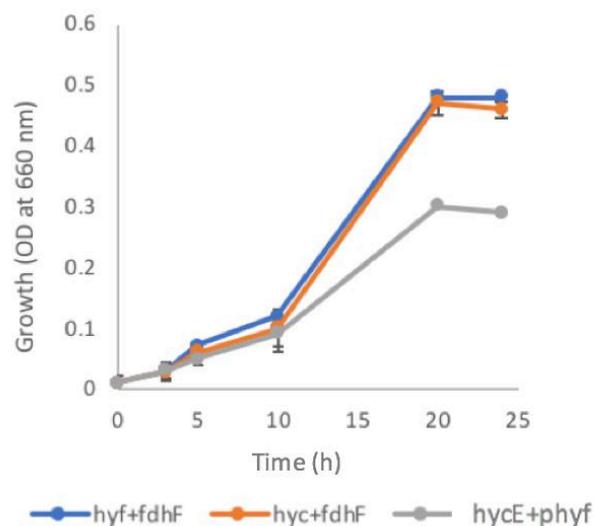
Thus, this result further confirms that dependence of both Hyf and Hyc H<sub>2</sub> evolution on Fdh-H, providing further support for the existence of two distinct FHL systems in *E. coli*.

The growth of pBAD<sub>rha</sub>-*hyf* plus pBAD<sub>ara</sub>-*fdhF* and pBAD<sub>rha</sub>-*hyc* plus pBAD<sub>ara</sub>-*fdhF* showed similar trend, which was higher than the control at 24 h by 1.5 fold (Fig. 6.18).

Another further experiment was performed in parallel that was identical to that above experiment, except that the one or other of the two inducers was excluded. When rhamnose was included for *hyf* induction but arabinose was excluded to limit *fdhF* induction, no hydrogen was produced in the  $\Delta hycE \Delta fdhF$  mutant carrying both pBAD<sub>rha</sub>-*hyf* and pBAD<sub>ara</sub>-*fdhF*. Also, under the opposite regime (induction of *fdhF* with arabinose and without *hyf* induction), again no hydrogen was produced by the double mutant strain (data not shown). Similar results were obtained with the  $\Delta hycE \Delta fdhF$  mutant carrying pBAD<sub>rha</sub>-*hyc* plus pBAD<sub>ara</sub>-*fdhF* (Fig 6.17). These findings thus allow the conclusion that Hyf and FdhF can form a second FHL (FHL-2) where formate oxidation to CO<sub>2</sub> is used for the reduction of protons to produce H<sub>2</sub>.



**Figure 6.17. Restoration of H<sub>2</sub> production in the  $\Delta hycE \Delta fdhF$  strain by complementation with *hyf* and *fdhF* expressing plasmids or with *hyc* and *fdhF* expressing plasmids.** The growths were as described in Fig. 6.8, but with Cm and Amp included as required, and rhamnose and arabinose at 0.02% each. Strains were: '*hyf+fdhF*', BW25113  $\Delta hycE \Delta fdhF$  with pBAD<sub>rha</sub>-*hyf* plus pBAD<sub>ara</sub>-*fdhF*; '*hyc+fdhF*', BW25113  $\Delta hycE \Delta fdhF$  with pBAD<sub>rha</sub>-*hyc* plus pBAD<sub>ara</sub>-*fdhF*; and '*hycE+phyf*', BW25113  $\Delta hycE$  with pBAD<sub>rha</sub>-*hyf*. Statistically significant difference was determined by Student's unpaired *T*-test ( $P < 0.05$ ). The experiment was performed three times and similar results obtained.



**Figure 6.18. Growth of  $\Delta hycE \Delta fdhF$  strain by complementation with *hyf* and *fdhF* expressing plasmids or with *hyc* and *fdhF* expressing plasmids.** The growths were as described in Fig. 6.13 in rich medium (section 2.5.3), but with Cm and Amp included as required, and rhamnose and arabinose at 0.02% each. Strains were: '*hyf+fdhF*', BW25113  $\Delta hycE \Delta fdhF$  with pBAD<sub>rha</sub>-*hyf* plus pBAD<sub>ara</sub>-*fdhF*; '*hyc+fdhF*', BW25113  $\Delta hycE \Delta fdhF$  with pBAD<sub>rha</sub>-*hyc* plus pBAD<sub>ara</sub>-*fdhF*; and '*hycE+phyf*', BW25113  $\Delta hycE$  with pBAD<sub>rha</sub>-*hyf*. The experiment was performed three times and similar results obtained.

## 6.6. Increased Hyf-FdhF-dependent H<sub>2</sub> production in the complemented $\Delta hycE \Delta fdhF$ double mutant

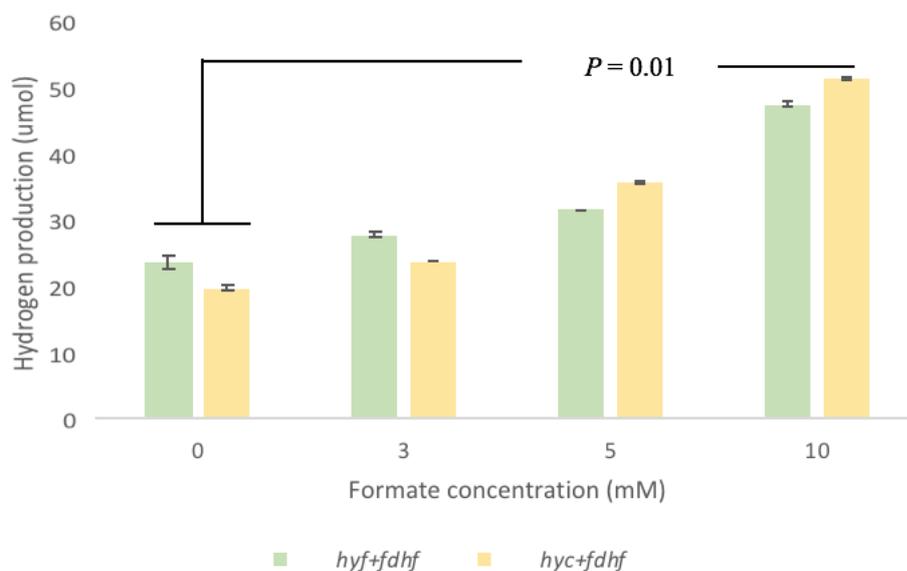
### 6.6.1. Effect of inducer levels on H<sub>2</sub> production

The above experiment clearly showed that the combined expression of *hyf* and *fdhF* results in the production of hydrogen, which suggests that Hyf with FdhF forms a second FHL. It was shown previously (Sawers *et al.*, 1985) that as the concentration of formate increases the production of hydrogen also increases during *E. coli* fermentation. The work described in chapter 5 further showed that both Hyc and Hyf dependent H<sub>2</sub> production could be stimulated by addition of formate, and that this effect was unlikely to be related to Fhl-1-mediated induction of *hyf* or *hyc* expression, under the conditions employed. However, the possibility that the formate-induced enhancement of H<sub>2</sub> production was caused by FhlA regulation of *fdhF* could not be discounted.

In the experiments below, the effects of both inducer levels and formate levels on H<sub>2</sub> production for pBAD-dependent FHL systems is explored in the  $\Delta hycE \Delta fdhF$  mutant carrying pBAD<sub>rha</sub>-*hyf* or -*hyc* plus pBAD<sub>ara</sub>-*fdhF*. Initially, the concentrations of arabinose and rhamnose used were increased from 0.02% w/v to 0.04 and 0.06%. Growths were in rich medium (0.01 M glucose and 120 mM MOPS, pH 6.5), as before, in syringes at 37 °C for 24-33 h. However, the results showed that raising the levels of the inducers, did not notably affect the amount of hydrogen produced (data not shown). Thus, either 0.02% inducer provides maximal expression with respect to 0.04-0.06%, or factors other than FHL protein levels limit the amount of hydrogen produced under the conditions tested.

### 6.6.2. Effect of formate addition on H<sub>2</sub> production

The effect of formate concentration (3, 5 and 10 mM) on H<sub>2</sub> production was also tested. The  $\Delta hycE \Delta fdhF$  double mutant with either pBAD<sub>rha</sub>-*hyf* and pBAD<sub>ara</sub>-*fdhF* or with pBAD<sub>rha</sub>-*hyc* and pBAD<sub>ara</sub>-*fdhF* were grown as above, with 0.02% w/v of both rhamnose and arabinose, together with the indicated levels of formate. The results clearly indicate that as formate concentration increases, hydrogen production also increases (Fig. 6.19). The results show that there is no significant difference between the results obtained with Hyf and Hyc in terms of hydrogen production in all cases (with or without formate). However, there is a clear difference in hydrogen production between 0 and 10 mM formate, with a 2 and 2.6 fold increase seen ( $P = 0.01$ ) with Hyf-FdhF and for Hyc-FdhF, respectively. This finding thus shows hydrogen production for both FHL-1 and -2 responds formate under conditions where formate is not expected to directly regulate the expression of the corresponding genes. This indicates that FHL enzyme activity responds directly to formate availability.



**Figure 6.19. Effect of formate concentration on hydrogen production by the plasmid encoded Hyf-FdhF and Hyc-FdhF FHL complexes in the  $\Delta hycE \Delta fdhF$  strain.** 0.02% w/v rhamose and arabinose were included, along with Cm and Ap. Statistically significant difference as determined by Student's unpaired *T*-test ( $P < 0.05$ ). Each value is the average of three technical replicates. The experiment was performed three times with similar results obtained.

### 6.7. Effect of the plasmid-specified Hyf-FdhF and Hyc-FdhF, FHL complexes on formate production and consumption in the $\Delta hycE \Delta fdhF$ mutant

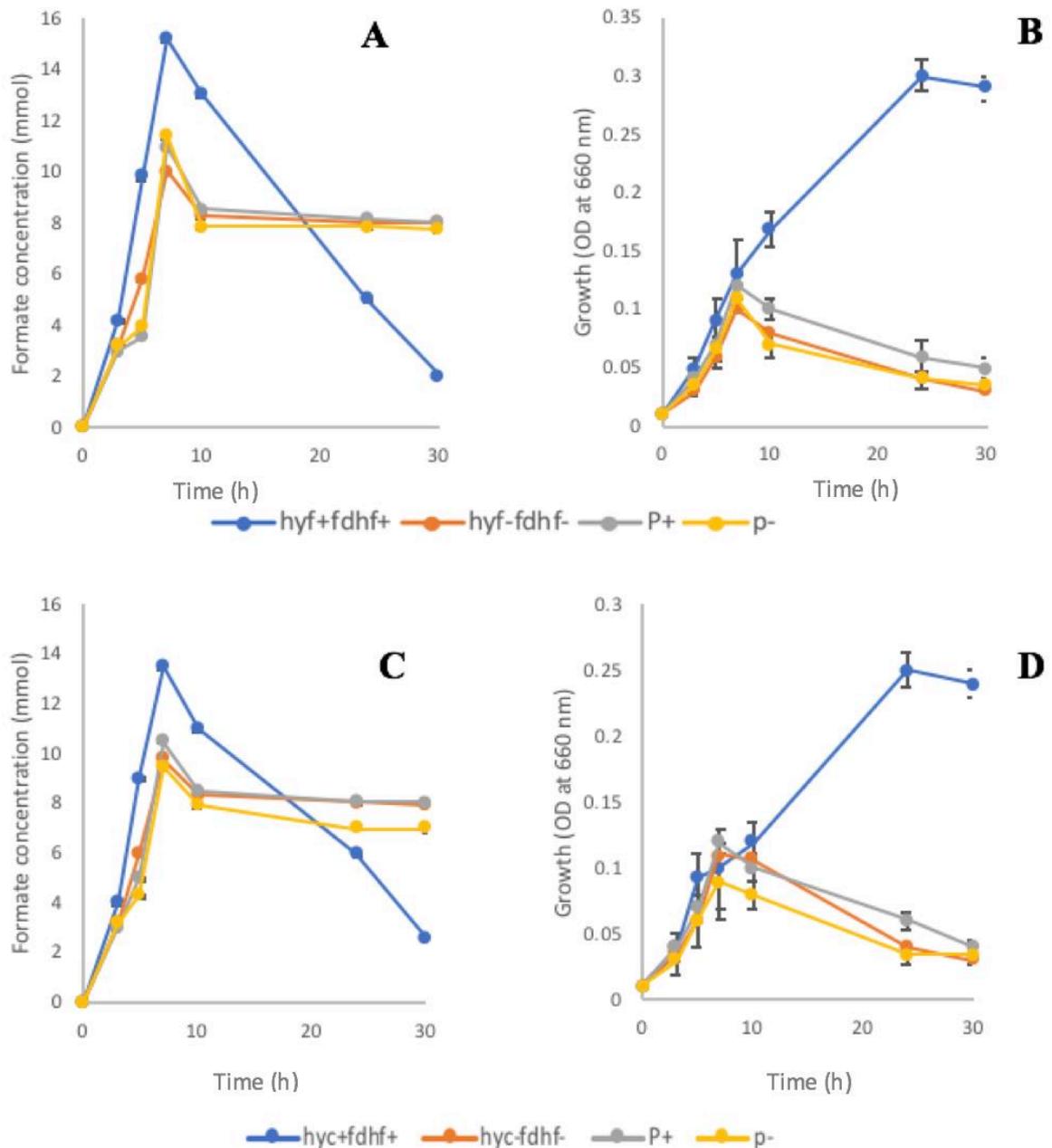
To confirm a role of *fdhF* with *hyf* in mediating the consumption of formate during anaerobic growth, the levels of formate were measured during the growth of the complemented double mutant in comparison to the vector control. Growth was in WM-medium (80 mM glucose and 0.8% trace elements with 120 mM MOPS at pH 6.5) in 20 ml syringes, with inducers and antibiotics.

The pattern of formate consumption observed was similar to that seen in chapter 5, except that the peak in formate production was at 7 rather than 10 h. A similar overall pattern of formate production was seen in all the vector controls and in the cultures lacking inducer: a rapid production of formate over the first 7 h to 9.4-11.5 mM, followed by a

slight reduction to ~8 mM formate over the next 3 h, which in turn is followed by a period where formate levels remain roughly constant until the end of the growth period (Fig 6.20). This indicates very little formate consumption by the vector controls or by the complemented strains in the absence of inducer. For the induced complemented strains, a similar rapid increase in formate levels was seen over the first 7 h, but the degree of formate production was greater than for the controls at 15.2 and 13.5 mM at 7 h for the induced Hyf-FdhF and Hyc-FdhF strains, respectively, cf. 11 mM for the induced vector controls, representing a 28-19% increase in formate production. Of particular note is the far greater degree of formate consumption observed for the induced Hyf-FdhF and Hyc-FdhF strains after 7 h (Fig. 6.20). The induced Hyf-FdhF and Hyc-FdhF strains reduced the formate concentration by 13.2 and 10.9 mM, respectively, over the 7-30 h growth period, whereas for controls the degree of reduction was far more modest (~2-3 mM). This strongly indicates that the Hyf-FdhF and Hyc-FdhF systems provide a major capacity for consuming exogenous formate, such that in the absence of these systems, *E. coli* is unable to consume exogenous formate effectively under fermentative conditions.

Induction Hyf-FdhF or Hyc-FdhF also had a major impact on growth of the  $\Delta hycE \Delta fdhF$  strain (Fig. 6.20B & D). This effect was particularly marked at and beyond the 10 h point. The vector controls showed a cessation in growth at ~7 h and a decline in cell density from then on, whereas the induced Hyf-FdhF or Hyc-FdhF strains continued to grow well. Indeed, by 30 h there was a major difference in cell density (~six and fivefold, respectively) between the induced Hyf-FdhF or Hyc-FdhF strains and the vector controls. Such a marked growth effect was not seen for the  $\Delta hycE$  single mutant (chapter 5) which suggests that the combined absence of both Fdh-H and Hyd-3 has a greater impact on fermentative growth than the absence of just Hyd-3. This suggests that the impaired growth of the  $\Delta hycE \Delta fdhF$  mutant shown here is presumed to relate to its inability to

consume exogenous formate resulting in a formate toxicity effect. Also, an additional role for Fdh-H in fermentative metabolism beyond its contribution to Hyc-dependent hydrogenase activity. Such a possibility could be tested by examining the growth of the *fdhF* mutant and by testing the effect on fermentative growth of complementing the double mutant with *fdhF* and *hyc* independently.



**Figure 6.20. Effect of plasmid-specified *fdhF* and *hyf*, or plasmid-specified *fdhF* and *hyc* induction on formate production and consumption in the  $\Delta hycE \Delta fdhF$  double mutant at pH 6.5.** Conditions are as for Fig. 6.15. **A** and **B**, double mutant containing pBAD<sub>rha</sub>-*hyf* and pBAD<sub>ara</sub>-*fdhF* or the corresponding vector controls (p), with (+) or without inducer (-). **C** and **D**, as A/B, except for the use of pBAD<sub>rha</sub>-*hyc* in place of the pBAD<sub>rha</sub>-*hyf*. **A** and **C** indicate formate levels, **B** and **D** indicate growth. Growths were performed in triplicate and standard deviations are indicated. The experiment was repeated twice, with similar results obtained.

## 6.8. Discussion

In this chapter, the role of FdhF in enabling Hyf activity was investigated. Both corresponding genes/operons were cloned into compatible, inducible vectors to allow their combined ability to mediate hydrogen production to be tested, as well as to test their ability to drive the consumption of formate. The results showed that neither FdhF nor Hyf can produce hydrogen alone, such production was only observed when both components were generated together. This is presumed to enable the formation of a novel FHL complex with the capacity to consume formate, with the consequent production of hydrogen.

Earlier experiments were conducted with C<sup>13</sup>-labelled formate in O<sup>18</sup> labelled water, showing that the FdhF/Fdh-H enzyme converts formate to CO<sub>2</sub> (Khangulov *et al.*, 1998). However, this reaction takes place only in combination with Hyc, because *hycBCDEFG* gene products are key components of the FHL enzyme (McDowall *et al.*, 2014). On the other hand, it was proposed that a single mutation in *hycB* is sufficient to prevent FdhF association with Hyd-3 and subsequently prevents hydrogen gas production (McDowall *et al.*, 2015). Furthermore, Jo and Cha (2015) found that in BL21(λDE3), a deletion in either *fdhF* or *hycE* resulted in cessation of all hydrogen production. This is in agreement with the findings here that absence of either *hyc* or *fdhF* results in failure to produce hydrogen gas in *E. coli*.

The capacity of the *hyf* encoded Hyd-4 to consume formate and yield H<sub>2</sub>, in combination with FdhF, had not been previously proven. Thus, the role of a second FHL complex in *E. coli* has been established here. The research here clearly shows that *hyf* in an *fdhF* mutant strain does not produce any hydrogen gas. The same finding applies to provision of the plasmid-encoded *hyc* to an *fdhF* mutant – the chromosomal and plasmid borne

copies of *hyc* failed to enable H<sub>2</sub> production in the absence of FdhF, as previously indicated (Soboh *et al.*, 2011).

A double  $\Delta hycE \Delta fdhF$  mutant was created and the inducible *hyf* and *fdhF* plasmids were co-introduced which re-enabled hydrogen production. Replacing the *hyf* plasmid with the *hyc* plasmid also allow in H<sub>2</sub> production, as expected, but no hydrogen was generated for the single transformants. Thus, when both *fdhF* and *hyf* are induced, detection of hydrogen production is clear established, while with expression of just one of these systems, no hydrogen gas is released. This result therefore provides strong evidence for the collaborative production of H<sub>2</sub> by Hyf together with FdhH, indicating that hydrogenase-4 activity is FdhF dependent.

It was not possible to increase the amount of hydrogen produced by increasing the levels of inducer used; it is unclear if this failure is related to a saturated expression level at standard arabinose/rhamnose levels (0.02%) or whether other factors limit FHL activity (e.g. Hyp factors or co-factor levels). However, although inducer could not be used to raise H<sub>2</sub> production further, addition of formate was effective, with a 2.6- and 2-fold increase in production seen with 10 mM formate for FHL-1 and -2, respectively (Fig. 6.15). This improvement in hydrogen production by formate addition was suggested to be caused by an increase in the availability of formate as an FHL substrate. Additionally, changes in exogenous levels of formate were measured with the double transformants (*fdhF* and, *hyf* or *hyc*) of the  $\Delta hycE \Delta fdhF$  double mutant. It was found that either *fdhF* and *hyf*, or *fdhF* and *hyc*, restored the consumption of released formate to levels similar to those observed in the wildtype (after the peak in formate production at 7 h) (Fig. 6.16). Indeed, the levels of formate at the end of the growth period were ~fourfold lower upon complementation than in the vector control or in the absence of inducer, which emphasises the role of the FdhF-Hyc and FdhF-Hyf systems in consumption of formate.

The combined absence of *hyc/hyf* and *fdhF* resulted in a major growth defect under fermentation conditions with respect to the induced complemented strains, indicating that the absence to FHL impairs fermentative growth under conditions where formate is generated.

## Chapter 7: Effect of metals and accessory genes on the activity of Hyf

### 7.1. Introduction

Three formate dehydrogenases (FDH-H, FDH-N and FDH-O) (section 1.7.1), and four hydrogenases (Hyd-1 to 4) are known to exist in *E. coli*, all of which are metallo-enzymes containing iron-sulphur [Fe-S] clusters. The dehydrogenases contain molybdenum and selenium, and the hydrogenases contain nickel and iron, located in their active sites (Lukey *et al.*, 2010; Shafaat *et al.*, 2013). Furthermore, the *hyp* operon in *E. coli* includes several genes which are required significantly for the synthesis of hydrogenases 1 to 3 (Jacobi *et al.*, 1992) and the *nik* operon is an important system needed for nickel transport for provision of this metal for hydrogenase synthesis (Navarro *et al.*, 1993).

### 7.2. Accessory genes

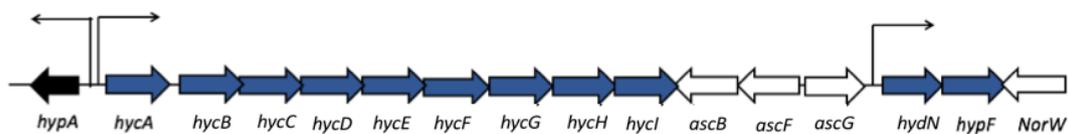
#### 7.2.1. *hyp* operon

The Hyp proteins are encoded by the *hypABCDE-fhlA* operon and the *hypF* gene (located between *hydN* and *norW*) (Fig 7.1), which are located together upstream of the *hyc* operon (Fig 7.2). The ‘*hyp*’ mnemonic represents ‘hydrogen pleiotropy’ to indicate genes that are significantly important in the biosynthesis of the [NiFe] cofactor. HypA with HycI are solely involved in the maturation of Hyd-3 (Bock and Sawers, 1996). HypB is a GTPase required for the formation of the metal centre of [NiFe] in the HycE precursor (pre-HycE). Maier *et al.* (1993) indicate that HypB acts as a nickel-donating system, where GTP hydrolysis releases HypB from pre-HycE after the metals have been incorporated. HypC is important in forming a complex with pre-HycE during the maturation process. This complex inhibits any nickel insertion and prevents any association of the small subunit (HycG) to pre-HycE during the maturation process

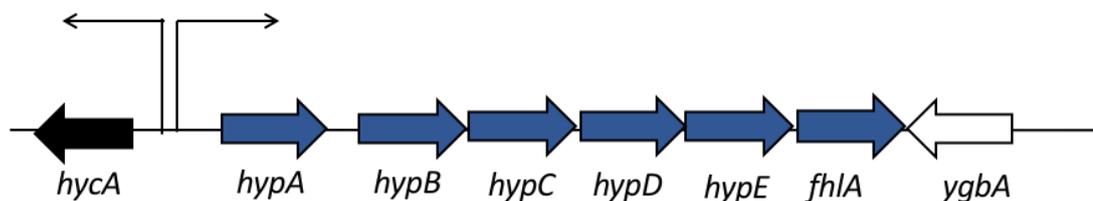
(Magalon and Böck, 2000). HypF also interacts with pre-HycE and is involved in inserting the non-proteinaceous ligands (CO and CN<sup>-</sup>) in the active site cavity of the large subunit (Casslot and Rousset, 2001); any defect in HypF will result in lack of activity for all hydrogenases. HypE is a partner of HypF in the maturation process.

The first step in the biosynthesis of [NiFe] is the synthesis of the non-proteinaceous ligand (CN). The HypF protein interacts with HypE. HypE has carbomoyl dehydrogenase activity using ATP hydrolysis to modify the bound carbomoyl to form either a bound thiocyanate (Tominaga *et al.*, 2013) or more possibly a bound iso-thiocyanate (Stripp *et al.*, 2015). The iso-thiocyanate bound to HypE contributes CN<sup>-</sup> to Fe(II) in [NiFe]-centre biosynthesis. The second non-proteinaceous ligand source in the active site is CO, which is less understood than CN<sup>-</sup> (Sargent, 2016). Usually, CO is derived from tyrosine in other biological systems (Kuchenreuther *et al.*, 2014). Soboh *et al.* (2013) suggested that the source of CO might be CO<sub>2</sub>, which can bind naturally to the Fe(II) ion held by the HypCD complex, which helps in building the first half of the [NiFe] cofactor. *E. coli* HypD contains a [4Fe-4S] cluster as a cofactor (Bloksech and Böck, 2006). The HypD protein binds also to at least one additional Fe ion that acts as a foundation for [NiFe] cofactor assembly and this Fe(II) may be associated with HypC in the HypCD complex (Watanabe *et al.*, 2012). This complex is the preliminary point for building [NiFe] by holding Fe-CO<sub>2</sub> species, which is probably reduced to Fe-CO (Soboh *et al.*, 2013). This reaction is in parallel to another reaction involving loading of CN<sup>-</sup> into the HypCD-Fe complex (Burstel *et al.*, 2012) where HypE carrying iso-thiocyanate interacts with the HypCD complex forming Fe-CO-2CN<sup>-</sup> (Stripp *et al.*, 2013). The whole process, as indicated above, generates the first half of the [NiFe] cluster in the HypCD complex (Fig 7.3). Next, is the insertion of this ‘first half’ into an empty apo-enzyme (Fig 7.4). Since HypC, and the homologous HybG, can cooperate directly with the undeveloped catalytic

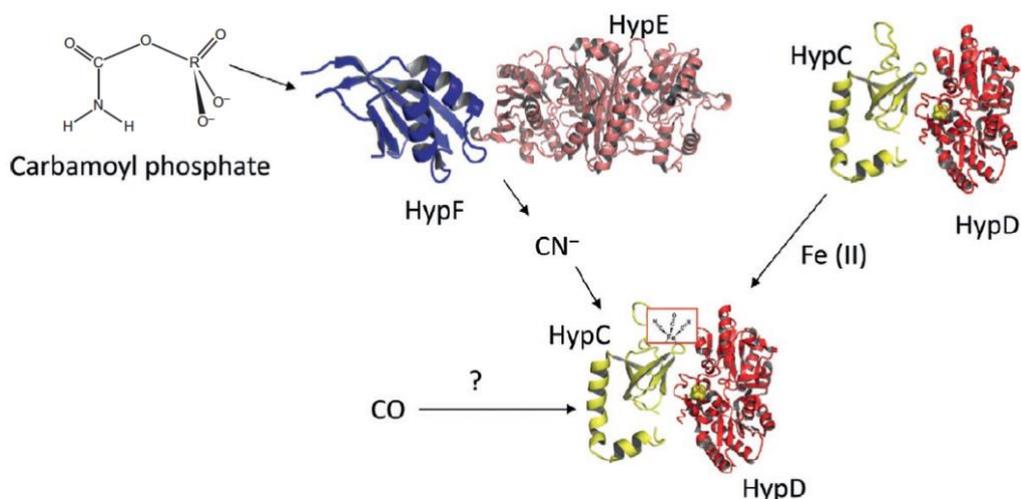
subunits of all three [NiFe]-hydrogenases, these proteins can act as a link between the biosynthesis of [NiFe] and the empty apo-enzymes (Blokesch *et al.*, 2001). The interaction between HypC/HybG and the empty subunit brings HypD, HypE and HypF proteins close to each other to allow an efficient assembly and interaction of the first half of the cofactor with the other half (Fig. 7.4). The continuing steps involve nickel processing and further cluster assembly.



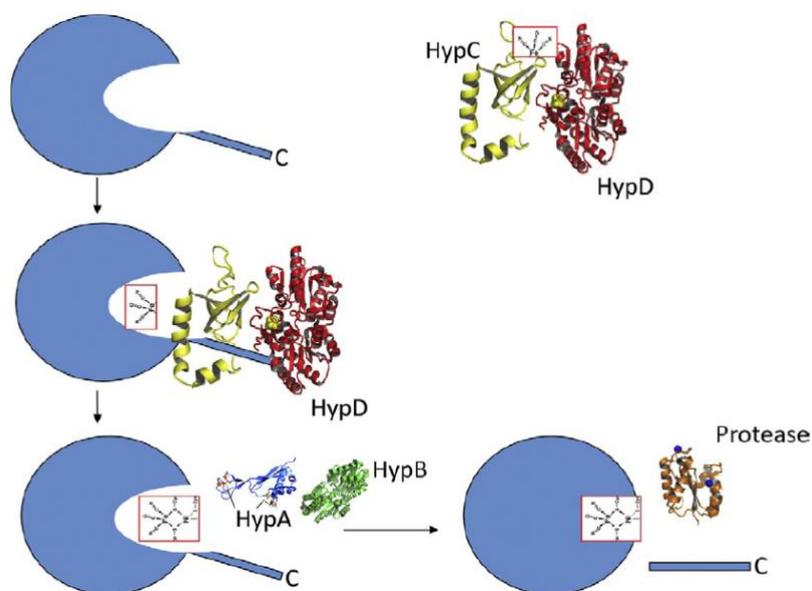
**Figure 7.1. The position of *hypF* close to the *hyc* operon in *E. coli*.** The *hycBEFG* genes encode the soluble domain of Hyd-3, *hycCD* encode the membrane domain and *hycAHI* are involved in gene transcription and enzyme formation (accessory genes). The *hyc* operon is bordered by the *hypA* operon and genes encoding a putative  $\beta$ -glucosidase (*ascBFG*). On the other side of the *hyc* operon (downstream) are two operons (bicistronic operon) which are FhlA-regulated *hydN-hypF* (Filenko *et al.*, 2007).



**Figure 7.2. The organisation of *hyp* locus in *E. coli*.** Genetic organisation around the *hypABCDE-fhla* operon. The *hypAB* gene products help in nickel processing, *hypCDE* are assigned to synthesis of the first half of the [NiFe] cofactor and *fhla* is the  $\sigma^{54}$ -dependent formate- responsive transcriptional regulator. The *hyp* operon is adjacent to the *hyc* operon at one end and the *ygbA* gene at the other, which is a part of the NO-NsrR regulon (Filenko *et al.*, 2007).



**Figure 7.3. The biosynthesis of the first part of the [NiFe] cofactor.** First, carbamoyl phosphate is modified by HypF and HypE to afford a source of  $\text{CN}^-$ . At the same time, the HypCD complex is loaded with  $\text{Fe(II)}$ , ready to receive  $\text{CO}$  and  $\text{CN}^-$ . The source of  $\text{CO}$  might be from tyrosine (as indicated above), which is the case for the [FeFe]-hydrogenase cofactor, or from reduction of  $\text{CO}_2$ . Finally, the first half of the [NiFe] cluster is assembled in the HypCD complex (Sargent, 2016).



**Figure 7.4. The biosynthesis of second half of the [NiFe] cofactor.** The HypCD complex contains a  $\text{Fe-CO-2CN}^-$  cluster, which interacts with the empty apoenzyme of a hydrogenase held in an open conformation due to the availability of the C-terminal peptide. After the insertion of the cluster, the HypCD complex is replaced by HypAB (nickel binding system). Finally, the whole process is condensed and the HypAB cleaved by a specific protease (Sargent, 2016).

### 7.2.2. *nik* operon

The *nik* system was named initially as the *hydC* locus. Wu and Mandrand-Berthelot (1986) found that mutation of the *hydC* locus results in lack of hydrogenase activity which can be recovered by adding 10-500  $\mu\text{g}$  of Ni(II) to the growth medium. Wu *et al.* (1989) found that under such conditions, Ni(II) is transported via the magnesium transport system instead. The whole *nikABCDE*R operon encodes as an ATP binding cassette (ABC) transporter for nickel (Wu *et al.*, 1991). NikA is a soluble periplasmic binding protein of this uptake system. NikB and NikC form the transmembrane part, allowing the passage of  $\text{Ni}^{2+}$  across the cytoplasmic membrane. NikD and NikE are ABC subunits, which hydrolyze ATP and couple this energy to  $\text{Ni}^{2+}$  uptake. Finally, NikR is nickel-responsive transcriptional regulator (Fig 1.14). Therefore, the response to excess  $\text{Ni}^{2+}$  is regulated by NikR, encoded by the last gene in the operon, where it represses transcription of the *nik* operon (de Pina *et al.*, 1991).

## 7.3. Metals required in hydrogenase activity

### 7.3.1. Nickel

Nickel is a base element of [NiFe] hydrogenases, which participates in redox reactions. Nickel itself is an uncommon metal in the environment. Under anaerobic conditions, the *nik* operon is expressed according to *E. coli* nickel requirements (Wu *et al.*, 1989). Thus, nickel requirement of *E. coli* is probably minimal when hydrogenases are not required. Cherrier *et al.* (2008) suggested, according to crystallographic evidence, that butane-1,2,4-tricarboxylate assists with nickel binding in NikA in *E. coli* and Chivers *et al.* (2012) found that nickel uptake is enhanced by the addition of L-histidine to the medium suggesting that  $\text{Ni}^{2+}$  is taken up as a histidine complex.

### 7.3.2. Iron

Iron is one of the main components of the [NiFe] cofactor, where the HypD protein binds to at least one iron ion to form a base for the [NiFe] complex (Soboh *et al.*, 2012). Iron is also required in the Fe-S cluster, which is critical component in [NiFe]–hydrogenase function (Sawers, 1995). In addition, iron is a major metal required in bacterial growth due to its use in haem, Fe-S cluster, and mono- and di-nuclear iron sites within a wide range of proteins. Therefore, iron is an essential nutrient for nearly all bacteria.

### 7.3.3. Molybdate

Pinsent (1953) revealed that traces of molybdate are essential for the formation of an active formate dehydrogenase by a strain of coliform. She found that molybdate has an effect on hydrogenase formation but not on growth. It was established that formate dehydrogenases required an effective and appropriate conversion of molybdate to molybdopterin and incorporation of this into appropriate site in the apo-protein. Uptake is mediated by the periplasmic binding-protein dependent ABC system, ModABC, that also mediates chlorate resistance (Maupin-Furlow *et al.*, 1995). Mutations in *modABC-F* can be suppressed by an increase in molybdate concentration in the growth medium (Lee *et al.*, 1990). Transcription of the *mod* genes is controlled by ModE (ModR). Synthesis of molybdopterin is performed by the products of the *moaABCDE* operon (McNicholas *et al.*, 1997; Leimkühler *et al.*, 2011), which is also controlled by ModE.

### 7.3.4. Selenium

Selenium is one of the metals needed in FdhH enzyme. Zinoni *et al.* (1987) found that a selenocysteine (Sec) residue, specified by an UGA codon associated with a Sec insertion sequence (SECIS), is essential for formate dehydrogenase activity in *E. coli*. The crystal

structure of FdhF in *E. coli* confirmed that selenocysteine is a direct ligand to the molybdenum atom at the active site (Boyington *et al.*, 1997). The *selABCD* genes are required for synthesis and incorporation of Sec into proteins during translation (Dobosz-Bartoszek *et al.*, 2016).

So, as indicated above, H<sub>2</sub> production by *E. coli* is dependent on four distinct metals (nickel, iron, molybdenum and selenium). Ni is specifically required for hydrogenase activity, Mo and Se are required for Fdh-H activity, and iron has a more general role in bacterial metabolism, in addition to its requirement in FHL activity.

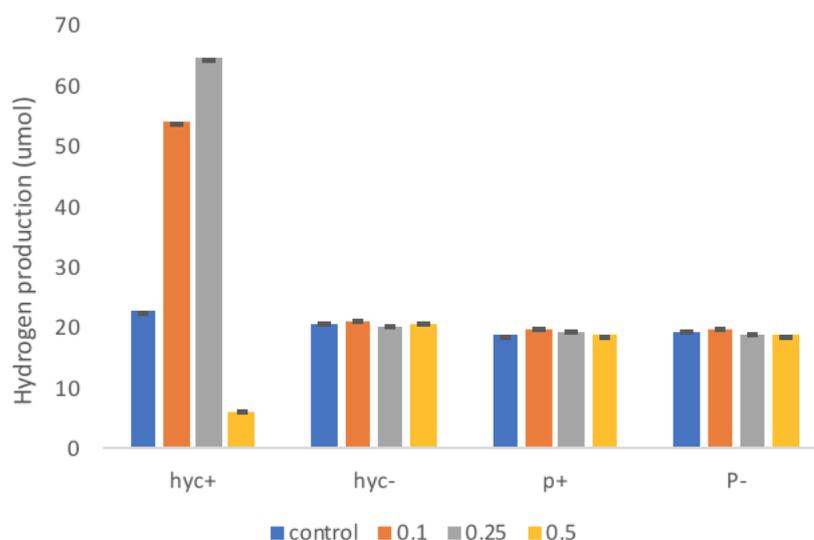
In this chapter, the effect of Ni, Fe, Mo and Se levels on Hyf-dependent H<sub>2</sub> production was tested. In addition, the effect of nickel addition in wildtype on H<sub>2</sub> production was also examined.

#### **7.4. Is there any increase in H<sub>2</sub> production with addition of nickel to wildtype *E. coli* with pBAD<sub>rha</sub>-*hyc* or -*hyf*?**

Previous work has shown an association between hydrogenase activity and Ni<sup>2+</sup> availability in *E. coli* (Rowe *et al.*, 2005). Thus, to enable the effect of Ni<sup>2+</sup> concentration on H<sub>2</sub> production to be determined for Hyf (and Hyc) when expressed from pBAD<sub>rha</sub>, BW25113 (wildtype) strains carrying pBAD<sub>rha</sub>-*hyc* and pBAD<sub>rha</sub>-*hyf* were propagated under a range of Ni<sup>2+</sup> concentrations. The strains were grown in rich medium (with 0.01 M glucose and 120 mM MOPS at pH 6.5) at 37 °C in syringes, as before. NiCl<sub>2</sub> was provided at 0.1, 0.25 and 0.5 mM.

The results show clearly (Fig 7.5) that as nickel levels rise there is little impact on H<sub>2</sub> production in the wildtype vector controls, or the wildtype with the *hyc* plasmid in absence of inducer. However, upon induction of the plasmid-encoded *hyc* operon, H<sub>2</sub> production increased by 2.5 fold with 0.1 mM nickel and by 2.9 fold with 0.25 mM nickel. However, at 0.5 mM nickel the amount of H<sub>2</sub> produced upon *hyc* induction was

dramatically reduced, by 4 fold with respect to that seen without nickel. The reason for this reduction is unclear but does not correlate to a reduction in growth (data not shown). The above finding was in agreement with Carrieri *et al.* (2008) who found that increase in nickel concentration in (*in vivo*), increases the H<sub>2</sub> production by 18 fold. Kim *et al.* (2010) revealed that increasing of nickel concentration up to 30  $\mu$ M, reduces the H<sub>2</sub> production by approx. 50%, but combination of Fe with Ni, increase the production of H<sub>2</sub>. Based on limited time available, further studies need to be done in the future.

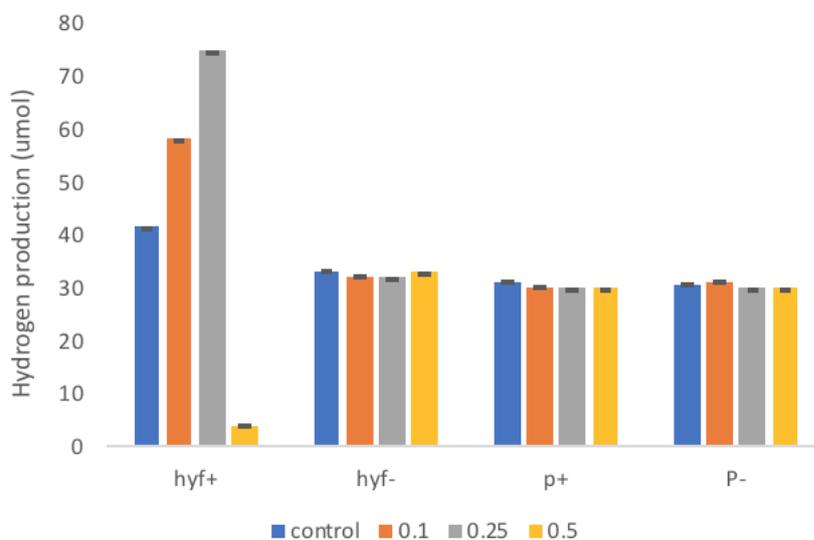


**Figure 7.5. Effect of different concentrations of nickel (II) on hydrogen production in *E. coli* BW25113 with pBAD<sub>rha</sub>-*hyc* induction.** ‘+’ with rhamnose induction and ‘-’ without rhamnose induction. ‘P’ indicate the vector control. The growth was in rich media with 0.01 M glucose at pH 6.5. Incubation was at 37 °C for 30 hours using syringes. Control indicates pBAD<sub>rha</sub>-*hyc* strain without addition of nickel. The numbers, 0.1, 0.25 and 0.5 are different nickel chloride concentrations added to the rich media (in mM). The Durans were washed by acid wash. Growths were performed in triplicate and standard deviations are indicated. The experiment was repeated twice, with similar results obtained.

A similar effect was shown when *hyf* was induced with 0.1 or 0.25 mM nickel in the wildtype, H<sub>2</sub> production was increased by 1.4 and 1.9 fold, respectively ( $P = 0.013$  and  $0.03$ ) (Fig 7.6).

The un-induced *hyf* (Fig 7.6) compared to the induced strain, addition of nickel didn't show any increase in the production of H<sub>2</sub>, as was the case with un-induced *hyc*. However, as before, 0.5 mM nickel resulted in a marked reduction in H<sub>2</sub> produced; again, the reason for this affect remains unclear although it is not related to any marked change in growth (data not shown).

In summary, the results indicate that raising Hyf or Hyc levels artificially results in an increased demand for Ni<sup>2+</sup> and that unless addition Ni<sup>2+</sup> is provided the raised Hyf/Hyc levels are insufficiently provisioned with nickel to enable the production of a fully active hydrogenase. The reason for the reduced H<sub>2</sub> production with 0.5 mM Ni<sup>2+</sup> for the *hyc/hyf* induced strains is unclear, but the effect is consistent for both hydrogenases and is not seen in the absence of *hyc/hyf* induction. This effect requires further investigation to determine any potential defect in the Hyp apparatus arising as a consequence of a combined high Ni<sup>2+</sup> level and excessive Hyf/Hyc production.



**Figure 7.6. Effect of different concentrations of nickel (II) on hydrogen production in *E. coli* BW25113 with pBAD<sub>rha</sub>-*hyf* induction.** ‘+’ with rhamnose induction and ‘-’ without rhamnose induction. ‘P’ indicate the vector control. The growth was in 20 ml rich media with 0.01 M glucose at pH 6.5. Incubation was at 37 °C for 30 hours using 20 ml syringes. Control indicates pBAD<sub>rha</sub>-*hyf* strain without addition of nickel. The numbers, 0.1, 0.25 and 0.5 are different nickel chloride concentrations added to the rich media (in mM). The Durans were washed by acid wash. Growths were performed in triplicate and standard deviations are indicated. The experiment was repeated twice, with similar results obtained.

### 7.5. Effect of metals on H<sub>2</sub> production and bacterial growth

In order to examine the effect of iron, nickel, molybdenum and selenium on H<sub>2</sub> production in the *hycE* mutant strain and wildtype complemented with *hyc* and *hyf*, in parallel to their growth, the transformants were grown in WM-medium (section 2.5.2) with 80 mM glucose and 120 mM MOPS, pH 6.5 and supplemented with 0.8% trace elements (section 2.5.2.1). Acid-wash sterile tubes were used to avoid any adventitious metal contamination. Cultures were incubated overnight at 37 °C. After 16 h, the bacterial culture was inoculated under anaerobic condition in the same fresh medium with the same conditions using 20 ml syringes with the addition of the indicated metals (16 μM Fe, 5

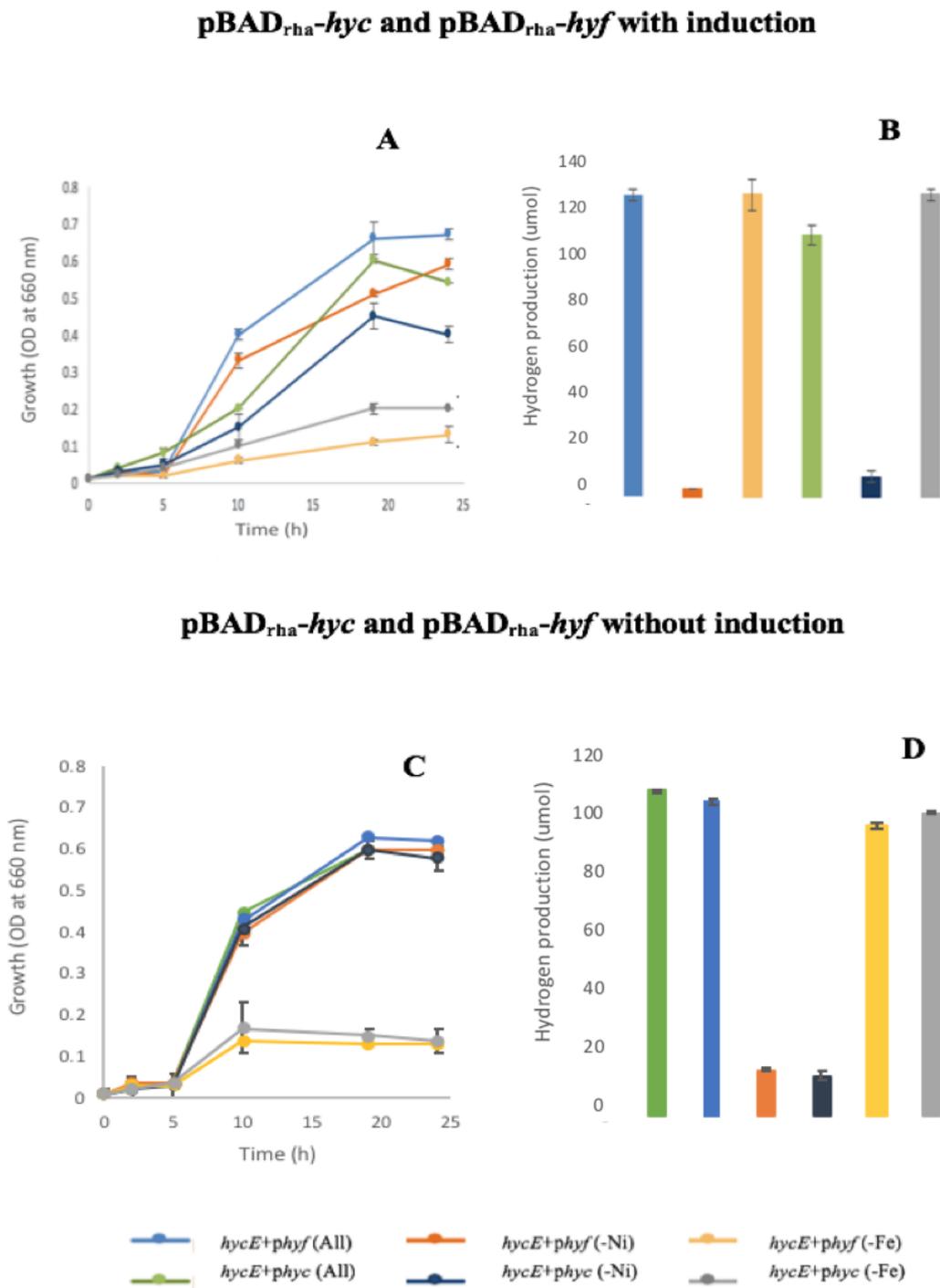
$\mu\text{M}$  Ni, 1  $\mu\text{M}$  Mo and 1  $\mu\text{M}$  Se), either all four in combination or three only in combination (Skibinski *et al.*, 2002). The incubation was for 24 h to detect the H<sub>2</sub> production and to measure the growth trend.

When all the metals were available during *hyc* and *hyf* induction in the wildtype (Fig. 7.7A), growth was stronger for the *hyf* strain than for the *hyc* strain, reaching an OD of 0.4 and 0.2, respectively, at 10 h, and 0.65 and 0.60, respectively, at 19 h. This difference was not seen for the un-induced strains (Fig. 7.7C) indicating that it arose as a consequence of *hyc* induction. In the absence of nickel, induction of either *hyf* or *hyc* resulted in a slight decrease in growth ( $\sim 0.2$  OD units lower than the growth with Ni<sup>2+</sup> at 20 h). However, no such effect was observed for the un-induced controls where the growth was unaffected by the absence of Ni<sup>2+</sup> (Fig. 7.8A). The absence of Ni<sup>2+</sup> had a profound effect on H<sub>2</sub> production for the wildtype with plasmid encoded *hyc* or *hyf*, with or without inducer (Figs 7.7B and 7.7D), with H<sub>2</sub> levels being reduced by up to 30 fold for the induced strains and by  $\sim 7$ -fold for the un-induced strains. The induced strains showed slightly higher ( $\sim 10\%$ ) H<sub>2</sub> production than the uninduced strains when all 4 metals were present, but a much lower level (2-4 fold) was seen when Ni<sup>2+</sup> was absent. This effect likely reflects the raised levels of Hyf/Hyc upon induction that would be expected to dilute the availability of intracellular Ni under Ni restriction and thus lower the level of Ni-replete hydrogenase. For lack of iron, growth of the wildtype (all four formats) was greatly impaired ( $\sim$ fivefold), but more so upon induction and more so for the *hyc* and for the *hyf* induced strain (as was the case with lack of Ni<sup>2+</sup>). However, H<sub>2</sub> production was not affected. This indicates that the reduced growth caused by lack of iron was not related to low hydrogenase activity. In summary, the results (Fig. 7.7) for the wildtype indicate that lack of Ni<sup>2+</sup> results in reduced growth for the *hyf* and *hyc* induced strains due to diminished hydrogenase activity but that lack of Ni<sup>2+</sup> has no such

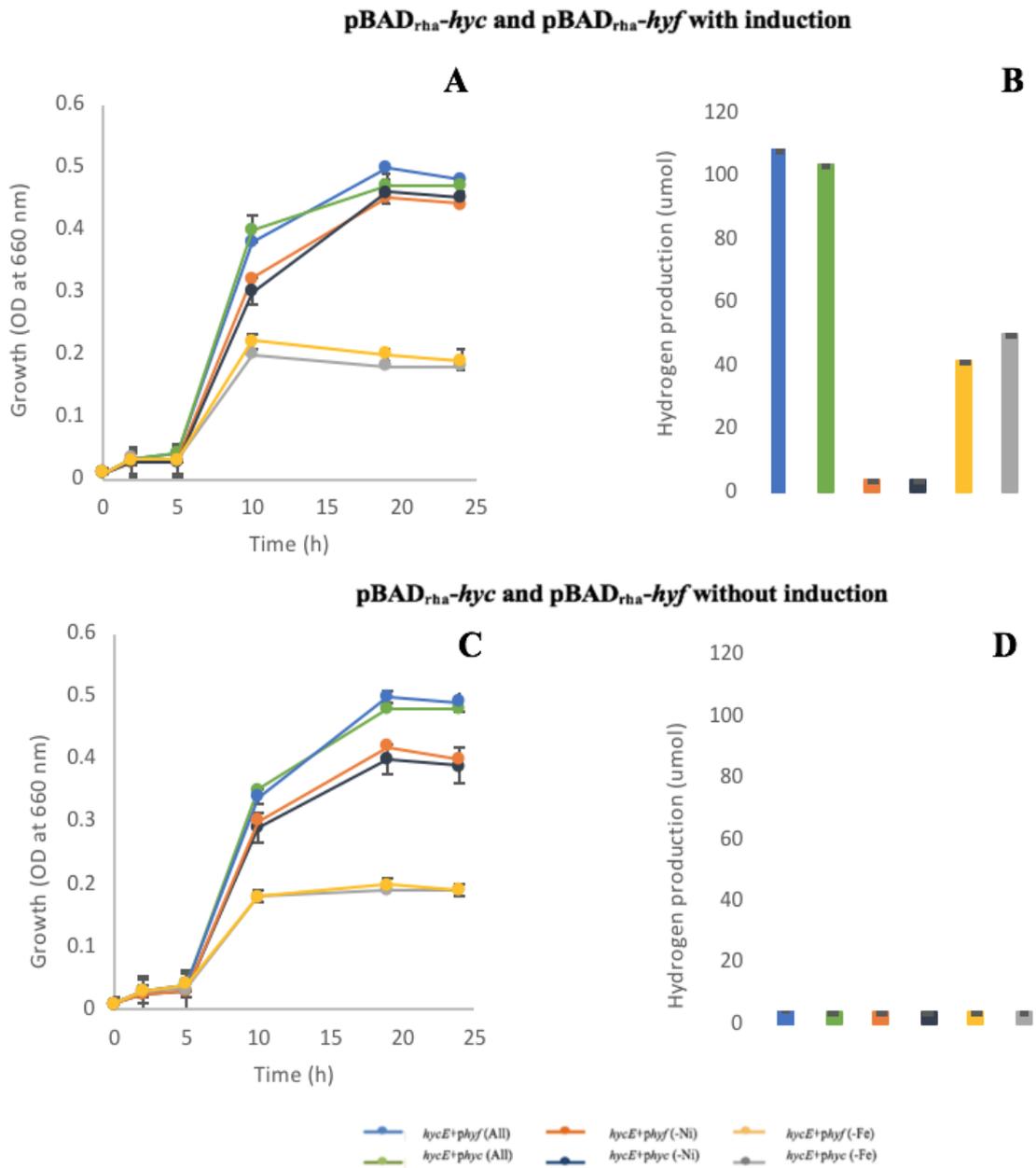
impact on growth when *hyf* and *hyc* remain un-induced, possibly due to the greater hydrogenase activity of the un-induced strains under  $\text{Ni}^{2+}$  limitation.

Results with the *hycE* mutant are shown in Fig. 7.8. It is clear that lack of inducer results in absence of any notable  $\text{H}_2$  production, as expected, no matter what metal regime is employed (Fig. 7.8D). In addition, the lack of  $\text{Ni}^{2+}$  resulted in lower growth (up to 0.1 OD unit) for induced and un-induced strains which is clearly unrelated to hydrogen-production (hydrogenase) activity since no  $\text{H}_2$  was generated in the absence of inducer. This result thus suggests that Ni is required to support the fermentative growth of *E. coli* in a fashion that is independent of  $\text{H}_2$  production. The growth reduction caused by lack of Ni was somewhat more marked for the un-induced than the induced strains; this is in contrast to the results obtained above for the wildtype (Fig. 7.7) where the un-induced strains showed no growth effect in the absence of  $\text{Ni}^{2+}$ . Lack of iron, as for the wildtype, resulted in a major growth reduction (~2.5 fold) in all cases (Fig. 7.9A & C) and in the presence of inducer a reduced level of  $\text{H}_2$  production (twofold) was observed (Fig. 7.8D). This reduction is likely related to the weaker growth under low iron conditions, although it should be noted that the weaker growth of the wildtype caused by low iron did not have any impact on  $\text{H}_2$  production (Fig. 7.7B & D).

In summary, absence of  $\text{Ni}^{2+}$  had a major negative impact on Hyf-dependent  $\text{H}_2$  production which strongly indicates that the Hyf system is nickel dependent (Fig 7.8B).



**Figure 7.7. Effect of iron and nickel on bacterial growth and hydrogen production in *wildtype* strains complimented with pBAD<sub>rha</sub>-hyc, pBAD<sub>rha</sub>-hyf.** A & B with induction, C & D without induction. The growth was in WM-medium supplemented with 0.8% trace elements with 80 mM glucose at pH 6.5. Incubation was at 37 °C for 24 h 20 ml using syringes. 5 μM of nickel chloride as a source of nickel and 1.6 μM ferric citrate as a source of iron were added. The Durans were washed with acid. Growths (A/C) were performed in triplicate and standard deviations are indicated. The experiment was repeated twice, with similar results obtained.

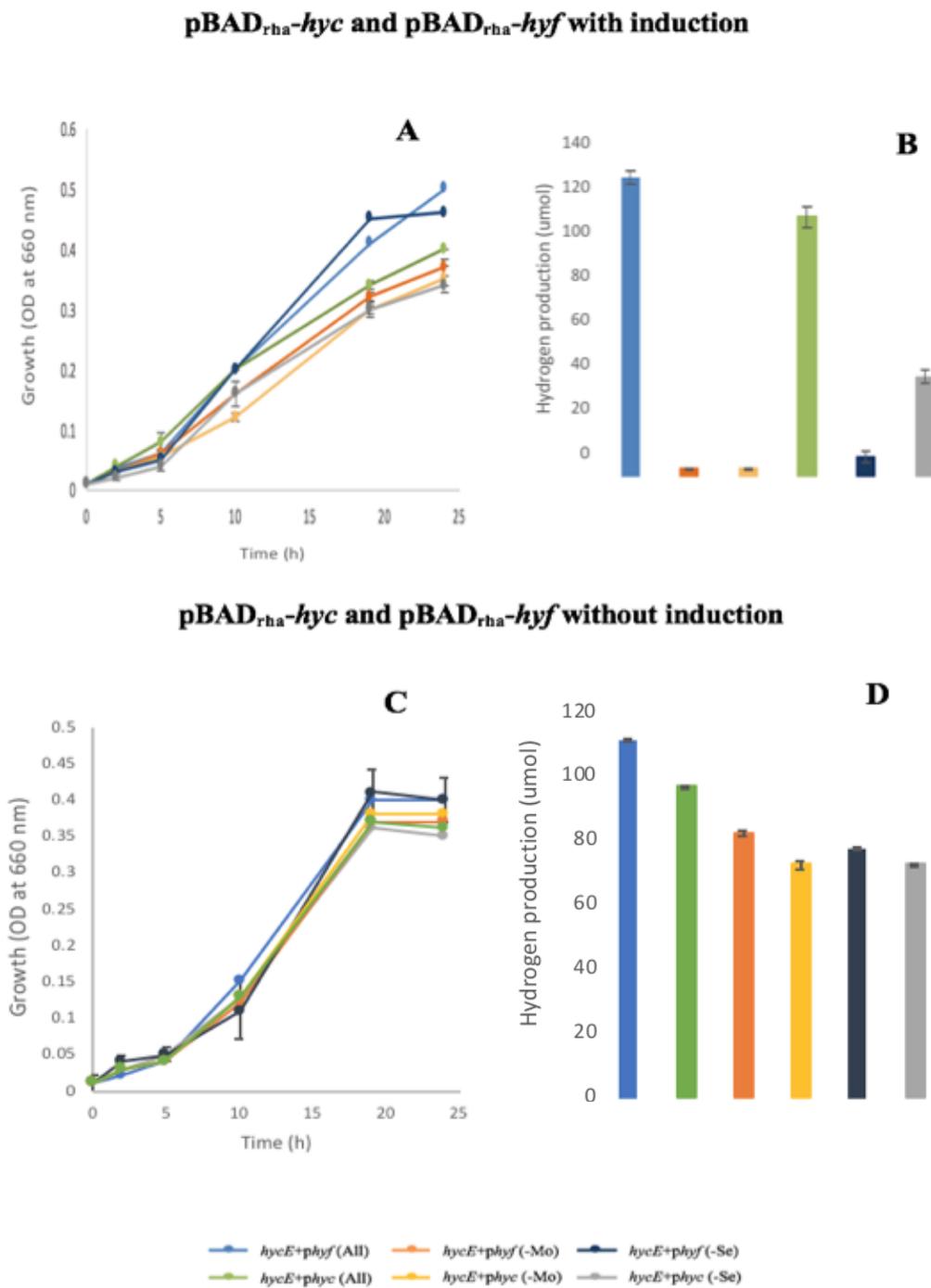


**Figure 7.8. Effect of iron and nickel on bacterial growth and hydrogen production in *hycE* mutant strains complemented with pBAD<sub>rha</sub>-*hyc*, pBAD<sub>rha</sub>-*hyf*.** Details are as for Fig. 7.7 except for the use of the *hycE* mutant in place of the wildtype.

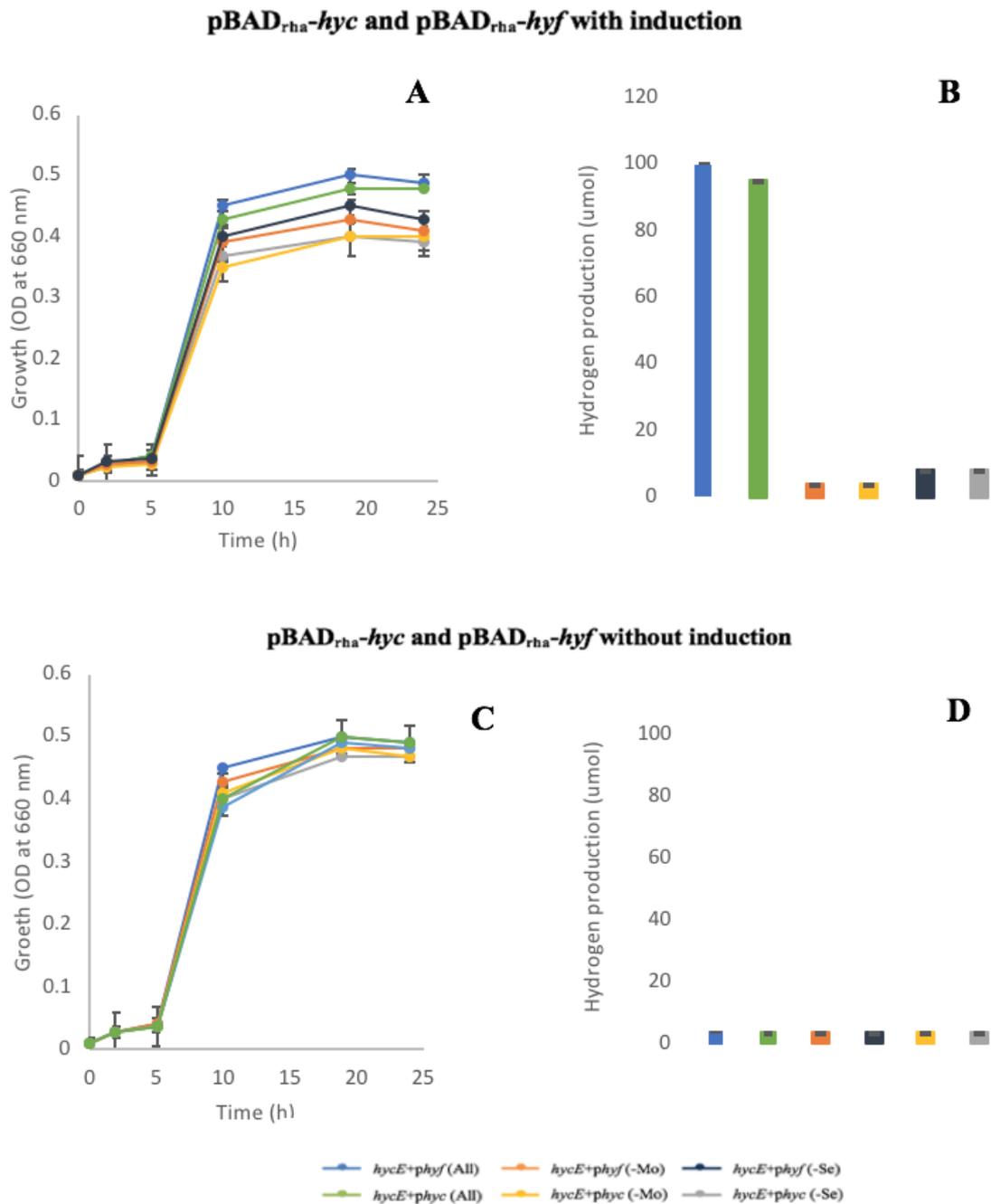
The above experiment was performed again but this time without Mo or Se. The results showed that growth of the un-induced wildtype is little affected by lack of Mo or Se, whereas for the induced wildtype there was a clear growth reduction for the *hyf* plasmid bearing strain when Mo was absent, although (surprisingly) there was an increase in growth when Se was absent (Fig. 7.9A). For the induced wildtype, H<sub>2</sub> production was markedly impaired by lack of Mo (~30 fold) or Se (3-12 fold), with *hyf* induction causing a more marked effect on H<sub>2</sub> production in absence of Se than *hyc* induction (12 fold cf. 3 fold). For the wildtype in the absence of inducer, lack of Mo or Se caused only a slight (10-15%) reduction in H<sub>2</sub>. Thus, high levels of Hyf or Hyc components result in a raised sensitivity to (an apparent higher demand) Mo and Se restriction with respect to H<sub>2</sub> production. This is a surprise since Mo and Se are required for Fdh-H activity, and are not considered as cofactors for Hyc or Hyf. Thus, increased Hyf or Hyc production results in a raised requirement for Mo and Se; the reason for this effect is unclear.

A similar effect was observed for the *hycE* mutant, with lack of Se or Mo resulting in complete absence of H<sub>2</sub> production in the induced condition (Fig. 7.10). This result demonstrates an Se and Mo requirement for Hyf function (as for Hyc) which is anticipated to be related to the dependence of Hyf activity on Fdh-H, which is an Mo (molybdopterin guanine dinucleotide) and Se (selenocysteine) dependent enzyme (chapter 1.section 1.8.3; Boyington *et al.*, 1997; Maupin-Furlow *et al.*, 1995). Thus, the Se and Mo dependence of Hyf activity is consistent with the formation of a second FHL complex by Hyf with Fdh-H. The lack of Mo or Se caused a slight reduction in growth for the induced *hycE* mutant (0.05-0.1 OD unit at 24 h; Fig 7.10A) but had no impact in the absence of inducer (Fig. 7.10C). Whether this growth reduction is related to lack of H<sub>2</sub> production or to a homeostatic imbalance effect caused by a combined low level of Se/Mo with high levels of Hyf/Hyc components is unclear.

In summary, the large reduction in Hyf-dependent H<sub>2</sub> production caused by lack of Se or Mo is consistent with the suggested dependence of Hyf activity on Fdh-H.



**Figure 7.9.** Effect of Molybdenum and Selenium on bacterial growth and hydrogen production in wildtype strains upon induction of pBAD<sub>rha</sub>-*hyc* and pBAD<sub>rha</sub>-*hyf*. Details are as for Fig. 7.7 except for absence of 1 μM sodium molybdate or 1 μM sodium selenite (rather than iron or nickel).



**Figure 7.10. Effect of Molybdenum and Selenium on bacterial growth and hydrogen production in *hycE* mutant strains upon induction of pBAD<sub>rha</sub>-*hyc* and pBAD<sub>rha</sub>-*hyf*.** Details are as Fig. 7.9 except for the use of the *hycE* mutant in place of the wildtype.

## 7.6. Discussion

This chapter focused on the effect of metals (Fe, Ni, Mo and Se) on H<sub>2</sub> production by the Hyf and Hyc systems. The metals selected are those playing an important role in Hyc activity, and therefore also expected to play a similar role for Hyf. Initially, the effect of addition of nickel, as a key component at the active sites of Hyc and Hyf, was tested. The results showed that in rich medium the addition of 0.1 and 0.25 mM Ni<sup>2+</sup> increased H<sub>2</sub> production by 1.8 and 2.8 fold, respectively, in the wildtype upon *hyc* or *hyf* induction. This finding supports the dependence of Hyf activity on the availability of Ni<sup>2+</sup>. The results also clearly showed that both Hyc- and Hyf-dependent hydrogenase activity requires Mo and Se. This is consistent with the proposed role of Fdh-H in combining with the Hyf complex to form a second FHL in *E. coli*, since Fdh-H is a selenocysteine and molybdopterin guanine dinucleotide (MGD) dependent enzyme (Sawers, 1994).

The relatively weak effect of low iron on H<sub>2</sub> production in contrast with the strong effect on growth suggests that the lack of iron impairs growth due to the reduced activity of other iron requiring systems in the cell. Indeed, it is well known that *E. coli* undergoes an iron rationing response under low-iron conditions whereby iron-dependent proteins (particularly those involved in respiration) are down regulated in order to allow a more efficient use of cellular iron resources (McHugh *et al.*, 2003). Since H<sub>2</sub> production remained relatively high under iron restriction, it seems likely that the Hyf and Hyc systems are not subject to the iron-rationing response.

Nickel is absolutely required for [NiFe] centre formation and H<sub>2</sub> production. In the synthesis of the iron centre, nickel is last metal inserted using metallochaperones (HypA, HypB) (Peters *et al.*, 2015; Lacasse and Zamble, 2016) and without the nickel insertion, the large subunit will not be complete and hydrogenase activity will be affected such that no H<sub>2</sub> will be produced. Indeed, it was displayed clearly in the results above that absence

of nickel affects the production of H<sub>2</sub> for both Hyc and Hyf, but had relatively little impact on growth.

Mo is an important component of formate dehydrogenase-H, which is a part of the FHL-complex that converts formate to H<sub>2</sub>. Consequently, lack of Mo affects the production of H<sub>2</sub> as observed here for both Hyc and Hyf. Selenium is also a key metal for FHL (Boyington *et al.*, 1997; George *et al.*, 2008 and Khangulov *et al.*, 1998). Indeed, all the three formate dehydrogenases (section 1.7.1) of *E. coli* are selenoenzymes, which have MGD as a cofactor. Thus, deficiency in Se will also affect FdhH and consequently affect H<sub>2</sub> production.

In summary, it was clearly reported that Ni, Mo and Se are requirements for Hyf and Hyc activity which supports the formation of the proposed FHL-2 complex between Fdh-H and Hyf.

## Chapter 8. General discussion

### 8.1. Introduction

Hydrogenases are metallo-enzymes containing nickel and/or other metals, such as iron, at their active sites. Also, they contain organometallic ligands such as carbon monoxide and cyanide, (Shafaat *et al.*, 2013). These metalloenzymes are used in energy metabolism and are responsible for catalysing the reversible cleavage of H<sub>2</sub> into protons and electrons (Hjersing, 2011; Shomura *et al.*, 2011). The activity of the enzyme was first reported in 1930, when it was shown that *E. coli* produces H<sub>2</sub> during their growth at lower pH under anaerobic conditions (Stephenson and Stickland, 1930).

Hydrogenases are classified, according to their active site (metal ion content), into three different types: nickel-iron [Ni-Fe]-, di-iron [Fe-Fe]-, and mono-iron (iron-only) [Fe]-hydrogenases (Lukey *et al.*, 2010). Generally, hydrogenases are sensitive to anaerobic conditions and this is applicable to *E. coli* hydrogenases (Abou Hamdan *et al.*, 2012), although Hyd-1 is considered oxygen tolerant. In *E. coli*, there are four types of [Ni-Fe] hydrogenases. These each contain a small subunit comprising iron-sulphur clusters, used in electron transfer, and a large subunit which contains the nickel-iron active site. Normally, hydrogenase 1 (Hyd-1) and hydrogenase 2 (Hyd-2) are responsible for the anaerobic oxidation of H<sub>2</sub> (linked to respiration), while hydrogenase 3 (Hyd-3) is a part of the formate hydrogen lyase (FHL) complex (Soboh *et al.*, 2011), which produces H<sub>2</sub> from formate in order to dispose of the excess formate generated during mixed acids fermentation. In *E. coli*, under anaerobic conditions, synthesis of Hyd-1 and Hyd-2 is stimulated, and during fermentation the third hydrogenase appears in response to formate build up. Surprisingly, the operon encoding hydrogenase 4 (Hyd-4) seems to be silent since the *hyf-focB* genes are not expressed at significant levels (self *et al.*, 2004; Skibinski *et al.*, 2002; Sanchez-Torres *et al.*, 2013). However, Hyd-4 shows similar sequence in the

majority of its predicted subunits to those of Hyd-3, but Hyd-4 has three additional subunits that render it even more similar to Complex I than Hyd-3. Hyd-4 is encoded by a twelve-gene operon (*hyfABCDEFGHIJR-focB-hyfR*) and is thought to form a ten-subunit, membrane-associated hydrogenase complex. The operon also encodes a formate- and sigma 54-dependent transcriptional activator (HyfR, homologous with FhlA), and a putative formate transporter channel (FocB homologous to FocA) (Skibinski et al., 2002; Andrews *et al.*, 1997).

The main objective in this study was to investigate and characterize the role of FocB in formate transport, in comparison to FocA. In addition, a major aim was to induce *hyf* operon expression to determine the function of the proposed Hyd-4 in H<sub>2</sub> production and to see if Hyd-4 forms a second FHL complex in *E. coli* providing a second route for formate consumption under fermentation conditions.

## **8.2. *hyf* is not fully conserved in *E. coli* strains**

Seven genes of the *hyf* (*hyfABCDEFGHIJR-focB*) operon are homologues to corresponding *hyc* genes (encoding Hyd-3) (Andrews *et al.*, 1997) (Fig 3.1). However, three genes (*hyfD*, *hyfE*, and *hyfF*) encode integral membrane subunits which have no direct equivalent in Hyd-3 (Skibinski *et al.*, 2002). An extensive set of phylogenetically distinct *E. coli*, *Shigella* and *Salmonella* strains were assembled into a phylogenetic tree using the RpoB protein amino acid sequence (Case *et al.*, 2007) in order to compare the phylogenetic relationship between *hyf*<sup>+</sup> and *hyf*<sup>-</sup> strains (Fig. 3.2). From the 94 representative strains analysed, 45 *E. coli* and 9 *Shigella* strains contained *hyf*, while 18 *E. coli* strains and all 22 *Salmonella* species were *hyf* free. The tree generated resembled trees previously produced, such as that of Fukushima *et al.* (2002) using *gyrB* gene sequences. Interestingly, the 18 *hyf*-free *E. coli* strains were found to form a single clad

that corresponds to the B2 phylogroup, identified previously (Sims and Kim, 2011). These strains are associated with urinary tract infection diseases (UPEC) isolates and avian pathogenic species (APEC) whereas most of the *hyf*-carrying strains are faecal isolates and the B2 phylogroup is the most basal *E. coli* phylogroup. This suggests that the *hyf* operon might provide a benefit for strains occupying the mammalian intestine, and thus may not be important for *E. coli* survival at other sites, such as the urinary tract. It is possible that the *hyf* operon was gained by the non-B2 phylogroup *E. coli* common ancestor by horizontal gene transfer as a genomic island (GI). Alternatively, *hyf* may have been present in the most recent common ancestor for all *E. coli* strains (since it is present in *E. fergusonii*), but was subsequently lost from the B2 branch as part of an adaptive process for life outside of the mammalian intestine.

In contrast to *hyf*, all *Salmonella*, *Shigella* and *E. coli* strains analysed, contain the *hyc* operon, which offers the capacity to eliminate 'toxic' formate under conditions of fermentation. It is supposed that Hyf offers the advantage over Hyc of providing energy (or more energy) from the formate release, under conditions where H<sub>2</sub> is maintained at relatively low levels (Andrews *et al.*, 1997). However, this supposition remains to be proven.

### **8.3. FocB as a formate transporter**

FocB is a putative formate channel homologous with FocA (50% amino acid sequence identity), encoded by the last gene of the *hyf* operon (Andrews *et al.*, 1997). Earlier studies showed that the *focA* mutant produces less formate than the wildtype by 50% (Suppmann and Sawers, 1994). From a simple bioinformatics analysis, it was shown here that FocB shares with FocA all of the important structural and functional elements known to be required for function as a formate channel. The sequence alignment showed

similarity in the formate binding ligands (Lys156, Glu208, His209 and Thr91), which form H-bonds between the formate substrate and the FocA/ FocB channel in *E. coli* and *Salmonella* (Waight *et al.*, 2010; Wiechert and Beitz, 2017). In addition, they have similar pore constriction residues at both the cytoplasmic and periplasmic sites. Furthermore, at the C-terminus, the residues considered important in pH-regulation of transport mode are fully conserved (Hunger *et al.*, 2017) (Fig. 4.1). These results suggested that FocB is pH, formate dependent, like FocA. A direct analysis of FocB function was achieved by cloning *focB* into an inducible vector (pBAD<sub>rha</sub>) followed by induced expression of *focB* in a *focA* mutant strain. This approach demonstrated that HP (toxic formate analogue) inhibition under anaerobic fermentation conditions was enhanced by expression of *focB*, as was the case for *focA* – such that 0.5 mM HP concentration was sufficient to inhibit growth (Fig. 4.10; Fig. 4.12). This effect was absent under aerobic condition (Fig. 4.9A). These findings are in agreement with those of Suppmann and Sawers (1994) where FocA was presumed to offer a portal for uptake of HP resulting in growth inhibition through interference with Pfl activity. The effect of induction of *focB* was also analysed at different pH levels (Fig. 4.13). This showed that induction of *focB* causes formate toxicity at pH 6 and 8, and prevents growth at pH 7 (with or without formate), and only restored HP sensitivity at pH 6 (Fig. 4.13). Interestingly, HP showed low toxicity at pH 8 in the wildtype, but toxicity was increased when plasmid-encoded *focA* was induced indicating that the low HP toxicity at pH 8 is related to poor chromosomal *focA* expression at this pH. Thus, FocA appears inactive at high pH.

Formate production was reduced by twofold by *focA* mutation, as observed previously (Beyer *et al.*, 2013). However, since formate was still released in the *focA* mutant (although in lower amount) this indicates a second formate exports channel, as previously reported (Beyer *et al.*, 2013).

Induction of *focA* or *focB* increased formate production in the *focA* mutant at pH 6.5 to at least partly compensate for the low formate production observed for the uncomplemented *focA* mutant. Induction of *focA* gave a stronger increase in formate production than did *focB* induction (Fig. 4.17). The reduced H<sub>2</sub> production of the *focA* mutant at pH 6.5 was reversed by induction of plasmid-encoded *focA* and also by non-induced *focB* (Fig. 4.19), and in addition induction of *focA* or *focB* at this pH promoted the growth of the *focA* mutant under fermentation conditions (Fig. 4.17D). The results thus show that, at pH 6.5, both FocA and FocB promote formate export, and support H<sub>2</sub> production and fermentative growth. At pH 7.5, chromosomally encoded *focA* appeared to have little function in formate export (possibly due to lack of expression), but when *focA* was plasmid-induced a reduced formate export was observed, suggesting a synthetic production of FocA reduces release of formate at pH 7.5.

#### **8.4. Hyf is an active H<sub>2</sub>-producing (and formate-consuming) hydrogenase, and can thus be designated as the fourth hydrogenase (Hyd-4) of *E. coli***

The ability of Hyd-4 (encoded by *hyf*) to generate H<sub>2</sub> and its formate dependency were investigated by cloning the 13 kb *hyf* operon into an inducible plasmid (pBAD<sub>rha</sub>); this plasmid allowed induction of *hyf* using rhamnose as an inducer. The pBAD<sub>rha</sub>-*hyf* plasmid was then used to complement a  $\Delta hycE$  strain (lacking Hyd-3 activity) and the transformant was tested for H<sub>2</sub> production at different pH levels. The results showed that H<sub>2</sub> production is restored when *hyf* is induced in the  $\Delta hycE$  strains, at all pH levels tested. The highest amount of H<sub>2</sub> production was at pH 6.5 for both *hyc* and *hyf* induction (Fig. 5.10). However, H<sub>2</sub> was not produced in the  $\Delta hycE$  strain carrying the vector control. This suggests that Hyd-4 is an H<sub>2</sub> producer (only when *hyf* is induced) as is Hyd-3 at acidic pH level

The effect of formate addition was tested to determine whether there is any relation between formate levels and H<sub>2</sub> production at pH 5.5-7.5. It was previously reported that as formate concentration increases, H<sub>2</sub> production also increases for *E. coli* grown fermentatively (Yoshida *et al.*, 2005; Mnatsakanyan *et al.*, 2002). The results generated here showed that at pH 5.5 and 7.5, very little H<sub>2</sub> is produced upon *hyc* and *hyf* induction. However, at pH 6.5, the H<sub>2</sub> was produced at relatively high levels, and these were increased further with addition of 10 mM formate by 50%. This effect was presumed to relate to a direct substrate-availability-stimulated increase in FHL activity. This finding suggests that Hyf is forming a second FHL, presumably by interacting with Fdh (Fig. 5.10).

Formate production and consumption were also measured under anaerobic fermentation growth conditions with *hyc* and *hyf* induction and it was found that formate build up in the medium (at pH 6.5) (Fig. 5.14) reached a peak at 10 h, with maximum production slightly higher in the wildtype than in the *hycE* mutant. At the end of the growth, the released formate was mostly consumed in the wildtype, but in the *hycE* mutant relatively little formate consumption occurred. This poor formate consumption of the *hycE* mutant was reversed by induction of plasmid-borne *hyc* or *hyf*, which suggests that Hyf, similar to Hyc, drives formate consumption. This provides further evidence that Hyf is a part of a novel FHL (FHL-2) that converts formate to H<sub>2</sub>. The induction of *hyf* produced even more H<sub>2</sub> than did induction of *hyc* at low pH (6 and 6.5). On the other hand, production of H<sub>2</sub> by *hyc* induction was greater than that of *hyf* at higher pH (6.5 and 7). This suggests that Hyc may function more effectively than Hyf at alkaline pH.

Experiments were also performed to test if formate can be used to promote the growth of a Hyc<sup>+</sup> or Hyf<sup>+</sup> strain under fermentation conditions (Fig. 5.14). Surprisingly, results suggested that Hyf (but not Hyc) promotes growth with formate as sole energy source in

both the *hycE* strain and wildtype. The unexpected result might be explained by the presence of rhamnose as induced that might be used as a carbon source to support growth of the Hyf<sup>+</sup> strains. This finding clearly demands further investigation to confirm the preliminary results obtained.

Hyd-4 and Hyc-3 activity was detected by activity staining using a method reported by Pinske *et al.* (2012) (Fig. 5.17), as diffuse bands at the top of the gel upon induction of pBAD<sub>rha</sub>-*hyf* or -*hyc* appeared. This suggests that Hyd-4 can oxidise hydrogen to reduce BV, further indicating its hydrogenase capacity.

### 8.5. FdhH is required for Hyf activity

The role of FdhH in supporting Hyf activity was examined in order to determine whether these two components might combine to form a second FHL. The *fdhF* gene and *hyf* operon were provided, separately, in inducible and compatible vectors (pBAD<sub>rha</sub> and pBAD<sub>ara</sub>) to test their capacity to enable H<sub>2</sub> production under fermentative conditions. The results showed that neither *fdhF* nor *hyf* alone are able to produce H<sub>2</sub> singly, when provided in a  $\Delta fdhF \Delta hycE$  mutant. However, when both were induced together, H<sub>2</sub> was released, which strongly suggests that Hyf interacts with FdhH to form FHL-2.

Previous studies suggested that the products of *hyc* and *fdhF* form FHL, which transforms formate into H<sub>2</sub> and CO<sub>2</sub>. McDowall *et al.* (2014) found that FHL is indeed a combination of Hyc and Fdh-H and that a mutation in *hycC* and *hycD* prevents the Hyc complex from interacting with FdhH to form the FHL complex. Therefore, formation of the second FHL, as strongly suggested by the data provided here, provides the first evidence for two FHL complexes in *E. coli* K-12, and by inference in all other *hyf*-bearing *E. coli* strains. Formate addition of 10 mM increased the Hyc-FdhH and Hyf-FdhH dependent H<sub>2</sub> production by 2.6 and 2 fold, showing that even in the absence of FhlA/HyfR dependent

formate induced expression of the corresponding genes, FHL-1 and -2 activity is regulated by formate (Fig. 6.16).

The effect of the combined induction of *fdhF* and *hyf*, or *fdhF* and *hyc*, on formate levels under fermentative conditions was tested, in the  $\Delta fdhF \Delta hycE$  mutant (Fig. 6.17). Formate production was increased somewhat in the early stages of growth (possibly due to a growth rate difference), but formate consumption was greatly enhanced (~fourfold) in the latter stages of growth, with respect to the vector controls (Fig. 6.17). In addition, the culture density was markedly increased in the latter stages of growth (~5-6 fold) indicating an important role for the FHL systems in limiting formate toxicity and/or utilisation of formate for energy. This finding highlights a role of FHL-2 in formate consumption, as well as H<sub>2</sub> production, similar to that for FHL-1.

### **8.6. Hyf (FHL-2) activity is Ni, Se and Mo dependent**

The effect of nickel addition to a wildtype was tested in H<sub>2</sub> production upon induction of *hyc* and *hyf* using inducible vectors. Previous study showed by Menon and Robson (1994), that addition of 25  $\mu$ M of nickel increases the hydrogenase activity sharply in *Azotobacter vinelandii*. Thus addition of different Ni amounts was tested with pBAD<sub>rha-hyc / hyf</sub> (Fig. 7.5; Fig. 7.6). The results were similar in both inductions and showed that addition of 0.1 and 0.25 mM nickel, increases the H<sub>2</sub> production gradually. However, higher nickel concentration (0.5 mM) affected the H<sub>2</sub> production in opposite manner, where very little H<sub>2</sub> was produced. The reason for this effect is unclear but suggests a dysregulation in FHL assembly possibly caused by Ni interference in the distribution of other metals required for FHL function.

The effect of addition of Ni<sup>2+</sup> on FHL-2 dependent H<sub>2</sub> production was tested during fermentation in a minimal medium. The addition of 5  $\mu$ M nickel increased H<sub>2</sub> production

by approx. 30 fold with *hyc* induction and *hyf* induction in the  $\Delta hycE$  strain (Fig. 7.7) compared to the strain with no Ni (Fig 7.8), indicating that the plasmid-borne *hyf* and *hyc* genes resulted in enhanced expression (as suggested by the activity staining results) of Hyf and Hyc components leading to insufficient  $Ni^{2+}$  availability during growth in the supplemented minimal medium. Also, Ni is a main metal found as [NiFe] hydrogenase (Volbeda and Fontecilla-Camps, 2003), and the absence of Ni will affect the large subunit of all hydrogenases. The requirements for Mo and Se in FHL-1 and -2 dependent  $H_2$  production were also tested. For both metals, lack of their addition to minimal medium resulted in a major reduction in FHL-1 and FHL-2 dependent  $H_2$  production when *hyf* or *hyc* were induced from the pBAD<sub>tha</sub> vectors (Fig. 7.8), although there was little effect on growth. This suggest that excess production of Fdh-H places excessive demands on Se and Mo requirements such that these metals became insufficient without supplementation. The results are consistent with Mo (Self and Shanmugam, 2000) and Se inclusion in the active site of FdhF as single Mo atom coordinated by a selenocysteine residue and two molybdopterin guanine dinucleotide moieties (Rosentel *et al.*, 1995; Boyington, J. C. *et al.*, 1997).

## 8.7. Conclusion

In summary, the work reported in this thesis shows that *hyf*-free *E. coli* strains grouped in the B2 phylogroup of the phylogenetic tree. Also, FocB was found to be as a second formate channel and as shown by its ability to promote HP sensitivity and affect both formate and  $H_2$  levels during fermentation. The most interesting finding was that expression of the *hyf* operon produces  $H_2$  and it was found as well that Hyf is formate dependent (like-Hyc). The activity of Hyf was shown to be FdhF dependent  $H_2$ , which indicates that Hyd-4 forms a second FHL with FdhF, which was a novel finding. Also, it

was found that H<sub>2</sub> production by FHL-2 is Ni, Mo and Se dependent, as is the case for FHL-1. However, the reason of the presence of two H<sub>2</sub>- producing and formate consuming FHL systems in *E. coli* remains unclear, as are the environmental conditions that promote *hyf* expressed in the wildtype and under which Hyf provides a growth advantage over Hyc.

## 8.8. Future work

1. Further proof is needed to confirm that FocB it is a formate channel by performing direct transport studies e.g. using liposomes or membrane vesicles. Direct <sup>14</sup>C-formate transport experiments could be used to demonstrate FocB as a formate channel and determine any role as a H<sup>+</sup>/formate symporter.
2. Differences in the transport function of FocA and FocB need to be established in order to understand their distinct physiological roles.
3. The potential exchange of Hyf and Hyc components should be explored to determine whether these two systems are entirely independent
4. Perform <sup>63</sup>Ni labelling to detecting labelled HyfG (Hyf large subunit) and to prove that Hyd-4 is Ni containing. This work, along with activity staining, should be done under a range of growth conditions and mutant backgrounds to examine the environmental conditions and accessory factors that affect Hyf activity.
5. The effect of *hyp* and *nik* mutations on Hyf activity should be determined to confirm the role of Ni uptake and hydrogenase processing on Hyf activity. The role of the HycI maturation endo-protease on Hyf hydrogenase large subunit (HyfG) processing should also be tested, since the *hyf* operon encodes no such equivalent
6. It will be important to directly compare the expression of the *fdhF*, *hyc* and *hyf* genes to determine any difference in the conditions that favour their expression,

to provide a clue to the specific roles of Hyc and Hyf. In particular, it will first be necessary to obtain natural expression of *hyf* within a wildtype *E. coli* strain.

This might be achieved by studying strains directly isolated from the intestine.

7. The possibility that Hyf supports growth with formate should be more thoroughly tested, under conditions that would favour energy generation from the FHL reaction.
8. More experiments are required on Hyf as an energy-conserving Fhl by continuing and expanding the experiment given in Fig 5.21.

## References

- Abaibou, H., J. Pommier, S. Benoit, G. Giordano and M. A. Mandrand-Brthelot. 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.
- Abou Hamdan, A., P Lieb Gott, V. Fourmond, O. Gutierrez-Senz, A. L. De Lacey, P. Infossi, M. Rousset, S. Dementin and C. Leger. 2012. Relation between anaerobic inactivation and oxygen tolerance in a large series of NiFe hydrogenase mutants. *Proceedings of the National Academy of Sciences of the United States of America*. **109**: 19916-19921.
- Achebach, S., Q. Hon Tran, A. Vlamis-Gardikas, M. Müllner, A. Holmgren and G. Uden. 2004. Stimulation of Fe-S cluster insertion into apoFNR by *Escherichia coli* glutaredoxins 1, 2 and 3 in vitro. *FEBS Letter*. **565**: 203-306.
- Adams, M. W. 1990. The structure and mechanism of iron-hydrogenases. *Biochimica et Biophysica Acta*. **1020**: 115-145.
- Addy, C., M. Ohara, F. Kawai, A. Kidera, M. Ikeguchi, S. Fuchigami, M. Osawa, I. Shimada, S. Y. Park, J. R. Tame and J. G. Heddle. 2007. Nickel binding to NikA: an additional binding site reconciles spectroscopy, calorimetry and crystallography. *Acta Crystallogr D Biol Crystallogr*. **63**: 221-229.
- Andersen, P. S., M. Stegger, M. Aziz, T. Contente-Cuomo, H. S. Gibbons, P. Keim, E. V. Sokurenko, J. R. Johnson and L. B. Price. 2013. Complete genome sequence of the epidemic and highly virulent CTX-M-15 producing H30-Rx subclone of *Escherichia coli* ST131. *Genome Announcement*. **1**: 1-2.
- Andrews, S. C., B. C. Berks, J. McClay, A. Andrew, M. A. Quail, P. Golby and J. R. Guest. 1997. A 12-cistron *Escherichia coli* operon (*hyf*) encoding a putative proton-translocating formate hydrogenlyase system. *Microbiology*. **143**: 3633-3647.
- Appenzeller, B. M. R., M. Batt'e, L. Mathieu, J.-C. Block, V. Lahoussine, J. Cavard, and D. Gatel. 2001. Effect of adding phosphate to drinking water on bacterial growth in slightly and highly corroded pipes. *Water Research*. **35**: 1100– 1105.
- Appenzeller, B. M. R., C. Yañez, F. Jorand and J. Block. 2005. Advantage provided by iron for *Escherichia coli* growth and cultivability in drinking water. *Applied and Environmental Microbiology*. **71**: 5621-5623.
- Argun, H. and F. Kargi. 2011. Bio-hydrogen production by different operational modes of dark and photo-fermentation: An overview. *International Journal of Hydrogen Energy* **36**: 7443-7459.
- Avasthi, T. S., N. Kumar, R. Baddam, A. Hussain, N. Nandanwar, S. Jadhav and N. Ahmed. 2011. Genome of multidrug resistant uropathogenic *Escherichia coli* strain NA114 from India. *Journal of Bacteriology*. **193**: 4272-4273.

## References

- Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner and H. Mori. 2006. Construction of *Escherichia coli* K-12 in-frame, single gene knockout mutants: the keio collection. *Molecular Systems Biology*. **2**: 1-11.
- Bach, P. B., M. J. Kelley, R. C. Tate and D. C. McCrory. 2003. Screening for lung cancer: a review of the current literature. *Chest*. **123**: 72S-82S.
- Badia, J., L. Baldoma, L. Aguilar and A. Boronat. 1989. Identification of the *rhaA*, *rhaB* and *rhaD* gene products from *Escherichia coli* K12. *FEMES Microbiology letters*. **65**: 253-258.
- Bagramyan, K., N. Mnatsakanyan, A. Poladian, A. Vassilian and A. Trchounian. 2002. The roles of hydrogenases 3 and 4, and the F0F1-ATPase, in H<sub>2</sub> production by *Escherichia coli* at alkaline and acidic pH. *Federation of European Biochemical Societies Letters*. **516**: 172-178.
- Ballantine, S. P., and D. H. Boxer. 1985. Nickel-containing hydrogenase isoenzyme from anaerobically grown *Escherichia coli* K-12. *Journal of Bacteriology*. **163**: 901-908.
- Ballantine S. P., and D. H. Boxer. 1986. Isolation and characterisation of a soluble active fragment of hydrogenase isoenzyme 2 from the membranes of anaerobically grown *Escherichia coli*. *European Journal of Biochemistry*. **156**: 277-84.
- Barrios, H., B. Valdeema and E. Morett. 1999. Compilation and analysis of sigma (54)-dependent promoter sequences. *Nucleic Acid Research*. **27**: 4305-4313.
- Bassegoda, A., C. Madden, D. W. Wakerley, E. Reisner and J. Hirst. 2014. Reversible interconversion of CO<sub>2</sub> and formate by a molybdenum-containing formate dehydrogenase. *Journal of the American Chemical Society*. **136**: 15473-15476.
- Batista, A. P., B. C. Marreiros and M. M. Pereira. 2013. The antiporter-like subunit constituent of the universal adaptor of complex I, group 4 membrane-bound [NiFe]-hydrogenases and related complexes. *Biological Chemistry*. **394**: 659-666.
- BC Bioenergy Strategy. 2010. <http://energybc.ca/biofuels.html>.
- Becker, A., K. Fritz-Wolf, W. Kabsch, J. Knappe, S. Schultz and A. F. Volker Wagner. 1999. structure and mechanism of the glycyl radical enzyme pyruvate formate-lyase. *Nature Structural Biology*. **6**: 969-975.
- Behera, S., R. Singh, R. Arora, N. K. Sharma, M. Shukla and S. Kumar. 2014. Scope of Algae as third generation biofuels. *Front Bioengineering and Biotechnology*. **2**: 1-13.
- Bekker, M., S. de Vries, A. Ter Beek, K. J. Hellingwerf and M. J. Teixeira Mattos. 2009. Respiration of *Escherichia coli* can be fully uncoupled via the nonelectrogenic terminal cytochrome *bd*-II oxidase. *Journal of Bacteriology*. **191**: 5510-5517.
- Benemann, J. R. 1998. Process analysis and economics of biophotolysis of water. IEA agreement on the production and utilization of hydrogen. *IEA/H2/10/TR2-98*. 1-35.

## References

- Benoit, S., H. Abaibou and M. A. Mandrand-Berthelot. 1998. Topological analysis of the aerobic membrane-bound formate dehydrogenase of *Escherichia coli*. *Journal of Bacteriology*. **180**: 6625-6634.
- Benoit, S., N. Mehta, G. Wang, M. Gatlin and R. J. Maier. 2004. Requirement of *hydD*, *hydE*, *hypC* and *hypG* genes for hydrogenase activity in *Helicobacter pylori*. *Microbiology Pathogen*. **36**: 153-157.
- Benemann, J. R. 1998. Process analysis and economics of biophotolysis of water. IEA agreement on the production and utilization of hydrogen. *IEA/H2/10/TR2-98*. 1-35.
- Berg, B. L., J. Li, J. Heider and V. Stewart. 1991. Nitrate inducible formate dehydrogenase 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., and V. Stewart. 1990. Structural genes for nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. *Genetics*. **125**: 691-702.
- Berggren, G.; Adamska, A.; Lambertz, C.; Simmons, T. R.; Esselborn, J.; Atta, A.; Gambarelli, S.; Mouesca, J.-M.; Reijerse, E.; Lubitz, W.; Happe, T.; Artero, V.; Fontecave, M. 2013. Biomimetic assembly and activation of [FeFe]-hydrogenases. *Nature*. **499**: 66–69.
- Beyer, L., C. **24**: Doberenz, D. Flake, D. Hunger, B. Suppmann and R. G. Sawers. 2013. Coordination of FocA and pyruvate formate-lyase synthesis in *Escherichia coli*. Demonstrates preferential translocation of formate over other mixed-acid fermentation products. *Journal of Bacteriology*. **195**: 1428-1435.
- Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acid Research*. 1513-1523.
- Blokesh, M., S. P. J. Albracht, B. F. Matzanke, N. M. Drapal, A. Jacobi and A. Böck. 2004. The complex between hydrogenase-maturation proteins HypC and HypD is an intermediate in the supply of cyanide to the active site iron of [NiFe]-hydrogenase. *Journal of Molecular Biology*. **344**: 155-167.
- Blokesch, M., A. Magalon and A. Böck. 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.
- Blokesch, M., S. P. Albracht, B. F. Matzanke, N. M. Drapal, A. Jacobi and A. Böck. 2004. The complex between hydrogenase-maturation proteins HypC and HypD is an intermediate in the supply of cyanide to the active iron of [NiFe]- hydrogenases. *Journal of Molecular Biology*. **344**: 155-167.
- Blokesch, M., and A. Böck, A. 2002. Maturation of [NiFe]-hydrogenase in *Escherichia coli*: thr HypC cycle. *Journal of Molecular Biology*. **324**: 287-296.
- Blokesch, M., and A. Böck, A. 2006. Properties of the [NiFe]-hydrogenase maturation protein HypD. *FEBS Letters*. **580**: 4065–4068.

## References

- Bock, A., and G. Sawers. 1996. Fermentation in *Escherichia coli* and *Salmonella*. *Cellular and Molecular Biology*. Pp. 262-282. 2<sup>nd</sup> eddition. Neidhardt, F. C. Washington, DC, USA. ASM Press.
- Böck, A. K. Forchhammer, J. Heider and C. Baron. 1991. Selenoprotein synthesis: an expression of the genetic code. *Trends of Biochemistry Sciences*. **16**: 463-467.
- Bohom, R., M. Sauter and A. Bock. 1990. Nucleotide sequence and expression of an operon in *Escherichia coli* coding for formate hydrogelyase components. *Molecular Microbiology*. **4**: 231-243.
- Boyington, J. C., V. Gladyshev, S. V. Khangulov, T. C. Stadtman and P. D. Sun. 1997. Crystal structure of formate dehydrogenase H: catalysis involving Mo, molybdopterin and an Fe<sub>4</sub>S<sub>4</sub> cluster. *Science*. **275**: 1305-1308.
- Boyington, J. C., V. N. Gladyshev, S. V. Khangulov, T. C. Stadtman and P. D. Sun. 1997. Crystal structure of formate dehydrogenase H: Catalysis involving Mo, molybdopterin, selenocysteine, and an Fe<sub>4</sub>S<sub>4</sub> cluster. *Science*. **275**: 1305–1308.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Analysis Biochemistry*. **72**: 248-254.
- Brinkmann, U., R. E. Mattes and P. Buckel. 1989. High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the *dnaY* gene product. *Gene*. **85**: 109-114.
- British Colombia (BC) bioenergy strategy. 2010. Ministry of Energy, Mines and Petroleum resources. Growing Our Natural Energy Advantage.
- Brøndsted, L., and T. Atlung. 1994. Anaerobic Regulation of hydrogenase-1 (*hya*) operon of *Escherichia coli*. *Journal of Bacteriology*. **176**: 5423-5428.
- Burstel, I., E. Siebert, G. Winter, P. Hummel, I. Zebger, B. Friedrich and O. Lenz. 2012. A universal scaffold for synthesis of the Fe(CN)<sub>2</sub>(CO) moiety of [NiFe] hydrogenase. *The Journal of Biological Chemistry*. **287**: 38845–38853.
- Brush, E. J., K. A. Lipsett and J. W. Kozarich. 1988. Inactivation of *Escherichia coli* pyruvate formate-lyase by hypophosphite: evidence for a rate-limiting phosphorus-hydrogen bond cleavage. *Biochemistry*. **27**: 2217-2222.
- Casalot, L., and M. Rousset. 2001. Maturation of the [NiFe] hydrogenases. *Trends in microbiology*. **9**: 228-237.
- Case, R. J., Y. Boucher, I. Dahllo, C. Holmstrom, W. F. Doolittle and S. Kjelleberg. 2007. Use of 16S rRNA and *rpoB* genes as molecular markers for microbial ecology studies. *Applied and Environmental Microbiology*. **73**: 278-288.

## References

- Carrieri, D., G. M. Ananyev, A. M. G. Costas, D. A. Bryant, G. C. 2008. Dismukes. Renewable hydrogen production by cyanobacteria: Nickel requirement for optimal hydrogenase activity. *International Journal of Hydrogen Energy*. **33**: 2014-2022.
- Chakraborty, A., V. Saralaya, P. Adhikari, S. Shenoy, S. Baliga and A. Hegde. 2015. Characterization of *Escherichia coli* phylogenetic groups associated with extraintestinal infection in south Indian population. *Annual Medical Health Science Research*. **5**: 241-246.
- Cherepanov, P. P., and W. Wackernagel. 1995. Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*. **158**: 9-14.
- Cherrier, M. V., C. Cavazza, C. Bochot, D. Lemaire and J. C. Fontecilla-Camps. 2008. Structural characterization of a putative endogenous metal chelator in the periplasmic nickel transport NikA. *Biochemistry*. **47**: 9937-9943.
- Chivers, P. T., E. L. Benanti, V. Heil-Chapdelaine, J. S. Iwig and J. L. Rowe. 2012. Identification of Ni-(L-His)<sub>2</sub> as a substrate for NikABCDE-dependent nickel uptake in *Escherichia coli*. *Metallomics*. **4**: 1043-1050.
- Christiansen, L., and S. Pedersen. 1981. Cloning, restriction endonuclease mapping and post-transcriptional regulation of *rpsA*, the structural gene for ribosomal protein S1. *Molecular Genetics and Genomics*. **181**:
- Christopher, J. S., O. Djaman, J. A. Imlay and P. J. Kiley. 2000. The cysteine desulfurase, IscS, has a major role in *in vivo* Fe-S cluster formation in *Escherichia coli*. *Proceedings of the National Academy of Sciences*. **97**: 9009-9014.
- Compan, I., and D. Touati. 1994. Anaerobic activation of *arcA* transcription in *Escherichia coli*: roles of Fnr and ArcA. *Molecular microbiology*. **11**: 955-964.
- Crowther, G. J., G. Kosály and M. E. Lidstrom. 2008. Formate as the main branch point for methylophilic metabolism in *Methylobacterium extorquens* AM1. *Journal of Bacteriology*. **190**: 5057-5062.
- Das, D., and T. N. Veziroğlu. 2001. Hydrogen production by biological processes. A survey of literature. *International Journal of Hydrogen Energy*. **26**: 13-28.
- Das, D., and T. N. Veziroglu. 2008. Advances in biological hydrogen production processes. *International Journal of Hydrogen Energy*. **33**: 6046-6057.
- Datsenko, K. A., and B. L. Wanner. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences USA*. **97**: 6640-6645.
- De Pina, K., V. Desjardin, M. A. Mandrand-Berthelot, G. Giordano and L. F. Wu. 1999. Isolation and characterization of the *nikR* gene encoding a nickel-responsive regulator in *Escherichia coli*. *Journal of Bacteriology*. **181**: 670-674.

## References

- Demirbas, A. 2009. Political, economic and environmental impacts of biofuels: A review. *Applied Energy*. **86**: S108-S117.
- Demirbas, M. F., and M. Balat. 2006. Recent advances on the production and utilization trends of bio-fuels: A global perspective. *Energy Conversion and Management*. **47**: 2371-2381.
- Dharmaraj, C. H., and S. AdishKumar. 2012. Economical hydrogen production by electrolysis using nano pulsed DC. *Journal homepage: www. IJEE. IEEFoundation. org*. **3**: 129-136.
- Dobosz-Bartoszek, M., M. H. Pinkerton, Z. Otwinowski, S. Chakravarthy, D. Söll, P. R. Copeland and M. Simonovic. 2016. *Nature Communication*. **7**: 1-11.
- Doucleff, M., J. G. Pelton, P. S. Lee, B. T. Nixon and D. E. Wemmer. 2007. Structural basis of DNA recognition by the alternative sigma-factor, sigma 54. *Journal of Molecular Biology*. **369**: 1070-1078.
- Drapal, N., and A. Btick. 1998. Interaction of the hydrogenase accessory protein HypC with hycE, the large subunit of *Escherichia coli* hydrogenase 3 during maturation. *Biochemistry*. **37**: 2941-2948.
- Dross, F., V. Geisler, R. Lenger, F. Theis, T. Krafft, F. Fahrenholz, A. Duchene, D. Tripier, K. Juvenal and A. Kroger. 1992. The quinone-reactive Ni/Fe hydrogenase of *Walinella succinogenes*. *European Journal of Biochemistry/FEBS*. **106**: 93-102.
- Drury, L. S., and R. S. Buxton. 1985. DNA sequence analysis of the dye gene of *Escherichia coli* reveals amino acid homology between the Dye and OmpR proteins. *Journal in Biology and Chemistry*. **260**: 4236- 4242.
- Dunker, A. M., S. Kumar & P. A. Mulawa. 2006. Production of hydrogen by thermal decomposition of methane in a fluidized-bed reactor—Effects of catalyst, temperature, and residence time. *International Journal of Hydrogen Energy*. **31**: 473-484.
- Earhat, C. F. 1996. Uptake and metabolism of iron and molybdenum. *Escherichia coli* and *Salmonella* Cellular and molecular Biology, 2<sup>nd</sup> edition. ASM Press. In F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Schaechter and H. E. Umbarger (editors). Washington, D. C.
- Efremov, R. G., and L. A. Sazanov. 2012. The coupling mechanism of respiratory complex-1 a structural and evolutionary prespective. *Biochimica et Biophysica Acta*. **1817**: 1785-1795.
- Egan, S. M., and R. F. Schleif. 1993. A regulatory cascade in the induction of rhaBAD. *Journal of Molecular Biology*. **234**: 87-98.
- Eitingner, T. and M. A. Mandrand-Berthelot. 2000. Nickel transport systems in microorganisms. *Archives of Microbiology*. **173**: 1-9.
- Enoch, H. G., and L. R. Lester. 1975. The purification and properties of formate dehydrogenase and nitrate reductase from *Escherichia coli*. *Journal of Biological Chemistry*. **250**: 6693-6705.

## References

- Escobar-Páramo, P., O. Clermont, A. Blanc-Potard, H. Bui, C. Le Bouguéneq and E. Denamur. 2004. A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. *Molecular Biology and Evolution*. **21**: 1085-1094.
- Fan, Z., L. Yuan and R. Chatterjee. 2009. Increased hydrogen production by genetic engineering of *Escherichia coli*. *PLOS One*. **4**: e4432.
- Falke, D., K. Schulz, C. Doberenz, L. Beyer, H. Lilie, B. Thiemer and R. G. Sawers. 2010. Unexpected oligomeric structure of the FocA formate channel of *Escherichia coli*: a paradigm for the formate-nitrite transporter family of integral membrane proteins. *FEMS Microbiology Letter*. **303**: 69-75.
- Fileenko, N., S. Spiro, D. F. Browning, D. Squire, T. W. Overton, J. Cole and C. Constantinidou. 2007. The NsrR regulon of *Escherichia coli* K-12 includes genes encoding the hybrid cluster protein and the periplasmic, respiratory nitrite reductase. *Journal of Bacteriology*. **189**: 4410–4417.
- Flint, D. H., J. F. Tuminello and T. J. Miller. 1996. Studies on the synthesis of the Fe-S cluster of dihydroxy-acid dehydrates in *Escherichia coli* crude extract. *The Journal of Biological Chemistry*. **271**: 16053-16067.
- Fontecilla-Camps, J. C., A. Volbeda, C. Cavazza and Y. Nicolet. 2007. Structure/function relationships of [NiFe] and [FeFe]-hydrogenases. *Chemical Reviews*. **107**: 4273-4303.
- Ford, D. C., P. M. Ireland, H. L. Bullifent, R. J. Saint, E. V. McAlister, M. Sarkar-Tyson and P. C. F. Oyston. 2014. Construction of an inducible system for the analysis of essential genes in *Yersinia pestis*. *Journal of Microbiological Methods*. **100**: 1-7.
- Forzi, L., and R. G. Sawers. 2007. Maturation of [NiFe]- hydrogenases in *Escherichia coli*. *Biometals*. **20**: 565-578.
- Fritsch, J., P. Scheerer, S. Frielingsdorf, S. Kroschinsky, B. Friedrich, O. Lenz and C. M. T. Spahn. 2011. The crystal structure of an oxygen-tolerant hydrogenase uncovers a novel iron-sulphur centre. 2011. *Nature*. **479**: 249-253.
- Fukushima, M., K. Kakinuma and R. Kawaguchi. Phylogenetic analysis of *Salmonella*, *Shigella* and *Escherichia coli* strains on the basis of the *gyrB* gene sequence. 2002. *Journal of Clinical Microbiology*. **40**: 2779-2785.
- Gennis, R. B., and V. Stewart. 1996. Respiration in *Escherichia coli* and *Salmonella typhimurium*: *Cellular and molecular Biology*, pp. 217-261. Edited by F. C. Neidhardt. Washington DC. American Society of Microbiology.
- George, G. N., M. Gnida, D. A. Bazylinski, R. C. Prince and I. J. Pickering. 2008. X-ray absorption spectroscopy as a probe of microbial sulfur biochemistry: the nature of bacterial sulfur globules revisited. *Journal of Bacteriology*. **190**: 6376-6383.
- Gibson, D. G., L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison and H. O. Smith. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Natural Methods*. **6**: 343-345.

## References

- Gladyshev, V. N., S. V. Khangulov, M. J. Axley, M. J. and T. C. Stadtman. 1994. Coordination of selenium to molybdenum in formate dehydrogenase H from *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. **91**: 7708–7711.
- Gladyshev, V. N., J. C. Boyington, S. V. Khangulov, D. A. Grahame, T. C. Stadtman and P. D. Sun. 1996. Characterization of crystalline formate dehydrogenase H from *Escherichia coli*-stabilization, EPR spectroscopy, and preliminary crystallographic analysis. *Journal of Biological Chemistry*. **271**: 8095-8100.
- Glick, B. R., W. G. Martin and S. M. Martin. 1980. Purification and properties of the periplasmic hydrogenase from *Desulfovibrio desulfuricans*. *Journal of Microbiology*. **10**: 1214-1223.
- Gray, C. T., and H. Gest. 1965. Biological formation of molecular hydrogen. *Science*. **148**: 186-192.
- Guest, J. R. 1992. Oxygene-regulated gene expression in *Escherichia coli*. *Journal of General Microbiology*. **138**: 2253-2263.
- Guzman, L. M., D. Belin, M. J. Carson and J. Beckwith. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology*. **177**: 4121-4130.
- Haddock, E. A., and C. W. Jones. 1977. Bacterial respiration. *Bacteriological Reviews*. **41**: 47-99.
- Haldimann, A., L. L. Daniels and B. L. Wanner. 1998. Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the *Escherichia coli* phosphate regulon. *Journal of Bacteriology*. **180**: 1277-1286.
- Hallenbeck, P. C., and J. R. Benemann. 2002. Biological hydrogen production: fundamentals and limiting processes. *International Journal of Hydrogen Energy*. **27**: 1185-1193.
- Hallenbeck, P. C. 2009. Fermentative hydrogen production: Principles, progress, and prognosis. *International Journal of Hydrogen Energy*. **34**: 7379-7389.
- Hallenbeck, P. C., and D. Ghosh. 2009. Advances in fermentative biohydrogen production: the way forward? *Trends in Biotechnology*. **27**: 287-297.
- Hallenbeck, P. C., and D. Ghosh. 2012. Improvements in fermentative biological hydrogen production through metabolic engineering. *Journal of Environmental Management*. **95**: S360-S364.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*. **166**: 557-580.
- Hartwig, S., C. Pinske and R. G. Sawers. 2015. Chromogenic assessment of the three molybdate-selenoprotein formate dehydrogenases in *Escherichia coli*. *Biochemistry and Biophysics Reports*. **1**: 62-67.

## References

- Hatchikian, E. C.; N. Forget, V. M. Fernandez, R. Williams and R. Cammack. 1992. Further characterization of the [Fe]-hydrogenase from *Desulfovibrio desulfuricans* ATCC7757. *European Journal of Biochemistry / FEBS*. **209**: 357-365.
- Hatchikian, E. C., V. Margo, N. Forget, Y. Nicolet and J. C. Fontecilla-Camps. 1999. Carboxy-terminal processing of the large subunit of [Fe] hydrogenase from *Desulfovibrio desulfuricans* ATCC 7757. *Journal of Bacteriology*. **9**: 2947-2952.
- Heinekey, D. M. 2009. Hydrogenase enzymes: Recent structural studies and active site models. *Journal of Organometallic Chemistry*. **694**: 2671-2680.
- Hennig, R. G. 2017. Science is linking fossil fuels and neurological damage. The John Mark Fund. Boston, MA 02109.
- Hiromoto, T.; E. Warkentin, J. Moll, U. Ermler and S. Shima. 2009. Iron-Chromophore Circular Dichroism of [Fe]-Hydrogenase: The Conformational Change Required for H<sub>2</sub> Activation. *Angewandte Chemie International Edition*. **49**: 9917–9921.
- Hjersing, C. 2011. Hydrogen production in *Escherichia coli* – Genetic engineering of the formate hydrogenlyase complex. *Institutionen för fysik, kemi och biology*. 1-58.
- Holtz, C., Kaspari, H., and Klemme, J.-H. 1991. Production and properties of xylanases from thermophilic actinomycetes. *Antonie van Leeuwenhoek*. **59**:1–7.
- Hong, E., M. Doucleff and D. E. Wemmer. 2009. Structure of the RNA polymerase core-binding domain of sigma (54) reveals a likely conformational fracture point. *Journal of Molecular Biology*. **390**: 70-82.
- Hube, M., M. Blokesch and A. Böck. 2002. Network of hydrogenase maturation in *Escherichia coli*: role of accessory proteins HypA and HybF. *Journal of Bacteriology*. **184**: 3879-3885.
- Hunger, D., C. Doberenz and R. G. Sawers. 2014. Identification of key residues in the formate channel FocA that control import and export of formate. *Biology and Chemistry*. **395**: 813-825.
- Hunger, D., M. Röcker, D. Falke, H. Lilie and R. G. Sawers. 2017. The C-terminal six amino acids of the FNT channel FocA are required for formate translocation but not homopentamer integrity. *Frontiers in Microbiology*. **8**: 1-11.
- Hunt, T. P., and B. Magasanik. 1985. Transcriptional of *glnA* by purified *Escherichia coli* components: Core RNA polymerase and the production of *glnF*, *glnG* and *glnL*. *Proceedings of the National Academy of Sciences*. **82**: 8453-8457.
- Iguchi, A., N. R. Thomson, Y. Ogura, D. Saunders, T. Ooka, I R. Henderson, D. Harris, M. Asadulghani, K. Kurokawa, P. Dean, B. Kenny, M. A. Quail, S. Thurston, G. Dougan, J. Parkhil and G. Frankel. 2009. Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. *Journal of Bacteriology*. **191**: 347-354.
- Illinois Department of Public Health. Division of Environmental Health. 2009. <http://www.idph.state.il.us/envhealth/factsheets/polycyclicaromatichydrocarbons.htm>.

## References

International conferences, 2018-2019. <https://www.omicsonline.org/fossil-fuels.php>.

Iuchi, S., Z. Matsuda, T. Fujiwara and E. C. C. Lin. 1990. The *arcB* gene of *Escherichia coli* encodes a sensor-regulator protein for anaerobic repression of the *arc* modulon. *Molecular Microbiology*. **4**: 715-727.

Jacobi, A., R. Rossmann and A. Böck. 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.

Jacquier, H., C. Zaoui, M. Sandon-le Pors, D. Mazel and B. Berçot. 2009. Translation regulation of integrons gene cassette expression by the *attC* sites. *Molecular Microbiology*. **72**: 1475-1486.

Jo, B. H., and H. J. Cha. 2015. Activation of formate hydrogen-lyase via expression of uptake [NiFe]-hydrogenase in *Escherichia coli* BL21(DE3). *Microbial Cell Factories*. **14**: 151.

Jones, S. A., F. Z. Chowdhury, A. J. Fabich, A. Anderson, D. M. Schreiner, A. L. House, S. M. Autieri, M. P. Leatham, J. J. Lins, M. Jorgensen, P. S. Cohen and T. Conway. 2007. Respiration of *Escherichia coli* in the mouse intestine. *Infection and Immunity*. **75**: 4891-4899.

Kane, R. D. 2006. Corrosion in petroleum refining and petrochemical operations, corrosion: environments and industries. *ASM Handbook, ASM international C*. **13**: 967-1014.

Khangulov, S. V., V. N. Gladyshe, G. C. Dismukes and T.C. Stadtman. 1998. Selenium containing formate dehydrogenase H from *Escherichia coli*: A molybdopterin enzyme that catalyzes formate oxidation without oxygen transfer. *Biochemistry*. **37**: 3518–3528.

Kim, S., E. Seol, S. M. Raj, S. Park, Y. Oh and D. D. Y. Ryu. 2008. Various hydrogenases and formate-dependent hydrogen production in *Citrobacter amalonaticus* Y19. *International Journal of Hydrogen Energy*. **33**: 1509-1515.

Kim, J. Y., B. H. Jo and H. J. Cha. 2010. Production of biohydrogen by recombinant expression of [NiFe]-hydrogenase 1 in *Escherichia coli*. *Microbial Cell Factories*. **9**: 1-10 Kirtay, E. 2011. Recent advances in production of hydrogen from biomass. *Energy Conversion and Management*. **52**: 1778-1789.

Kirtay, E. 2011. Recent advances in production of hydrogen from biomass. *Energy Conversion and Management*. **52**: 1778-1789.

Komeda, H., M. Kobayashi and S. Shimizu. 1997. A novel transporter involved in cobalt transport. *Proceedings of the National Academy of Sciences of the United States of America*. **94**: 36-41.

Kotay, S. M., and D. Das. 2008. Biohydrogen as a renewable energy resource- Prospects and potentials. *International Journal of Hydrogen Energy*. **33**: 258-263.

Kovarik, B. 1998. Henry Ford, Charles F. Kettering and the fuel of the future. *Automotive History Review*. **32**: 7-27.

## References

- Krause, D. O., A. C. Little, S. E. Dowd and C. N. Bernstein. 2011. Complete genome sequence of adherent invasive *Escherichia coli* UM146 isolated from ileal Crohn's disease biopsy tissue. *Journal of Bacteriology*. **163**: 583.
- Kuchenreuther, J. M., W. K. Myers, D. L. Suess, T. A. Stich, V. Pelmeshnikov, S. A. Shiigi, S. P. Cramer, J. R. Swartz, R. D. Britt and S. J. George. 2014. The HydG enzyme generates an Fe(CO)<sub>2</sub>(CN) synthon in assembly of the FeFe hydrogenase H-cluster. *Science*. **343**: 424–427.
- Kuniyoshi, T. M., A. Balan, A. C. G. Schenberg, D. Severino and P. C. Hallenbeck. 2015. Heterologous expression of proteorhodopsin enhances H<sub>2</sub> production in *Escherichia coli* when endogenous Hyd-4 is overexpressed. *Journal of Biotechnology*. **206**: 52-57.
- Lacasse, M. J., and D. B. Zamble. 2016. [NiFe]-hydrogenase maturation. *Biochemistry*. **55**: 1689-1701.
- Lederberg, E. M., and S. N. Cohen. 1974. Transformation of *Salmonella typhimurium* by plasmid deoxyribonucleic acid. *Journal of Bacteriology*. **119**: 1072-1074.
- Lee, J. H., J. C. Wendt and K. T. Shanmugam. 1990. Identification of a new gene, *molR*, essential for utilization of molybdate by *Escherichia coli*. *Journal of Bacteriology*. **172**: 2079-2087.
- Lee, H., W. F. J. Vermaas and B. E. Rittmann. 2010. Biological hydrogen production: prospects and challenges. *Trends in Biotechnology*. **28**: 262-271.
- Lee, R. A., and J. Lavoie. 2013. From first- to third-generation biofuels: Challenges of producing a commodity from a biomass of increasing complexity. *Animal Frontiers*. **3**: 6- 11.
- Leinfelder, W., K. Forchhammer, F. Zinoni, G. Sawers, M. A. Mandard-Berthelot and A. Böck. 1988. *Escherichia coli* genes whose products are involved in selenium metabolism. *Journal of Bacteriology*. **170**: 540-546.
- Leimkühler, S., M. M. Wuebbens and K. V. Rajagopalan. 2011. The history of the discovery of the molybdenum cofactor and novel aspects of its biosynthesis in bacteria. *Coordination Chemistry Reviews*. **255**: 1129-1144.
- Leonhartsberger, S., I. Korsá and A. Bock. 2002. The molecular biology of formate metabolism in *Enterobacteria*. *Journal of Molecular Microbiology and Biotechnology*. **4**: 269-276.
- Levanon S. S., K. Y. San and G. N. Bennett. 2005. Effect of oxygen on the *Escherichia coli* ArcA and FNR regulation systems and metabolic responses. *Biotechnological Bioengineering*. **5**: 556-564.
- Levin, D. B., L. Pitt and M. Love. 2004. Biohydrogen production: prospects and limitations to practical application. *International Journal of Hydrogen Energy*. **29**: 173-185.
- Liang, Y., N. Sarkanyet, and Y. Cui. 2009. Biomass and lipid productivities of *Chlorella vulgaris* under autotrophic, heterotrophic, and mixotrophic growth conditions. *Biotechnology Letters*. **31**:1043–1049.

## References

- Lin, E. C. C., and S. Luchi. 1991. Regulation of gene expression in fermentative and respiratory system in *Escherichia coli* and related bacteria. *Annual Review of Genetics*. **25**: 361-378.
- Lindenstrauß, U., P. Skorupa, J. S. McDowall, F. Sargent and C. Pinske. 2017. The dual-function chaperone HycH improves assembly of the formate hydrogenlyase complex. *Biochemical Journal*. **474**: 2937-2950.
- Liu, Z., and P. Hu. 2002. A density functional theory study on the active center of Fe-only hydrogenase: characterization and electronic structure of the redox state. *Journal of the American Chemical Society*. **124**: 5175-5182.
- Liu X., and P. De Wulf. 2004. Probing the ArcA-P modulon of *Escherichia coli* by whole genome transcriptional analysis and sequence recognition profiling. *Journal of Biology and Chemistry*. **279**:12588 –12597.
- Lu, C., C. Liu, Y. Huang, C. Liao, L. Teng, J. D. Turnidge and P. Hsueh. 2011. Antimicrobial susceptibilities of commonly encountered bacterial isolates to fosfomycin determined by agar dilution and disk diffusion methods. *Antimicrobial Agents Chemother*. **55**: 4295-4301.
- Lü, W., J. Du, T. Wacker, E. Gerbig-Smentek, S. L. Andrade and O. Einsle. 2011. pH-dependent gating in a FocA formate channel. *Science*. **332**: 352-354.
- Lü, Wei., J. Du, N. J. Schwarzer, E. Gerbig-Smentek, O. Einsle and S. L. A. Andrade. 2012. The formate channel FocA exports the products of mixed-acid fermentation. *Proceedings of the National Academy of Sciences*. **109**: 13254-13259.
- Lü, W., J. Du, N. J. Schwarzer, T. Wacker, S. L. Andrade and O. Einsle. 2013. The formate/nitrite transporter family of anion channels. *Biology and Chemistry*. **394**: 715–727.
- Luchi, S., and E. C. Lin. 1993. Adaptation of *Escherichia coli* to redox by gene expression. *Molecular Microbiology*. **9**: 9-15.
- Lüke, I., G. Butland, K. Moore, G. Buchanan, V. Lyall, S. A. Fairhurst, J. F. Greenblatt, A. Emili, T. Palmer and F. Sargent. Biosynthesis of the respiratory formate dehydrogenases from *Escherichia coli*: characterization of the FdhE protein. *Archives of Microbiology*. **190**: 685-696.
- Lukey, M. J., A. Parkin, M. M. Roessler, B. J. Murphy, J. Harmer, T. Palmer, F. Sargent and F. A. Armstrong. 2010. How *Escherichia coli* is equipped to oxidize hydrogen under different redox conditions. *The Journal of Biological Chemistry*. **285**: 3928–3938.
- Lutz, S., A. Jacobi, V. Schlensog, R. Bohm, G. Sawers and A. Blick. 1991. Molecular characterisation of an operon (*hyp*) necessary for the activity of the three hydrogenases isoenzyme in *Escherichia coli*. *Molecular Microbiology*. **5**: 123-135.
- Lv, Hiaoying, H. Liu, M. Ke and H. Gong. 2013. Exploring the pH-dependent substrate transport mechanism of FocA using molecular dynamics simulation. *Biophysical Journal*. **105**: 2714-2723.

## References

- Magalon, A., and A. Bock. 2000. Dissection of the maturation reactions of the [Ni-Fe] hydrogenase 3 from *Escherichia coli* taking after nickel incorporation. *FEBS Letters*. **473**: 254-258.
- Maheux, A. F., D. K. Boudreau, M. G. Bergeron and M. J. Rodriguez. 2014. Characterization of *Escherichia fergusonii* and *Escherichia albertii* isolated from water. *Journal of Applied Microbiology*. **117**: 597-609.
- Maier, T., A. Jacobi, M. Sauter and A. Btick. 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., U. Binder and A. Böck. 1996. Analysis of the *hydA* locus of *Escherichia coli*: two genes (*hydN* and *hypF*) involved in formate and hydrogen metabolism. *Archive of Microbiology*. **165**: 333-341.
- Maier R. j., and E. W. Triplett. 1996. Toward more productive efficient and competitive nitrogen fixing symbiotic bacteria. *Crit Revs Plant Science*. **15**: 191-234.
- Malpica, R., B. Franco, C. Rodriguez, O. Kwon and D. Georgellis. 2004. Identification of a quinone-sensitive redox switch in the ArcBsensor kinase. *Proceedings of the National Academy of Sciences*. **101**: 13318-13323.
- Manish, S., and R. Banerjee. 2008. Comparison of biohydrogen production processes. *International Journal of Hydrogen Energy*. **33**: 279 – 286.
- Mantsakanyan, N., B. Karine and T. Armen. 2004. Hydrogenase 3 but not hydrogenase 4 is major in hydrogen gas production by *Escherichia coli* formate hydrogenlyase at acidic pH and in the presence of external formate. *Cell Biochemistry and Biophysics*. **41**: 357-366.
- Martin, S., and W. Griswold. 2009. Human health effects of heavy metals. *Environmental Science and Technology Briefs for Citizens*. **15**: 1-6.
- Masuda, N., and G. M. Church. 2003. Regulatory network of acid resistance genes in *Escherichia coli*. *Molecular Microbiology*, **48**: 699–712.
- Matsumoto, T. 2001. Urinary tract infections in the elderly. *Current Urology Reports*. **2**: 330-333.
- Maupin, J. A., and K. T. Shanmugam. 1990. Genetic regulation of formate hydrogenlyase of *Escherichia coli*: role of the *fhla* gene product as a transcriptional activator for a new regulatory gene, *fh1B*. *Journal of Bacteriology*. **172**: 4798 4806.
- Maupin-Furlow, J. A., J. K. Rosental, J. H. Lee, U. Deppenmeier, R. P. Gunsalus, K. T. Shanmugam. 1995. Genetic analysis of modABCD (molybdate transport) operon of *Escherichia coli*. *Journal of Bacteriology*. **177**: 4851-4856.
- Mazloomi, K., N. B. Sulaiman and H. Moayedi. 2012. Electrical efficiency of electrolytic hydrogen production. *International Journal of Electrochemical Science*. **7**: 3314-3326.

## References

- McCusker, M. P., K. Hokamp, J. F. Buckley, P. G. Wall, M. Martins and S. Fanning. 2014. Complete Genome Sequence of *Salmonella enterica* Serovar Agona Pulsed-Field Type SAGOXB.0066, Cause of a 2008 Pan-European Outbreak. *Genome Announcement*. **2**: e01219-13.
- McDowall, J. S., B. J. Murphy, M. Haumann, T. Palmer, F. A. Armstrong, F. A. and F. Sargent. 2014. Bacterial formate hydrogenlyase complex. *Proceedings of the National Academy Sciences of the United States of America*. **111**: E3948–E3956.
- McDowall, J. S., M. C. Hjersing, T. Palmer, T and F. Sargent. 2015. Dissection and engineering of the *Escherichia coli* formate hydrogenlyase complex. *FEBS Letters*. **589**: 3141–3147.
- McHugh, J. P., F. Rodriguz-Quinones, H. Abdul-Tehrani, D. A. Svistunenko, R. K. Poole, C. E. Cooper and S. C. Andrews. 2003. *Journal of Biological Chemistry*. **278**: 29478-29486.
- McNicholas, P. M. S. A. Rech and R. P. Gunsalus. 1997. Characterization of the ModE DNA-binding sites in the control regions of modABCD and moaABCDE of *Escherichia coli*. *Molecular Microbiology*. **23**: 515-524.
- Meeda, T., V. Sanchez-Torres and T. K. Wood. 2007. *Escherichia coli* hydrogenase 3 is a reversible enzyme processing hydrogen uptake and synthesis activities. *Applied Microbiology Biotechnology*. **76**: 1035-1042.
- Mendel, R. R., and S. Leimkuhler. 2015. The biosynthesis of the molybdenum cofactors. *Journal of Biological Inorganic Chemistry*. **20**: 337–347.
- Menon, N. K., J. Robbins, H. D. Peck, Jr., C. Y. Chatelus, E. S. Choi and A. E. Przybyla. 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, A. L., and R. L. Robson. 1994. In vivo and in vitro nickel-dependent processing of the [NiFe] hydrogenase in *Azotobacter vinelandii*. *Journal of Bacteriology*. **176**: 291-295.
- Miyake, J., M. Miyake and Y. Asada. 1999. Biotechnological hydrogen production: research for efficient light energy conversion. *Journal of Biotechnology*. **70**: 89-101.
- Mnatsakanyan, N., A. Vassilian, L. Navasardyan, K. Bagramyan and A. Trchounian. 2002. Regulation of *Escherichia coli* formate hydrogenlyase activity by formate at alkaline pH. *Current Microbiology*. **45**: 281-286.
- Mnatsakanyan, N., K. Bagramyan and A. Trchounian. 2004. Hydrogenase 3 but not hydrogenase 4 is major in hydrogen gas production by *Escherichia coli* formate hydrogenlyase at acidic pH in the presence of external formate. *Cell Biochemistry Biophysics*. **41**: 357-366.
- Mollet, C., M. Drancourt and D. Raoult. 1997. rpoB sequence analysis as a novel basis for bacterial identification. *Molecular Biology*. **26**: 1005-1011.
- Mori, H., K. Isono, T. Horiuchi and T. Miki. 2000. Functional genomics of *Escherichia coli* in Japan. *Research in Microbiology*. **151**: 121-128.

## References

- Mukherjee, A., and W. D. Donachie. 1990. Different translation of cell division proteins. *Journal of Bacteriology*. **172**: 6106-6111.
- Mulder, D. W., E. S. Boyd, R. Sarma, R. K. Lange, J. A. Endrizzi, J. B. Broderick and J. W. Peters. 2010. Stepwise [FeFe]-hydrogenase H-cluster assembly revealed in the structure of HydA (DeltaEFG). *Nature*. **465**: 248-251.
- Mulder, D. W., E. M. Shepard, J. E. Meuser, N. Joshi, P. W. King, M. C. Posewitz, J. B. Broderick and J. W. Peters. 2011. Insight into [FeFe]-hydrogenase structure, mechanism, and maturation. *Structure*. **19**: 1038-1052.
- Müller, V. 2001. Bacterial fermentation. *Encyclopaedia of Life Sciences*. Nature publishing group. 1-7.
- Mulrooney, S. B., and R. P. Hausinger. 2003. Nickel uptake and utilization by microorganisms. *FEMS Microbiology Reviews*. **27**: 239-261.
- Navarro, C., W. Long-Fei and M. Marie-Andree. 1993. The *nik* operon of *Escherichia coli* encodes a periplasmic binding-protein-dependent transport system for nickel. *Molecular Microbiology*. **9**: 1181-1191.
- Nicolet, Y., C. Cavazza and J. C. Fontecilla-Camps. 2002. Fe-only hydrogenases: structure, function and evolution. *Journal of Inorganic Biochemistry*. **91**: 1-8.
- Nitschke, W., and M. J. Russell. 2009. Hydrothermal focusing of chemical and chemiosmotic energy, supported by delivery of catalytic Fe, Ni, Mo/W, Co, S and Se, forced life to emerge. *Journal of Molecular Evolution*. **69**: 481-496.
- Noguchi, K., D. P. Riggins, K. C. Eldahan, R. D. Kitko and J. L. Slonczewski. 2010. Hydrogenase0-3 contributes to anaerobic acid resistance of *Escherichia coli*. *Public Library of Science One*. **5**: e10132.
- Nordstrom, K., and S. J. Austin. 1989. Mechanisms that contribute to the stable segregation of plasmids. *Annual Review of Genetics*. **23**: 37-69.
- Novick, R. P. 1987. Plasmid incompatibility. *Microbial Revision*. **4**: 381-395.
- O'Connor, P. A. 2010. Energy transitions. *The Pardee Papers*. **12**: 1-50.
- Ogata, H., W. Lubitz and Y. Higuchi. 2009. [NiFe] hydrogenases: structural and spectroscopic studies of the reaction mechanism. *Dalton Transactions*. 7577-7587.
- Ogata, H., S. Hirota, A. Nakahara, H. Komori, N. Shibata, T. Kato, K. Kano and Y. Higuchi. 2005. Activation process of [NiFe] hydrogenase elucidated by high-resolution X-ray analysis: conversion of the ready to the unready state. *Structure*. **13**: 1635-1642.
- Oliver, R. I., and D. A. King. 2009. Quo Vadis Biofuels. *Energy and Environmental Science*. **2**: 343.

## References

- Ooka, T., Y. Ogura, K. Katsura, K. Seto, H. Kobayashi, K. Kawano, E. Tokuoka, M. Furukawa, S. Harada, S. Yoshino, J. Seto, T. Ikeda, K. Yamaguchi, K. Murasa, Y. Gotoh, N. Imuta, J. Nishi, T. Gomes, L. Beutin and T. Hayashi. 2015. Defining the genome features of *Escherichia alberti*, an emerging enteropathogen closely related to *Escherichia coli*. *Genome Biology and Evolution*. **7**: 3170-3179.
- Oshima, K., H. Toh, Y. Ogura, H. Sasamptp, H. Morita, S. Park, T. Ooka, S. Iyoda, T. Taylor, T. Hayashi, K. Itoh and M. Hattori. 2008. Complete genome sequence and comparative analysis of the wild-type commensal *Escherichia coli* strain SE11 isolated from a healthy adult. *DNA Research*. **15**: 375-38.
- Pecher, A., H. P. Blaschkowski, J. Knappe and A. Bock. 1982. Expression of pyruvate formate-lyase of *Escherichia coli* from the cloned structural gene. *Archives of Microbiology*. **132**: 365-371.
- Pecher, A., F. Zinoni, C. Jatisatienr, R. Wirth, H. Hennecke and A. Böck. 1983. On the redox control of synthesis of anaerobically induced enzymes in enterobacteriaceae. *Archives of Microbiology*. **136**: 131-136.
- Peck, H. D., Jr and Gest. 1957. Formic dehydrogenase and the hydrogenlyase enzyme complex in coli-aerogenes bacteria. *Journal of Bacteriology*. **73**: 706-721.
- Pekgöz, G., U. Gündüz, I. Eroğlu and G. Rákhely. 2010. Removal of the effect of ammonium on the regulation of nitrogenase enzyme in *Rhodobacter Capsulatus* DSM1710 for improved hydrogen production. *Proceedings World Hydrogen Energy Confrence*. **78**: 234-240.
- Peris-Bondle, F., E. Muraille and L. V. Melderren. 2013. Complete genome sequence of the *Escherichia coli* PMV-1 strain, a model extraintestinal pathogenic *E. coli* strain used for host-pathogen interaction studies. *Genome Announcement*. **1**: 1-2.
- Peters, J. W., G. J. Schut, E. S. Boyd, D. W. Mulder, E. M. Shepard, J. B. Broderick, P. W. King and M. W. W. Adams. 2015. [FeFe]- and [NiFe]-hydrogenase diversity, mechanism and maturation. *Biochemica and Biophysica Acta*. **1853**: 1350-1369.
- Phillips, C. M., E. R. Schreiter, C. M. Stutz and C. L. Drennan. 2010. Structural basis of low-affinity nickel binding to the nickel-responsive transcription factor NikR from *Escherichia coli*. *Biochemistry*. **49**: 7830-7838.
- Pinsent, J. 1953. The need for selenite and molybdate in the formation of formic dehydrogenase by members of the *Coli-aerogenes* group of bacteria. *Medical Research Council Unit for Bacterial Chemistry*. **57**: 10-15.
- Pinske, C., M. Jaroschinsky, F. Sargent and G. Sawers. 2012. Zymographic differentiation of [NiFe]-hydrogenase1, 2 and 3 of *Escherichia coli* K-12. *BioMedical Central of Micribiology*. **12**: 1-12.
- Pinske, C., J. S. Mcdowall, F. Sargent and R. G. Sawers. 2012. Analysis of hydrogenase 1 levels reveals an intimate link between carbon and hydrogen metabolism in *Escherichia coli* K-12. *Microbiology*. **158**: 856-868.

## References

- Pinske, C., and R. G. Sawers. 2012. Delivery of iron-sulfur clusters to the hydrogen-oxidizing [NiFe]-hydrogenases in *Escherichia coli* requires the A-Type carrier proteins *erpA* and *IscA*. *PLoS One*. **7**: e31755.
- Pinske, C., and R. G. Sawers. 2014. The importance of iron in the biosynthesis and assembly of [NiFe]-hydrogenases. *Biomolecular Concepts*. **5**: 55-70.
- Pinske, C., M. Jaroschinsky, S. Linek, C. L. Kelly, F. Sargent and R. G. Sawers. 2015. Physiology and Bioenergetics of [NiFe]- hydrogenase 2-catalyzed H<sub>2</sub>- consuming and H<sub>2</sub>-producing reactions in *Escherichia coli*. *Journal of Bacteriology*. **197**: 296-306.
- Plunkett, G., V. Burland, D. L. Daniels and F. R. Blattner. 1993. Analysis of the *Escherichia coli* genome. III. DNA sequence of the region from 87.2 to 89.2 minutes. *Nucleic Acid Research*. **21**: 3391-3398.
- Pollak, N., C. Dolle and M. Ziegler. 2007. The power to reduce: pyridine nucleotides-small molecules with a multitude of functions. *Biochemistry Journal*. **402**: 205-218.
- Pravettoni, R. 2012. World biofuels production trends. UNEP Report. <https://www.grida.no/resources/6187>
- Przybyla, A. E., Robbins, J., Menon, N., and Peck, H. D., J. 1992. Structure-function relationships among the nickel-containing hydrogenases. *FEMS Microbiology Reviews*. **8**: 109–35.
- Raaijmakers, H. C., and M. J. Romao. 2006. Formate-reduced *E. coli* formate dehydrogenase H: the reinterpretation of the crystal structure suggests a new reaction mechanism. *Journal of Biological Inorganic Chemistry*. **11**: 849-854.
- Rangarajan, E. S., A. Asinas, A. Proteau, C. Munger, J. Baardsnes, P. Iannuzzi, A. Matte and M. Cygler. 2008. Structure of [NiFe] hydrogenase maturation protein HypE from *Escherichia coli* and its interaction with HypF. *Journal of Bacteriology*, **190**: 1447–1458.
- Reed, W. L., and R. F. Schleif. 1999. Hemiplegic mutations in AraC protein. *Journal of Molecular Biology*. **294**: 417-425.
- Reissmann, S., E. Hochleitner, H. Wang, A. Paschos, F. Lottspeich, R. S. Glass and A. Böck. 2003. Taming of a poison: Biosynthesis of the NiFe-hydrogenase cyanide ligands. *Science*. **299**: 1067–1070.
- Richard, D. G., G. Sawers, F. Sargent, L. McWalter and D. H. Boxer. 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., D. H. Boxer, M. A. Mandrand-Berthelot and L.F. Wu. 1996. Requirement for nickel of the transmembrane translocation of NiFe-hydrogenase 2 in *Escherichia coli*. *FEBS Letters*. **392**: 81-86.

## References

- Rosentel, J. K., F. Healy, J. A. Maupin-Furlow, J. H. Lee and K. T. Shanmugam. 1995. Molybdate and regulation of *mod* (molybdate transport), *fdhF* and *hyc* (formate hydrogenlyase) operons in *Escherichia coli*. *Journal of Bacteriology*. **177**: 4857-4864.
- Rossmann, R., G. Sawers and A. Bock. 1991. Mechanism of regulation of the formate-hydrogenlyase pathway by oxygen, nitrate and pH: definition of the formate regulon. *Molecular and Microbiology*. **5**: 2807-2814.
- Rossmann, A., M. Sauter, F. Lottspeich and A. Böck. 1994. Maturation of the large subunit (HycE) of *Escherichia coli* hydrogenase 3 requires nickel incorporation followed by C-terminal processing at Arg537. *European Journal of Biochemistry*. **220**: 377-384.
- Rowe, J. L., G. L. Starnes and P. T. Chivers. 2005. Complex transcriptional control links NikABCDE-dependent nickel transport with hydrogenase expression in *Escherichia coli*. *Journal of Bacteriology*. **187**: 6317-6323.
- Ruiz, J. A., A. de Almeida, M. S. Godoy, M. P. Mezzina, G. N. Bidart, B. S. Mèndez, M. J. Pettinari and P. I. Nikel. 2012. *Escherichia coli* redox mutants as microbial cell factories for the synthsis of reduced biochemical. *Computational and Structural Biotechnology Journal*. **3**: 1-10.
- Saier, M. H., B. H. Eng, S. Fard, J. Garg, D. A. Haggerty, W. J. Hutchinson, D. L. Jack, E. C. Lai, H. J. Liu, D. P. Nusinew, A. M. Omar, S. S. Pao, I. T. Paulsen, J. A. Quran, M. Sliwinski, T. T. Tseng, S. Wachi and G. B. Young. 1999. Phylogenetic characterization of novel transport protein families revealed by genome analyses. *Biochemica et Biophysica Acta*. **1422**: 1-56.
- Salomone-Stagnia, M., F. Stellatob, C. M. Whaley, S. Vogtd, S. Morante, S. Shimad, T. B. Rauchfuss and W. Meyer-Klaucke. 2010. The iron-site structure of [Fe]-hydrogenase and model systems: An X-ray absorption near edge spectroscopy study. *Dalton Trans*. **39**: 3057-3064.
- Salmon KA., S. P. Hung, N. R. Steffen, R. Krupp, P. Baldi, G. W. Hatfield and R. P. Gunsalus. 2005. Global gene expression profiling in *Escherichia coli* K12: effects of oxygen availability and ArcA. *Journal of Biology and Chemistry*. **280**:15084– 15096.
- Sambrook J. 1989. *Molecular Cloning; A Laboratory Manual* (3<sup>rd</sup> Edition). *Cold Spring Harbor Laboratory Press. New York*.
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbour Laboratory Press, Woodbury, NY.
- Sambrook J., Fritsch F., and Maniatis T. 2001. *Molecular cloning; A Laboratory Manual*, *Cold Spring Harbor Laboratory Press. New York*.
- Santos, D. M. 2011. Genetic engineering. Recent development in applications. *Edn Santos, D. M. Apple Academic PRESS*.

## References

- Sanchez-Torres, V., M. Z. M. Yusoff, C. Nakano, T. Maeda, H. I. Ogawa and T. K. Wood. 2013. Influence of *Escherichia coli* hydrogenases on hydrogen fermentation from glycerol. *International Journal of Hydrogen Energy*. **38**: 3905-3912.
- Sargent, F. 2016. The model [NiFe]-hydrogenases of *Escherichia coli*. *Advances in Microbial physiology*. **68**: 443-507.
- Sauter, M., R. Bohm and A. Bock. 1992. Mutational analysis of the operon (*hyc*) determining hydrogenase 3 formation in *Escherichia coli*. *Molecular Microbiology*. **6**: 1523-1532.
- Savaliya, M. L., B. D. Dhorajiya and B. Z. Dholakiya. 2013. Recent advancement in production of liquid biofuels from renewable resources: a review. *Research on Chemical Intermediates*. 1-35.
- Sawers, G. 1994. The hydrogenases and formate dehydrogenases of *Escherichia coli*. *Antonie Van Leeuwenhoek*. **66**: 57-88.
- Sawers, G. 2005. Evidence for novel processing of the anaerobically inducible dicistronic *focA-pfl* mRNA transcript in *Escherichia coli*. *Molecular Microbiology*. **58**: 1441-1453.
- Sawers, G., and A. Böck. 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, R. G., S. P. Ballantine and D. H. Boxer. 1985. Differential expression of hydrogenase isoenzymes in *Escherichia coli* K-12: evidence for a third isoenzyme. *Journal of Bacteriology*. **164**: 1324-1331.
- Sawers, G., J. Heider, E. Zehelein and A. Böck. 1991. Expression and operon structure of *sel* genes of *Escherichia coli* and identification of a third selenium-containing formate dehydrogenase isoenzyme. *Journal of Bacteriology*. **173**: 4983-4993.
- Sawitzki, J. A., L. C. Thomason, M. Bubunencko, X. Li, N. Costantino and D. L. Court. 2011. Recombineering: using drug cassette to knock out genes *in vivo*. *Court Lab protocol*.
- Schleif, R. 2010. AraC protein, regulation of the L-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. *FEMES Microbiology revision*. **34**: 779-796.
- Schlenso, V., A. Birkmann and A. Bock. 1989. Mutations in trans which affect the anaerobic expression of a formate dehydrogenase (*fdhF*) structural gene. *Archives of Microbiology*. **152**: 83-89.
- Schnackenberg, J., H. Ikemoto, and S. Miyachi. 1996. Photosynthesis and hydrogen evolution under stress conditions in a CO<sub>2</sub>-tolerant marine green alga, *Chlorococcum littorale*. *Journal of Photochemistry and Photobiology b-biology*. **34**: 59-62.
- Schwartz, C. J., O. Djaman, J. A. Imlay and P. J. Kiley. 2000. The cysteine desulfurase, IscS, has a major role in *in vivo* Fe-S cluster formation in *Escherichia coli*. *Proceedings of the National Academy of Sciences*. **97**: 9009-9014.

## References

- Self, W. T., and K. T. Shanmugam. 2000. Isolation and characterization 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.
- Self, W. T., A. Hasona and K. T. Shanmugam. 2001. N-terminal truncations in the FhlA protein result in format and MoeA-independent expression of the *hyc* (formate hydrogenlyase) operon of *Escherichia coli*. 2001. *Microbiology*. **147**: 3093-3104.
- Self, W. T., A. Hasona and K. T. Shanmugam. 2004. Expression and regulation of a silent operon, *hyf*, coding for hydrogenase 4 isoenzyme in *Escherichia coli*. *Journal of Bacteriology*. **186**: 580-587.
- Shafaat, H. S., O. Rüdiger, H. Ogata and W. Lubitz. 2013. [NiFe] hydrogenases: a common active site for hydrogen metabolism under diverse conditions. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*. **1827**: 986-1002.
- Sharma, L. K., J. Lu and Y. Bai. 2009. Mitochondrial respiratory complex 1: structure, function and implication in human diseases. *Current Medical Chemistry*. **16**: 1266- 1277.
- Shepard, E. M., S. E. Mcglynn, A. L. Bueling, C. S. Grady-Smith, S. J. George, M. A. Winslow, S. P. Cramer, J. W. Peters and J. B. Broderick. 2010. *Proceedings of the National Academy of Sciences*. **107**: 10448-10453.
- Shima, S., O. Pilak, S. Vogt, M. Schick, M. S. Stagni, W. Meyer-Klaucke, E. Warkentin, R. K. Thauer and U. Ermler. 2008. The crystal structure of [Fe]-hydrogenase reveals the geometry of the active site. *Science*. **321**: 572-575.
- Shima, S., and U. Ermler. 2010. Structure and function of [Fe]-hydrogenase and its iron-guanlylpyridinol (FeGP) cofactor. *European Journal of Inorganic Chemistry*. S 963-972.
- Shomura, Y., K. S. Yoon, H. Nishihara and Y. Higuchi. 2011. Structural basis for a [4Fe-3S] cluster in the oxygen-tolerant membrane-bound [NiFe]-hydrogenase. *Nature*. **479**: 253-256.
- Skibinski, D. A. G., P. Golby, Y. Chang, F. Sargent, R. Hoffman, R. Harper, J. R. Guest, M. M. Attwood, B. C. Berks and S. C. Andrews. 2002. Regulation of the hydrogenase-4 operon of *Escherichia coli* by the  $\sigma^{54}$ -dependent transcriptional activators FhlA and HyfR. *Journal of Bacteriology*. **184**: 6642- 6653.
- Sims, G. E., and S. Kim. 2011. Whole-genome phylogeny of *Escherichia coli* / *Shigella* group by feature frequency profiles (FFPs). *Proceedings of the National Academy of Sciences*. **108**: 8329-8334.
- Soboh, B., C. Pinske, M. Kuhns, M. Waclawek, C. Ihling, K. Trchounian, A. Trchounian, A. Sinz and G. Sawers. 2011. The respiratory molybdo-selenoprotein formate dehydrogenases of *Escherichia coli* have hydrogen: benzyl viologen oxidoreductase activity. *Biomedical central Microbiology*. **11**: 173.

## References

- Soboh, B., S. T. Stripp, E. Muhr, C. Granich, M. Braussemann, M. Herzberg, J. Heberle and R. G. Sawers. 2012. [NiFe]-hydrogenase maturation: Isolation of a HypC-HypD complex carrying diatomic CO and CN<sup>-</sup> ligands. *FEBS Letters*. **586**: 3882–3887.
- Soboh, B., S. T. Stripp, C. Bielak, U. Lindenstraus, M. Braussemann, M. Javaid, M. Hallensleben, C. Granich, M. Herzberg, J. Heberle and R. G. Sawers. 2013. The [NiFe]-hydrogenase accessory chaperones HypC and HybG of *Escherichia coli* are iron- and carbon dioxide-binding proteins. *FEBS Letters*. **587**: 2512–2516.
- Stephenson, M., and L. H. Stickland. 1930. Hydrogenase: A bacterial enzyme activating molecular hydrogen. The properties of the enzyme. *The Biochemical Laboratory, Cambridge*.
- Stripp, S. T., B. Soboh, U. Lindenstraus, M. Braussemann, M. Herzberg, D. H. Nies, R. G. Sawers and J. Heberle. 2013. HypD is the scaffold protein for Fe-(CN)<sub>2</sub>CO cofactor assembly in [NiFe]-hydrogenase maturation. *Biochemistry*. **52**: 3289–3296.
- Stripp, S. T., U. Lindenstraus, R. G. Sawers, R. G and B. Soboh. 2015. Identification of an isothiocyanate on the HypEF complex suggests a route for efficient cyanyl-group channeling during [NiFe]-hydrogenase cofactor generation. *PLoS ONE*. **10**: e0133118.
- Suppmann, B., and G. Sawers. 1994. Isolation and characterization of hypophosphite-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.
- Taylor, B. L., and I B. Zhulin. 1999. PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiology and Molecular Biology*. **63**: 479–506.
- Thakker, C., I. Martinez, K. San and G. N. Bennett. 2012. Succinate production in *Escherichia coli*. *Biotechnology Journal*. **7**: 213-224.
- Theodorou, M. K., B. A. Williams, M. S. Dohanoa, A. B. McAllan and J. France. 1994. A simple gas production method using a pressure transducer to determine the fermentation kinetics of ruminant feeds. *Animal Feed Science and Technology*. **84**: 185-197.
- Thomas, C. M. 2014. Plasmid incompatibility. *Molecular Life Sciences*. 1-3.
- Thome, R., A. Gust, R. Toci, R. Mendel, F. Bittner, A. Magalon and A. Walburger. 2012. A sulfurtransferase is essential for activity of formate dehydrogenase in *Escherichia coli*. *The Journal of Biological Chemistry*. **287**: 4671- 4678.
- Thomason, L. C., N. Costantino and D. L. Court. 2016. Examining a DNA replication requirement for bacteriophage λ Red and Rac prophage Rec-ET- promoted recombination in *Escherichia coli*. *Microbiology*. **7**: e01443-16.
- Tominaga, T., S. Watanabe, R. Matsumi, H. Atomi, T. Imanaka, T and K. Miki. 2013. Crystal structures of the carbamoylated and cyanated forms of HypE for [NiFe] hydrogenase maturation. *Proceedings of the National Academy of Sciences of the United States of America*. **110**: 20485–20490.

## References

- Tomiyama, M., M. Shiotani, K. Sode, E. Tamiya and I. Karube. 1991. Nucleotide sequence analysis and expression control of *hydA* in *Escherichia coli*. In *Third International Conference on the Molecular Biology of Hydrogenases*. Troia, Portugal. 24-25.
- Tosatto, S. C. E., G. M. Giacometti, G. Valle and P. Costantini. 2006. Functional insights from the structural modeling of a small Fe-hydrogenase. *Biochemical and Biophysical Research Communications*. **339**: 277-283.
- Trchounian, K., and A. Trchounian. 2009. Hydrogenase 2 is most and hydrogenase 1 is less responsible for H<sub>2</sub> production by *Escherichia coli* under glycerol fermentation at neutral and slightly alkaline. *International Journal of Hydrogen Energy*. **34**: 8839-8845.
- Trchounian, K., and A. Trchounian. 2013. *Escherichia coli* hydrogenase 4 (*hyf*) and hydrogenase 2 (*hyb*) contribution in H<sub>2</sub> production during mixed carbon (glucose and glycerol) fermentation at pH 7.5 and pH 5.5. *International Journal of Hydrogen Energy*. **38**: 3921-3929.
- Trchounian, K., and A. Trchounian. 2014. Different role of *focA* and *focB* encoding formate channels for hydrogen production by *Escherichia coli* during glucose or glycerol fermentation. *International Journal of Hydrogen Energy*. **39**: 20987- 20991.
- Trchounian, K., S. Mirzoyan, A. Poladyan and A. Trchounian. 2017. Hydrogen production by *Escherichia coli* growing in different nutrient media with glycerol: Effects of formate, pH, production kinetics and hydrogenases involved. *International Journal of Hydrogen Energy*. **42**: 24026-24034.
- Uden, G., S. Becker, J. Bongaerts, J. Schirawski and S. Six. 1994. Oxygen regulated gene expression in facultatively anaerobic bacteria. *Antonie Van Leeuwenhoek*. **66**: 3-22.
- Uden, G., and J. Bongaerts. 1997. Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors. *Biochemistry Biophysics Acta*. **1320**: 217-234.
- Uden, G., S. A. Achebach, G. Holighaus, H. Q. Tran, B. Wackwitz and Y. Zeuner. 2002. Control of FNR function of *Escherichia coli* by O<sub>2</sub> and reducing conditions. *Journal of Molecular Microbiology and Biotechnology*. **4**: 263-268.
- United Nations Environment Programme (UNEP) report. 2012. World biofuels production trends. <http://www.grida.no/resources/6187>.
- Urbaniec, K., A. Friedl, D. Huisingh and P. Claassen. 2010. Hydrogen for a sustainable Global economy. *Journal of Cleaner Production*. **18**: S1-S3.
- Verhees, C.H., S. W. Kengen, J. E. Tuininga, G. J. Schut, M. W. Adams, W. M. De Vos, and J. van der Oost. 2003. The unique features of glycolytic pathways in Archaea. *Biochemistry Journal*. **375**: 231-246.
- Vignais, P. M., B. Billoud and J. Meyer. 2001. Classification and phylogeny of hydrogenase. *FEMS Microbiology Reviews*. **25**: 455-501.

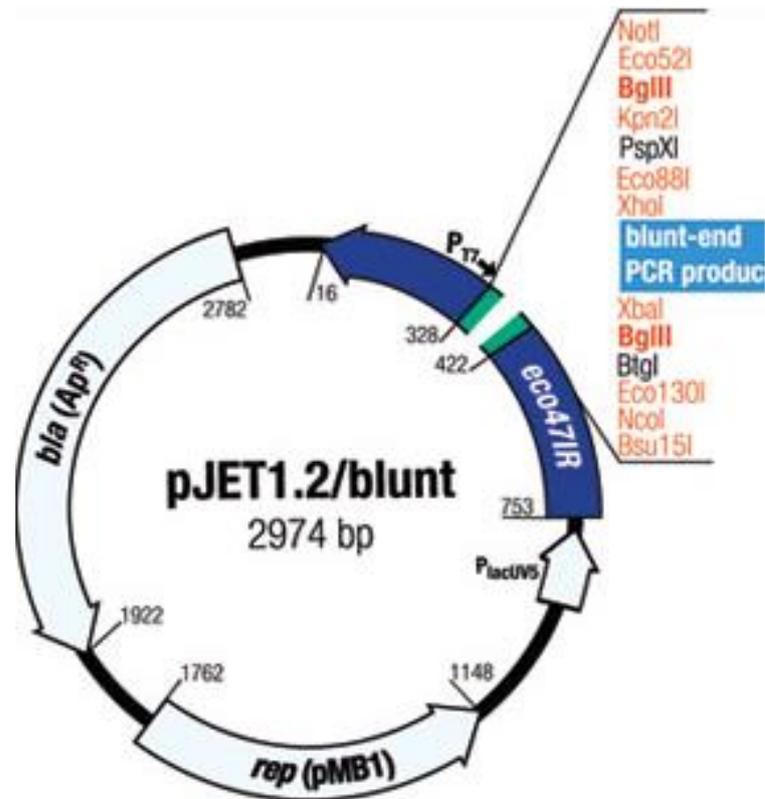
## References

- Vignais, P. M., and B. Billoud. 2007. Occurrence, classification, and biological function of hydrogenases: an overview. *Chemistry Review*. **107**: 4206-4272.
- Vignais, P. M. and Colbeau, A. 2004. Molecular biology of microbial hydrogenases. *Current Issues of Molecular Biology*. **6**: 159–188.
- Vivijis, B., L. U. Haberbeck, V. Baiye Mfortaw Mbong, K. Bernaerts, A. H. Geeraerd, A. Aertsen and C. W. Michiels. 2015. Formate hydrogen lyase mediates stationary-phase deacidification and increases survival during sugar fermentation in acetoin-producing enterobacteria. *Frontiers in Microbiology*. **6**: 150.
- Volbeda, A. and J. C. Fontecilla-Camps. 2003. The active site and catalytic mechanism of [NiFe] hydrogenases. *Dalton Transactions*. **21**: 4030–4038.
- Volbeda, A., P. Amara, C. Darnault, JM. Mouesca, A. Parkin, MM. Roessler, FA. Armstrong and JC. Fontecilla-Camps. 2012. X-ray crystallographic and computational studies of the O<sub>2</sub>-tolerant [NiFe]-hydrogenase 1 from *Escherichia coli*. *Proc Natl Acad Sci USA*. **109**: 5305-5315.
- Waight, A. B., J. Love, and D. N. Wang. 2010. Structure and mechanism of a pentameric formate channel. *Nature Structural & Molecular Biology*. **17**: 31-37.
- Wang, Y., Y. Huang, J. Wang, C. Cheng, W. Huang, P. Lu, Y.N. Xu, P. Wang, N. Yan, and Y. Shi. 2009. Structure of the formate transporter FocA reveals a pentameric aquaporin-like channel. *Nature*. **462**: 467-472.
- Waston, J. D., T. A. Baker, S. P. Bell, A. Gann, M. Levine and R. Losick 2003. Molecular biology of the gene. *Edn Kindle Edition 7<sup>th</sup>*. Cold Spring Harbor Laboratory PRESS. USA.
- Watanabe, S., T. Arai, R. Matsumi, H. Atomi, T. Imanaka and K. miki. 2009. Crystal structure of HypA, a nickel-binding metallochaperone for [NiFe] hydrogenase maturation. *Journal of Molecular Biology*. **394**: 448-459.
- Watanabe, S., R. Matsumi, H. Atomi, T. Imanaka, T and K. Miki. 2012. Crystal structures of the HypCD complex and the HypCDE ternary complex: Transient intermediate complexes during [NiFe] hydrogenase maturation. *Structure*. **20**: 2124–2137.
- Waugh, R., and D. H. Boxer. 1986. Pleiotropic hydrogenase mutants of *Escherichia coli* K12: growth in the presence of nickel can restore hydrogenase activity. *Biochimie*. **68**: 157-166.
- Wedel, A., and S. Kustu. 1995. The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. *Genes*. **9**: 2042-52.
- Wegerer, A., T. Sun and J. Altenbuchner. 2008. Optimization of an *E. coli* L-rhamnose-inducible expression vector: test of various genetic module combinations. *BMC Biotechnology*. **8**: 1-12.
- Wiechert, M., E. Beitz. 2017. Mechanism of formate-nitrite transporters by dielectric shift of substrate acidity. *The EMBO Journal*. **36**: 949-958.

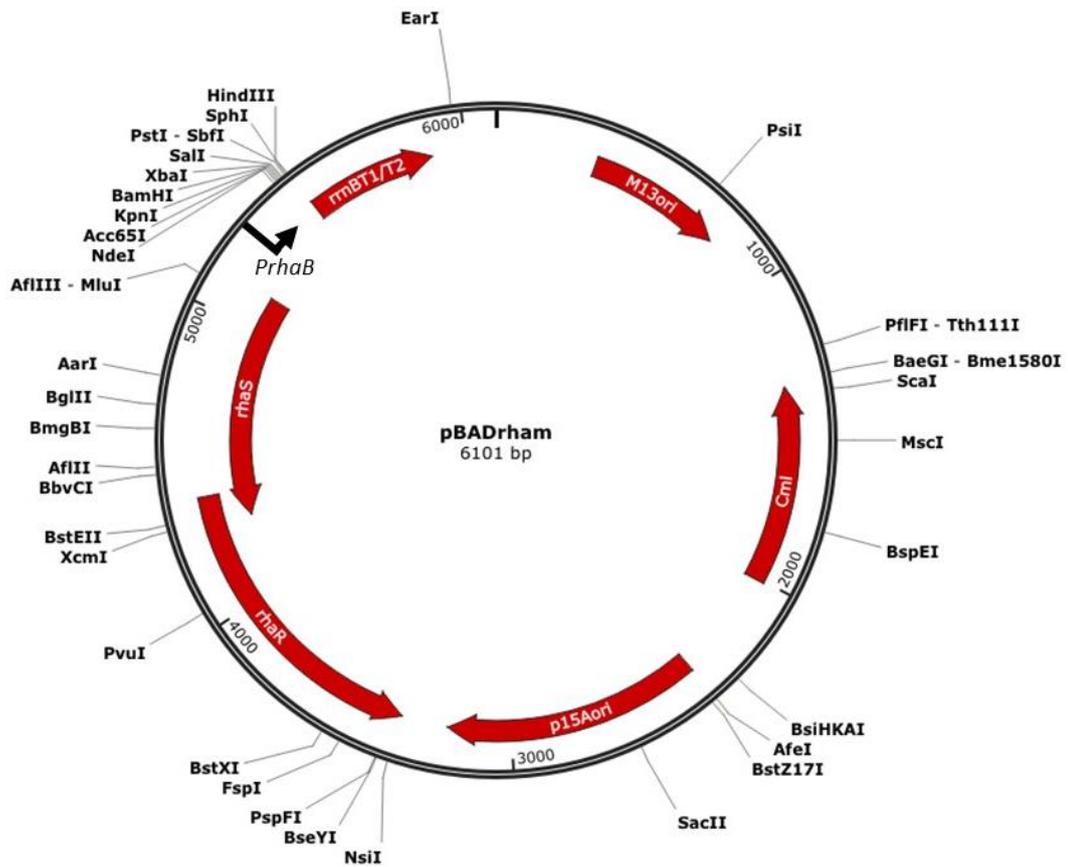
## References

- Wong, C., Y. Tintut and J. Gralla. 1994. The domain structure of Sigma54 as determined by analysis of a set of deletion mutants. *Journal of Molecular Biology*. **236**: 81-90.
- Woods, D. D. 1936. Hydrogenlyases: the synthesis of formic acid by bacteria. *Journal of Biochemistry*. **30**: 515-527.
- Wu, L. F., and M. A. Mandrand-Berthelot. 1986. Genetic and physiological characterization of new *Escherichia coli* mutants impaired in hydrogenase activity. *Biochimie*. **68**: 167–179.
- Wu, L. F., M. A. Mandrand-Berthelot, R. Waugh, C. J. Edmonds, S. E. Holt and D. H. Boxer. 1989. Nickel deficiency gives rise to the defective hydrogenase phenotype of *hydC* and *fnr* mutants in *Escherichia coli*. *Molecular Microbiology*. **3**: 1709–1718.
- Wu, L. F., C. Navarro, C and M. A. Mandrand-Berthelot. 1991. The *hydC* region contains a multi-cistronic operon (*nik*) involved in nickel transport in *Escherichia coli*. *Gene*. **107**: 37–42.
- Yu, D., M. E. Hilary, L. E-Chiang, A. J. Nancy, G. C. Neal and L. C. Donald. 2000. An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proceedings of the National Academy of Sciences*. **97**: 5978-5983.
- Yu, J., and P. Takahashi. 2007. Biophotolysis-based Hydrogen Production by Cyanobacteria and Green Microalgae. *Communicating Current Research and Educational Topics and Trends in Applied Microbiology*. **1**: 79-89.
- Yudkin, M. D. 1970. Catabolic repression of the *lac* operon, effect of mutations in the *lac* promoter. *Biochemistry Journal*. **118**: 741-746.
- Yoshida, A., T. Nishimura, H. Kawaguch, M. Inui and H. Yukawa. 2005. Enhanced hydrogen production from formic acid by formate hydrogen lyase-overexpressing *Escherichia coli* strains. *Applied and Environmental Microbiology*. **71**: 6762-6768.
- Zhou, S., T. B. Causey, A. Hasona, K. T. Shanmugam and L. O. Ingram. 2003. Production of optically pure D-Lactic acid in mineral salts medium by metabolically engineered *Escherichia coli* W3110. *Applied and Environmental Microbiology*. **69**: 399-407.
- Zhou, Z., A. McCann, E. Litrup, R. Murphy, M. Cormican, S. Fanning, D. Brown, D. S. Guttman, S. Brisse and M. Achtman. 2014. Neutral genomic microevolution of a recently emerged pathogen, *Salmonella enterica* Serovar Agona, *Public Library of Science-Genetics*. **9**: e1003471.
- Zinoni, F., A. Birkmann, T. C. Stadtman and A. Böck. 1986. Nucleotide sequence and expression of the seleno-cysteine containing polypeptide of formate dehydrogenase(formate-hydrogen-lyase-linked) from *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*. **83**: 4650–4654.
- Zinoni, F., A. Birkmann, W. Leinfelder and A. Böck, A. 1987. Cotranslational insertion of selenocysteine into formate dehydrogenase from *Escherichia coli* directed by a UGA codon. *Proceedings of the National Academy of Sciences of the United States of America*. **84**: 3156–3160.

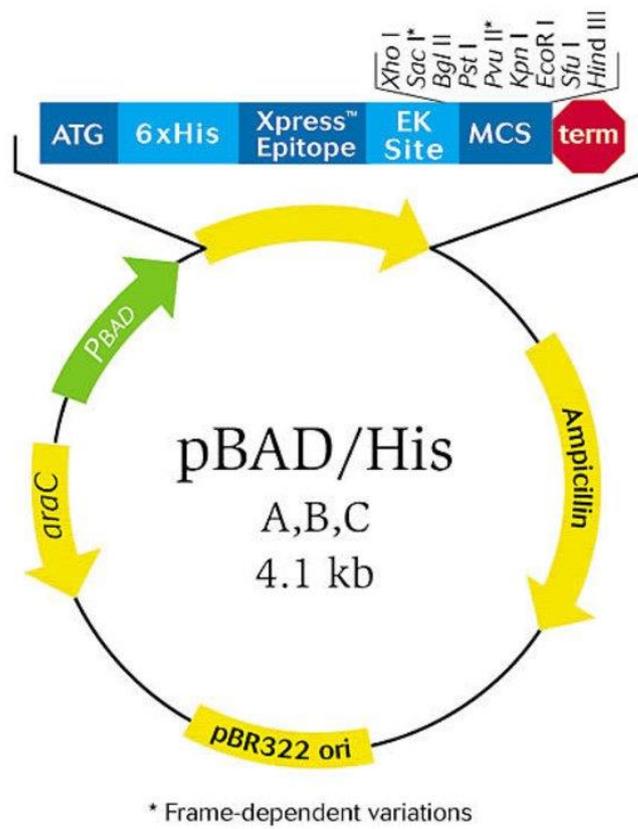
## Appendix



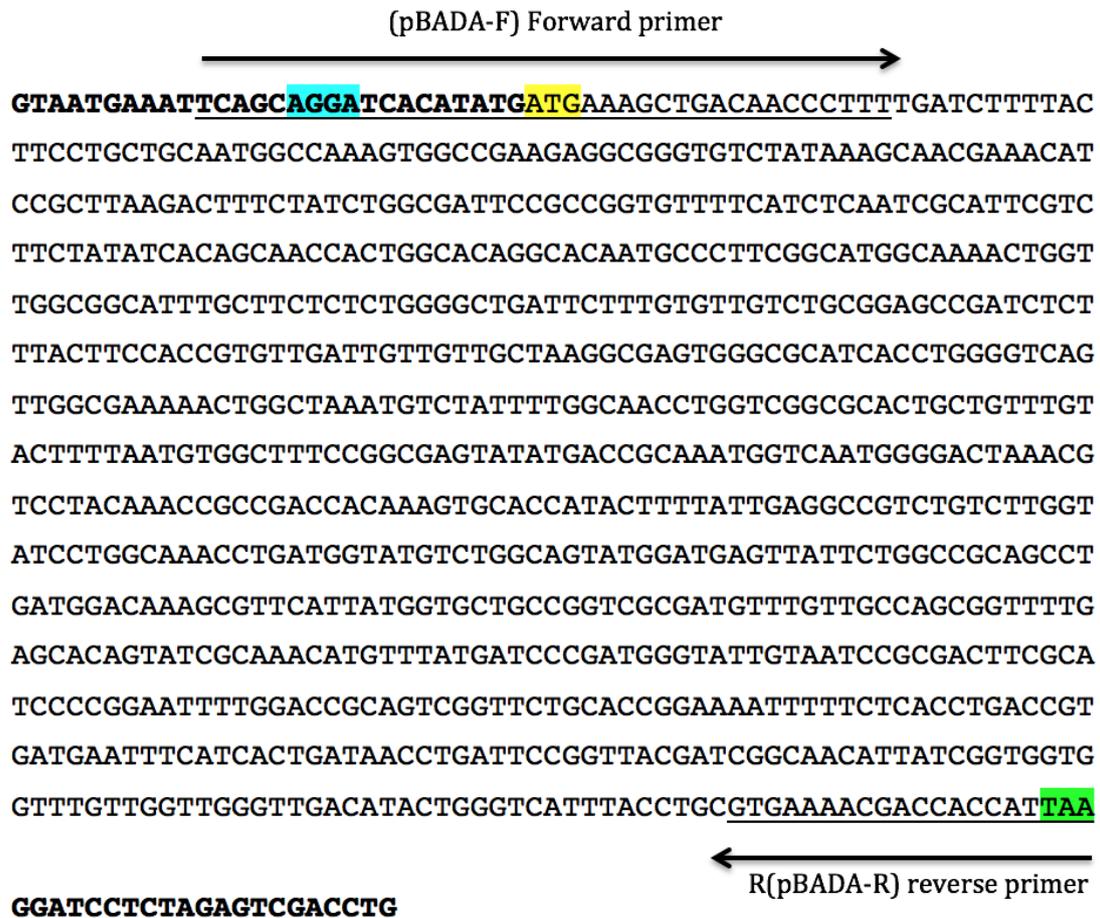
1. **Restriction map of pJET1.2/blunt**, plasmid used for cloning PCR fragments of *fdhF* to be sent for sequencing. Source; [http://www.bioinfo.pte.hu/f2/pict\\_f2/pJETmap.pdf](http://www.bioinfo.pte.hu/f2/pict_f2/pJETmap.pdf).



2. Map of rhamnose inducible plasmid pBAD<sub>rham</sub>, cloning vector with rhamnose inducible promoter (P<sub>rham</sub>). Source: Genebank.



3. **Map of arabinose inducible plasmid pBAD<sub>his</sub>**, cloning vector with arabinose inducible promoter (P<sub>ara</sub>). Source: Thermo Fisher Scientific.



4. Location of pBAD<sub>rham</sub>-*focA* primers. pBADF, forward primer and pBADR reverse primers are underline and indicated by arrows. Azure highlight indicates Ribosomal binding site, yellow highlight indicate the start codon of *focA*, green highlight indicates stop codon of *focA* and gray highlight indicates rhamnoseS (RhaS) promoter region. Bold nucleotides are indicating the pBAD<sub>rham</sub> plasmid side.

GGCCATTTTCCTGTCAGTAACGAGAAGGTCGCGAATTCAGGCGCTTTTATAGACTGGTC  
 (pBADB-F) Forward primer  
 GTAATGAAATTCAGCAGGATCACATATGATGAGAAACAAACTCTCTTTCGACTTGCAGT  
 TGAGCGCCAGAAAAGCGGCAATCGCTGAACGGATTGCCGCCATAAAATGCCCCGAGT  
 AAAGTGTCCGTCTTTTTAATGGCGATGTCCGCTGGCGTGTTTATGGCGATCGGATTTAC  
 TTTTACCTTTCCGTTATCGCCGATGCCCCGTCTTACAGGCATTAACCCATCTGGTGG  
 GCGGCCTTTGCTTTACACTCGGCTTTATTTTGCTGGCGGTTTGCGGCACCAGCCTGTTC  
 ACCTCGTCCGTAATGACGGTGATGGCAAAAAGTCGGGGCGTTATTAGTTGGCGAACTTG  
 GCTGATTAACGCACTTCTGGTGGCCTGCGGTAATCTGGCAGGTATTGCCGTGTTTCAGTT  
 TGTTAATCTGGTTTTCCGGGCTGGTGATGAGTGAAAACGCGATGTGGGGAGTCGCGGTT  
 TTACTGCGCCGAGGGCAAATGCATCATAACATTTACTGAATCTGTCAGCCTCGGCAT  
 TATGTGCAATCTGATGGTTTTGCCGCTGTGGATGAGTTATTGCCGGCGTTTCGTTAT  
 GCGACAAAATCGTCGCCATGATTTTGCCATCACCTGTTTGTGCGCCAGTGGCTTTGAG  
 CACTGTATCGCCAATTTGTTTGTGATTCCGTTTCGCCATTGCCATTTTCGCCCC  
 TCCCCCTTCTGGCAGCTGGCGCACAGTAGCGCAGACAATTTTCCGGCACTGACGGTCA  
 GCCATTTTATTACCGCCAATCTGCTCCCGGTGATGCTGGGTAATATTATCGGCGGTGCG  
 GTGCTGGTGAGTATGTGTTATCGGGCTATTTATTTACGTCAGGAACCC**TGAGGATCCTC**  
**TAGAGTCGACCTG**  
 (pBADB-R) Reverse primer

5. Location of pBAD<sub>rham</sub>-*focB* primers. pBADF, forward primer and pBADR reverse primers are underline and indicated by arrows. Azure highlight indicates Ribosomal binding site, yellow highlight indicate the start codon of *focB*, green highlight indicates stop codon of *focB* and gray highlight indicates rhamnoseS (RhaS) promoter region. Bold nucleotides are indicating the pBAD<sub>rham</sub> plasmid side.

## Appendix



6. Location of pBAD<sub>rham-hyf</sub> primers used in infusion cloning. pBAD forward primer and pBAD reverse primers are bold and indicated by arrows. Azure highlight indicates Ribosomal binding site, yellow highlight indicate the start codon of *hyfA* and green highlight indicates stop codon of *focB*.

Appendix

TAAAAACGTCGATTTTTCAAGATACAGCGTGAATTTTCAGGAAATGCGGTGAGCATCACAT  
Forward 1  
CACCACAATTCAGCAAATTGTGAACATCATCACGTTTCATCTTTCCCTGGTTGCCAATGGCCC  
ATTTTCCTGTTCAGTAACGAGAAGGTCGCGAATTCAGGCGCTTTTACTAGACTGGTCGTAATGAA  
ATTCAGCAGGATTCACATATGATCAACCGCTTTGTGGTGGCCGAACCACTGTGGTGTACAGGATGTAAT  
ACCTGTCTCGCTGCCTGTTGGACGTGCATAAAACGCAAGGTTTACAGCAACACCCGCGCCTGGCCCTGGC  
GAAGACGTCAACAATCACTGCCCCTGTCGTGTGTCATCACTGTGAGGAAGCCCCCTGCCTGCAGGTCTGCC  
CGGTCAATGCCATCTCTCAGAGGGATGATGCCATCCAACCAACGAAAGCCCTCTGTATTGGCTGCAAGCTT  
TGCGCCGTGGTCTGCCCATTTGGCGCAATCAGCGCTTCAGGAAGCCGTCGCGTGAATGCCCATGCCAATA  
Reverse 17  
TGTTTTTCAGGCTGAAGGCTCACTCAAAGACGGCGAAGAAAACCGCCCAACACAACATGCTTTGCTGGCT  
GGGAACCTGGTGTCCAGACCGTCGCGGTGAAATGCGACCTGTGTGATTTCTTGGCAGAAGGTCGGCCTGC  
GTTCCGGCTTGGCCGAATCAGGCGTTACGGGTGATCACCCTGATAGCCTGCACGTCAGATGAAAGAAAA  
ACAGCCCTTGGCCGAAGCTGGTTTGGCAATGGCGGGGAGGATCCCTTTCCCTCACTCAGGAGCAACGCT  
AATGGATGCCCTGCAATTATTAACCTGGTCGCTGATTCTCTATCTGTTTGTAGTCTGGCTTCGCTGTTTT  
TACTCGGTCTGGACAGACTGGCTATTAAGCTTTCCGGCATCACATCGCTGGTGGCGCGCTGATTGGCATC  
Forward 2  
ATCAGCGGAATTACGCAATTACATGCTGGTGAACCTTTAGTCGCCCGTTTTGCCCCCTTTTGAATTTGC  
CGATTTAACCTGCGAATGGATAGCCTCTCGGCATTTATGGTGTGGTTATCTCCTTGCTGGTGGTGGTTT  
GTTCTCTCTATTCACTTATATGCGCGAATACGAGGGCAAAGCGCGCGCGGATGGGCTTCTTTATG  
AATATTTTCATCGCATCGATGGTTGCCCTGCTGGTGTGGACAACGCTTTTTGGTTCATCGTGTGTTGA  
AATGATGTCGCTGTCTTCTGGTTTCTGGTCATTGCCAGGCAGGATAAAACGTCGATCAACGCTGGCATGC  
TCTACTTTTTATCGCCCACCGGATCGGTGCTGATAATGATCGCTTCTTGTGATGGGGCGGAAAGC  
Reverse 16  
GGCAGCCTCGATTTTGGCAGTTTCCGCACGCTTTCACTTTCTCCGGGGCTGGCGTGGCGGCTGTTCTCTGCT  
GGCCTTTTTCGGTTTTGGCCGAAAGCCGGATGATGCCGTTGCACAGCTGGTTGCCGCGCGCTCACCTG  
CCGCACCATCGCACGCTTCCGGCTTGTGCTGGCGTAATGGTCAAAATAGGTATTTTCCGCATCCTGAAA  
GTAGCGATGGATCTGCTGGCGCAAACGGGTTGCCTCTGTGGTGGGGCATTCTGGTGTGGCGATCGGGC  
AATCTCCGCGCTCCTGGCGTGTATATGCCCTGGCGGAACAGGATATCAAACGGCTGCTGGCCTGGAGTA  
CCGTCGAAAACGTCGGCATTATTTTGTGGCAGTCGGTGTGGCGATGGTCCGCTGTGCTACTGCACGACCCG  
Forward 3  
CTGCTCACCGTGGTTGGACTGCTCGGCGCACTGTTTCATCTGCTCAACCATGCGCTGTTCAAAGGCTGCT  
ATTTCTCGGCGCGGAGCGATTATTTCCGCTTTCATACCCACGACATGGAAAAATGGGGGCACTAGCGA  
AACGGATGCCGTGGACAGCCGACGATGCCGATTGGTTGCCCTCGCGATATCAGCCATTCTCCGCTGAAT  
GGTTTTATCAGCGAATGGTACACCTGGCAGTCGCTGTTCTCACTAAGTCGTGTGAAGCCGTAGCGCTACA  
ACTTGGGGTCTCTATTGCTATGGTAATGCTGGCAGTCACTGGTGGGCTGGCAGTAATGTGCTTCTGTA  
Reverse 15  
TGTACGGTATTACTTTCTGTGGTGCGCCGCGCAGTACACACGCTGAAGAGGCACAGGAAGTGCCAAATACG  
ATGATCGTCGCCATGCTACTGCTCGCGGCACTCTGCGTATTAATTGCGCTTAGTGCCAGTTGGCTGGCACC  
GAAGATAATGCATATTGCCCATGCGTTTACCAATACCCCTCCCGCCACTGTGCCAGCGGAATAGCACTTG  
TACCCGGCAGTTTTATACACAGGTCACCCCTCATTACTGTTGCTGTTACTACTGGCGATGCCTTTGCTG  
CCTGGCCTTTACTGGCTGTGGTGTGCTTCCGCGCGCGCAGCGTTTCGTGCGCACAGGAGATGCCTGGCGATG  
CGGCTACGGCTGGGAAAATGCGATGGCCCCGTGAGGCAATGGCGTGTGAGCCGCTGCGTGTGGTCTTTT  
Forward 4  
CTGCGCTATTTGCTCTACGACAACAGCTCGACCCTACGCTGAGGCTAAATAAAGGTTCTTGGCGACGTCACC  
GCCAGGGCTCAGAGCACAGAACCCTTCTGGGATGAGCGGGTGTCCGCCCCATCGTGAGCGCCACCCAAAG  
GCTGGCCAAAGAAATACAGCATCTGCAAAGCGGCGACTTTCTGCTCTATTGCCCTGTATGTGGTCCGCGCAC  
TGTTTGTGCTGTAATCGCTATTGCCGCTTAAGGAAATCACCATGAGACAAACTCTTTGCGACGGATATCT

## Appendix

GGTCATTTTTGCGTTAGCACAGGCCGTGATTCTGCTGATGCTAACCCCACTTTTTACGGGTATTTCCCGGC  
AGATACGCGCGCGTATGCACTCCCGCCGCGGGCCGGGATCTGGCAGGATTATCGCGATATCCACAAACTG

Reverse 14

TTTAAACGCCAGGAAGTTGCGCCGACATCTTCAGGTCTGATGTTCCGCCTGATGCCGTGGGTATTAATCAG  
CAGCATGCTGGTGTGCGGATGGCCTTACCAGTGTATTACCCTTTCCCTTTTTCGGGGCGGGCGGATC  
TGATCACCTTATCTATCTTCTTGGCCTGTTTCGTTTTTTCTTTGCTCTTTCGGGCTGGATACCGGAAGT  
CCGTTTGGGGAGTCGGTGCCAGTCGCGAGTTGACGCTCGGCATTCTGGTCGAACCAATGCTTATTCTCTC  
ACTGCTGGTATTGGCGCTGATAGCAGTTCACGCATATCGAGATGATCAGCAATACGCTGGCGATGGGCT  
GGAACTCGCCGCTAACCCCGTACTGGCCTTACTGGCCTGTGTTTTGCTGCTTCATTGAGATGGGAAAA

Forward 5

ATTCCCTTTGATGTTGCTGAAGCAGAACAGGAATTACAGGAAGGCCCGCTGACCGAATATTCGGGTGCCGG  
GCTGGCGCTAGCGAAATGGGGGCTGGGGCTGAAACAGGTCTGATGGCATCACTGTTTGTGGCCCTGTTTC  
TGCCCTTTGGGGCGCGCAAGAAGTCTCTCGCCTGCCTGCTGACTTCACTTGTCTGTACGCTGCTCAAG  
GTTTTGCTGATTTTTGTAAGTGGCCTCAATCGCAGAAAACACGCTGGCACGCGGGCGTTTTTTACTCAATCA  
CCATGTGACCTGGCTTGGCTTACGCTTGTGCGCTTGCATGGGTCTTCTGGTTAACCGGTCTGTAAGGAG

Reverse 13

CACTGACGGAATATGGAAAATCTTGCTCTGACGACGTTATTGCTGCCTTTTATCGGGCGCACTGGTCTGTTTC

GTTTTCGCCACAACGTCGGGCCGCGCAATGGGGGTTTTGTTCCGCCGCTGACCACGCTGTGCATGTTGT  
CACTGATCTCCGCGTTTTATCAGGCCGATAAAGTTGCCGTACGTTGACGTTGGTCAACGTTGGGGGATGTG  
CGTTGTTTTGGCCTGGTCAATGATCGCGTACGCTGATTCTGTTTTGTTGGTGGTGTCTCTCGGTTTTGCT  
GGTACGATCTACTCCACGGGTTATCTGACGGATAAAAATCGCAGAACACCCGCATAACGGCACGAATCGTT  
ATTACGATTTTTTACTGGTGTATTATCGGGCGGATGGCGGGACTGGTACTCTCTCGACGCTGCTCGGTGAG

Forward 6

TTGTTGTTTTTTGAAATTACAGGCGGCTGCTCCTGGGCGTTGATCAGTTATTACCAGAGCGATAAAGCGCA  
GCGTTCAGCACTAAAAGCGTTACTTATCACTCATATCGGCTCGTTGGGGTTGATCTTGCCGCCGCCACGC  
TGTTTTTGACAGACCGGAACGTTTGGCCTTAGCGGATGAGCGAGTTACACGGCGACGCAGTTATCTGGTT  
TATGGCGGCATCCTGTTTCCCGCGTGGGGGAAATCGGCGGAGCTACCGATGCAAGCGTGGTACCGGACGC  
AATGGAAGCGCAACACCGATCAGCGCCTATCTCCACGCCGATCGATGGTAAAAGTGGCGTTTTACATTT  
TTGCCCGGCTATTATCGACGGCGCAATATCCCGCATGTGATTGGCGGCGTTGGCATGGTCAATGGCACTG

Reverse 12

GTCACCATTCTTTATGGCTTTCTGATGATTTGCCACAGCAGGATATGAAGCGGTTGCTAGCCTGGTCCGAC  
CATCACTCAACTTGGCTGGATGTTCTTCGGCTTGTGCTCTCCATCTTCGGCTCGCGGCTGGCGCTGGAGG  
GCAGCATCGCCTACATCGTCAACCACGCGTTCGCTAAAAGCCTGTTTTTCTTGTAGCAGGTGCGCTGAGT  
TACAGTCCGCGCACCGCTTGTTCGCCGCTGCGGTGGCGTATTGCACACCCTGCCGTTGCCAGCGTGGG  
TTTCTGCGTGGCAGCGCTGGCGATTACCGGCGTCCCGCGTTCAACGCGTTCTTCAGTAAATTCGGCTGT  
TTGCTGCGGTTTTTGGCTTGTGAGTGGTACTGGATCCTGCTGCCCGCATGATTCTTCTGATGATTGAA

Forward 7

TCGGTCGCCAGTTTTCCCTGGTTTTATTCGCTGGTTGGTCCGCTTGTGCTGGCAAACCGAGCGAGGCCGT  
CGCCGATGCCGCACCGCTGCCAGGATCAATGCCCTGGTGTGATTGTAAGTGTGATGTCGCTGATTT  
CCAGCGTAATCGCCCGACCTGGTTGCAGTAAGGAGATGATGAATGACCGGTTCTATGATCGTAAATAATC  
TGGCGGACTGATGATGCTGACATCGCTGTTTGTGATTAGCGTCAAAGCTATCGCCTGTCATGCGGATTT  
TACGCCTGCCAGTCACTGGTGTCTATTTTCGCCACTCTCTCGTGCCTGTTCCGCCGAGCAACT

Reverse 11

GCTGATCTGGTCCGCCAGCGCCTTTATCACCAAAGTGTGCTGCTGGTACCCTTAATCATGACTTACGCTGCAC

GAAATATTTCCCGAGAACATCCCGGAAAAAGCGTTATTCCGGTCCGGCAATGATGGCACTGCTCGCGGCGTTA  
ATTGCTCTGCTTTGGCATTGTCGTTTCCAGCCGTTGAAGCTACCGATGGCTACCGGGCTGAAACCGGGCGT  
GGCGGTAGCGTTAGGTCAATTTCTGCTTGGCCTGCTGTGATTGTCAGCCAGCGCAATATCCTGCGGCAAA  
TTTTTGGTTACTGCCTGATGAAAAACGGCTCCCATCTGGTGTGCTGGCGCTTCTTGCCTGGCGAGCACCGGAA  
CTGGTGGAAATAGGTATCGCTACCGACGCCATACGGCACGCTGGACGTGAACAACCTTGACCGCGCTGAAGG

Appendix

Forward 8  
→

GATAATGAGATGAGTTATTCTGTGATGTTGCGCTTACTCCTGCTCACGCCGCTGCTTTTTTCGCTGCTCTG  
TTTTGCCTGCCGAAACGGAGACTTCTGCGACTCGCACGGTGACCGTATTACATAGCTTAGGGATCACAC  
TGCTGCTGATTCTGGCACTCTGGGTGGTCCAAACTGCCGCTGATGCAGGAGAAATATTCGCTGCCGGACTG  
TGGCTTCATATTGATGGTCTGGGCGGTTTGTTCCTCGCCATTCCTGGTGTGATTGGCTTCTCACCGGTAT  
TTACTCGATTGGCTACATGCCATGAAGTGGCACACGGCGAGCTTTCACCCGTTACGCTGTGCGATTACT  
ACGGTTTCTCCATCTGTTTTTGTTCACCATGCTGGTGTGTTACCAGCAATAACCTGATTGTGATGTGG

Reverse 10  
←

CGGGCGATCGAAGCCACCACCTTAAGCTCGGCGTTTTCTGGTAGGCATTTACGGTCAGCGTTCATCGCTGGA  
AGCTGCATGGAAGTACATCATTATTTGACTGTTGGTGTGCGCTTTTGGTCTGTTCCGTACCGTCTGGTAT  
ACGCCAACCGCCGACTTATGCCGAGGCAAAATGGCGATATTCTGGAGCGAGGTTCTTAAGCAATCG  
TCCTTGCTTGACCCAACATTAATGCTGTTGGCCTTTGTGTTTTTGTAAATTGGCTTTGGTACCAAAAACCGG  
GCTATTTCCCATGCACGCCTGGCTGCCGATGCTCACAGTGAAGCGCCGAGTCCGGTCAGCGCCCTGCTCT  
CCGCCGATTGCTGAACTGCCGCTGTTGGTGTGATTGCTATTACATCATTATTTGCCAAGCCATCGGC

Forward 9  
→

AGCGATTTCCCCAACCGGTTGTGCTCATCTTCGGCATGTTGTGCGTTGCCGTGGCGGCATTTTTTCATTCT  
GGTACAGCGGGACATTAAGCGTCTGCTGGCGTACTCCAGCGTGGAGAACATGGGGCTGGTCCGCGTGGAGC  
TAGGCATTGGCGGGCGCTGGGAATTTTTGCCGCGCTGCTGCACATCTTAAACCACAGTCTGGCAAAAACG  
CTGCTGTTCTGCGGTTCCGGCAATGACTGCTCAAGTACGGCACGGCGGATCTCAACGTCGTCTGTGGGAT  
GCTCAAAATCATGCCATTTACCGCCGCTGCTGTTTGGCGCGGTGCGCTGGCGCTGGCAGGGATGCCGCCCT

Reverse 9  
←

TCAACATTTTTCTTAGCGAATTTATGACCATTACCGCCGACTGGCACGTAATCACCTGCTGATTATCGTC  
CTGCTGTTATTGCTGTTAACGCTGGTGTGCGGGCCCTGGTACGGATGGCTGCGCGGCTGTTAATGGCGAA  
ACCGCCGAGGCCGTTAACGGGGTGATCTCGGCTGGTTGACCACCTCGCCAATGGTATTCTGCTGGTCA  
TGATGCTGGCGATGGGAACGCATATCCACAACCTGTATCAGGATCCTGGCGGGCGCTTCCACTATAGTC  
CTCTCAGGACCGCACGATCTGCCTGCACAACGTAGCACCTGGCATGATTTTTTGCCTTCAGGCACCGCATC  
TGTTTCGGAGAAACACAGTGAACGTTAATTCATCGTCAAAATCGTGGCGAAGCGATTCTCGCCGCCCTGAAA

Forward 10  
→

ACGCAGTTCGCCGCGCGGTGCTGGATGAAGAGCGACAAACGCCCTGAACAGGTCACCATTACGGTGAAAAAT  
CAATCTGCTGCCTGACGTTGTACAGTATCTTTATTATCAACATGATGGCTGGCTTCCGGTCTGTTTGGCA  
ACGACGAGCGGACACTTAACGGTCATTACGCGTTTTATTATGCCCTTTCAATGGAAGGGCCGAAAAATGC  
TGGATTGTTGGAAGGCGCTGGTGCATGCCGACAGTCCGGAGTTTCCGTGAGTACACACCGCGCTCCCTGC  
CGCGGTTGCGGGCGAGCGAGAAATTCGCGATATGTACGGGCTGATTCCGGTTGGCTGCCGGATCAGCGTC  
GCCTGGTGTGCCCGATGACTGGCCGGAAGATATGCATCCGCTGCCGAAAGATGCCGATGATTATCGACTG

Reverse 8  
←

CGCCCTGAACCGACGACTGATTCGGAACGATCCGTTTATCAATGAGGGCAACAGCGATGCGCGGGTAT  
CCCTGTCGCCCCGCTGCATATCACCTCCGATGAACCGGGTCACTTCCGCTTGTGTTGGATGGCGAGCAAA  
TTGTCGATGCTGATTACCGCCTGTTTATGTCCATCGCGGCATGGAGAAACTGGCAGAAACCGGGATGGGC  
TACAACGAAGTGACCTTCTTATCGGACCGCGTGTGTGGGATTTGCGGTTTTGCCACAGTGTGGCCTATAC  
CAATTCGGTTGAAAATGCACTGGGATGAGGTGCCGCAACGAGCACATACTATTCGCTCGATTCTGCTGG  
AAGTGAACGGCTACACAGTCATTTGCTTAACCTTGGCCTCTCCTGCCATTTGTTGGTTTTGATACCGGC

Forward 11  
→

TTTATGCAATTTTTCCGCGTGGCGGAAAAGTCGATGACGATGGCGGAATTGCTGATCGGGTCCGCTAAAAC  
CTACGGTCTGAATCTGATTGGTGGTGTTCGCCGCGATATTCTCAAAGAGCAACGCTCTGCAAACGCTGAAAC  
TGGTGGCGGAGATGCCGCGCCGACGTGTCCGAGCTGGTAGAGATGCTGCTTGTACGCCGAATATGGAACAA  
CGCACTCAGGGCATTGGCATTCTCGACCGACAAATCGCCCGTGAATTTGCGCTTTGATCACCCCTACGCCGA  
CTACGGCAATATTCAAAAACACTGTTTACCTTTACCGCGGCGGATGTTTTCTCCCGGTGATGGTCCGTCG  
TCAAAGAGACGTTTTGATTGCTGGCAATGCTGGAATTTGCCCTCGACAACATGCCGGATACCCACTGCTG

Appendix

Reverse 7

**ACCGAAGGCTTTA**GCATATAAACCTCACGCATTCGGCGCTGGGCTTTGTTGAAGCGCCACGCGGTGAAGACGTT  
GCACTGGAGCATGCTCGGTGATAAACCAAAAATTGTTCCGCTGGCGCTGCCGTGCCGCCACTACGCCAACT  
GGCCGGTGTTCGTTACATGCTGCGCGGCAATACCGTTTCTGACGCACCGCTGATTATCGGTAGCCTTGAT  
CCCTGCTACTCCTGTACCGACCGTGTGACGCTGGTAGATGTGCGCAAGCGCCAGTCAAAAACCGTGCCGTA  
TAAAGAGATCGAACGCTACGGCATTGATCGTAACCGTTCCGCGCTGAAGTAAGGACAGAAGATGCTGAAGT

TACTGAAAACATATTATGCGCGCCGGAACCGCGACGGTGAATATCCCTTCCGCGCCACTGGAGGTGAGCCCT

Forward 12

**GGCTTTCGGGAAAAC**CGGACCTGATGCCAGCCAATGTATTGCCTGCGGTGCCCTGCGCTGTGCTTGTCC  
GGCAAATGCGCTGACTATCCAGACCGACGACCAGCAAAAATTCGCGCACCTGGCAGCTCTATCTGGGGCGTT  
GTATTTACTGCGGACGTTGTGAAGAAGTGTGCCCGACAGAGCCATCCAGCTTACCAATAACTTTGAACTG  
ACCGTCAACAAATAAGCCGATCTCTATACCCGCGGACGTTCCATCTACAACGTTGCAGCCGTTGCGAACC

Reverse 6

CCCGTTTCCCCGCAAAAACCATCGCACTGGCTGCTGAATT**TTAGCACAGCAACAAAATGCGCCACAAA**

**ACCGCGAAAATGTTGTGGGCGCAAGCGAGCGTCTGCCCGGAATGCAAAACAACGCGCGACGCTGATCAACGAC**  
GATACAGATGTACTGCTGGTGGCTAAGGAGCAGCTATGAGTCCAGTGCTTACACAACATGTGAGCCAGCCC  
ATCACGCTGGACGAGCAAAACGCAAAAGATGAAGCGGCATTGCTACAGGATATCCGTCGCTCGGCTTACGT  
TTATCGCGTTCGATTGCGCGCGCTGCAACGCCTGTGAAATCGAAAATTTTGTGCTGCCATTACACCAGTATTCG  
ACGCAAGAACGTTTTGGCATTAAAGGTTGTTTCATCACCGCGTACGCGCGATATTTGTTATTTACTGGCGCA  
GTCACCCGGGCGATGCGTATGCCTGCCTTCGGGCGTATGAGTCTGCCCGGATCATAAAAATTTGTGTTTC

Forward 13

CTACGGCGCGTGGCGTGTGCGCG**GGGTTATTTCCACGATC**CTACAGCGTCTGGGGCGGTAGCGACACCA  
TTGTCCCATTTGATGTTTGGATCCCCGGCTGCCCGCAACACCGCGCCACCATTACCGGTTTCCGCGTG  
GCGCTCGGTTTGTGCAACAGAAAGATTACGCTGTGGATTATCGCGATCCACCGGGGTGACTATGCAACC  
GTTGTGGCGCAGATCCCGCCATCACAGCGTATCGCCATTGAGCGAGAAGCGCGCGGCTGGCGGGCTATC  
GTCAGGGGCGAGAAATTTGCGATCGGCTCCTGCGCCATTTAAGCGACGATCCTACAGGAAATCGGGTTAAC  
ACCTGGTTGCGCGATGCCGACGATCCACGCTCAATAGTATCGTTTCAGCAACTCTTTCGCGTACTCCGGG

Reverse 5

**GTTACATGACTGAAGAGT**GCGGGGAAATTTTGTGACGCTGCCAAAAAGTTTGTGCCAGTAGCGAC  
GAGATGCCGGAACACAGCTCTCAGGTAATGTATTACTCGCTAGCTATCGGCCATCACGTTGGCGTGATTGA  
TTGTCTGAATGTGCCCTTCCGCTGCCCACTGACGGAATACGAAGATTGGCTTGCCTGGTGAAGAGGAGC  
AAGCCCGACGTAAGATGCTGGGGGTGATGACTTTTGGTGAGATTGTTATTGACGCCAGCCACACCGCCCTG  
TTGACCCGGGCATTCGCGCCACTGGCGGATGACGCGACGCTCTGTGTGGCAGGCGGCTAGCATTCAATTCAT

TCATCTGTTGGATGAAATTTGTGCAGGAACCGCCATCTATCTGATGGCCAGAAAAAT**TCGGTGAGAAGGAT**

Forward 14

**TTCTCATT**AATAAGGACTGTTGATGGCTATGCAGACGAGGCGATGTTTGGCCCGCCACAAGGAATAACAAT  
TGAAGCGGTAACGGAAATGCTCGCGGAGCGGTTAGCACAGAAACACGGCAAGGCGTCTTTATTACGCGCCT  
TCATCCCGCTGCCCGCCCGTTTCAGCCCGGTACAACCTTATTGAACTGCATGTTCTCAAAAGCAACTTCTAT  
TACCGCTACCATGATGATGGCAGCGATGTGACGGCAACAACAGAGTATCAGGGCGAGATGGTCGATTATTC  
CGGTACGCCCGTCTTCTCGGCAGTAGTGAATGGCGGAGCTACGCTTTATTTCGCCACCCACGGCAGTCGTT

Reverse 4

TTACTTCCCAGGAT**TGCACACTGTTTAACTGG**CTGGCGCGGATAATCACCCCGGTTCTGCAATCATGGCTC  
AATGATGAAGAACAGCAGGTGGCGCTGCGTTTGGTGGAGAAAGATCGCGATCATCATCGGGTACTGGTTGA  
TATTACTAATGCAGTGTGTACATCTTGATCTCGACGATCTGATCGCTGACGTCGCTCGTGAGATCCATC  
ATTTTTTCGGTCTGGCTTCAGTCAGTATGGTACTGGGCGATCATCGAAAGAACGAGAAGTTCAGCCTGTGG  
TGCAGCGATCTTCTGCCTCACATTGTGCGTGTCTGCCACGCTGTATGCCTGGCGAAAGTGTATTGCTGAC  
ACAAACGCTACAAACCCGACAACCGACCTTGACGCACCGTGCAGATGATCTGTTTCTCTGGCAACGGGAC

## Appendix

Forward 15  
→

CGTTATTACTCTTACTTGCATCTAACGGCTGC**GAATCTGCGCTCCTTATA**CCGCTTACCTTTGGCAACCAT  
 ACACCGGGTGCATTGTTGCTGGCGCATACTCTTCCACTCTCTTTAGTGAGGAAAACCTGCCAGCTACTACA  
 ACACATAGCCGATCGCATCGCTATTGCCGTTGGCAATGCCGATGCCTGGCGTAGCATGACCGATTTGCAGG  
 AAAGTTTGCAGCAAGAAAACCACCAGCTTAGCGAGCAGCTCCTTTGGAATCTGGGCATCGGTGACATTATC  
 TATCAAAGCCAGGCAATGGAAGACCTGCTCCAGCAGGTAGATATTGTGGCGAAGAGCGACAGTACGGTGT  
 GATTTGTGGTGAAACCGGAACGGCAAAGAGGTGATCGCCAGAGCGATCCATCAACTTAGCCCGCGACCG

Reverse 3  
←

**ACAAGCCGCTGGTCAAAA**TCAACTGCGCTGCCATCCCCGCCAGTCTTCTGGAAAGTGAGTTATTCGGTCAT  
 GACAAAGGGGCGTTTACTGGTGCATTAATACCCATCGTGGTCTGTTTTGAAATTGCCGATGGCGGCACGTT  
 GTTTCTCGATGAAATTGGCGATCTGCCGTTAGAACTTACGCCTAAACTGCTGCGCGTATTGCAGGAGCGGG  
 AGATTGAGCGTCTCGGCGGAGTAGAACGATCCCCGTGAATGTCAGAGTCATTGCCGCCACCAACCGTGAT  
 TTGTGGCAAATGGTTGAAGATCGCCAGTTTCGCAGCGATCTCTTTATCGCCTGAATGTCTTCCCACTGGA

Forward 16  
→

**GCCATATGAAT**CGCGCAATTGACGCCATCCCGACCGAGGCACTACGCCAGTTGATGTCTGTTGGGATTGGCCG  
 GGCAACGTGCGCGAGCTGGAAAACGTGATTGAGCGGGCGGTACTGTTGACTCGTGGTAACAGTCTGAATTT  
 ACATCTAAATGTCCGACAAAGCCGTTTACTGCCGACGCTAAATGAAGATTCAGCGCTTCGCCAGTTCAATGG  
 CGCAGTTACTGCACCCGACGACGCCAGAGAATGACGAAGAAGAACGTGAGCGCATTGTTCCAGGTATTGCCA  
 GAAACCAATGGCATTGTTGCCGGGCCCGTGGCGCAGCGACGCGATTAGGGATGAAGCGCACCAACGCTGCT

Reverse 2  
←

GTCACGAATGCAGCGGCTGGGGATCTCGGTTTCG**GAGGTGTTGTAATCTGCT**TTTGCAGGAGTATGCATGA  
 GAAACAACTCTCTTTTCGACTTGCAGTTGAGCGCCAGAAAAGCGGCAATCGCTGAACGGATTGCCGCCAT  
 AAAATTGCCCGCAGTAAAGTGTCCGTCTTTTTAATGGCGATGTCGCGTGGCGTGTATGGCGATCGGATT  
 TACTTTTTACCTTTCCGTTATCGCCGATGCCCCGTTCTCACAGGCATTAACCCATCTGGTGGCGGCCCTTT  
 GCTTTACACTCGGCTTTATTTTGTGGCGGTTTTCGGCACCAGCCTGTTACCTCGTCCGTAATACGGTGA  
 TGGCAAAAAGTCCGGGCGTTATTAGTTGGCGAACTTGGCTGATTAACGCACTTCTGGTGGCCTGCGGTAAT

Forward 17  
→

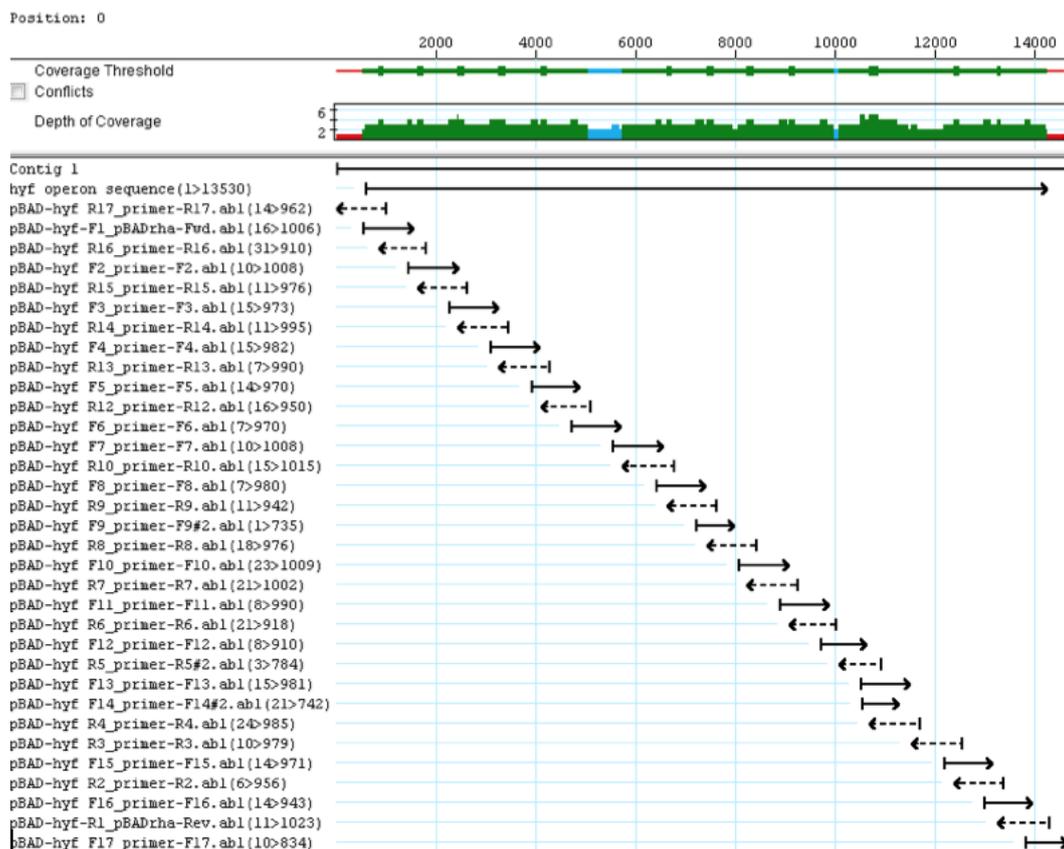
CTGGCAGGTATTGCCTGTTTCAGTTTGT**TAATCTGGTTTTCCGGGCT**GGTGATGAGTGA AAAACCGGATGTG  
 GGGAGTCCGGGTTTTACTGCGCGGAGGGCAAAATGCATCATACTTACTGAATCTGTGACGCTCGGCA  
 TTATGTGCAATCTGATGGTTTGCCTGGCGCTGTGGATGAGTTATTGGGGCGTTCTGTTATGGACAAAATC  
 GTCGCCATGATTTTGCCCATCACCTGTTTGTCCGAGTGGCTTTGAGCACTGTATCGCCAATTTGTTTGT  
 GATTCGGTTCGCCATTGCCATTGCCATTTCCGCCCTCCCCCTTCTGGCAGCTGGCGCACAGTAGCGCAG  
 ACAATTTTCCGGCACTGACGGTCAGCCATTTTATTACCGCCAATCTGCTCCCGGTGATGCTGGGTAATATT  
 ATCGGCGGTGCGGTGCTGGTGTGATGTTTATCGGGCTATTTATTTACGTCAGGAACCC**TGA**AAAATCAG

Reverse 1  
←

CCCCGGCAAAACAGTTCGTCCGGCTGAAGATTATATGCCATGGTACCCGGGGATCCTCTAGAGT**CGACCTGC**

**AGGCATGCAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGA**

7. **Location of pBAD<sub>rham</sub>-*hyf* primers for sequence.** pBADF, 17 forward primer and pBADR 17 reverse primers are bold and indicated by arrows. Azure highlight indicates Ribosomal binding site, yellow highlight indicate the start codon of *hyf* operon (*hyfA*) and green highlight indicates stop codon of *hyf* operon (*focB*).



8. Strategy view of *hyf* operon after sequencing. Solid lines indicates forward primers that has been used in the sequence and dash lines indicates the reverse primers.

## Appendix

```

AGGAGTGTTTATCGTCAGCATGAATGTAAAAGAGATAACATCATCACGTTTCATCTTTCCCTGGTTGCCAAT
GGCCCATTTTCCTGTCAGTAACGAGAAGGTCGCGAATTCAGGCGCTTTTGTAGACTGGTCGTAATGAAATTC
pBADrha-hyc Forward primer
----->
AGCAGGATCACAATGACTATTTGGGAAATAAGCGAGAAAGCCGATTACATCGCACAGCGGCATCGTCGCCT
TTTGATGGTGAGGAGACTGCGTGACTGACGTTTTACTCTGTGTTGGCAATAGCATGATGGGCGATGATGGC
GCAGTCCGCTGCTGGCGGAAAAGTGCGCCGCCGCCGAAAGGTAAGTGGGTGGTGATTGACGGCGGTAG
-----
----- hyc operon -----
-----
CGCACCGGAAAAACGACATCGTCGCTATCCGTGAACTGCGCCCCGACACCGACTGCTGATTGTCGACGCCACGG
ATATGGGGCTAAACCCCGCGAGATCCGCATCATCGACCCGGATGATATCGCCGAGATGTTTATGATGACT
ACCCATAACATGCCGTTGAATTACCTTATCGACCAGTTGAAAGAAGATATTGGCGAAGTGATTTTCCTCGG
CATTGAGCCGGATATCGTCGGCTTTTACTACCCGATGACCCAGCCGATTAAGATGCGGTAGAAACCGTTT
-----
-----<
pBADrha-hyc reverse primer
ATCAACGACTGGAAGGCTGGGAAGGAAATGGCCGGCTTCGCGGCAGTTAGCGGTGGAAGAAGAGTAGGGATCC
-----
TCTAGAGTCGACCTG

```

9. **Location of pBAD<sub>rham</sub>-hyc primers used in infusion cloning.** pBAD forward primer and pBAD reverse primers are bold and indicated by arrows. Azure highlight indicates Ribosomal binding site, yellow highlight indicate the start codon of *hycA* and green highlight indicates stop codon of *hycI*.

# Appendix

CACCACAATTCAGCAAATTTGTAACATCATCACGTTTCATCTTTCCCTGGTTGCCAATGGCCCATTTTCCTG  
Forward 1  
TCAGTACGAGAAGGTCGCGAATTCAGGCGCTTTTTAGACTGGTCGTAATGAAATTCAGC**CAGGATCACAATG**  
→  
**ACTA**TTTGGGAATAAGCGAGAAAGCCGATTACATCGCACAGCGGCATCGTCGCCTACAGGACCAGTGGCAC  
ATCTACTGCAATTCGCTGGTTCAGGGGATCACGTTATCGAAAGCGCGCTGCATCACGCCATGAGCTGCGC  
GCCGACAAAGAACTCTGTTTCGTCCTTTTTGAACATTTTCGCATTTACGTACCCCTGGCGGATGGCTTTA  
ACAGCCACACCATCGAGTATTACGTGAAACAAAAGATGGCGAAGACAAACAGCGGATTGCGCAGGCGCAA  
CTGAGCATTGACGGCATGATTGATGGCAAGGTCAACATCCGCGATCGCGAACAGGTTCTGGAACACTATCT  
Reverse 11  
←  
CGAAAAATCGCTGGCGTTTACG**ACAGCTTATACACC**GCTATTGAAAAAATGTGCCGGTGAATTTAAGCC  
AACTGGTAAAGGACAAAGCCCGCAGCATGAGCTGAGGCTTTGCCCGTTTTCAGGCGTTACGCCCTGTTT  
GGGATGGCGCTGTCGATGAGTGTGAAAAATGACATTTTCATCGGCATGTTTTCGTCAAAAAATGACAATCAC  
CTGAGGAATGCCGGTGAATCGTTTTGTAATTGCTGACTCCACGCTCTGTATCGGCTGCCACACTTGTGAG  
GCCGCTGTTTCAGAGACGCATCGCCAGCACGGCTGCAATCAATGCCGCGCTGAGAGTGTGCTGAATGA  
AAAAGAATCTGCCCGCAGCTCTGTCAACTGTGAAGATGCACCCTGCCGGTGGTCTGCCCGTTAAGC  
Forward 2  
→  
CCATCACCCCGCTCG**ATGGGCGCGTGCAGTTGA**ATGAAAGCCTGTGCGTAAGCTGCAAGCTGTGCGGCATC  
GCCTGCCCGTTTGGCGCAATFGAATTTTCCGGCAGCCGTCGCTGGATATTCGGCAAACGCCAATACCCC  
GAAAGCGCCACCGGCACCGCTGCTCCGGCGCGTGTGACACATTTGCTTGACTGGGTGCCAGGTATTGCGG  
CGATCGCCGCTCAATGTGACCTTTGTAGCTTTGATGAAACAGGTCCGGCCTGCGTGGCGATGTGCCGACT  
AAAGCCCTGCATCTGGTGGATAACACCGATATCGCCCGCTCAGCAAACGTAAGCGTGAGCTGACCTTTAA  
Reverse 10  
←  
CACGGACTTTGGCGATCTCACCTTGTTCAGCAGGCTCAAAG**TGGAGAGGCTAAATGAG**CGCAATTTCCCT  
GATCAATAGCCGGTGGCGTGGTTTGTGCGCGCGCTGTTCTGGCATTCTCTTTCTTTTCAAAAAGCGT  
TAAGTGGCTGGATAGCTGGAATTTGGCGCGCGGTTGGTAGTCTGTATACGGCAGCCGCGGGCTCACTGTA  
CTGACTGGCGCGGTGGCGTGAGCGGTGCGCTGTGCGTGGTAAGCTACGATGTGCAAATCTCTCCGCTTAA  
CGCGATTTGGCTGATTACGCTCGGTCTGTGCGGTCTGTTGTGACCTCTACAACATTGACTGGCATCGCC  
ACGGCAGGTGAAGTGAACCGCTTGCAGATCAATATGTTGATGGCTGCCCGCTGCGCCCTCATTGCC  
Forward 3  
→  
AGCAACCTCGGCATGTTTCGTGGTAATGGCCGAAAT**CAATGGCCCTGTGCGCGG**TGTTCCCTACCAGCAACAG  
CAAAGAGGGCAAACCTGTGGTTTGGCGTGGGGCGTCTTGGCACTCTGCTGCTGGCGATTGCTTGTGGCTGC  
TGTGGCAGCGTTACGGCACGCTGGATCTGCGCCTGTGGATATGCGTATGCAACAGCTGCCGCTCGGTTCC  
GATATCTGGCTGCTCGGAGTATTGGCTTTGGCCTGCTGGCCGGGATTATTCCGCTGCACGGCTGGGTGCC  
GCAGGCACATGCGAACGCCTCTGCGCCAGCTGCCGCGTTGTTTTCTACGGTAGTCATGAAAATTTGGCCTGC  
Reverse 9  
←  
TGGGCATTTTAAACCCTGTCACTGCTGGCGGTAATGCACCCTGTGGTGGGGATCGCGCT**GCTGGTGCTC**  
**GGCATGAT**CACCGCTTTGTCGGTGGTCTGTATGCGCTGGTGGAGCACAACATCCAGCGCTGCTGGCTTA  
CCACACCCTGGAAAATATCGGCATCATCTGCTGGGGCTGGGCGCTGGCGTAACGGGTATCGCGCTCGAAC  
AACCGCGCTGATTGCTCTTGGCCTGGTGGTGGTCTGTACCATCTGCTTAACCATAGCCTGTTCAAAAGC  
GTACTGTTCCCTCGGGGCGGGAGCGTCTGGTCCGTACCGGTATCGCGATATCGAAAACCTGGTGGTAT  
TGGCAAGAAAATGCCGTTATCTCCATCGCCATGTTAGTCCGGCTGATGGCAATGGCTGCGCTGCCGCGC  
Forward 4  
→  
TGAATGGTTTTGCCGGGAATGGGTTATCTATCAATCATTTTTCAAAC**TGAGCAATAGTGGCGGTTGTT**  
GCCGCTGCTGGGGCGCTGCTCGCTGTGGGGCTGGCAATTACCGGTGCGCTGGCGGTGATGTGTATGGC  
GAAAGTCTATGGCGTCACGTTCCCTCGGCGCGCCGCGCACAAAGAAGCCGAAAACGCCACCTGTGCGCCGC  
TCCTGATGAGCGTAAGCGTAGTGGCACTGGCGATTTGCTGCGTAATTGGCGGTGTTGCTGCGCCGTGGCTA  
CTGCCGATGCTCTGCTGCTGTACCTCTGCCGCTGGAGCCTGCTAACACCACCGTTTCTCAACCGATGAT  
CACGTTGCTGCTGATTGCCTGCCGCTGCTGCCATTATGCGGATTTGCAAAGCGCATCGTTTGC

Appendix

Reverse 8  
←

CATCGCGTTCCCGCGGTGCGGCCTGGGT**TGTGCGGTTACGACCACG**AAAAATCAATGGTGATTACCGCTCAC  
GGTTTTGCCATGCCGGTGAACAGGCGTTTGGCCCGGTGCTGAAACTACGCAAATGGCTGAATCCGGTGT  
TCTGGTGCCGGGCTGGCAGTCCGAGGGGAGTGCCTGCTGTTCCGCCGGATGGCGCTGGTTGAACTGGCGG  
TACTGGTGGTGATTATTGTTTACAGGAGCCTGAGAATGAGTGTTTTATATCCGTTAATTCAGGCGCTGG  
TGTTATTTGCCGTTGCGCCGCTGCTCTCCGGTATAACCCGCGTGGCGCGCGCCGCTTGCATAACCGTCCG  
GGCCCGGGCTGTTCAGGAGTATCCGACATTATCAAACCTGCTGGGCGCTCAGAGCGTCGGCCCGGATGC

Forward 5  
→

**CTCCGGCTGGGTGTTCCGC**CTGACGCCGATGTGATGGTGGGCGTCATGCTGACTATCGCTACTGCGCTGC  
CGGTGGTGACCGTCCGGTCTCCGCTGCCGCAACTGGGTGATTTGATCACCTTACTGTATCTCTTTGCCATC  
GCGCGTTTCTCTTTGCCATTTCTGGTCTGGATACCCGGTAGCCCGTTTACCGCTATCGGCGCGAGCCGTTGA  
AGCGATGCTTGCGGTGCTGGTCGAACCGATGCTGCTGCTTGGTCTGTGGGTTGCCGCACAGGTTGCCGCT  
CCACCAACATCAGCAACATCACCGACACCGTTTACTGCTGGCCGCTGAGCCAGAGCATCCCGCTGGTACTG

Reverse 7  
←

GCGCTTTGTGCCTGTGCGTTCGCCACCT**TTATCGAAATGGGCAAACT**GCCGTTCCGACCTGGCGGAAGCCGA  
GCAGGAGCTGCAGGAAGGCCCGCTCTCTGAATACAGCGGCAGCGGCTTTGGCGTCATGAAATGGGGTATCA  
GCCTGAAACAGCTGGTGGTGTTCAGATGTTCTGTCGGGGTGTTTATTCGGTGGGGACAAATGGAAACCTTC  
ACCGCCGGTGGACTGCTGCTGGCGCTGGTGATTGCCATCGTAAACTGGTGGTCCGCGTCTGGTTATCGC  
GCTGTTCCAAAACAGCATGGCCCGTCTGCGTCTTGATATTACTCCGCGCATTACCTGGGCTGGGTTTGCT  
TTGCATTTTTAGCGTTCGTCTCCTTGCTGGCGCGGTGATTAAGAGAGTTTGAGCATGTCTGAAGAAAAAT

Forward 6  
→

TAGGTCAACATTATCTCGCC**GCGCTGAATGAGGCATTT**CCGGGCGTCTGCTGGACCACGCCCTGGCAGACC  
AAAGATCAGCTGACTGTCACCGTAAAGGTGAACCTACCTGCCGGAAGTGGTGGAGTTTCTTTACTACAAACA  
GGGTGGCTGGCTGTCCGGTCTGTTGGTAACGACGAACGCAAACCTGAATGGTCATTACGCCGTTTACTACG  
TGCTGTCGATGAGAAGGGCACTAAGTGTGGATTACGGTTCCGCTCGAAGTTGACGCCAACAAACCGGAA  
TATCCGTCGGTGACGCCCGCGCTTCCGGCGCGGTGTGGGGCGAGCGTGAAGTCCGCGATATGTACGGTTT

Reverse 6  
←

GATTCCGGTTGGTCTGCCGGATGAACGTCGTCTGGTCTGCCGGATG**ACTGGCCGGATGAACTTT**TATCCGC  
TGCGTAAAGACAGCATGGATTATCGTCAGCGTCCGGCACCGACCACCGATGCTGAAACCTACGAGTTTCTC  
AACGAACTGGGCGACAAGAAAAACAACGTCGTGCCGATTTGGTCCGCTGCACGTCACCTCTGATGAACCGGG  
CCACTCCGTCCTGTCGTGATGGCGAAAACATTATCGACGCCGACTACCGTCTGTTCTACGTCATCCGCG  
CGATGGAAAAACTGGCGGAAACCCGATATGGGTTATAACGAAGTGACCTTCTCTGACCGTGTGTGCGGG  
ATCTGCGGCTTTGCCACAGCACCGCTACACCACGTCGGTGGAAAACGCGATGGGTATTCAGGTGCCAGA

Forward 7  
→

ACGTGCGCAGATGATCCGCGCCATTCTGCTGGAGGTAGAA**ACGCTTGCACTCGCATCT**GCTCAACCTTGCC  
TGGCCTGTCACTTTACCGGCTTCGACTCCGGCTTTATGCAGTCTTCCGCGTGCCTGAAACCTCCATGAAA  
ATGGCAGAGATCCTTACCGGTGCCGTAACCTACGGCCTGAACTTGATCGGCGGGATTCTGTCGCGATCT  
GCTGAAAGACGACATGATCCAGACCCGCCAGCTGGCACAACAGATCCGTCGTGAAGTGCAGGAGCTGGTGG  
ATGTGCTGCTGAGCACTCCGAACATGGAACAGCGCACTGTCCGCATTGGTCTGTTGGACCCGGAATCGCT

Reverse 5  
←

**GTTTGTGCGGCTAT**GGCCTGCTGCCAATGGAAGTCCACAGCGAGCGGCTGCGACGTTATTTCCCGTCTGA  
AAGTGCATCAACGAAGTCTATACCGCGCTGAACATGATCGACTACGGTCTGGATAACCTGCCGGGTGCC  
CCACTGATGGTGAAGGCTTTACCTACATTCGCCACCGCTTTGCCGCTGGGCTTTGCCGAAGCGCGCGCGG  
CGATGATATCCACTGGAGCATGACCGGCGACAACCAGAAGCTGTACCCTGGCGCTGCCGTGCCGCGACCT  
ACGCGAACTGGCCGACCCTGCGTACATGCTGCCGGCAACACCGTTTCCGATGCGCCGCTGATTATCGGT

Forward 8  
→

AGCCTCGACCCTTGCTACTCCTGTACCGACCGCATGACCGTGGTTCGATGTGCGTAAGA**AGAAGAGCAAAGT**

## Appendix

**GGTGCCGTACAAAGAACTCGAGCGTTACAGCATTGAGCGTAAAAACTCGCCGCTGAAATAAGGAATCGCCA**  
 TGTTTACCTTTTATCAAAAAAGTCATCAAAACCGGCACGGCGACCTCGTCTTATCCGCTGGAGCCGATTGCG  
 GTTGATAAAAACTTCCGTGGTAAGCCAGAGCAGAACCCTCGCAGCAGTGCATCGGCTGCGCGCCCTGCGTCAA  
 TGCTGCCCCGTCAAACGCCTTAAACGGTTGAAACTGACCTCGCCACAGGAGAGCTTGCTGGGAGTTTAAATC  
 TTGGGCACTGCATCTTCTGTGGACGCTGCGAAGAAGTCTGCCCGACGGCGCGGATCAAACCTGTGCAAGAG

Reverse 4

TACGAACTGGCGGT**GTGGAAGAAAGAAGACTT**CCTGCAACAGTCCCGCTTCGCGCTGTGCAACTGCCGCGT  
 CTGCAATCGTCTTTTCGCCGTCCAGAAAGAGATCGACTACGCCATTGCGCTGCTTAAGCACAACGGCGACA  
 GCCGCGCGGAAAACCCGCGAAAGCTTTGAGACTTGCCCGGAATGTAAGCGCCAGAAATGCCTGGTGCCG  
 TCCGACCGTATTGAACTGACTCGCCATATGAAAGAGGCCATCTGATGAGCAATTTATTAGGCCCCCGTGAC  
 GCCAACGGCATTCCGGTCCCCATGACGGTGGATGAATCCATCGCCAGCATGAAGGCGTGGTTACTGAAAAA  
 AATCAAACGTTCTGCCTATGTTTACCGCGTGGACTGCGCGCGCTGCAACGGTTGCGAAATCGAAATTTTCG

Forward 9

CCACGCT**TTTCGCCGCTGTTTGATG**CAGAACGCTTCGGCATTAAAGTCGTTCCCTCACCGCGTCATGCGGAT  
 ATTTTACTGTTTACCGCGCGGTCACCCGTGCAATGCGATCCCTGCGCTGCGTGGCAGTCCGCGCC  
 GGACCCGAAAATTTGTATCTCCTACGGTGCCTGCGGTAAACAGTGGCGGGATCTCCACGATCTCTACTGCG  
 TGTGGGGCGGTACGGATAAAAATTGTCCCTGTGGATGTTTATATCCCTGGCTGCCCGCAAACGCCTGCCGCC  
 ACGCTGTACGGCTTTGCAATGGCGCTCGGCCTGCTGGAGCAGAAAATTCACGCCGCTGGGCCGGGTGAACT

Reverse 3

GGATGAACAACCGCGGAGATCCTGCATGGT**GATATGGTGCAGCCGCTGCG**CGTAAAGTGGATCGCGAAG  
 CACGTCGCCTGGCGGGTTATCGTTACGGTCGTGAGATTGCCGATGATTACCTTACACAGTTAGGGCAGGGC  
 GAAGAACAGGTTGCACGCTGGCTGGAAGCGGAAAACGATCCGCGTCTGAACGAGATTGTCAGCCATCTGAA  
 TCATGTTGTTGAAGAGCGCGTATCCGATGAGTGAAAAGTGGTGTTCAGTCAACTGAGCCGTAATTTTAT  
 TGATGAGAACGATGCCACGCCCGCGAGGCGCAGCAGTGGTCTATTACAGCCTGGCGATTGGTCACCACC  
 TTGGGGTTATCGATTGCCTGGAAGCGCGCTCACCTGCCCGTGGGATGAATATCTGGCATGGATTGCCACT

Forward 10

CTGGAGGCAGGCAGTGAAGCCCGCC**GCAAAATGGAAGCGTGC**CGAAATATGGTGAGATCGTCATCGACAT  
 TAACCATGTGCCGATGCTGGCCAACGCATTTCGATAAAGCCCGGCAGCGCAAACCTTCGCAGCAGCAGGAAT  
 GGAGTACAATGCTGTTAAGTATGCTGCATGATATTCATCAGGAAAACGCCATCTATTGATGGTGAGGAGA  
 CTGCGTGACTGACGTTTTACTCTGTGTTGGCAATAGCATGATGGGCGATGATGGCGCAGGTCCGCTGCTGG  
 CGAAAAGTGGCGCCCGCGCGGAAAAGGTAACCTGGTGGT**GATTGACGGCGGTAGCGC**ACCGGAAAACGAC

Reverse 2

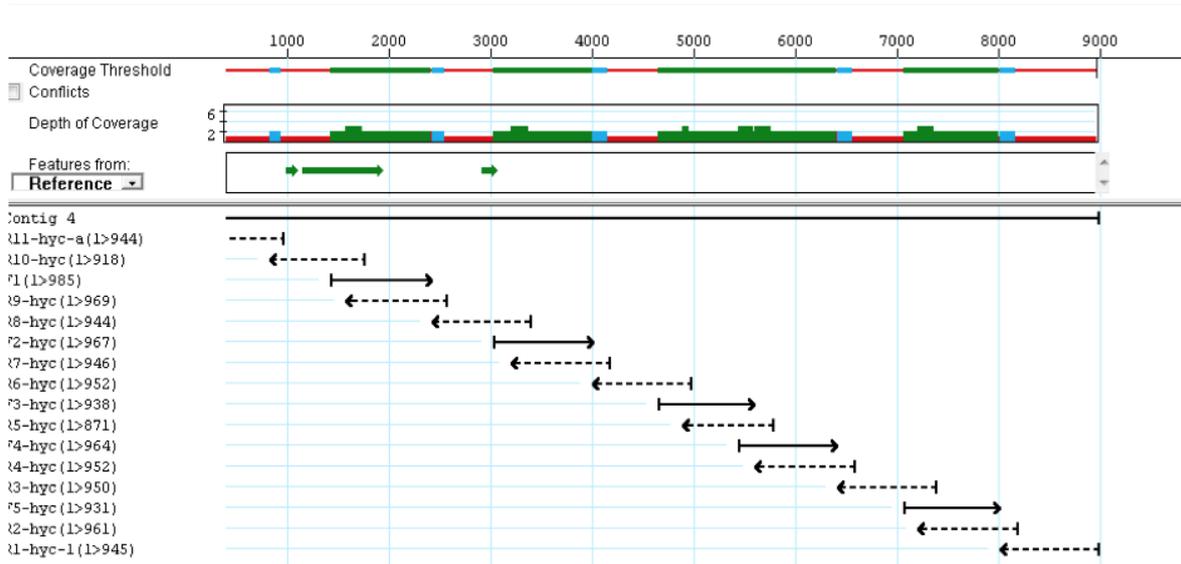
ATCGTCGCTATCCGTGAACTGCGCCCGACACGACTGCTGATTGTGACGCC**ACGGATATGGGGCTAAACCC**  
 CGCGGAGATCCGCATCATCGACCCGGATGATATCGCCGAGATGTTTATGATGACTACCCATAACATGCCGT  
 TGAATTACCTTATCGACCACTTGAAGAAGATATTGGCGAAGTGATTTTCCCTCGGCATTCAGCCGGATATC  
 GTCGGCTTTTACTACCCGATGACCCAGCCGATTAAGATGCGGTAGAAAACCGTTTTATCAAACGACTGGAAGG  
 CTGGGAAGGAAATGGCGGCTTCGCGCAGTTAGCGGTGGAAGAAGAG**TAC**TATGCCATGGTACCCGGGGATC  
 CTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATAC  
 AGATTAATCAGAACCGAGAAGCGGTCTGATAAAACAGAAATTTGCCTGGCGGCAGTAGCCGGTGGTCCCA  
 CCTGACCCCATGCCGAATCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCATGCGAG  
 AGTAGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAAGACTGGGCCCTTTCGTTTTATCTGT  
 TGTTTGTGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACG  
 GCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAAGTCCAGGCATCAAATTAAGCAGAAGGCCATCCTGA  
 CGGATGGCCTTTTTGCGTTTTCTACAAACTCTTTTGTATTTTCTAAATACATTCAAATATGTATCCGCT

Reverse 1

**CATGAGACAATAACCCGTG**ATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTA

10. Location of pBAD<sub>rham</sub>-*hyc* primers for sequence. pBADF, 10 forward primer and pBADR 11 reverse primers are bold and indicated by arrows. Azure highlight indicates Ribosomal binding site, yellow highlight indicate the start codon of *hyc* operon (*hycA*) and green highlight indicates stop codon of *hyc* operon (*hycI*).

## Appendix

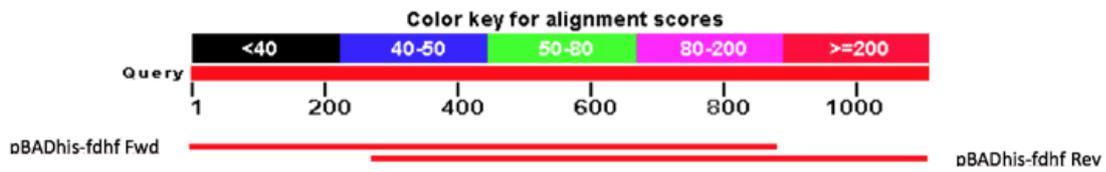


**11. Strategy view of *hyc* operon after sequencing.** Solid lines indicates forward primers that has been used in the sequence and dash lines indicates the reverse primers.

AAGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCCTGCGTCTTTTACTGGCTCTTC  
 TCGCTAACCAAACCGGTAACCCCGCTTATTAAGCATTCTGTAACAAAGCGGGACCAAAGC  
 CATGACAAAACCGGTAACAAAAGTGTCTATAATCACGGCAGAAAAGTCCACATTGATTATTT  
 GCACGGCGTCACACTTTGCTATGCCATAGCATTTTTATCCATAAGATTAGCGGATCCTACCT  
 GACGCTTTTTATCGCAACTCTCTACTGTTTTCTCCATACCCGTTTTTTGGGCTAACAG**GAGGA**  
**ATTAACCATGGATGAAAAAGTCGTCACGGTTGCCC**CCTATTGCGCATCAGGTTGCAA  
 AATCAACCTGGTCGTCGATAACGGCAAATCGTCCGGGCGGAGGCAGCGCAGGGGAA  
 AACCAACCAGGGTACCCTGTGTCTGAAGGGTTATTATGGCTGGGACTTCATTAACGA  
 TACCCAGATCCTGACCCCGCGCCTGAAAACCCCATGATCCGTCGCCAGCGTGGCGG  
 CAACTCGAACCTGTTTTCTGGGATGAGGCACTGAATTACGTTGCCGAGCGCCTGAG  
 CGCCATCAAAGAGAAGTACGGTCCGGATGCCATCCAGACGACCGGCTCCTCGCGTGG  
 TACGGGTAACGAAACCAACTATGTAATGCAAAAATTTGCGCGCGCCGTTATTGGTAC  
 CAATAACGTTGACTGCTGCGCTCGTGTCTGACACGGCCCATCGGTTGCAGGTCTGCA  
 CCAATCGGTCCGTAATGGCGCAATGAGCAATGCTATTAACGAAATTGATAATACCGA  
 TTTAGTGTTTCGTTTTTCGGGTACAACCCGGCGGATTCCCACCCAATCGTGGCGAATCA  
 CGTAATTAACGCTAAACGTAACGGGGCGAAAATTATCGTCTGCGATCCGCGCAAAAT  
 TGAAACCGCGCGCATTTGCTGACATGCACATTGCACTGAAAACGGCTCGAACATCGC  
 GCTGTTGAATGCGATGGGCCATGTCATTATTGAAGAAAATCTGTACGACAAAGCGTT  
 CGTCGCTTCACGTACAGAAGGCTTTGAAGAGTATCGTAAAATCGTTGAAGGCTACAC  
 GCCGGAGTCGGTTGAAGATATCACCGGCGTCAGCGCCAGTGAGATTTCGTCAGGCGGC  
 ACGGATGTATGCCAGGCGAAAAGCGCCGCCATCCTGTGGGGCATGGGTGTAACCCA  
 GTTCTACCAGGGCGTGGAAACCGTGCCTCTCTGACCAGCCTCGCGATGCTGACCGG  
 TAACCTCGGTAAGCCGCATGCGGGTGTAAACCCGGTTCGTGGTTCAGAACAAACGTTCA  
 GGGTGCCTGCGATATGGGCGCGCTGCCGGATACGTATCCGGGATACCA**GTACGTGAA**  
**AGATCCGGCTAAAAGCTTGGCTGTTTT**GGCGGATGAGAGAAGATTTTCAGCCTGATACAGA  
 TTAAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTG  
 GTCCCACCTGACCCCATGCCGAACTCAGAAGTAAAACGCCGTAGCGCCGATGGTAGTGTGGG  
 GTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAA

12. Location of pBAD<sub>his</sub>-*fdhF* primers used in infusion cloning. pBAD<sub>ara</sub> forward primer and pBAD reverse primers are bold and indicated by arrows. Azure highlight indicates Ribosomal binding site, yellow highlight indicate the start codon of *fdhF* and green highlight indicates stop codon of *fdhF*.

## Appendix



**13. BLAST view of *fdhf* gene after sequencing.** Solid lines indicates forward and reverse primers that has been used in the sequence.