CHARACTERISATION OF A NOVEL NITROGEN-FIXING, POLYHYDROXYALKANOATE-PRODUCING BACTERIUM, *NOVOSPHINGOBIUM NITROGENIFIGENS* Y88^T

By

Anne-Marie Smit

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy
in Cell and Molecular BioSciences

Victoria University of Wellington
2008

Abstract

The novel sphingomonad *Novosphingobium nitrogenifigens* Y88^T (Y88^T) is an obligate aerobe able to grow in nutrient-imbalanced environments where nitrogen is naturally limiting, but carbon is found in abundance. Due to its ability to fix atmospheric nitrogen and produce the bioplastic polyhydroxyalkanoate (PHA), Y88^T is well-suited for growth in a nitrogen-limited but carbon-enriched environment. Because of these metabolic abilities, Y88^T is of interest as a model organism for PHA production unconstrained by nitrogen-limiting conditions. Growth profiles and PHA production profiles were determined for Y88^T under conditions of carbon enrichment, nitrogen sufficiency and depletion to investigate carbon and nitrogen utilisation as well as PHA production in this organism. Also, since the nitrogenase enzyme required for nitrogen fixation is oxygen labile, the effect of DO concentration and the relationship between aerobic metabolism and the nitrogen-fixing and PHA-producing abilities of Y88^T was investigated. This study demonstrated: that glucose is the preferred growth substrate for Y88^T; that no direct relationship exists between nitrogen fixation and PHB accumulation in Y88^T; that Y88^T can reliably produce in excess of 80 % of its dry weight as polyhydroxybutyrate (PHB), a type of PHA, from glucose under nitrogen-limiting conditions. Proteomic signatures were determined for the various physiological responses of Y88^T to growth, nitrogen utilisation, PHB production and exposure to different levels of DO. More than 250 unique proteins, including the core nitrogen-fixation, PHB-synthetic and glycolytic proteins were identified. Y88^T apparently converts glucose to PHB via three interrelated glucose catabolic pathways and proteins likely involved in these pathways were identified. This study revealed that, regardless of growth conditions and despite decreased abundance of the Y88^T nitrogenase enzyme, growth and PHB synthesis were not inhibited at DO^{high} concentrations. Proteomic characterisation of the Y88^T phasin, a PHA granule-associated protein,

identified an amino-terminal, low complexity alanine and proline rich segment found only in other sphingomonads. The expression level of the Y88^T phasin correlated well with PHB yields, suggesting the use of this protein as a biomarker to optimise PHB yield in a production environment. Y88^T has the potential to be a useful production strain in pure culture, utilising its natural and robust propensity to metabolise glucose to preferentially produce PHB. Targets for biotechnological improvement and the potential for application of Y88^T to biofuel production are discussed.

Acknowledgements

In Africa, we have a proverb: *"Not everyone who chased the zebra caught it, but he who caught it chased it."* In other words, many chase after their dreams but they do so only half-heartedly. Those who "caught the zebra," stayed focused and let nothing divert them from their path until they had achieved their goal.

My supervisor at Scion, Gareth Lloyd-Jones, reminded me of this proverb several months ago when I was finding it a somewhat daunting task to overcome the writer's block that precipitates the final leap of faith that allows one to overcome the mental anguish of completing a PhD write-up. At that time, it was becoming increasingly obvious that my first draft was anything but complete. Suffice it to say, a PhD is ALL it is made out to be! To sum it up in the understated words of Bill Jordan, my supervisor at VUW: "Writing a thesis is a terrible business." I can now confirm this with utter conviction. Therefore, it is only appropriate at this point to thank the many people who turned this "terrible business" into a somewhat "bearable business".

First and foremost I would like to thank my mentor, Tim Strabala, who did this all a long time before I did and whose pearls of wisdom, support and endless discussions on the biochemical aspects of my project constantly benefited me throughout my intrepid journey. Without his willingness to part with his prized copy of "Zubay" for endless periods of time, and his constant belief in my ability to conquer any self-doubt, I really do doubt that I would be anywhere near journey's end. I have to admit, being my partner-in-crime on the daily coffee-run, did wonders for my morale if not my bank balance! Also, without his wonderful foresight in obtaining a nearly completed Y88^T genome sequence, my task would have been that much more difficult.

Next I would like to thank both Gareth and Bill for their guidance and input into my project, each in their own field of expertise. I am most grateful for their valuable advice and encouragement, which was particularly useful in the writing-up stage. Their continued support is highly appreciated.

A special thanks to Nicki Reid who was instrumental in selecting Y88^{T} as a potential organism of interest long before I arrived at Scion. Without her perseverance in the "old" days together with the work of Sarah Addison who isolated Y88^{T}, there would have been no Y88^{T} project to begin with. Also, sharing an office and many lighthearted moments have been a welcome respite from some very "terrible" PhD "business"!

Furthermore, special thanks must go to the following people who each in their own way, contributed to some aspect of my project and to whom I am most grateful: Pisana Rawson for her technical expertise in 2-D DIGE and running some of the repeat gels at a time when her own workload was heavy; Lifeng Peng for her willingness to run many of my samples on the mass spec and making the time in her own busy schedule to do so; Danyl McLauchlan for enabling me to have access to the FTP server at Victoria University of Wellington (VUW) so that I could access my files remotely as well as converting files for me when required; Shaun Taylor, Daniel van de Pas and Sheree Anderson who each inherited the PHA analysis job at Scion at different times and who were always willing to carry out any PHA analysis when required; the ladies at Veritec, Kaye Eason and Carli van Zyl for some of the carbon and nitrogen analyses; Stefan Hill (Scion) and Michael Schmitz (Auckland University) for the ^{13}C-NMR analysis; Armin Thumm for the GPC analysis; Sarah Addison for maintaining requisite lab supplies; Lucy Macdonald for helping me with the Y88^{T} Mascot database; Lucy and Vincent Liu for their great work on the Y88^{T} sequence annotations that expedited my protein discoveries carried out at Scion; Lloyd Donaldson for the Y88^{T} TEM photos and the team at

John Morris Scientific who provided technical support for the Bioflo110 during installation and subsequent replacement of probes and control units.

Financial support for this project including travel to and from Scion and VUW as well as attendance at conferences, were provided by Scion and VUW. Many thanks especially to Trevor Stuthridge and Bill Jordan who were instrumental in making this possible.

Thank you to Dad and Mom for all your love, guidance and support always. I treasure you more than you could possibly know. You truly are the best parents anyone could wish for. Last but never least, thank you to Bronwyn, Tarryn and Travis, my amazing children, for their enduring love, support and encouragement. They were as committed to me completing my PhD as I was, and I consider myself the most fortunate mother in the world to have such wonderful children as they have been, and continue to be. Bron, Taz, Trav - you make me so proud and you make everything totally worthwhile. I know you will be happy to know that, finally, I think I may have caught my zebra.

Table of Contents

Abstract ..... **ii**

Acknowledgements ..... **iv**

Table of Contents ..... **vii**

List of Tables ..... **xiii**

List of Figures ..... **xv**

Abbreviations ..... **xix**

Chapter 1 : Introduction **1**

1.1 Overview *1*

1.2 Sphingomonads *2*

1.2.1 *Novosphingobium nitrogenifigens* Y88^T sp. nov.: a newly characterised sphingomonad as a model organism for low nitrogen waste stream remediation 5

1.3 Bacterial utilisation of nitrogen and carbon *6*

1.3.1 Biological nitrogen fixation 6

1.3.2 Ammonium assimilation 9

1.3.3 PHA biosynthesis and bioplastics 13

1.3.3.1 Overview and commercial application 13

1.3.3.2 Bacterial PHA and the biosynthetic pathway 14

1.3.4 Carbon utilisation and partitioning 20

1.3.4.1 Central metabolism and its relationship to PHA biosynthesis 20

1.3.4.2 Carbon flux and the regulation of PHA biosynthesis 21

1.4 Proteomic approaches for protein expression profiling *23*

1.5 Aims *28*

Chapter 2 : Materials and Methods **30**

2.1 Microbiology *30*

2.1.1 Description of Y88^T 30

2.1.2 Other bacterial isolates 31

2.1.2.1 *Azotobacter vinelandii* ICMP4036, *Pseudomonas putida* mt2 and *Sphingomonas sp.* WP01 (DSM 19371)....31
2.1.3 Growth conditions....31
2.1.3.1 Cell stocks, plate inoculation and pre-culture inocula....31
2.1.3.2 Shaker flask cultivation....32
2.1.3.3 Bioreactor cultivation....34
2.1.4 Optical density as an indicator of growth....36
2.1.4.1 Batch flask cultivation....36
2.1.4.2 Batch bioreactor cultivation....36
2.1.5 Dry weights as an indicator of biomass....36
2.1.6 Nitrogen utilisation assay....37
2.1.6.1 Determination of ammonium assimilation....37
2.1.7 Acetylene reduction assay....38
2.1.8 Glucose utilisation....40
2.1.8.1 $Y88^T$ growth on glucose and acetate....40
2.1.8.2 Comparative growth of $Y88^T$, *A. vinelandii* and *P. putida* mt2 on glucose under ammonium-free, ammonium-supplemented or ammonium-depleted conditions....41
2.1.8.3 Comparative growth of $Y88^T$ and WP01 during PHA accumulation in the presence of a single carbon substrate....42
2.1.9 Total organic carbon as a measure of dissolved organic carbon....42
2.1.10 Enzyme assays....43
2.1.10.1 Glucose utilisation assay....43
2.1.10.2 Acetic acid formation assay....43
2.1.11 PHA determination....44
2.1.11.1 PHA content in $Y88^T$ cells....44
2.1.11.2 Determination of polymer characteristics....44
2.2 $Y88^T$ whole genome sequencing and curation....44
2.3 Proteomics....45
2.3.1 Generation of samples for 2-D DIGE and LC-MS/MS analysis....45
2.3.1.1 2-D Gel Experimental Design 1....45
2.3.1.2 2-D Gel Experimental Design 2....49
2.3.1.3 2-D Gel Experimental Design 3....51
2.3.1.4 2-D Gel Experimental Design 4....51
2.3.2 Sample preparation....54
2.3.3 Protein extraction....55
2.3.4 Protein quantitation....56
2.3.4.1 BCA protein assay....56
2.3.5 Gel electrophoresis....56
2.3.5.1 Assessment of protein by 1-D SDS-PAGE....56
2.3.5.2 Gel washing, staining and destaining of 1-D gels....57

2.3.5.3 2-DE 58
2.3.5.3.1 Preparative gel sample preparation 58
2.3.5.3.2 Sample rehydration and loading onto IPG strip 58
2.3.5.4 2-D DIGE 58
2.3.5.4.1 Sample preparation 58
2.3.5.4.2 Sample labeling 59
2.3.5.4.3 Sample mixing and loading onto IPG strip 59
2.3.5.4.4 Sample rehydration 60
2.3.5.4.5 First dimension IEF 60
2.3.5.4.6 Protein Reduction 60
2.3.5.4.7 Protein Alkylation 62
2.3.5.4.8 Second dimension protein separation 62
2.3.6 Gel image acquisition 62
2.3.6.1 Gel Image analysis 63
2.3.7 Staining 2-D DIGE gels 64
2.3.8 Protein spot picking 64
2.3.9 Protein preparation for LC-MS/MS analysis 64
2.3.9.1 Destaining of excised gel spots 64
2.3.9.2 In-gel reduction and alkylation of protein prior to LC-MS/MS analysis 65
2.3.9.3 In-gel enzymatic digestion of protein 65
2.3.9.4 Extraction of peptides following protein digestion 66
2.3.9.5 C_{18} membrane binding of peptides for reversed phase LC-MS/MS 66
2.3.10 LC-MS/MS 67
2.3.10.1 Instrumentation 67
2.3.10.2 Data analysis and protein identification 68

Chapter 3 : Microbiology of Y88^T 70

3.1 Introduction 70

3.2 Results 70
3.2.1 Y88^T cultivation in batch flasks 70
3.2.1.1 Growth in a glucose-enriched minimal nitrogen medium 70
3.2.1.1.1 Modeling Y88^T growth to the Gompertz model 70
3.2.1.1.2 Biomass produced by Y88^T 71
3.2.1.2 Y88^T growth and carbon substrate utilisation 77
3.2.1.2.1 Single substrate utilisation: glucose vs. acetate 77
3.2.1.2.2 Mixed substrate utilisation: glucose and acetate 79
3.2.1.2.3 Y88^T growth in high and low glucose-enriched minimal nitrogen medium 81
3.2.1.2.4 Composition of the PHB polymer produced by Y88^T from glucose 84
3.2.2 Comparison between Y88^T, *A. vinelandii* and *P. putida* mt2 86

3.2.2.1 Growth and nitrogen fixation in NH_4^+-free and NH_4^+-supplemented glucose-enriched media 86
3.2.2.2 Growth, biomass and PHB production in NH_4^+-supplemented media 92
3.2.2.2.1 Acetylene reduction assay 97
3.2.2.3 Comparison between $Y88^T$ and WP01 98
3.2.2.3.1 Growth and PHB production 98
3.2.2.4 Carbon and nitrogen utilisation by $Y88^T$ 103
3.2.2.4.1 Glucose assay results 103
3.2.2.4.2 Total dissolved organic carbon (DOC) in spent medium 103
3.2.2.4.3 Ammonium assimilation 106
3.2.2.4.4 Change in carbon to nitrogen (C:N) ratio. 109
3.2.2.4.5 Acetic acid assay 111
3.2.2.4.6 Conversion of glucose to PHB 111

Chapter 4 : "Proteomic signatures" for the Fix^{on}/Fix^{off} physiology of $Y88^T$ during PHA synthesis 115

4.1 Introduction 115

4.2 Results 116
4.2.1 1DE 116
4.2.2 2-D/2-D DIGE 118
4.2.2.1 Differential protein abundance between the Fix^{on} and Fix^{off} physiologies of $Y88^T$ during PHA accumulation 118
4.2.2.2 Differential protein abundance at different optical densities for the Fix^{on} and Fix^{off} physiologies of $Y88^T$ during PHA accumulation 122
4.2.2.2.1 Nitrogen-fixing and associated proteins 125
4.2.2.2.2 PHA synthesis and granule-associated proteins 134
4.2.2.2.3 Stress-related proteins 135
4.2.2.2.4 Other proteins 138

4.3 Summary 142

Chapter 5 : The paradox of $Y88^T$ growth at high oxygen concentration 144

5.1 Introduction 144

5.2 Results 144
5.2.1 Cultivation of $Y88^T$ in a bioreactor at DO^{low} or DO^{high} concentrations 144
5.2.1.1 Growth and nitrogen utilisation 144
5.2.1.2 PHB production and glucose utilisation 150
5.2.1.3 Comparison of protein profiles obtained for DO^{low} or DO^{high} conditions at high cell densities (OD_{600}=2.0) 154
5.2.1.3.1 Acidic proteins 154

5.2.1.3.1.1 Nitrogen-fixing proteins 157
5.2.1.3.1.2 PHA-accumulating proteins 166
5.2.1.3.1.3 Other proteins 167
5.2.1.3.2 Basic proteins 173
5.2.1.3.2.1 Nitrogen fixation proteins 175
5.2.1.3.2.2 PHB synthesis proteins 175
5.2.1.3.2.3 Other 177
5.2.1.4 Comparison of protein profiles obtained for DO^{low} or DO^{high} conditions at high cell densities (OD_{600}=2.0), pH controlled 177
5.2.1.4.1 Acidic proteins 177
5.2.1.4.1.1 Linear gradient pH 4-7 range 178
5.2.1.4.1.2 Nitrogen fixation proteins 181
5.2.1.4.1.3 Nitrogen assimilation 186
5.2.1.4.2 Basic proteins 189
5.2.1.4.2.1 Nitrogen-fixing proteins 189
5.2.1.4.2.2 PHB synthesis proteins 189
5.2.1.4.2.3 Other proteins 189
5.2.1.5 Comparison of protein profiles obtained for DO^{low} or DO^{high} conditions at low cell densities (OD_{600}=0.5), pH controlled 192
5.2.1.5.1 Acidic proteins 192
5.2.1.5.2 Basic proteins 194
5.3 Summary 198

Chapter 6 : Bioinformatic analysis of post-translational modifications and primary and secondary structure of selected Y88^T nitrogenase and PHB synthesis proteins 200

6.1 Introduction 200
6.2 Results 201
6.2.1 Amino acid sequence analysis of the nitrogenase protein of Y88^T 201
6.2.1.1 NifH 201
6.2.1.2 NifK and NifD proteins 210
6.2.2 Protein and peptide analysis of the Y88^T phasin 218
6.2.2.1 Proteomics of the three Y88^T phasin isoforms 218
6.2.2.2 Phasin protein sequence alignments 222

Chapter 7 : Summary and Discussion 231

7.1 Y88^T: Microbiology, Nitrogen Fixation and Carbon Utilisation 231
7.2 Y88^T growth profiles 232
7.3 Nitrogen metabolism 235
7.3.1 Assimilation of ammonium 235
7.3.2 Nitrogen fixation in Y88^T 237

7.4 Central metabolism, energy and cellular redox state 239

7.5 PHB production 243

7.6 Amino acid and nucleic acid metabolism proteins 247

7.7 Outer membrane proteins 247

7.8 Other 249

7.9 Future prospects 249

Chapter 8 Appendices 253

8.1 Appendix A Additional proteins (following six pages) not all discussed in chapters 4-7 identified by Mascot-based searches of LC-MS/MS data. 253

8.2 Figure A.1 A representative 2-D preparative gel showing additional proteins identified in the pH 4-7 linear range and 6-100 kDa range using LC-MS/MS and not all discussed in chapters 4-7... 260

8.3 Appendix B Additional protein identifications not all discussed in Chapters 4-7. 261

8.4 Figure B.1. A representative 2-D DIGE gel showing additional proteins identified in the pH 6-11 linear range using LC-MS/MS and not all discussed in chapters 4-7. 263

8.5 Appendix C Additional protein identifications not all discussedin Chapters 4-7. 264

References 266

List of Tables

Table 2.1. Growth media and corresponding cell physiological state with regard to nitrogen fixation defined for Y88^{T} growth conditions. 33

Table 2.2. Physiological state of Y88^{T} cells defined for a specific set of growth conditions. 47

Table 2.3. Protocols for first dimension IEF focusing of protein 61

Table 3.1 Gompertz equations for Y88^{T} replicate growth curves 73

Table 3.2 Dry biomass (cumulative over 72 hrs), actual change in dry biomass (for each time interval separately), rate of change in dry biomass produced (for each time interval separately) and yield of biomass by Y88^{T} in a glucose-enriched minimal nitrogen medium. 75

Table 3.3. Glucose consumed by Y88^{T} during 72 hr growth in a glucose-enriched minimal nitrogen medium 113

Table 4.1 Y88^{T} proteins showing altered abundance in for the Fix^{off} state (OD_{600} of 0.5 and 1.0) compared to the Fix^{on} state (OD_{600} of 2.0), identified using LC-MS/MS 127

Table 4.2 LC-MS/MS peptide analysis of Y88^{T} proteins showing differential abundance for the Fix^{off} state (OD_{600} of 0.5 and 1.0) compared to the Fix^{on} state (OD_{600} of 2.0) 129

Table 5.1 Carbon and nitrogen content of Y88^{T} cells grown in glucose-enriched, nitrogen-limiting medium 148

Table 5.2 Oxygen input to meet the demand of Y88^{T} cells and maintain set DO levels constant during Y88^{T} growth in a glucose-enriched, nitrogen-fixing medium in bioreactors 155

Table 5.3 2-D DIGE protein spot analysis (pH 4-7) and LC-MS/MS analysis of protein from cells grown at DO^{low} or DO^{high} 159

Table 5.4 2-D DIGE protein spot analysis showing statistical values and LC-MS/MS identification of proteins with altered abundance at high OD600 (2.0), and pH not controlled at DOlow or 176

Table 5.5 2-D DIGE protein spot analysis (pH 4-7) and LC-MS/MS analysis of protein from cells grown at DO^{low} or DO^{high} at constant pH and high OD_{600} (2.0) 183

Table 5.6 2-D DIGE protein spot analysis (pH 6-11) and LC-MS/MS analysis of protein from cells grown at DO^{low} or DO^{high} at constant pH 191

Table 5.7 2-D DIGE spot analysis (pI 4-7) and LC-MS/MS of Y88^{T} protein from cells grown to a low OD_{600} under ammonium-depleted conditions at DO^{low} or DO^{high} at constant pH 7 195

Table 5.8 2-D DIGE protein spot analysis (pI 6-11) and LC-MS/MS analysis of Y88^{T} protein from cells grown to a low OD_{600} under ammonium-depleted conditions at DO^{low} or DO^{high} at constant pH 7 197

Table 6.1 Bacterial species names and the corresponding abbreviated names in the NifH alignment with Y88^T (Fig. 6.1).....202
Table 6.2 Bacterial species names and the corresponding truncated names in the order of appearance in the list of phasin sequences in the alignment with Y88^T (Fig. 6.2).....223
Table 7.1 Y88^T enzymes involved in protein and nucleic acid synthesis that showed unaltered states irrespective of the growth conditions. 248

List of Figures

Figure 1.1 A schematic of the energy-demanding reaction catalysed by the nitrogenase complex. 8
Figure 1.2 Schematic representation of adenylylation/deadenylylation of GS in response to ammonium status. 12
Figure 1.3 General structure of PHAs 16
Figure 1.4. Biosynthetic pathway for P(3HB) 19
Figure 1.5 Schematic of NADH and NADPH production and utilisation via catabolic vs. anabolic pathways 24
Figure 2.1. Schematic representation of workflow to investigate the microbiology of Y88^T growth, nitrogen assimilation, carbon utilisation and PHA accumulation under different growth conditions. 35
Figure 2.2. Schematic representation of workflow to investigate the proteome of Y88^T under different growth conditions. 46
Figure 2.3. 2-D gel experimental design 1 48
Figure 2.4. 2-D gel experimental design 2 50
Figure 2.5. 2-D gel experimental design 3 52
Figure 2.6. 2-D gel experimental design 4 53
Figure 3.1 Y88^T growth data fitted to the Gompertz model 72
Figure 3.2 OD_{600} and dry biomass produced by Y88^T in a glucose-enriched minimal nitrogen medium over 72 hr 74
Figure 3.3 Plot of Y88^T biomass versus OD_{600} 76
Figure 3.4. Total dry biomass and PHB mass produced by Y88^T after 24 hr growth in a minimal nitrogen medium with various carbon substrates 78
Figure 3.5 Comparative growth of Y88^T in high and low glucose minimal nitrogen media. 82
Figure 3.6 PHB and dry biomass produced by Y88^T in high and low glucose minimal nitrogen medium 83
Figure 3.7 ^{13}C-NMR spectrum of PHB produced by Y88^T. 85
Figure 3.8 Transmission electron micrograph (TEM) of Y88^T cells containing PHB accumulated up to 48 hr growth. 87
Figure 3.9 Y88^T growth in minimal nitrogen and nitrogen-free media. 88
Figure 3.10 Ammonium utilisation by Y88^T 90
Figure 3.11 Comparative growth of Y88^T, *A. vinelandii*, and *P. putida* mt2 91

Figure 3.12 Comparative PHB mass and total dry biomass production by Y88^T and *A. vinelandii* 93
Figure 3.13 Comparative PHB production by Y88^T and *A. vinelandii* at 24, 48 and 72 hr. 94
Figure 3.14 Acetylene reduction assays of Y88^T and *A. vinelandii* under various growth conditions 95
Figure 3.15 Growth of Y88^T and WP01 in a glucose-enriched, acetate-enriched and carbon-free medium 100
Figure 3.16 PHB and biomass produced by Y88^T and WP01 grown for 24 hr in glucose-enriched nitrogen-fixing medium, acetate-enriched nitrogen-fixing medium, acetate-enriched 1x BSM and acetate-enriched variant 1x BSM with no nitrogen source 101
Figure 3.17 Glucose depletion by Y88^T 104
Figure 3.18 PHA production by Y88^T during glucose depletion 105
Figure 3.19 Total dissolved organic carbon (DOC) change during Y88^T growth 107
Figure 3.20 Change in C:N ratio during ammonium assimilation by Y88^T 108
Figure 3.21 Y88^T glucose depletion: change in C:N ratio and rate of glucose depletion 110
Figure 3.22 Schematic representation of glucose conversion to PHB and ammonium assimilation by Y88^T 112
Figure 4.1 Coomassie-stained 1-DE gel showing comparison of Y88^T total protein extracts 117
Figure 4.2.Total protein gel of Y88^T at various stages of culture growth 119
Figure 4.3 Overlay image of Y88^T protein profiles generated from minimally labeled pooled Fix^{on} and Fix^{off} samples 121
Figure 4.4 2-D DIGE profiles of Y88^T cells for the Fix^{off} physiological state of the cells (A and B) 123
Figure 4.5. 2-D DIGE profiles of Y88^T cells for the Fix^{off} (A) and Fix^{on} (B) physiological state of the cells 124
Figure 5.1 OD_{600}, biomass production and nitrogenase assays of Y88^T grown at DO^{low} or DO^{high} for 24 hrs in a glucose-enriched, minimal nitrogen medium 146
Figure 5.2 Total nitrogen (Kjeldahl) determined in the spent medium after removal of Y88^T cells grown at DO^{low} or DO^{high} 147
Figure 5.3 Change in total biomass nitrogen content with change in biomass during growth in nitrogen-depleted medium DO^{low} (maroon bars) and DO^{high} (blue bars) during 24 hr growth. Error bars represent standard deviations, (n=2). 149
Figure 5.4 PHB produced by Y88^T at DO^{low} and DO^{high} during changing C:N ratios 151
Figure 5.5 Glucose utilisation by Y88^T and change in C:N ratio 152
Figure 5.6 Glucose consumption by Y88^T showing inefficiency of glucose utilisation at DO^{high} 153
Figure 5.7 Overlay image of Y88^T protein profiles generated from minimally labeled samples showing differential protein abundance under DO^{low} or DO^{high}, uncontrolled pH conditions in the linear gradient pH 4-7 range 156

Figure 5.8 2-D DIGE images (pI 4-7) showing protein profiles of Y88^{T} cells grown to high OD_{600} (2.0) under ammonium-depleted conditions at DO^{low} or DO^{high} concentrations, pH not controlled. 158
Figure 5.9 Magnification of the protein spots from 2-D DIGE pH 4-7 gels that showed altered abundance between DO^{low} and DO^{high} under constant pH 7 conditions in the growth medium. 169
Figure 5.10 2-D DIGE images (pI 6-11) showing protein profiles of Y88^{T} cells grown to high OD_{600} (2.0) at DO^{low} or DO^{high} concentrations, pH not controlled..... 174
Figure 5.11 Overlay image of Y88^{T} protein profiles generated from minimally labeled samples showing differential protein abundance under DO^{low} or DO^{high}, constant pH 7 conditions in the linear gradient pH 4-7 range 179
Figure 5.12. 2D DIGE images (pI 4-7) showing protein profiles of Y88^{T} cells grown to a high OD_{600} (2.0) under ammonium-depleted conditions at DO^{low} or DO^{high} at constant pH..... 180
Figure 5.13 Magnification of the protein spots from 2-D DIGE pH 4-7 gels that showed altered abundance between DO^{low} and DO^{high} under constant pH 7 conditions in the growth medium. Arrows denote increased abundance for the protein under DO^{low} or DO^{high} conditions..... 182
Figure 5.14 2-D DIGE images for the linear gradient pI 6-11 range showing protein profiles of Y88^{T} cells grown to a high OD_{600} at DO^{low} or DO^{high} at constant pH 7..... 190
Figure 5.15 2-D DIGE images (pI 4-7) showing protein profiles of Y88^{T} cells grown to a low OD_{600} under ammonium-depleted conditions at DO^{low} or DO^{high} at constant pH 7 193
Figure 5.16 2-D DIGE images for the linear gradient pI 6-11 range showing protein profiles of Y88^{T} cells grown to a high OD_{600} under DO^{low} or DO^{high} conditions at constant pH 7..... 196
Figure 6.1 CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} NifH protein sequence with 20 NifH sequences from other bacteria (Table 6.1) 203
Figure 6.2 CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} NifD and NifK protein sequences with the *A. vinelandii* NifD and NifK sequences..... 211
Figure 6.3 CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} phasin protein with representative phasin sequences randomly selected from Genbank (Table 6.2) 219
Figure 6.4 Neighbour-joining bootstrap analysis of the Y88^{T} phasin with 44 phasins from Genbank. 226
Figure 6.5 Secondary structure predictions for the Y88^{T} and *R. eutropha* phasins. 227
Figure 6.6 Helical wheel predictions of the Y88^{T} phasin 228
Figure 6.7 Helical wheel predictions of the *R. eutropha* phasin..... 229
Figure 7.1 Schematic of proposed carbon and reducing equivalent pathways relating to central metabolism, nitrogen fixation and PHB accumulation in Y88^{T}. 241

Figure 7.2 Schematic of proposed PHB pathway in Y88T including possible alternative routes to PHB production, based on proteomic identification of enzymes (red) involved in the relevant pathways. 246

Abbreviations

2-DE	two-dimensional electrophoresis
2-D DIGE	two-dimensional difference in-gel electrophoresis
ACN	acetonitrile
BVA	Biological Variation Analysis
DO,	dissolved oxygen
DO^{low}, DO^{high}	low dissolved oxygen, high dissolved oxygen
DTT	dithiothreitol
ED	Entner-Doudoroff pathway
EM	Embden-Meyerhof pathway
Fix^{off}	non-nitrogen-fixing physiological state
Fix^{on}	nitrogen-fixing physiological state
GC	gas chromatography
GDH	glutamate dehydrogenase
GOGAT	glutamine oxoglutarate aminotransferase
GS	glutamine synthetase
IEF	isoelectric focusing
LC-MS/MS	liquid chromatography tandem mass spectrometry
MS	mass spectrometry
MQ	MilliQ
OPP	oxidative pentose phosphate pathway
ORF	open reading frame
PHA	polyhydroxyalkanoate
PHB	poly(3-hydroxybutyric) acid
SDS	sodium dodecylsulfate
SDS-PAGE	sodium dodecylsulfate-polyacrylamide gel electrophoresis
TCA	tricarboxylic acid cycle
TFA	trifluoroacetic acid

Chapter 1: Introduction

1.1 Overview

Effluent from pulp and paper mills is a carbon-rich, nitrogen-poor material which must be remediated prior to its release into the environment. Currently, the remediation process requires the addition of nitrogen to allow naturally occurring bacteria to maximally metabolise and detoxify the effluent. Nitrogen supplementation adds significantly to the cost of the remediation process and presents environmental difficulties of its own, such as eutrophication of lakes that receive much more nitrogen than they require. Researchers at Scion (Rotorua, New Zealand), are exploring the application of bacteria capable of fixing atmospheric nitrogen via biochemical reduction to ammonium ion as a potential lower-cost alternative to nitrogen supplementation in the bioremediation process (Slade et al., 2004). Combining nitrogen fixation with other bacterial metabolic capabilities to convert the waste carbon into useful raw materials would be a highly desirable further enhancement to the process.

Global warming, increased energy demand and increased petroleum costs have been major drivers toward the search for sustainably produced, biologically based, biodegradable plastics as substitutes for petroleum-based plastics. To this end, many microbes have been identified that naturally accumulate carbon in the form of PHA, a class of storage polymers (metabolically equivalent to fats in animals) that has properties similar to conventional plastics, but is completely biodegradable. Despite such promise, attempts to commercially produce PHAs as substitutes for petroleum-derived polymers have been challenging and considerable effort has focused on producing greater amounts of PHA at reduced cost. Thus, there is a great potential in utilising industrial waste streams as low-cost sources for PHA production.

To study the possibility of further enhancement of the value proposition of pulp mill waste remediation from cost-neutrality to profitability, Scion has isolated

and identified nitrogen-fixing bacteria that also naturally accumulate PHA from remediating pulp mill effluent. Through such bacteria, much of the remediable carbon could be converted to PHA, which could be harvested and used to make biodegradable consumer products. Both nitrogen fixation and PHA synthesis are energy-intensive and require significant amounts of reducing equivalents, thus competing for intracellular resources. The emphasis of this thesis has been on understanding the metabolic dynamics between nitrogen fixation and PHA accumulation to aid in identifying bacteria from pulp and paper wastewater that carry out these activities in an optimal manner. To conduct these studies, a model organism was required.

Y88^T, a novel Gram-negative bacterium was isolated from a mixed culture bioreactor at Scion. DNA sequence analysis has shown this strain to possess the *nif*H$^+$ gene required for nitrogen fixation, and the strain can produce a positive C_2H_2 reduction assay, both of which are indicative of an ability to fix nitrogen. Furthermore, this strain possesses the *pha*C$^+$ gene required for PHA synthesis. The combined ability of Y88^T to fix nitrogen and synthesise PHA makes it particularly suitable as a model organism to develop novel and potentially less costly industrial applications. This study aims to elucidate underlying pathways of PHA synthesis and nitrogen fixation in Y88^T, using microbiological and proteomic techniques to identify proteins and enzymes differentially altered under specific opposing growth conditions to clarify the relationship between nitrogen fixation and PHA accumulation in this species.

1.2 Sphingomonads

Sphingomonads are a group of organisms belonging to the α-4 subclass of the proteobacteria (Kawahara et al., 1999; Yabuuchi et al., 2002). The nomenclature of the commonly used appellation, α-proteobacteria, as proposed by Garrity et al., (2005) is Alphaproteobacteria and has subsequently been validated as such (Euzéby, 2006). An extensive body of literature describes the phylogenetic affiliation of sphingomonads to closely related genera based on 16S rDNA sequence analysis as well as structural and morphological distinctions

(Takeuchi et al., 1994; Takeuchi et al., 2001; Yabuuchi et al., 2002; Pal et al., 2006; Yang et al., 2006). The type species *Sphingomonas paucimobilis* was previously considered to be a member of the genus *Pseudomonas*, reflecting some of the confusion that has prevailed in the classification of this group of Gram-negative bacteria. Sphingomonads have been shown to have similar properties to phylogenetically closely related taxa such as *Zymomonas* and *Erythrobacter* (Kosako et al., 2000), but on the other hand they have unique traits such the ability to utilise homospermidine as the major polyamine that separate them from other Gram-negative bacteria (Busse et al., 2005). Members of these genera contain some unusual structural features such as glycosphingolipids (Kawahara et al., 1999; Fujii et al., 2003; Gupta & Mok, 2007) instead of the signature lipopolysaccharide profile characteristic of Gram-negative bacteria (Sriram et al., 2005) and possess ubiquinone 10 as their major respiratory quinone (Takeuchi et al., 2001). At present, sphingomonads are considered a homogeneous group of five phylogenetically distinct genera within the Alphaproteobacteria, viz., *Sphingomonas sensu stricto, Sphingobium, Sphingosinicella, Sphingopyxis* and *Novosphingobium* (Takeuchi et al., 2001). Historically, the genus *Novosphingobium* has been controversially classified in the genus *Sphingomonas* (Yabuuchi et al., 2002), but more recent phylogenetic and chemotaxonomic studies have confirmed that *Novosphingobium* is appropriately considered a separate genus (Takeuchi et al., 2001; Kampfer et al., 2002; Liu et al., 2005; Tiirola et al., 2005). The genus *Novosphingobium* comprises a diverse group of bacteria that now includes nineteen species with the most recent additions of *N. naphthalenivorans* (Suzuki & Hiraishi, 2007), *N. nitrogenifigens* (Addison et al., 2007) and *N. acidiphilum* (Glaeser et al., 2009).

Sphingomonads display diverse ecological as well as physiological properties such as heteropolymer assimilation and biofilm formation (Koskinen et al., 2000; Barrios et al., 2006; Pal et al., 2006). They are further characterised as a group of bacteria known to possess a number of traits such as the ability to biodegrade artificial compounds (Nakagawa et al., 2002) and utilise

contaminants as a growth and energy source (Kawai, 1999; Bastiaens et al., 2000). Their physiological properties enable them to exploit a variety of habitat types from water supplies (Koskinen et al., 2000; Kelley et al., 2004) to soil (Leys et al., 2004) and subsurface sediments (Fredrickson et al., 1995). They also occur in association with humans and plants as causative agents of disease (Yabuuchi et al., 1990; White et al., 1996; Kelley et al., 2004).

Their versatility and scavenging behaviour as well as their ability to accumulate carbon confers a distinct biological advantage to the sphingomonads that inhabit naturally carbon-enriched, but otherwise nutrient-limited environments such as activated sludge (Neef et al., 1999) as well as soil environments where they are commonly associated with the roots of plants (Takeuchi et al., 1995; Xie & Yokota, 2006). Studies on their unique traits have focused on their unusual membrane structures and lipid compositions, as well as their abilities to induce corrosion, cause disease in humans (Kaplan, 2004) and plants, produce polymers and biodegrade refractory pollutants such as polycyclic aromatic hydrocarbons (Dilworth et al., 1993). Their biosynthetic and biodegradative capabilities have been widely exploited for a range of biotechnological applications from producing extracellular polymers to bioremediation of environmental contaminants. Although their extraordinary ability to biodegrade recalcitrant carbon sources in particular has led to several genome sequencing projects (Leys et al., 2004; Thiel et al., 2005; Zhou et al., 2006) with the completed genomes of *Novosphingobium aromaticivorans*, *Sphingopyxis alaskensis* (Kwon & Kim, 2007) and *Sphingomonas wittichii* RW1 (JGI) as well as other genome sequencing projects such as *Sphingomonas elodea* (Hiran College), *Sphingobium chungbukense* DJ77 (Chungbuk National University, Korea) and *Sphingomonas* sp. SKA58 underway (Kwon & Kim, 2007), more widespread investigations are needed to understand the reasons for their unusual biology and versatility. In particular, more comprehensive studies on diazotrophy and biopolymer synthesis are still lacking within this group. Members within the sphingomonads such as *S. chilensis*, *S. adhaesiva*, *S. macrogoltabida*, *S. alaskensis*

and *N. aromaticivorans* are known to produce PHA. More recently, *N. aromaticivorans* was shown to possess the PHA granule-associated phasin gene (Xie & Yokota, 2006). *Sphingomonas azotifigens* is the first known diazotrophic sphingomonad (Sprent & Sprent, 1990). Diazotrophy and biopolymer synthesis may therefore be more common among members of this group than is currently known. Given the widespread occurrence of sphingomonads that biodegrade refractory organic compounds (White et al., 1996; Oshiman et al., 2007), nitrogen-deficient environments could be augmented with diazotrophic sphingomonads that have the ability to degrade the carbon in these environments and store it in the form of polymer. Thus the identification and characterisation of additional sphingomonads that meet these criteria is of immediate practical interest from both a bioremediation and a biotechnological perspective.

1.2.1 *Novosphingobium nitrogenifigens* Y88^T sp. nov.: a newly characterised sphingomonad as a model organism for low nitrogen waste stream remediation

The bacterial strain originally designated Y88 was isolated under nickel-enriched growth conditions from a naturally nitrogen-limited but carbon-enriched pulp and paper wastewater mixed culture bioreactor at Scion (Addison et al., 2007). Y88 is a diazotrophic, Gram-negative, rod-shaped, non-sporulating obligate aerobe that is capable of storing carbon in the form of a polymer, primarily PHB. Although Y88 shares distinct characteristics with other members of the sphingomonads, biochemical and chemotaxonomic data distinguish Y88 from other genera within this group and 16S rRNA sequence analysis confirms its phylogenetic affiliation to members of the genus *Novosphingobium.* Based on its genotypic (with a characteristic signature 21 base pair gap in the 16S rRNA sequence) and chemotaxonomic characteristics as well as its phenotypic traits, Y88 has recently been classified as the type strain of a new species, *Novosphingobium nitrogenifigens* Y88^T sp. nov. (Addison et al., 2007), hereafter referred to as Y88^T. DNA sequence analysis has confirmed that Y88^T contains the *nif*H$^+$ gene required for nitrogen fixation and reproducible nitrogen

fixation was confirmed biochemically via C_2H_2 reduction assays. A more definitive measurement of biological nitrogen fixation by Y88^T was subsequently obtained using enriched $^{15}N_2$ to trace the incorporation of nitrogen into Y88^T cells grown under reduced nitrogen-limiting conditions (Addison *et al.*, in preparation). Within the *Novosphingobia*, the combined ability to utilise dinitrogen as a nitrogen source and to produce PHA currently appears to be unique to Y88^T (Addison et al., 2007).

The combined abilities to fix nitrogen and accumulate carbon as a metabolic reserve made Y88^T ideal as a model sphingomonad for studying PHA accumulation in carbon-rich, nitrogen-poor growth conditions that necessitate nitrogen fixation. Therefore, this strain was chosen for detailed proteomic analysis to better understand the interrelationship between the metabolically competing activities of nitrogen fixation and PHA accumulation during growth in a carbon-rich environment. An improved understanding of the dynamic interplay between these two important metabolic capabilities has applicability for all similar sphingomonads, which will assist in optimising waste stream remediation as well as polymer production under controlled conditions in pure culture.

1.3 Bacterial utilisation of nitrogen and carbon

1.3.1 Biological nitrogen fixation

Biological nitrogen fixation is the reduction of atmospheric nitrogen to ammonia that is catalysed by the nitrogenase enzyme complex (Rees et al., 2005). Atmospheric nitrogen is available in the form of dinitrogen. The reduction of dinitrogen to ammonium ion is an ATP-dependent reaction involving the production of two moles of ammonia from one mole of molecular nitrogen at the expense of 16 moles of ATP, with a supply of electrons and protons. The process can be represented by the following equation:

$$N_2 + 8\,H^+ + 8\,e^- + 16\,ATP \xrightarrow[\text{nitrogenase}]{} 2\,NH_3 + H_2 + 16\,ADP + 16\,P_i$$

Dinitrogen, as it exists in the biosphere, is only useful to most organisms once it is either biologically or abiotically fixed into ammonia and becomes part of the nitrogen cycle. Biological nitrogen fixation is an intensively studied process (Burris & Roberts, 1993; Mancinelli, 1996; Rees et al., 2005), with virtually every aspect of the nitrogen cycle biology (Rudolf & Kroneck, 2005), the structure and role of nitrogenase enzymes (Burris, 1991; Zhang et al., 1995; Zhang, 1997; Fisher et al., 2006) and the genes encoding nitrogen-fixing proteins (Auman et al., 2001) described in detail. Also, nitrogen-fixing organisms involved in host-microbe interactions (Oke & Long, 1999), root nodule metabolism (Zakhia et al., 2006), symbiotic relationships (Steenhoudt & Vanderleyden, 2000) and bacterial endophytic associations (Sandhiya et al., 2005) have been characterised. Activities of the proteins involved in nitrogen fixation are also described, as are the conditions under which these enzymes and proteins are active (Davis et al., 1964; Dilworth et al., 1988; Joerger et al., 1991; Fisher et al., 2000; Fisher et al., 2006).

The enzyme nitrogenase is composed of two metalloproteins, the molybdenum-iron (MoFe) protein as well as the iron (Fe) protein (Rees et al., 2005). The MoFe protein provides the active site for substrate reduction whereas the Fe protein couples ATP hydrolysis to electron transfer (Rees et al., 2005). Nitrogenase uses these electrons to reduce atmospheric nitrogen to ammonia (Fig. 1.1). Nitrogenase is highly oxygen-labile and bacteria that fix nitrogen are either obligate anaerobes or utilise mechanisms to protect nitrogenase from exposure to oxygen (Sabra et al., 2000). In *Azotobacter vinelandii*, a "switch-off" mechanism is postulated to counteract the effects of O_2 stress on nitrogenase (Kuhla & Oelze, 1988; Linkerhagner & Oelze, 1995; Sabra et al., 2000). Nitrogenase "switch-off" has also been postulated for other bacteria such as *Azospirillum brasilense*, but ammonium ion rather than oxygen is considered to be the "switch-off" agent (Klugkist & Haaker, 1984; Tsagou et al., 2003; Klassen et al., 2005). Conflicting reports have proposed that nitrogenase activity is inhibited

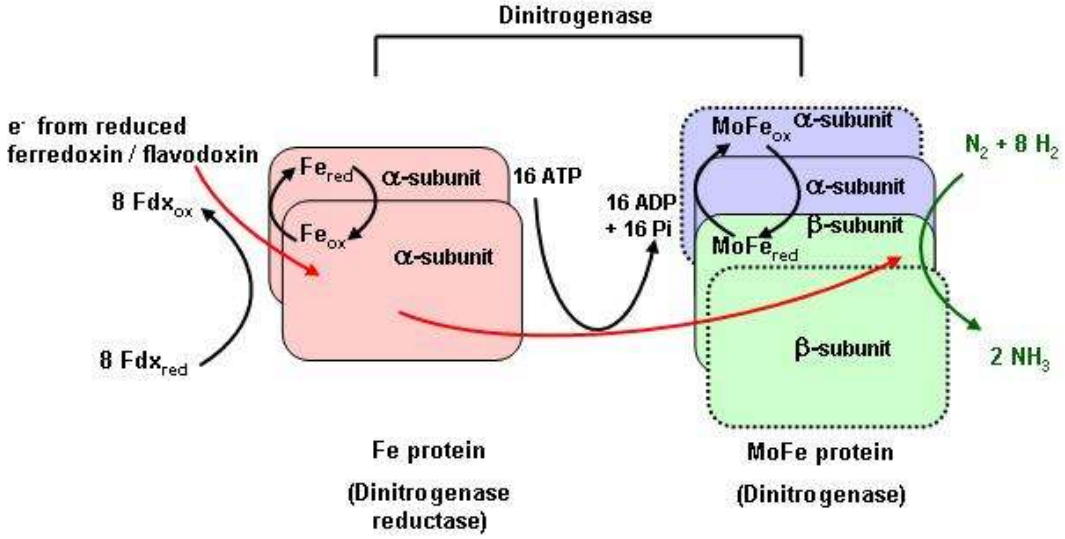


Figure 1.1 A schematic of the energy-demanding reaction catalysed by the nitrogenase complex. Repeated cycles of association and dissociation of dinitrogenase reductase (homodimeric component II α-subunits, red oblongs) with dinitrogenase (heterotetrameric component I α- and β-subunits, mauve and green oblongs respectively) result in the transfer of 8 e^- from ferredoxin (Fdx) (one per cycle) to the iron (Fe) cluster of NifH, thereby reducing NifH. After the binding of 16 ATP to NifH (two per cycle) and the subsequent hydrolysis of ATP, dinitrogenase reductase transfers e^- to the MoFe core of dinitrogenase. The reduction of one molecule of N_2 to form two molecules of NH_3 therefore requires the hydrolysis of 16 molecules of ATP and eight cycles of binding/dissociation of dinitrogenase reductase and dinitrogenase for electron (e^-) transfer, with the concomitant formation of eight molecules of H_2.

by ammonium ion (Drozd et al., 1972; Laane et al., 1980; Eady, 1981; Gordon et al., 1981; Hartmann et al., 1986; Hartmann & Burris, 1987) with Klugkist & Haaker (1984) contextualising the differences observed in ammonium ion inhibition of nitrogenase and the reasons for this. Other aspects of nitrogenase activity have been considered (Fu & Burris, 1989; Fu et al., 1989) with evidence for the use of the three oxygen sensitive nitrogenases, MoFe-nitrogenase, vanadium-iron (V-Fe) nitrogenase and iron-iron (Fe-Fe) nitrogenase, containing different active-site central metals in *A. vinelandii* (Dilworth et al., 1988). The utilisation of any of these nitrogenases is reported to depend on, among other conditions, the specific metal ions present with MoFe preferentially used (Fisher et al., 2006). The V-Fe and Fe-Fe nitrogenases are apparently only utilised when molybdenum is not available. In addition to being able to reduce dinitrogen, all three nitrogenases are able to reduce C_2H_2 to C_2H_4 although with varying efficiency. Additionally, an oxygen-insensitive nitrogenase that is unable to reduce C_2H_2 to C_2H_4 is known to exist only in *Streptomyces thermoautitrophicus* (Ribbe et al., 1997). The possibility that other alternative nitrogenases exist is under investigation (Zhao et al., 2006)

1.3.2 Ammonium assimilation

Reduced nitrogen in natural ecosystems is often scarce and can be a growth-limiting factor. As a result, bacteria have evolved highly sophisticated mechanisms to selectively assimilate nitrogen sources when these are available. These mechanisms include the ability to fix molecular nitrogen or utilise assimilatory pathways, such as ammonium assimilation. Irrespective of which pathway is used, the result is the synthesis of the major intracellular nitrogen donors, glutamate and glutamine (Reitzer, 2003).

Nitrogen fixation and ammonium metabolism are both regulated in a coordinated response by the type and amount of nitrogenous compound available. This response involves a complex process with a network of proteins which are regulated by enzymes that modify these proteins post-translationally in response to nitrogen availability. In free-living nitrogen-fixing bacteria, an

abundance of fixed nitrogen in the form of ammonium or other compounds such as nitrate or amino acids, is reported to affect nitrogen fixation by repressing either the synthesis of nitrogenase or its activity (Klugkist & Haaker, 1984; Klassen et al., 2005). Excess ammonium can also inhibit the expression or activity of other pathways that utilise nitrogenous compounds, for instance glutamine synthetase (GS), a key enzyme that is utilised in the assimilation of ammonium (Fei et al., 2006). Notably, since ammonia is considered to be the centre of nitrogen assimilation, extracellular changes in the amount of available ammonium affect the intracellular concentration of glutamine which in turn, regulates GS activity.

Since the presence of ammonium in the growth medium and its effect on nitrogenase activity have been well studied, it is known that exogenous ammonium rapidly and reversibly inhibits nitrogenase activity in the cells of a variety of nitrogen-fixing microorganisms. Most microorganisms assimilate ammonium by one of two pathways, either via the glutamate dehydrogenase (GDH) pathway or the glutamine synthetase/glutamine oxoglutarate aminotransferase (GS-GOGAT) pathway. GOGAT is a key enzyme in the early stages of ammonia assimilation in bacteria and catalyses the reductive transamidation of the amido nitrogen from glutamine and α-Ketoglutarate (AKG, also known as 2-Ketoglutarate or 2-Oxoglutarate) to form glutamate. The ammonia that enters this pathway can be supplied either by internal metabolic processes such as amino acid catabolism or through the reduction of external nitrogen sources such as atmospheric dinitrogen. In the presence of excess ammonium, the GDH pathway is preferentially induced and GDH assimilates ammonium, synthesising glutamate. Ammonium assimilation via GDH is associated with the repression and deactivation of GS, which is modified at the transcriptional and post-translational level in response to the activity of nitrogen sensor proteins.

GS occurs in two forms, the adenylylated and deadenylylated forms, which have opposing effects on its activity. Both forms of GS affect the production of glutamine. When ammonium occurs in excess, GS is adenylylated in response to the activity of the nitrogen sensor, P_{II}, in its native form (Fig. 1.2). The P_{II} protein is regulated by reversible uridylylation (Ninfa & Atkinson, 2000) and itself occurs in two forms, the native form and the modified (uridylylated) form. Adenylylation of GS is catalysed by adenylyltransferase, the activity of which is governed by P_{II}. Under conditions of excess nitrogen, the native form is expressed, whereas the modified form is only expressed under nitrogen-limiting conditions. Uridylylated P_{II} therefore stimulates deadenylylation of GS under nitrogen-fixing conditions whereas unmodified P_{II} stimulates adenylylation of GS in cells not fixing nitrogen when ammonium is present. An increase in the degree of adenylylation of GS also correlates with the repression of nitrogenase synthesis and the coordinated induction of GDH synthesis. Upon deadenylylation of GS, these events are reversed and GDH is repressed (Kleiner & Kleinschmidt, 1976). The post-translational modification of GS is governed by the nitrogen regulatory proteins NtrC and P_{II} in response to the uridylylation of P_{II} by uridylyltransferase. Uridylyltransferase is activated by AKG but inhibited by glutamine. An increase in the degree of adenylylation correlates with the repression of nitrogenase synthesis and an induction of GDH synthesis.

In the presence of low concentrations of ammonium or growth with a nitrogen source other than ammonium, for instance molecular nitrogen, the GDH pathway is repressed and the GS-GOGAT pathway is induced. In this pathway, GS assimilates ammonium whereas GOGAT synthesises glutamate. Therefore, under conditions of nitrogen fixation, nitrogen starvation results in the induction of the GS-GOGAT pathway which is associated with the derepression and activation of GS as well as the derepression of GOGAT. Consequently, maximal synthesis of glutamine occurs in a coordinated response to nitrogen limitation with the concomitant induction of proteins involved in transport and the catabolism of nitrogen sources.

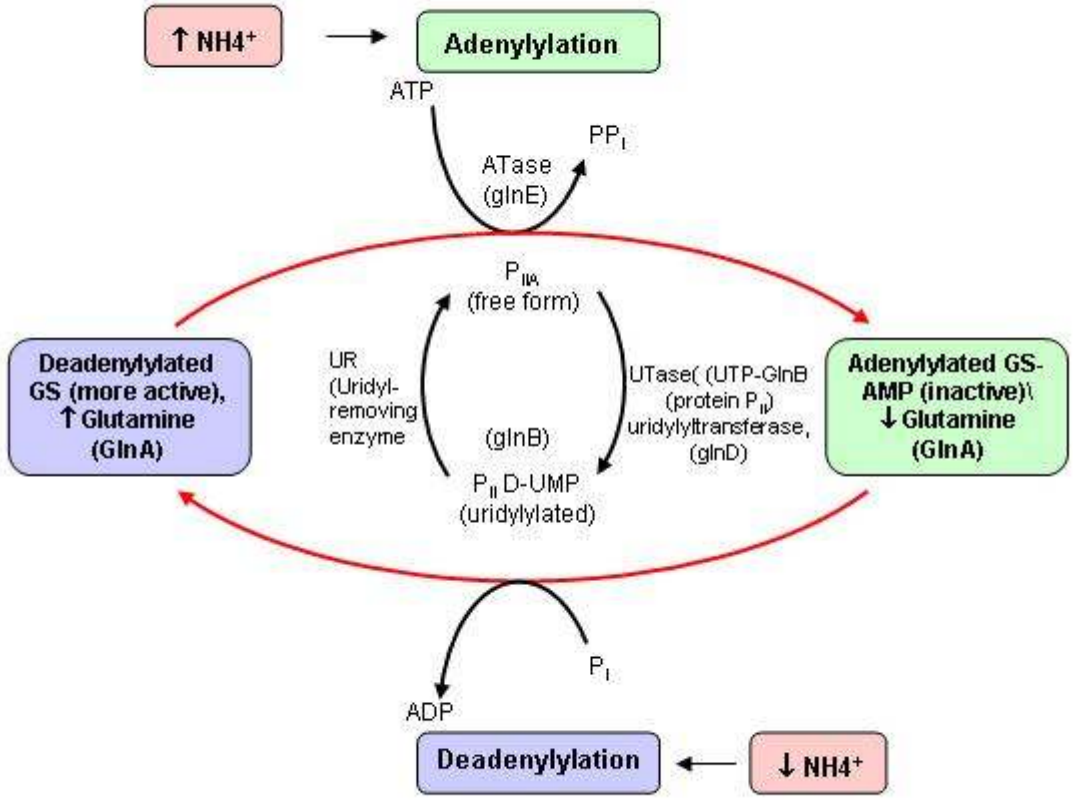


Figure 1.2 Schematic representation of adenylylation/deadenylylation of GS in response to ammonium status.

The adenylylation state of glutamine synthetase (GS) regulates its activity. The nitrogen regulating protein P_{II} stimulates adenylylation of GS by means of adenylyltransferase (ATase) which catalyses the adenylylation and deadenylylation of GS. P_{II}D-UMP stimulates the reverse reaction. Adapted from Foor et al., (1975)

1.3.3 PHA biosynthesis and bioplastics

1.3.3.1 Overview and commercial application

Plastics are considered to be the most widely used man-made materials worldwide but have environmentally unfriendly properties. For this reason, the search for biodegradable plastics to replace conventional plastics has been particularly compelling (Anderson & Dawes, 1990; Shi et al., 1999; De Oliveira et al., 2004; Sharma et al., 2007). The ability to produce "bioplastics" such as PHAs, is common to many bacteria and was first discovered in *Bacillus megaterium* by Lemoigne in 1925 (Lemoigne, 1926; Dawes & Senior, 1973). Since then, more than 250 microorganisms have been identified that can produce PHAs (Lee, 1996; Steinbüchel & Fuchtenbusch, 1998).

PHAs represent a class of natural polyesters that are biodegradable and have a broad range of uses (Anderson et al., 1990). These thermoplastics are widely considered to be of primary economic importance in a "sustainability" context, because of their similarities to conventional plastics and their complete biodegradability (Steinbüchel & Fuchtenbusch, 1998). Although plants (Poirier et al., 1995; Mittendorf et al., 1999; Poirier et al., 1999) and yeast (Poirier et al., 2002) have been genetically modified in attempts to produce high yields of PHA, the focus for production applications has mainly been on bacteria (both wild type and genetically modified) because of their presumed greater potential for high productivity with growth on low-cost carbon substrates (Byrom, 1992). Despite the first industrial production of PHA in 1982 (Anderson & Dawes, 1990), efforts to produce PHAs as substitutes for conventional polymers have historically been costly (Anderson & Dawes, 1990; De Oliveira et al., 2004). Therefore, efforts are now focused on isolating microbes that are able to produce greater amounts of PHA at lower costs using more efficient fermentation and recovery processes (Shi et al., 1999).

During the last few decades, increased awareness of environmental pollution through anthropogenic disturbances, such as discarded petrochemical plastics, has resulted in an increasing appreciation of the potential use of microorganisms to address problematic environmental issues. For this reason, the role of PHA-producing bacteria in their natural environment has come under the spotlight. Although the biological properties and biodegradability of PHAs are important, their production is based on renewable resources derived from complex carbon compounds in environments such as waste stream effluents containing mostly cellulose, but also polycyclic aromatic hydrocarbons and lignins (Peng et al., 1998). These environments are carbon-rich but otherwise nutrient-poor. Bacteria such as Y88^T that may naturally inhabit these waste streams, and are capable of remediating carbon compounds to produce bioproducts such as bioplastics, could therefore be used to enhance bioremediation.

1.3.3.2 Bacterial PHA and the biosynthetic pathway

Research to date suggests that bacterial PHA synthesis takes place in nutrient-imbalanced environments, for instance in the presence of surplus carbon when elements such as N, P, S, O or Mg are present in limiting quantities (Anderson & Dawes, 1990; Steinbüchel & Fuchtenbusch, 1998). Microorganisms investigated for their ability to produce PHA under carbon-excess, but otherwise limiting conditions, include *Ralstonia eutropha*, methylotrophs and pseudomonads. Less well documented are studies on microorganisms that do not require nutrient limitation but produce PHA during growth such as *Alcaligenes latus*, *A. vinelandii* and recombinant *Escherichia coli* (Wang & Lee, 1997). Although only a few enzymes are considered to be specifically involved in PHA synthesis (Madison & Huisman, 1999) there is a range of metabolic activities that influences the amount of PHA produced. Numerous factors, such as the carbon substrate, and its metabolic partitioning of carbon as well as the amounts of nutrients available under defined growth conditions, determine the type, quantity and quality of PHA that is synthesised (Shi et al., 1999).

PHAs are synthesised in the form of granules that accumulate as cytoplasmic inclusion bodies in bacteria (Lee, 1996). These compounds apparently serve as carbon and energy reserves, as well as sinks for reducing equivalents (Dawes & Senior, 1973; Pötter et al., 2004). Additionally, evidence supports PHA not simply as an inert storage polymer confined to bacteria, but as an interacting, solvating compound involved in various physiological functions, such as oligo-3-hydroxybutyrates (5-80 subunits) that are found in human plasma as 'solubilising' oligomers (Reusch et al., 1992).

The physical properties of PHAs depend on their monomer composition and molecular weights. PHAs are polymers made up of linear monomers that possess the general formula illustrated in Fig. 1.3, where m can be 1-4 (Madison & Huisman, 1999). However, most of the known PHAs identified to date comprise 3-hydroxyacids, where m=1. Differences in the alkyl substituents ("R" in Fig.1.3) are known, e.g. epoxidated or halogenated forms have been reported as substituents (Madison & Huisman, 1999). Variations in the alkyl substituents give rise to short-(SSC), medium-(MSC) and long-(LSC) chain length PHAs from hydroxyalkanoic acid monomers (Steinbüchel et al., 1992). The variability in the alkyl substituents, the position of the hydroxyl group or the modification of the side chains allows considerable variability in the structure of PHA. It is this diversity in structure that has enabled the varied exploitation of PHAs in commercial applications.

In order to better utilise PHA for commercial and industrial applications, a greater understanding of the structural aspects of these polymers, their physical state and the metabolic processes involved in regulation of their synthesis in the bacterial cell is needed (Anderson & Dawes, 1990). The metabolism of PHA and the enzymes directly involved in its synthesis have been well described in many bacteria e.g. *Ralstonia eutropha*, *Zoogloea ramigera*, *Aeromonas caviae, Azotobacter beijerinckii and Methylosinus trichosporium* (Senior & Dawes, 1973; Anderson & Dawes, 1990; Kichise et al., 1999; Madison & Huisman, 1999).

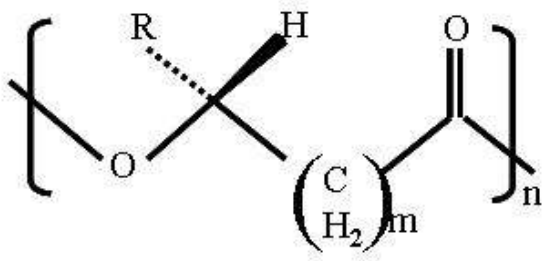


Figure 1.3 General structure of PHAs

For m=1 and alkyl substituent R equals one of the following, hydrogen, methyl, ethyl, propyl, pentyl or nonyl, then the polymer formed is either poly(3-hydroxypropionate), poly(3-hydroxybutyrate), poly(3-hydroxyvalerate), poly(3-hydroxyhexanoate), poly(3-hydroxyoctanoate) or poly(3-hydroxydodecanoate) respectively. For m=2 and alkyl substituent R equals hydrogen, then the polymer formed is poly(4-hydroxybutyrate). For m=3 and alkyl substituent R equals hydrogen, then the polymer formed is poly(5-hydroxyvalerate). Adapted from Madison & Huisman (1999).

Investigations by Ritchie *et al*. (1971) followed by demonstrations that oxygen limitation rather than nitrogen limitation resulted in PHA accumulation in bacterial cells, gave impetus early to studies on the more complex underlying regulatory pathways of PHA (Senior & Dawes, 1973).

The best understood PHA biosynthetic pathway was first described for a common form of this polymer containing poly(3-hydroxybutyric) acid (PHB) characterised in the chemolithoautotrophic bacterium *R. eutropha* (Sharma et al., 2007). This microorganism probably has the most well-studied metabolism of any organism that is capable of producing bioplastic (Pötter et al., 2004) and has undergone several name changes in the course of its characterisation due to its reclassification to different taxa (Reinecke & Steinbüchel, 2009). Initially isolated as a member of the genus *Hydrogenomonas* and named *H. eutropha*, Strain H16 was subsequently renamed *Alcaligenes eutropha* H16 (Davis et al., 1969) followed by *Ralstonia eutropha* H16 (Yabuuchi et al., 1995). *R. eutropha* was later reclassified as *Wautersia eutropha* (Vaneechouette et al., 2004) and finally renamed *Cupriavidus necator* (Vandamme & Coenye, 2004). *R. eutropha* and *C. necator* are currently the most commonly used appellations, with the former more frequently used (Reinecke & Steinbüchel, 2009) and will therefore be referred to here as *R. eutropha*. Other bacteria known to possess similar plastic-producing properties to *R. eutropha* include members of the cyanobacteria (Sharma et al., 2007), methane-oxidising bacteria (Chu & Alvarez-Cohen, 1998), halophilic bacteria (Quillaguaman et al., 2005) and nitrogen-fixing bacteria (Mandon et al., 1998; Pettinari et al., 2003). Although numerous PHA pathways have been elucidated, the current consensus is that four PHA synthetic pathways characterise the biosynthesis of all known bacterial PHAs (Steinbüchel & Schlegel, 1991; Choi & Lee, 1999).

The biosynthetic pathway of the PHB class of biopolymers is a three-step pathway utilising three major enzymes encoded by the *phb*CAB operon: β-Ketothiolase (PhbA) that catalyses the condensation of two molecules of acetyl-

CoA to acetoacetyl-CoA, a NADPH-dependent acetoacetyl-CoA reductase (PhbB) and the key enzyme PHB synthase (PhaC) (Fig. 1.4). Although the generic PHA biosynthetic pathways are well described (Senior et al., 1972; Choi & Lee, 1999), the particular enzymes, proteins and mechanisms of regulation implicated in PHA synthesis for different bacteria are the subjects of ongoing characterisation efforts (Lee et al., 1995; Sudesh et al., 2002; De Oliveira et al., 2004; Noda et al., 2005; Uchino & Saito, 2006). More specifically, enzymes involved in both the synthesis and the degradation of PHA have gained increased attention as well as enzymes active at the PHA granule surface (Liebergesell et al., 1994; Pötter et al., 2004). Of particular interest are the PHA synthases of which the *R. eutropha* PhaC is known to be constitutively expressed (Kolibachuk et al., 1999; Chen et al., 2004), the PHA depolymerases (Kim et al., 2005), the phasins (Pötter et al., 2004) and the transcription factors involved in the regulation of the relevant genes (Pötter & Steinbüchel, 2005). In *R. eutropha*, PHA synthase is currently the subject of intensive bioengineering efforts to generate enzymes with altered properties (Taguchi et al., 2002; Rehm, 2003; Taguchi & Doi, 2004; Reinecke & Steinbüchel, 2009). Additionally, the properties of phasins, which are PHA granule-associated proteins (Neumann et al., 2008) are of interest since they are known to constitute between at least three and five percent of the total cellular protein of PHB-accumulating bacteria (Reinecke & Steinbüchel, 2009). Their major function is considered to involve the stabilisation of PHA granules, preventing coalescence of the granules within the cytoplasm as well as preventing hydrophobic proteins binding to the granules. PHA synthases, phasins, PHA depolymerases, and other proteins such as 3-hydroxybutyrate dehydrogenases all bind to the granules and are considered to be involved in PHA granule biogenesis or degradation (Pötter & Steinbüchel, 2005).

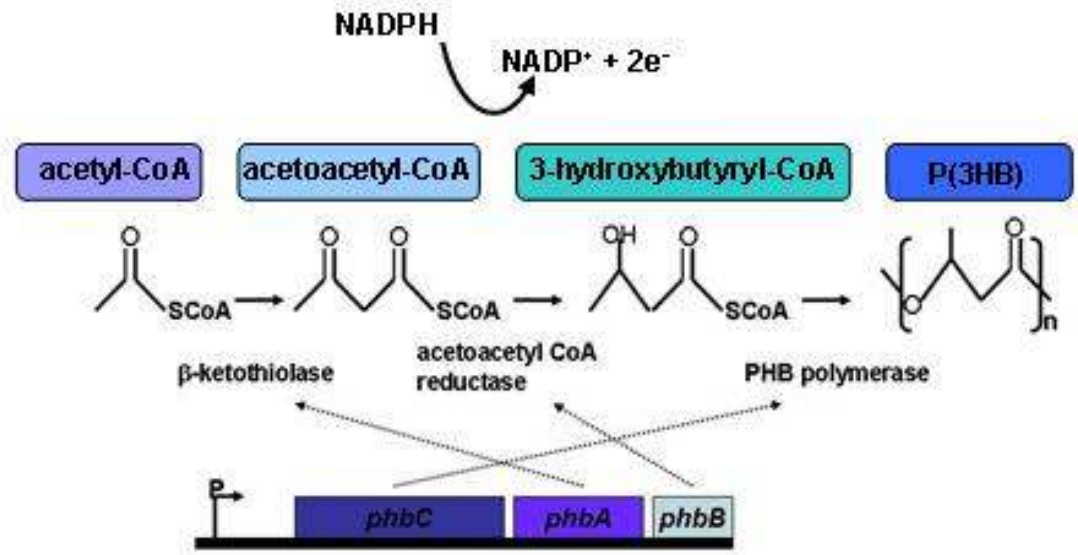


Figure 1.4. Biosynthetic pathway for P(3HB)

P(3HB) is synthesised in a 3-step pathway by the successive action of β-Ketoacyl-CoA-thiolase (β-Ketothiolase, acetyl-CoA C-acetyltransferase or PhbA), NADPH-dependent acetoacetyl-CoA reductase (PhbB) and P(3HB) polymerase (PhbC). The 3 enzymes are encoded by the genes of the phbCAB operon. A promoter upstream of phbC transcribes the complete operon. Adapted from Madison & Huisman (1999).

1.3.4 Carbon utilisation and partitioning

1.3.4.1 Central metabolism and its relationship to PHA biosynthesis

The tricarboxylic acid (TCA) cycle is a core amphibolic pathway that plays two central roles in metabolism. First, the cycle is responsible for the complete oxidation of acetyl-CoA to CO_2 and reducing equivalents used for ATP production and, second, it provides numerous chemical intermediates that are used in the biosynthesis of amino acids, among other key metabolites and biochemical "building blocks". Aerobic respiration, which involves TCA cycle activities, is central to the carbon flux that supports PHA biosynthesis via the key TCA cycle metabolite acetyl-CoA. Acetyl-CoA is produced from the glycolytic product pyruvate. Acetyl-CoA can enter the TCA cycle and is eventually oxidised to CO_2. During this process, reducing equivalents are generated via the conversion of NAD^+ to NADH. Since NADH is generated from both glycolysis and the TCA cycle, an excess of NADH occurs, which is then converted back to NAD^+ via electron transport and oxidative phosphorylation to produce ATP. Ubiquinones and cytochromes are components of the electron transport chain involved in oxidative phosphorylation and O_2 is the final electron acceptor which is reduced to form H_2O. Alternatively, when ATP production is not at a premium acetyl-CoA and NADH can be channeled toward the production of other metabolites.

Bacteria can use at least three interrelated glucose catabolic pathways to metabolise glucose: the Embden-Meyerhof (EM) or main glycolytic pathway, the Entner-Doudoroff (ED) pathway (an exclusively prokaryotic alternative glycolytic pathway) and the oxidative pentose phosphate (OPP) pathway (used to generate pentose sugars and NADPH). The OPP and ED pathways differ from the EM pathway in that although they involve oxidation of glucose, their primary role is anabolic rather than catabolic. The NADPH yielded by the OPP and ED pathways carry electrons from the catabolic pathways to biosynthetic processes. The EM pathway yields two ATP and two NADH (equivalent to 6

ATP, 3 per NADH at ΔG^0, total energy yield therefore equivalent to 8 ATP) for every glucose molecule consumed. In comparison, the OPP and ED pathways yield reducing equivalents instead of ATP, with either two NADPH (total energy yield equivalent to 6 ATP, 3 per NADPH at ΔG^0), or one ATP plus one NADH and one NADPH (equivalent to 6 ATP, 3 per NADH and 3 per NADPH at ΔG^0, total energy yield therefore equivalent to 7 ATP) for every glucose molecule consumed in the OPP and ED pathways, respectively. The use of the ED and OPP pathways are therefore more costly in terms of energy production but provide NADPH for biosynthesis as well as the prevention of oxidative stress. In fact, even though the standard reduction potential of NAD^+/NADH and $NADP^+$/NADPH couples are equivalent, in the cell, the metabolic value of NADPH is higher because [NADPH]>[$NADP^+$] whereas [NAD^+]>[NADH] resulting in inequalities in the cellular reduction potentials of these two pyridines. The NADPH/$NADP^+$ couple is therefore a better electron donating system than the NAD^+/NADH couple (Garret & Grisham, 1999). Maintaining catabolic *vs.* anabolic processes via the production of intermediates and substrates for these activities requires interrelated metabolic processes many of which are species-specific and therefore remain the subject of continued investigation.

1.3.4.2 Carbon flux and the regulation of PHA biosynthesis

The regulation of PHA synthesis and the role of the key regulatory metabolite, acetyl-CoA, in the PHA biosynthetic pathways has been the subject of many studies (Moskowitz & Merrick, 1969; Senior et al., 1972; Mandon et al., 1998; Madison & Huisman, 1999). Acetyl-CoA can either be oxidised by the TCA cycle, or it can serve as a substrate for PHA synthesis (Oeding & Schlegel, 1973). The ultimate fate of the acetyl-CoA formed depends on factors such as the availability of nutrients including excess carbon or limiting oxygen or nitrogen. These factors also influence the ratio of NADH/NAD^+ produced by the TCA cycle. NADH, NAD^+, NADPH and $NADP^+$ are involved in energy metabolism, reductive biosynthesis and oxidation. Whereas NADH is the electron donor for

the electron transport chain that generates ATP, NADPH is a key component in cellular antioxidant systems and is the reducing agent in anabolism (Rich, 2003). Although NADH seems to be the main electron donor, under conditions of oxidative stress NADPH-producing pathways are induced to replenish the NADPH pools (Vlamis-Gardikas, 2008).

In *R. eutroph*a the $NADPH/NADP^+$ ratio increases when limiting amounts of nitrogen, phosphorous or oxygen are available, thus promoting PHB synthesis. Under these conditions, cell growth is restricted (Choi & Lee, 1999) and the increase in $NADPH/NADP^+$ ratio inhibits citrate synthase and isocitrate dehydrogenase. Consequently, acetyl-CoA does not enter the TCA cycle, but is instead available for conversion to acetoacetyl-CoA by β-Ketothiolase, the first enzyme in the PHB biosynthetic pathway. Conversely, when nitrogen, phosphorus and oxygen are sufficient, cell growth is no longer restricted and NAD(P)H can donate its reducing equivalents to the oxidative phosphorylation pathway, which results in a lower $NAD(P)H/NAD(P)^+$ ratio. Furthermore, CoASH concentration increases as it is released by the entry of acetyl-CoA into the TCA cycle, which is inhibitory to β-Ketothiolase, thereby reducing the rate at which PHB is synthesised (Oeding & Schlegel, 1973; Lee et al., 1995; Choi & Lee, 1999).

Senior and Dawes (1971) proposed that PHB is not only a source of carbon and energy, but is also a sink for reducing equivalents and therefore could be considered a redox regulator within the cell. Shi *et al*., (1999) have demonstrated the importance of increased concentrations of, for instance, available NADPH as well as acetyl-CoA in achieving maximum PHB yields and redirecting metabolic fluxes in *R. eutropha*. More acetyl-CoA can be made available for consumption by effectively "blocking" the entry of acetyl-CoA into the TCA cycle. Thus by redirecting the carbon fluxes to another pathway, an increase in NADPH recycling can be achieved (Shi et al., 1999).

Unlike non-diazotrophic PHA producers, Y88^T provides an opportunity to study metabolism under conditions of nitrogen fixation as well as PHA synthesis. These studies can elucidate underlying catabolic (TCA) *vs.* anabolic (PHA) pathways, and the role of culture conditions in influencing the utilisation of carbon (Fig. 1.5). For example, nitrogen fixation has been shown to compete with PHA accumulation in bacteroids of *Rhizobium* and *Bradyrhizobium japonicum* (Povolo et al., 1994). Y88^T therefore provides an opportunity to determine how a nitrogen fixer that accumulates PHA is able to maintain its core metabolic requirements yet balance competing metabolic activities such as nitrogen fixation and PHA accumulation. Understanding the interplay of the regulatory pathways of nitrogen fixation, carbon utilisation and PHA synthesis in Y88^T can elucidate how a diazotroph copes with varying environments. Glucose, with its central position in carbon and energy metabolism, is a useful carbon source to initiate the analysis of carbon flux in Y88^T to produce metabolites specific to its PHA biosynthetic pathway and the type of PHA produced under nitrogen-fixing conditions.

Metabolic carbon fluxes are central to PHA biosynthesis in that the carbon source determines the particular precursor metabolites that are produced (Ramsay et al., 1989; Muller et al., 1999; Wendisch et al., 2000; El-Mansi, 2005; Kolkman et al., 2005; Zhang et al., 2006). These precursors in turn determine the type and amount of PHA that is synthesised (Zhao & Shimizu, 2003; Wolfe, 2005; El-Mansi et al., 2006). An understanding of these processes in pure culture may enable Y88^T, or other sphingomonads to be exploited in a production environment to produce higher yields of PHA from renewable resources.

1.4 Proteomic approaches for protein expression profiling

Proteomics had its inception with the advent of a protein separation technique introduced independently by both O'Farrell (1975) and Klose (1975) and its subsequent use in the study of bacterial physiology (Herendeen et al., 1979).

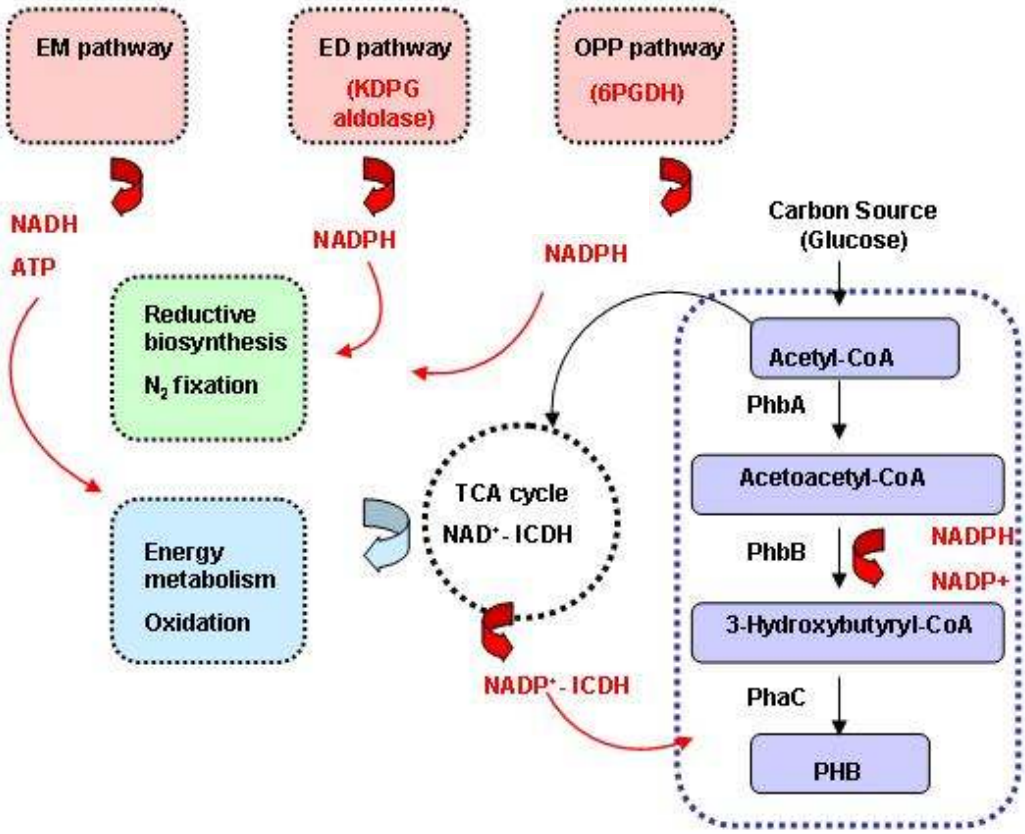


Figure 1.5 Schematic of NADH and NADPH production and utilisation via catabolic vs. anabolic pathways

NADH is produced and again consumed in catabolic pathways. NADPH couples catabolic to anabolic processes by carrying the reducing power released in certain catabolic sequences such as the two alternative glycolytic pathways, the ED and OPP pathways, to reductive biosynthetic pathways such as N_2 fixation and PHB production.

This separation technique was adopted as the classical proteomic approach and is based on protein separation using 2-D gel electrophoresis, in which proteins are resolved in the first dimension based on their isoelectric points and in the second based on their molecular masses, followed by protein identification and characterisation using MS. 2-D gel-based applications have since been shown to be particularly useful in studying bacterial proteomes under different growth conditions (Linn & Losick, 1976; Reeh et al., 1977; Agabian & Unger, 1978; Voigt et al., 2006).

Despite early advances in the use of proteomic applications, it was not until 1995 when the first genome sequence of the bacterium *Haemophilus influenzae* (Fleischmann et al., 1995) was published, that whole-genome protein identifications and functional genomics became possible. Since then, proteomics has advanced rapidly and has become an increasingly powerful tool used in molecular cell biology supported by the development of sophisticated technology such as multi-dimensional chromatography, MS and bioinformatics (Steen & Mann, 2004). Since proteins are expressed in different combinations and levels during various stages of the life cycle of cells and under changing environmental conditions, the proteome, unlike the (relatively) static genome, must be dynamic and in constant flux. The use of proteomics throughout the past decade has assisted in characterising these differences and has therefore provided an improved understanding of cellular physiology (Hecker & Volker, 2004). Whereas the genome sequence provides the "blueprint" of the cellular properties of a living cell, it is the proteome that brings a broader understanding of functional complexity to that sequence (Brotz-Oesterhelt et al., 2005; Kohler et al., 2005). Proteomics can therefore be used to investigate the biology of cells or organisms.

The term "proteomics" is loosely defined as the study of the set of proteins that is expressed throughout the lifetime of a cell and includes modifications of proteins that occur post-translationally (Takahashi et al., 2003). A more apt

description of proteomics ideally involves the study of the identity and function of the proteins from a cell, tissue or organism at a given time or times, thus providing a snapshot of an expressed group of proteins for a given set of conditions. Since the proteome is highly dynamic, its composition reflects the physiological state of the specific sample studied (Jurat-Fuentes & Adang, 2007).

Proteomic studies are now well advanced for diverse bacteria with extensive proteome coverage achieved for *Escherichia coli* (Han & Lee, 2006) and *Mycoplasma pneumoniae* (Hecker & Volker, 2004). Additionally, proteomic studies are increasingly being used to complement other molecular techniques such as microarrays (Hu et al., 2006; Uttamchandani et al., 2006) to gain additional information on cellular processes. The use of proteomic techniques, particularly in protein expression profiling (Ellmark et al., 2006), provides a powerful tool to elucidate the expression patterns of proteins expressed under different environmental conditions. Moreover, the degree of expression of individual proteins can be determined (Kohler et al., 2005). Protein expression profiling is supported by 2-D gel-based proteomics used in combination with MS (Brotz-Oesterhelt et al., 2005) and can provide a proteomic view of cellular activities that can ultimately contribute information on basal cellular processes (Hecker, 2003; VanBogelen, 2003; Hecker & Volker, 2004). As such, proteomics has proven to be useful in predicting the physiological states of bacterial cells, including their growth conditions, based on their "proteomic signatures" (VanBogelen, 2003; Voigt et al., 2004). These proteomic signatures can be used as indicators/predictors of the growth phase and related activities of actively growing cells. Additionally, predictions of other aspects of metabolism and physiology can also be made from the proteomic signatures of proteins involved in unrelated metabolic or regulatory activities.

Gel-free proteomic approaches have recently become popular (Linsen et al., 2006; Bosch et al., 2008) as they identify many proteins that are beneath the

threshold of detection of gel-based methods (Roe & Griffin, 2006). Gel-free proteomics is emerging as the preferred approach for quantitative proteomics (Haqqani et al., 2008) that uses methods such as stable-isotope labeling (Chen et al., 2008) and spectral counting (Bosch et al., 2008). Although the use of gel-free proteomic alternatives is particularly useful in identifying membrane (Wang et al., 2008) or low-abundance proteins as well as proteins with extreme pIs and/or molecular weights, 2-D gel-based approaches still remain popular as they complement gel-free alternatives (Hahne et al., 2008; Hecker et al., 2008).

The particular appeal of 2-D gels lies not only in the visualisation of proteins that have altered abundance, but in their ability to resolve post-translationally modified proteins into separate spots that can be individually measured. Conventional 2-D gel-based approaches can resolve thousands of proteins in a biological sample and provide fairly broad coverage of proteins from pI 3 to 11 and in the molecular weight range of approximately 1000 to 100,000. Despite the appeal of this technique to visualise altered protein abundance, its limitations include the under-representation of highly acidic or basic proteins as well as poor representation of hydrophobic proteins. Arguably, a key limitation is considered to be poor reproducibility across gels that hinders the comparison and interpretation of results (Clark & Gutstein, 2008; Fuxius et al., 2008). For this reason, the use of 2-D DIGE, which aims to improve reproducibility, is often used instead of the conventional 2-D gel-based approach, including when comparing response to changes in the extracellular environment. Usually, the identification of approximately 300 cytosolic proteins using 2-D gel-based approaches combined with MS can be considered a reasonable starting point for physiological studies, providing the potential to analyse the regulation of metabolic pathways such as glycolysis or TCA cycle activities (Voigt et al., 2004). 2-D gel-based proteomics will therefore continue to be valuable for comparative physiological proteomics involving multiple samples, particularly where a quantitative measure of altered protein abundance is required.

1.5 Aims

Since Y88^T grows largely in environments where nitrogen is naturally limiting but carbon is found in excess, it is potentially an ideal organism for the study of PHA production in a nitrogen-limiting environment. The first aim of this study was therefore to investigate the microbiology of Y88^T to establish reproducible growth profiles for this organism under conditions of carbon-enrichment and both nitrogen-supplemented and nitrogen-depleted conditions in pure culture, batch flask mode. The effect of growth conditions on nitrogen fixation and PHA accumulation were examined.

The second aim of this study was to examine, the Y88^T proteome to identify proteomic signatures of the physiological response of Y88^T cells to nitrogen-supplemented and nitrogen-depleted conditions (representing the non-nitrogen-fixing and nitrogen-fixing physiological states of Y88^T respectively) in carbon-enriched media during PHA production. The aim was to identify the Y88^T proteins involved in nitrogen fixation and PHA accumulation and map any altered protein abundance in the underlying metabolic pathways to a given set of growth conditions. This would allow predictions to be made about which growth conditions promoted optimal PHA production in nitrogen-fixing Y88^T pure cultures.

The physiological response of Y88^T cells to different levels of DO (DO^{low} and DO^{high}) at constant and changing pH was also examined during growth, biomass production, nitrogen fixation and PHA accumulation in nitrogen-supplemented and nitrogen-depleted, carbon-enriched media. Since the nitrogenase enzyme is oxygen labile and consequently irreversibly inactivated by high levels of oxygen, the effect of oxygen concentration on the ability of Y88^T to fix nitrogen and accumulate PHA was examined. Proteomic signatures were used to predict the impact, if any, of different oxygen concentrations on the proteins involved in nitrogen fixation and PHA production.

This study is the first combined microbiological and proteomic investigation of a novel, nitrogen-fixing, PHA-producing *Novosphingobium* strain and contributes to an understanding of the relationship between nitrogen fixation and PHA production in a family of bacteria that are proving useful in bioremediation efforts. Y88^T provides a suitable model to study the interplay of cellular metabolism, PHA synthesis and nitrogen fixation. Although many biotechnologically-based studies have been conducted on PHA-producing organisms, and many metabolic pathways have been described, most of these involved using mutant organisms with selectively engineered properties that enable optimisation of PHA yield. Y88^T provides a unique opportunity to study a novel organism that produces PHA under natural conditions necessitating nitrogen fixation. The findings of this study will also contribute to assessing optimal PHA accumulation in pure culture in a production environment.

Chapter 2: Materials and Methods

2.1 Microbiology

The microbiology of Y88^T was investigated to establish a reproducible set of growth parameters before predictions about growth, nitrogen fixation and PHA accumulation could be tested. From these parameters, appropriate sampling times corresponding to the physiological state of the cells under desirable conditions of PHA accumulation and nitrogen fixation were established. This was done to promote consistency across all replicates and allow interpretation of the perturbations of the growth conditions that affect PHA production, to enable its optimisation. Therefore, minimal growth media with an excess of carbon substrate were formulated to determine the physiological response of Y88^T cells to either nitrogen-depleted or nitrogen supplemented conditions with sufficient carbon to support primary metabolism as well as PHA accumulation.

2.1.1 Description of Y88^T

Y88^T is a mesophilic, Gram-negative, obligate aerobe possessing diazotrophic properties that can accumulate the biopolymer PHA intracellularly. This bacterium belongs to the sphingomonad group of the Alphaproteobacteria yet is distinctly different from other *Novosphingobia* within this group due to its nitrogen-fixing capability. Y88^T was isolated at Scion in Rotorua, New Zealand from a pulp and paper-mill effluent culture bioreactor operating under nitrogen-limiting conditions (C:N ratio of 140:1) (Addison et al., 2007). Y88^T cells are non-sporulating, non-motile rods that form off-white to pale yellow colonies after 2-4 days on nutrient agar. The morphology of Y88^T colonies appears entire, convex and shiny with the centre of the colonies developing a deeper yellow pigmentation with time. Y88^T cells grow optimally on nutrient agar at 30 °C and in minimal nitrogen-supplemented medium containing excess glucose as a carbon source.

2.1.2 Other bacterial isolates

2.1.2.1 *Azotobacter vinelandii* ICMP4036, *Pseudomonas putida* mt2 and *Sphingomonas sp.* WP01 (DSM 19371)

In addition to Y88^T, three other strains of bacteria were used for different aspects of the growth experiments. Two phylogenetically distant Gammaproteobacteria, *Azotobacter vinelandii* ICMP4036, a known nitrogen fixer, and *Pseudomonas putida* mt2, known to be incapable of nitrogen fixation, were selected for an interspecific comparison to Y88^T. Both *A. vinelandii* and *P. putida* mt2 are known to synthesise PHA. The growth profiles of all three species were compared under conditions conducive to nitrogen fixation or ammonium assimilation, and PHA accumulation. The nitrogen-fixing and PHA-accumulating abilities of all three species were therefore investigated in parallel. The sphingomonad WP01, like Y88^T, also belongs to the Alphaproteobacteria and was selected for an in-group comparison to Y88^T, due to its similar ability to synthesise PHA. WP01, designated *Sphingomonas* sp. strain WP01 (DSM 19371) was isolated from a polycyclic aromatic hydrocarbon-contaminated soil at Waipa Mill, Rotorua, in New Zealand. No nitrogen-fixing genes have been identified in the recently sequenced WP01 genome (Strabala, Macdonald & Liu, unpublished results). The growth profiles and PHA-accumulating abilities of both Y88^T and WP01 were examined.

2.1.3 Growth conditions

All experiments in this study were performed under sterile conditions with bacterial cells cultivated in batch mode, i.e. either in shaker flasks or in bioreactors. Batch cultivation makes use of a closed system in which all nutrients are in excess at the start of the cultivation relative to later phases of the growth process.

2.1.3.1 Cell stocks, plate inoculation and pre-culture inocula

Cells for all bacterial strains were maintained as freezer stocks on microbeads (Pro-lab Diagnostics, Nelson, New Zealand) placed in 2 mL cryogenic vials and

stored at -80°C. When required, one vial was transferred from -80° C storage and kept on ice during inoculation of the cells using the microbeads. For Y88^T, *P. putida* mt2 and WP01 inoculation, a single bead was transferred onto a nutrient agar plate (5 g/250 mL, Difco, Becton Dickinson & Co., Franklin Lakes, NJ, USA) using an inoculation loop. For *A. vinelandii*, a single bead was transferred onto a plate count (PC) agar plate (5 g/250 mL, Difco, Detroit, MI, USA). Inoculated nutrient agar plates were incubated at 30° C for 3 days to grow colonies of the appropriate size for inoculation into a nutrient growth broth (1 colony/5 mL growth medium with a final concentration of 2.5 g tryptone, 0.5 g glucose, 1.75 g yeast extract per litre of MQ H_2O). Single colonies were individually selected and inoculated into a 50 mL polypropylene Falcon tube (Greiner Bio-one, Frickenhauser, Germany) containing 5 mL nutrient growth broth and cultivated for approximately 24 hr in capped tubes in an orbital shaker incubator at 150 rpm.

2.1.3.2 Shaker flask cultivation

Batch cultivation was carried out in sterilised 250 mL (for pre-cultures) or 1 L Ehrlenmeyer flasks set up in replicate. Following overnight growth, a 2% (v/v) inoculum was transferred to triplicate flasks by transferring 2 mL Y88^T cells from the starter culture in 50 mL polypropylene Falcon tubes to 100 mL minimal nitrogen medium (defined in Table 2.1 with respect to nitrogen source used) enriched with carbon substrate (0.4 g KH_2PO_4, 0.1 g K_2HPO_4, 0.2 g $MgSO_4$, 0.1 g NaCl, 5 g glucose, 0.05 g yeast extract, 0.1 g NH_4Cl, 10 mg $FeCl_3$, 2 mg $Na_2MoO_4.2H_2O$ per litre of MQ H_20, pH 7.2)(Cote & Gherna, 1994). Cultures were cultivated for 24 hr to equilibrate the cells in the minimal nitrogen medium before scaling up to the larger volume 1 L flasks. After 24 hr of growth, a 2% (v/v) inoculum was transferred to triplicate flasks (for independent samples) per experimental growth condition being tested by transferring 10 mL of the Y88^T pre-culture to each of the replicate 1 L Ehrlenmeyer flasks containing 500 mL of minimal nitrogen medium enriched with carbon substrate as before. The batch cultures were grown aerobically to the required OD_{600} obtained by monitoring growth throughout a 24 hr growth period or 72 hr,

Growth medium	Available nitrogen in growth medium	Physiological state of the cell
NH_4^+-free	None. No NH_4Cl supplemented	Fix^{off} (negative acetylene reduction)
NH_4^+-supplemented	0.1 g/L NH_4Cl supplemented, not depleted by time of sampling	Fix^{off} (negative acetylene reduction)
NH_4^+-depleted	0.1 g/L NH_4Cl supplemented, depleted by time of sampling	Fix^{on} (positive acetylene reduction) or
		Fix^{off} * (decreased acetylene reduction but similar to Fix^{on} in terms of cell productivity)

Table 2.1. Growth media and corresponding cell physiological state with regard to nitrogen fixation defined for $Y88^T$ growth conditions.

depending on the experimental design. Y88^T cells were cultivated at 30° C, 150 rpm agitation in an orbital shaker incubator.

2.1.3.3 Bioreactor cultivation

Y88^T bioreactor batch cultivation was carried out in a Bioflo100 bioreactor (New Brunswick Scientific Co., Inc., New Brunswick, NJ, USA) consisting of two 1 L vessels set up in parallel. Each vessel contained 1 L of minimal nitrogen but carbon-enriched medium with a headspace of approximately 300 mL. Y88^T cells were cultivated under a specific set of growth conditions according to the experimental design (Fig. 2.1). Each experiment was carried out in triplicate. A 2% (v/v) inoculum was grown to comparable turbidity, measured with a PC-controlled GBC spectrophotometer, model UV/VIS918 after removing 2 mL of culture at each sampling period over 24 hr and diluting appropriately before transferring to a cuvette (1 cm light path) with a syringe (see section 2.1.6.1). Ten mL of bacterial culture was removed and discarded prior to removing the 2 mL sample to prevent using cells considered to be the dead volume in the sampling tube. Bioreactor parameters controlled throughout the 24 hr growth period were: agitation (150 rpm), temperature (30° C), dissolved O_2 (DO) (10% DO, 50% DO) for the first set of oxygen experiments to determine differences between selected characteristics of Y88^T cells grown under DO^{low} and DO^{high} conditions where the medium was allowed to acidify. For the second set of oxygen experiments to determine differences between selected characteristics of Y88^T cells grown to low (OD_{600}=0.5) and high (OD_{600}=2.0) under different DO conditions (10% DO, 70% DO,) at a constant pH, the bioreactor parameters controlled throughout 24 hr growth were: agitation (150 rpm), temperature (30° C), DO (10% DO, 70% DO) and pH = 7.2 ±0.2. The Bioflo110 bioreactor system was used in conjunction with the BioCommand Plus BioProcessing Software Version 3.30 for all bioreactor experiments.

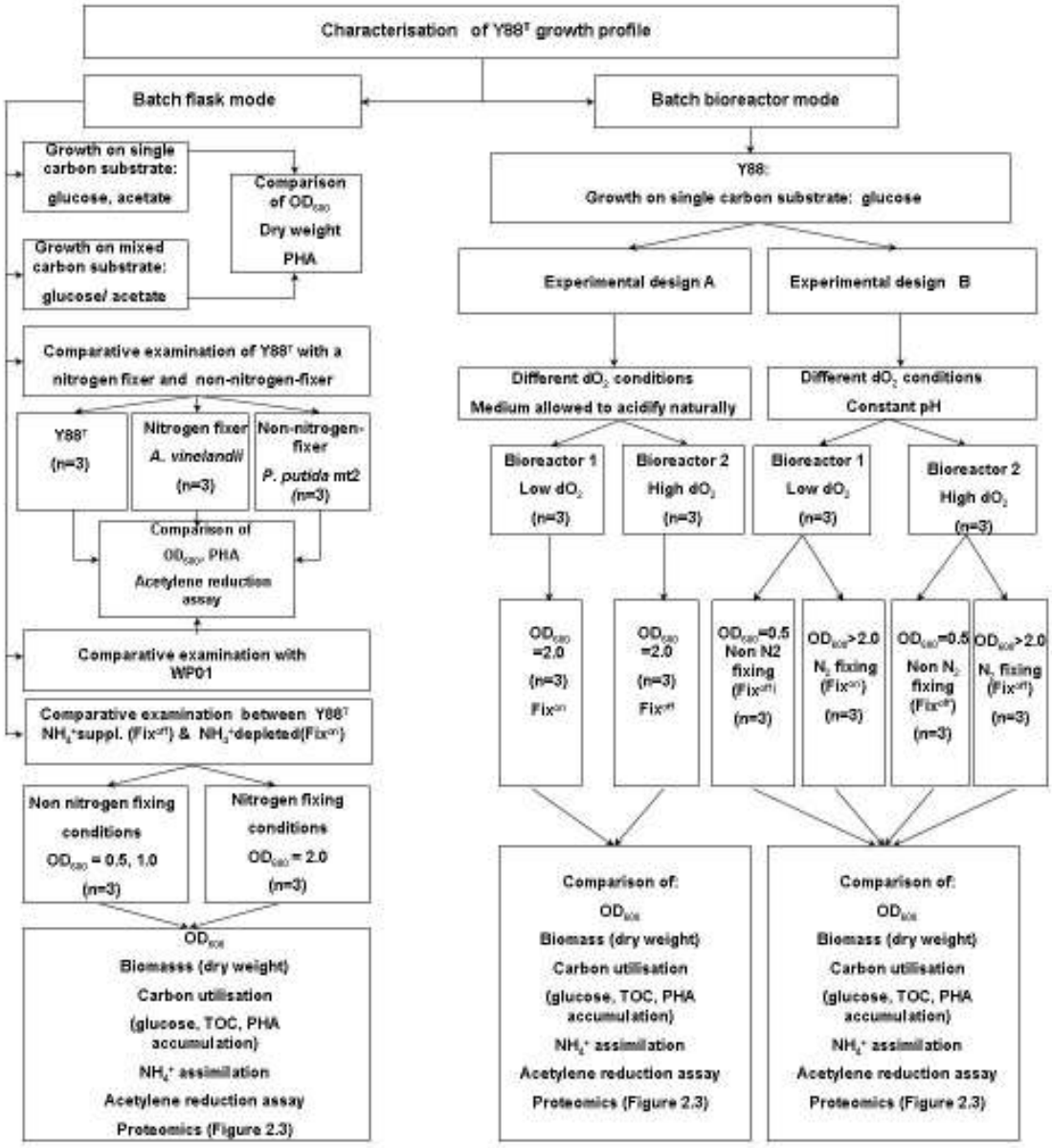


Figure 2.1. Schematic representation of workflow to investigate the microbiology of Y88^T growth, nitrogen assimilation, carbon utilisation and PHA accumulation under different growth conditions.

2.1.4 Optical density as an indicator of growth

2.1.4.1 Batch flask cultivation

To evaluate growth, optical density (OD_{600}) was monitored over time. $Y88^T$ was grown to lower cell densities (OD_{600}=0.5, 1.0) and high cell density (OD_{600} =2.0) in batch flasks during a 24 hr period. Each experiment was performed in triplicate. Two mL of $Y88^T$ culture was taken at each sampling for each replicate and transferred to a 3 mL cuvette tube (1 cm light path). OD_{600} measurements after appropriate dilution to assay growth, were carried out using a PC-controlled GBC spectrophotometer, model UV/VIS918. A blank, containing medium from a sterile stock was used to set the baseline reading. All OD_{600} measurements for each sample were carried out in duplicate and the average of each duplicate pair was determined. When the cells had reached the desired OD_{600}, all the culture in a flask was centrifuged in 15 mL Falcon tubes.

2.1.4.2 Batch bioreactor cultivation

To evaluate growth, OD_{600} was determined in bioreactor cultivation. $Y88^T$ was grown to high cell density (OD_{600} =2.0) under DO^{low} (10% DO) or DO^{high} (50% or 70% DO) conditions during a 24 hr period. Three independent biological replicates for each growth condition were generated. Culture sampling and OD_{600} measurements were carried out as described in section 2.1.4.1. When the cells had reached the desired OD_{600}, all the culture in a flask was centrifuged into separate 15 mL Falcon tubes.

2.1.5 Dry weights as an indicator of biomass

The dry weight of each sample was determined as an indicator of biomass at specific stages of growth coinciding with Fix^{on} and Fix^{off} states of the cells (Table 2.1) under selected growth conditions. Fifteen mL samples of cultivated $Y88^T$ cells were centrifuged at 1814 × g, 4 °C, for 30 min and the spent medium supernatant was decanted carefully so as not to dislodge any cells. The cells were then centrifuged to ensure a tight pellet and 15 mL of MQ H_2O was added to the tube to rinse. The cells were again centrifuged and the MQ H_2O was

decanted. Another 15 mL of MQ H_2O was added to the tube, which was centrifuged again. Following the second centrifugation, 14 mL of the supernatant was removed and 1 mL retained in the tube to resuspend the cells and facilitate transfer to aluminium foil containers. The foil containers had been tared and their masses documented, following which the containers were pre-dried in an oven at 105 °C and stored in a desiccator to prevent moisture condensing on the containers and confounding the initial container weights. After transferring the wet cells to the aluminium foil containers, they were immediately placed in an oven at 105 °C. The containers were consecutively weighed in a microbalance (Mettler Toledo International Inc., UK) over several days until the weight had stabilised and at least three successive, constant weight values per sample were obtained. The third constant measured weight value for each sample was accepted as the final combined weight of the sample and foil. The tare of the foil container was subtracted from the final weight to derive the total dry weight of the cells per sample.

2.1.6 Nitrogen utilisation assay

2.1.6.1 Determination of ammonium assimilation

A Megazyme ammonia (rapid, manual format) assay kit (Megazyme International, Wicklow, Ireland) was used to determine the amount of ammonia in the culture medium of the growing $Y88^T$ at specific time points (every 4 hr) over a 24 hr period of growth. The principle of the assay is based on the presence of GDH and reduced NADPH in the medium. In the presence of GDH and NADPH, ammonia (in the form of NH_4^+) reacts with 2-oxoglutarate to form L-glutamic acid and $NADP^+$. The amount of $NADP^+$ formed is stoichiometric with the amount of ammonia and the assay therefore measures the decrease in absorbance at 340 nm as the consumption of NADPH.

The assay was carried out according to the manufacturer's instructions with all samples measured in duplicate. Fifteen mL culture samples were centrifuged at 1814 × g to pellet the cells and the spent medium was decanted into another

Falcon tube before filtering into a third Falcon tube using a 0.2 μm Whatman filter (Global Science). The medium was subsequently filtered again using a fresh 0.2 μm filter to ensure that no residual cells could contaminate the medium and give confounding values for the amount of NH_4^+ present in the medium at each time point. Before carrying out the assay, each sample was diluted with nine volumes of MQ H_2O to one volume of filtered sample. On completion of the assay, the results were multiplied by the dilution factor to determine the actual final amounts of NH_4^+ in the medium at each time point. The final values were then subtracted from the known quantity of starting NH_4^+ supplemented as NH_4Cl (0.1 g/L) in the medium, to establish how much NH_4^+ had been assimilated throughout a 24 hr growth period.

2.1.7 Acetylene reduction assay

Acetylene (C_2H_2) reduction is routinely used to estimate the activity of nitrogenase, as it is a good substrate for nitrogenase (Ludden, 2001). When C_2H_2 is reduced, it produces ethylene (C_2H_4) that is relatively easy to quantify (Li, 2002) using the H_2-flame ionisation gas chromatography system (Hardy et al., 1968). C_2H_2 is a highly soluble gas that readily crosses the cell membrane. Although a much less soluble gas than C_2H_2, C_2H_4 still readily crosses the cell membrane and can accumulate in the gas phase above the cell suspension where it can be easily sampled and quantified by flame ionisation gas chromatography. Direct correlations can be made between C_2H_4 formation and nitrogenase activity as an indicator of nitrogen fixation. Theoretically, as few as 2 to 3 cells can produce sufficient C_2H_4 for detection by the H_2-flame ionisation system. In *Azotobacter*, ethylene formation measured 0.02 pmol per hour per *Azotobacter* cell over a 1000-fold range of cell concentration (Hardy et al., 1968). Due to the extreme sensitivity by H_2-flame ionisation to detect C_2H_4 formation, this technique was used to determine nitrogenase activity in $Y88^T$ cells assayed for C_2H_2 reduction.

The equation for the reaction forming C_2H_4 gas is as follows:

$$\underset{(C_2H_2)}{HC{\equiv}CH + 2\,H^+ + 2\,e^-} \xrightarrow[\text{nitrogenase}]{} \underset{(C_2H_4)}{H_2C{=}CH_2}$$

Y88^T cells were collected in 15 mL Falcon tubes and centrifuged at 4 °C for 30 min. Two replicates per sample were prepared for each C_2H_2 reduction assay as well as a positive (pure C_2H_4) and two negative (blank) controls. Cells were resuspended in 5 mL of culture supernatant and resuspended cells were transferred to 10 mL glass vials containing caps with rubber septums. The two blanks were prepared by using 5 mL of MQ H_2O instead of Y88^T cells. A gas bladder was prepared by evacuating as much air as possible. Approximately 200 mL of tap water was added to a 1 L conical flask and approximately 5 mL of calcium carbide (CaC_2) was added to generate C_2H_2 by the following reaction:

$$CaC_2 + 2\,H_2O \rightarrow Ca(OH)_2 + C_2H_2$$

After initiation of the reaction, the first 30 sec of gas formed was vented. The bladder was then quickly connected to the rubber bung by means of silicone tubing and filled with C_2H_2 by attaching a syringe needle to the end of the tubing and pushing the needle through the bladder seal. The glass vials were prepared by sparging with argon for about 30 sec to displace any O_2 in the vial. After sparging with argon, 1 mL O_2 was injected into the vial to provide some O_2 for the aerobic Y88^T. To start the assay, 0.1 mL of C_2H_2 was then injected with a syringe into the vials. Five mL of resuspended Y88^T cells was subsequently added to the vials. Sealed vials were incubated in a 30^o C water bath for 1 hr. Following incubation, approximately 6 mL of headspace gas was removed with a syringe and the syringe needle immediately plunged into a rubber bung to keep the sample sealed before injection into the gas chromatograph. C_2H_4 formed from C_2H_2 was analysed with a Pye Unicam Series 105 gas chromatograph using a Flame Ionisation Detector (FID) and a

glass column (1.5 m x 6.35 mm outer diameter and 4 mm inner diameter) filled with Porapak N (80/100 mesh), operated at 120 ^{o} C and carrier gas nitrogen (N_2) at 20 mL/min. An HP 3393A integrator was used for quantitative analysis.

Headspace gas from a blank sample was injected into the chromatograph prior to the injection of any Y88^{T} headspace gas sample and checked for a negative signal resulting from non-reduced C_2H_2. A second blank sample was injected at the end of the experiment to check for any false positive signal caused by carryover of residual C_2H_4 produced in any of the preceding Y88^{T} samples. The experimental results were only accepted if both blanks produced the expected negative C_2H_4 results. A positive control of 6 mL pure C_2H_4 gas from a C_2H_4 gas cylinder was injected to produce a positive C_2H_4 peak.

2.1.8 Glucose utilisation

2.1.8.1 Y88^{T} growth on glucose and acetate

Y88^{T} growth was investigated in batch flasks with glucose or acetate as carbon sources. PHA production, OD_{600} and biomass produced were monitored throughout the growth period. Y88^{T} growth curves were generated from OD_{600} measurements of cells grown in the presence of glucose or acetate as a single carbon source or glucose and acetate in combination. For Y88^{T} growth using either glucose or acetate separately in a minimal nitrogen medium, equivalent amounts of carbon were used for each compound. Acetate concentrations were 5 mM, 10 mM, 20 mM, 27 mM or 81 mM. The equivalent amount of carbon in glucose was 1.9 mM, 3.3 mM, 6.8 mM, 9 mM or 27 mM. This allowed an evaluation of these two carbon sources at low or high concentrations, independent of each other and comparable in terms of the amount of initial carbon present in the medium. Biomass and PHA accumulation were also determined.

Additionally, Y88^{T} growth was examined in batch flasks on combinations of glucose and acetate as a mixed carbon source to determine whether any change

in growth, cell yield and PHA accumulation could be detected as a consequence of growing in the presence of the mixed substrate, as compared to growth in the presence of either of these substrates alone. Since the purpose of growing the cells in the presence of a mixed substrate of glucose and acetate was to determine whether small amounts of glucose affected growth on acetate in any way, the mixed combinations of initial carbon substrate were as follows: 1.9 mM glucose/5 mM acetate, 3.3 mM glucose/10 mM acetate, 1.9 mM glucose/20 mM acetate and 3.3 mM glucose/20 mM acetate. An additional 27 mM glucose/27 mM acetate combination was used to compare the effect of excess glucose on acetate growth. All experiments were conducted in duplicate for each set of conditions being tested. The pH of the growth medium was determined at 0, 8, 12 and 24 hr post inoculation.

2.1.8.2 Comparative growth of Y88^T, *A. vinelandii* and *P. putida* mt2 on glucose under ammonium-free, ammonium-supplemented or ammonium-depleted conditions

Comparative growth of Y88^T, *A. vinelandii* and *P. putida* mt2 was examined in parallel in batch flasks (3 per species) containing minimal nitrogen medium enriched with glucose as the sole carbon source. The ability of Y88^T to fix nitrogen and accumulate PHA was investigated and compared with that of *A. vinelandii*. These experiments were carried out in two phases. Phase 1 investigated the ability of Y88^T to grow in nitrogen-free or nitrogen-supplemented, carbon-enriched (5 g/L) medium. Since no data were available for Y88^T growth in the presence of limiting glucose, triplicate controls containing the same medium but with less carbon (1 g/L), were included to determine whether this lower concentration could be considered limiting in terms of growth and PHA production for Y88^T. The activity of nitrogenase was examined under carbon excess, nitrogen-limiting and ammonium-supplemented growth conditions and the masses of PHA accumulated under these conditions were determined. *A. vinelandii* was used as a positive nitrogen-fixing, PHA-accumulating control and *P. putida* mt2 as a negative control. OD_{600} was measured at 0, 8, 12 and 24 hr post-inoculation. C_2H_2 reduction assays were

carried out on all three species when their growth reached maximum OD_{600}. Nitrogenase activity in the cells was measured in the C_2H_2 reduction assays as an indicator of the ability to fix molecular nitrogen under conditions appropriate for nitrogen fixation.

Phase 2 investigated $Y88^T$ growth in batch flasks in duplicate for each set of conditions being tested to examine the ability of $Y88^T$ to fix nitrogen in the presence of excess carbon and determine the amount of PHA accumulated. Since it was uncertain when exactly in the growth phase $Y88^T$ accumulated PHA, two sampling points were selected at different cell densities representative of low or high cell density. Cells were harvested for assays of nitrogenase activity at OD_{600} of approximately 0.5 and 1.0 as well as an OD_{600} greater than 2.0 (undiluted). For these experiments, the masses of PHA, dry weights as an indicator of biomass, glucose utilisation and total organic carbon (TOC), acetic acid formation and ammonium assimilation were determined over a 72 hr growth period. C_2H_2 reduction assays for nitrogenase activity were carried out at high cell densities at 22-24 hr post-inoculation. An initial (0 hr post-inoculation) and final (72 hr post-inoculation) pH were recorded.

2.1.8.3 Comparative growth of $Y88^T$ and WP01 during PHA accumulation in the presence of a single carbon substrate

$Y88^T$ and WP01 were grown in batch flasks in triplicate to examine PHA accumulation in a minimal nitrogen medium with either glucose or acetate as the sole carbon source. For these experiments, the masses of PHA, dry weights as an indicator of biomass and OD_{600} were determined over a 72 hr growth period.

2.1.9 Total organic carbon as a measure of dissolved organic carbon

An Elementar HiTOC analyser (Elementar, Hanau, Germany) and auto-injector was used to determine the total organic carbon (TOC) remaining in $Y88^T$ culture media after 24 hr of growth. Each sample was filtered twice with a 0.2 μm filter to remove possible contaminating cells/substances. A stock solution of 1.083 g

potassium hydrogen phthalate and 4.412 g of sodium carbonate was made up in 1000 mL of MQ H_2O and diluted to the appropriate standards (500 g/L, 250 g/L, 125 g/L, 25 g/L). To ensure that the results would be within the acceptable range of the standards used, all samples were diluted 1:4 prior to analysis. Samples were added to 25 mL glass scintillation vials with a 10 mm magnetic stirrer that was placed in the bottom of each vial. All samples were analysed in duplicate. The difference between the total carbon (TC) and the total inorganic carbon (TIC) was calculated as the DOC present in each sample. The DOC amount was subtracted from the initial amount of total carbon supplemented in the growth medium at the start of the experiment. The difference between these two values was calculated to be the amount of carbon utilised in the presence of $Y88^T$ cells during growth over a 24 hr period. A standard curve was generated from the standards used in the analysis.

2.1.10 Enzyme assays

2.1.10.1 Glucose utilisation assay

Glucose was measured using a Megazyme glucose assay (Megazyme International, Wicklow, Ireland) based on oxidation of glucose and measurement of the product H_2O_2 in a reaction catalysed by the enzyme peroxidase. The culture supernatant was filtered twice after centrifugation (1814 × g, 4 °C, 30 min) through a 0.2 μm filter and aliquots were incubated with the assay reagent at 40 °C for 15 min after 9x dilution with MQ H_2O. Following incubation, spectrophotometric measurements were carried out on the reacted product at 510 nm.

2.1.10.2 Acetic acid formation assay

Acetic acid formation was measured using a Megazyme (Wicklow, Ireland) Acetic Acid kit based on the use of acetyl-coenzyme A synthetase (ACS) to catalyse the reaction of acetic acid, ATP and CoA to give acetyl-CoA, AMP and pyrophosphate at room temperature. The culture supernatants used in the assays were prepared similarly to those in 2.1.10.1 above. The acetyl-CoA

formed in the reaction reacts with oxaloacetate and H_2O in the presence of citrate synthase to form citric acid and CoASH followed by catalysis of NAD^+ dependent L-malate by L-malate dehydrogenase to form oxaloacetate and reduced NADH, the production of which is measured by absorbance at 340 nm.

2.1.11 PHA determination

2.1.11.1 PHA content in Y88^T cells

Duplicate 15 mL cultures were centrifuged at 1814 × g, 4 °C to sediment the cells which were then dried overnight and subjected to PHA gas chromatographic analysis (Riis & Mai, 1988) to determine the amount and type of polymer that accumulated.

2.1.11.2 Determination of polymer characteristics

A CelLytic B Plus kit (Sigma) was used to extract PHA polymer from Y88^T cells for ^{13}C-NMR analysis. The kit lysis buffer working solution (made up as specified in the manufacturer's instructions for protein extraction), readily disrupted the inclusion bodies in the cells, precipitating the inclusion body compound in the insoluble material. This compound was manually removed from the cell debris and subjected to ^{13}C-NMR, infrared spectroscopy and gas chromatography (Riis & Mai, 1988).

2.2 Y88^T whole genome sequencing and curation

The Y88^T genome was shotgun sequenced using 454 Life Sciences (Branford, CT, USA) pyrosequencing to generate sequence reads with an average read length of 150 bases and an average sequence accuracy of >99%, resulting in approximately 99% genome coverage. Contig assembly was performed at 454 Life Sciences using the Newbler contig assembly programme. The Y88^T genome was estimated to be approximately 4.5 Mb of which 4.18 Mb have been sequenced.

The Y88^T genomic sequence was subjected to a BLAST search of the GenEMBL non-redundant database at Scion, Rotorua. Gene annotations were assigned to

identified open reading frames and used to create an in-house database that was used for protein identification from MS/MS data interpreted using MASCOT (Matrix Science, London, UK), (Strabala, Macdonald & Liu, unpublished results).

2.3 Proteomics

A 2-D DIGE proteomic approach (Fig. 2.2) was used to examine the response of $Y88^T$ to conditions conducive to nitrogen fixation (represented by the Fix^{on} physiological state of the cell) or no nitrogen fixation (represented by the Fix^{off} physiological state of the cell) in a carbon-enriched environment. MS of trypsinised protein from spots excised from Coomassie-stained 2-D DIGE gels, allowed the identification of $Y88^T$ proteins with altered abundance under different growth conditions. Using this approach, the key physiological parameters of nitrogen fixation, growth, carbon utilisation and PHA accumulation were correlated with their corresponding proteomic signatures. Initially, a linear gradient in the pH 3-10 range was used to gain an overview of the total distribution of $Y88^T$ proteins under the tested growth conditions. For increased resolution and a more detailed overview of acidic and basic protein distribution, linear gradient pH 4-7 and pH 6-11 strips were used.

2.3.1 Generation of samples for 2-D DIGE and LC-MS/MS analysis

2.3.1.1 2-D Gel Experimental Design 1

$Y88^T$ samples for each 2-D DIGE experiment were cultivated in batch mode (flasks) in triplicate for each set of conditions and harvested at two optical densities corresponding to Fix^{off} (OD_{600}=0.5, 1.0) and Fix^{on} (OD_{600}>2.0,) states (Table 2.2). Two preparative and three analytical 2-D DIGE gels were used for each sample (Fig. 2.3). Each preparative gel comprised pooled samples for the three replicates representing the Fix^{off} state of the cell (OD_{600}=0.5) (Fig. 2.3, A.1) and the Fix^{on} state of the cell (OD_{600}>2.0, undiluted) (Fig. 2.3, A.2). The 2-D DIGE gels (Fig. 2.3) contained three samples, two of them randomly selected and representing a different Fix^{on} or Fix^{off} state of the cell with the third

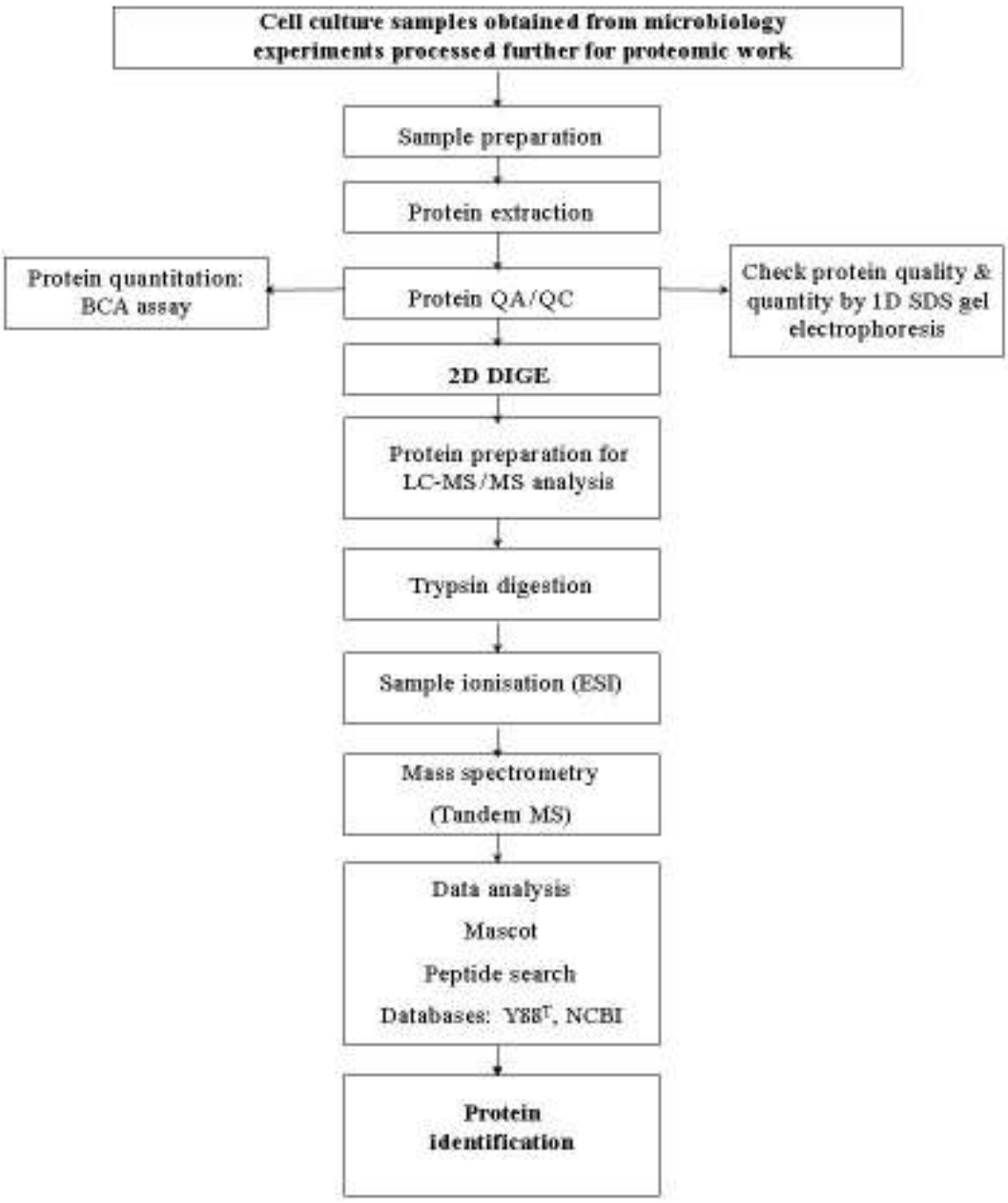


Figure 2.2. Schematic representation of workflow to investigate the proteome of Y88^{T} under different growth conditions.

Chapter reference	Experimental Design	Growth sampling phase/condition						
		NH_4^+-free OD_{600}=0.5	NH_4^+-supplemented OD_{600}=1.0	NH_4^+-depleted OD_{600}=2.0	NH_4^+-supplemented OD_{600}=0.5 DO^{low}	NH_4^+-depleted OD_{600}=2.0 DO^{low}	NH_4^+-supplemented OD_{600}=0.5 DO^{high}	NH_4^+-depleted OD_{600}=2.0 DO^{high}
		Physiological state of cells						
Chapter 4, section 4.2.2.1	1	Fix^{off}		Fix^{on}				
Chapter 4, section 4.2.2.2	2	Fix^{off}	Fix^{off}	Fix^{on}				
Chapter 5, section 5.2.1.3	3 (pH not controlled)				Fix^{off}	Fix^{on}		
Chapter 5, section 5.2.1.4-5.2.1.6	4 (pH controlled)				Fix^{off}	Fix^{on}	Fix^{off}	Fix^{off*}

Table 2.2. Physiological state of $Y88^T$ cells defined for a specific set of growth conditions.

Fix^{off}=negative acetylene reduction; Fix^{on}=positive acetylene reduction; Fix^{off*}=decreased acetylene reduction but similar to Fix^{on} in terms of cell productivity.

Note in chapter 5 and thereafter, Fix^{off*} will be referred to as Fix^{off}.

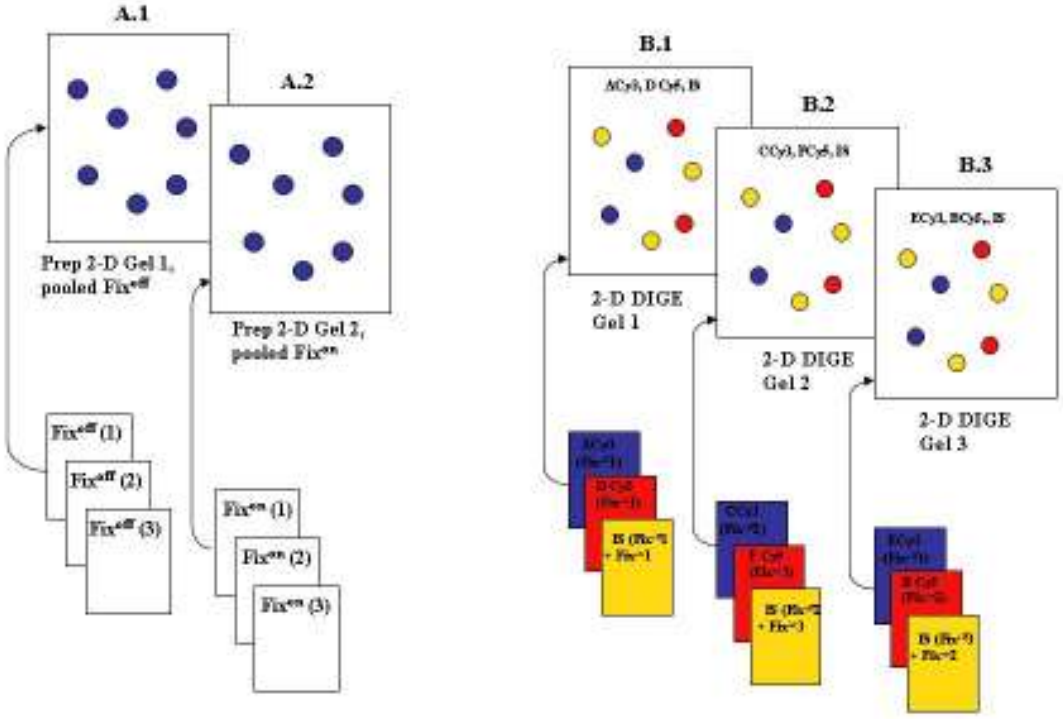


Figure 2.3. 2-D gel experimental design 1

(A) Preparative 2-D gels representing pooled replicate protein samples for Y88^{T} cells grown in ammonium-free medium in which the cells did not fix nitrogen (Fix^{off} state of the cells) (A.1) and a second gel representing pooled replicate protein samples for Y88^{T} cells grown in medium conducive to nitrogen fixation (Fix^{on} state of the cells) (A.2); (B) 2-D DIGE gels representing combinations of protein samples comprising, per gel, one sample (Fix^{off}) labeled with CyDye 3 and the other sample (Fix^{on}) with CyDye 5. An internal standard consisting of equal quantities of Fix^{off} and Fix^{on} was labeled with CyDye 2. The two different states of Fix^{on} and Fix^{off} resulted in 6 biological replicates (3 for each condition) resulting in 3 (B.1-B.3) 2-D DIGE analytical gels and 2 preparative 2-D gels (A.1, A.2).

being a pooled sample of equal quantities of Fix^{on} and Fix^{off} samples representing the internal standard. The combination of samples for each of the two states of the cell were included on a gel with an internal standard allowing a comparison at the protein level of the physiological state (Fix^{off} and Fix^{on}) of the cell under ammonium-free and ammonium-supplemented growth conditions respectively. Coomassie-stained protein spots were excised, reacted with trypsin to digest protein and analysed by MS to identify as many proteins as possible as well as to determine which, if any of these, were involved in nitrogen fixation and PHA accumulation under the selected growth conditions. Protein spots occurring on the gels at predicted isoelectric points and molecular weights for the nitrogenase enzymes and PHA accumulating proteins (based on relevant literature on these proteins from other bacteria) were targeted for analysis. Also selected were additional highly abundant proteins as well as low molecular weight basic and acidic protein spots since these are often of interest as bioactives. All spots picked were further processed for LC-MS/MS analysis.

2.3.1.2 2-D Gel Experimental Design 2

$Y88^T$ samples for each 2-D DIGE experiment were cultivated in batch mode (1 L Ehrlenmeyer flasks, working volume 500 mL) in triplicate for each set of conditions and harvested at two optical densities corresponding to Fix^{off} (OD_{600}=0.5), Fix^{off} (OD_{600}=1,0) and Fix^{on} (OD_{600}>2.0, undiluted) states (see conditions for Fix^{off} and Fix^{on} state of the cells for experimental design 2, Table 2.2). Three preparative and five analytical 2-D DIGE gels were generated (Fig. 2.4). Each preparative gel comprised pooled samples for the three replicates representing i) the Fix^{off} state of the cell (OD_{600}=0.5, Fig. 2.4, A.1), ii) the Fix^{off} state of the cell (OD_{600}=1.0, Fig. 2.4, A.2) and the Fix^{on} state of the cell (OD_{600}>2.0, undiluted, Fig. 2.4, A.3). The 2-D DIGE gels (Fig.2.4, B.1-B.5) contained three samples, two of them randomly selected and representing a different Fix^{on} or Fix^{off} state of the cell as defined (Table 2.2, Experimental Design 2) with the third being a pooled sample of equal quantities of Fix^{on} and Fix^{off} samples representing the internal standard. The combination of samples for each of the two states of the cell plus the internal standard included per 2-D

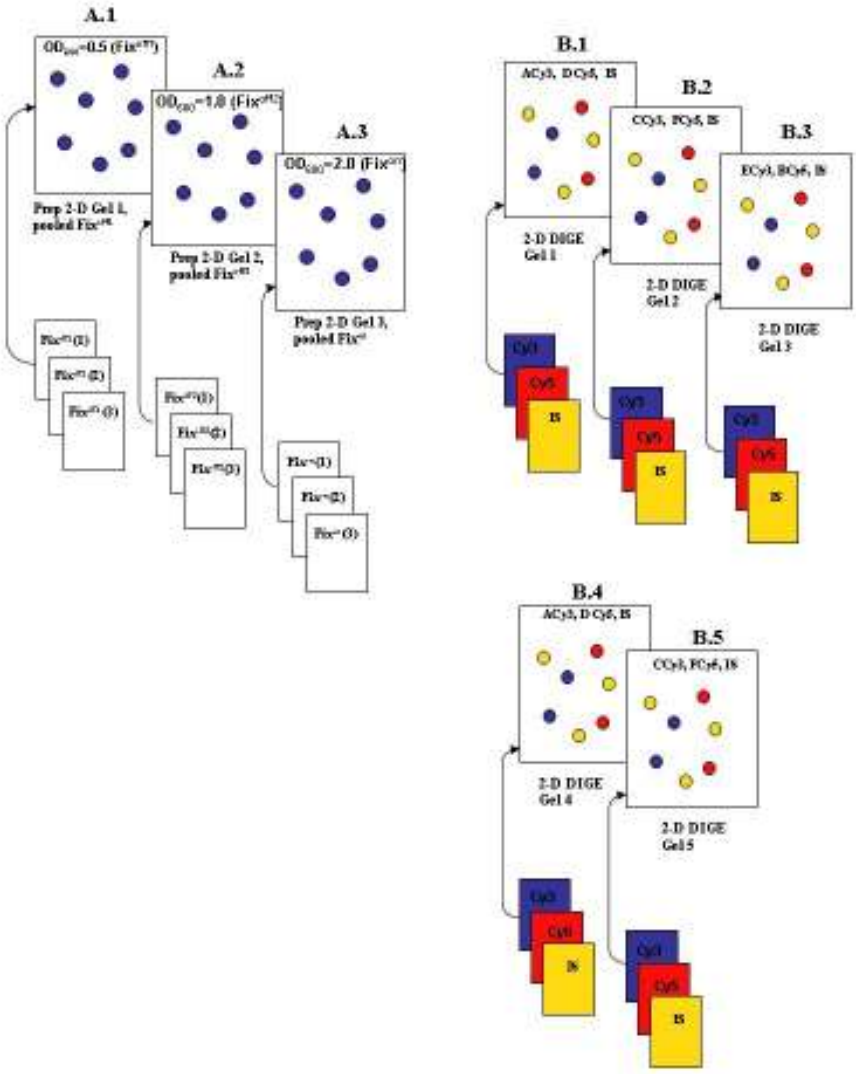


Figure 2.4. 2-D gel experimental design 2

(A) Preparative 2-D gels representing pooled replicate protein samples for cells grown to an OD_{600}=0.5 in ammonium-supplemented medium in which cells did not fix nitrogen ($Fix^{off}1$) (A.1), to an OD_{600}=1.0 in ammonium supplemented medium in which cells did not fix nitrogen (Fix^{off}) (A.2) and to an OD_{600}=2.0 (A.3) in nitrogen-depleted medium in which cells fixed nitrogen. (B) 2-D DIGE gels representing combinations of two randomly selected samples, one from each state, per gel labeled with either CyDye 3 and or CyDye 5. An internal standard (IS) representing equal quantities of sample for each state of the two states of the cell was labeled with CyDye 2. Thus, for an experiment with three different conditions, 9 biological replicates were generated (3 for each condition) resulting in 5 (B.1-B.5) 2-D DIGE analytical gels and 3 preparative 2-D gels (A.1, A.2, A.3).

DIGE gel, allowed a comparison at the protein level of the physiological state (Fix^{off} and Fix^{on}) of the cell under ammonium-supplemented conditions during which $Y88^T$ cells did not fix nitrogen and ammonium-depleted conditions during which $Y88^T$ cells fixed nitrogen. Protein spots excised from the 2-D gels were processed as before (Section 2.3.1.1).

2.3.1.3 2-D Gel Experimental Design 3

$Y88^T$ samples for each 2-D DIGE experiment were cultivated in parallel 1 L bioreactor vessels in batch mode for 24 hr and repeated three times to generate three independent samples for each set of growth conditions representing the Fix^{off} and Fix^{on} state of $Y88^T$ cells (see Table 2.2, Experimental Design 3). Two preparative and three analytical 2-D DIGE gels were generated (Fig. 2.5). Each preparative gel comprised pooled samples for the three replicates representing i) the Fix^{off} state of the cell at high growth, DO^{high} with pH not controlled and ii) the Fix^{on} state of the cell state at high growth, DO^{low} with pH not controlled. The 2-D DIGE gels (Fig.2.5, B.1-B.3) contained three samples, two of them randomly selected and representing a different Fix^{on} or Fix^{off} state of the cell (Table 2.2, Experimental Design 3) with the third being a pooled sample of equal quantities of Fix^{on} and Fix^{off} samples representing the internal standard. The combination of samples for each of the two states of the cell plus the internal standard included per 2-D DIGE gel, allowed a comparison at the protein level of the physiological state of the cell at high growth (OD_{600}=2.0) under ammonium-depleted conditions at DO^{high} and DO^{low} respectively with pH not controlled. Protein spots excised from the 2-D gels were processed as before (Section 2.3.1.1).

2.3.1.4 2-D Gel Experimental Design 4

$Y88^T$ samples for each 2-D DIGE experiment, were cultivated in parallel 1 L bioreactor vessels in batch mode and repeated three times to generate three independent samples for each set of growth conditions representing the Fix^{off} or Fix^{on} state of $Y88^T$ cells (see Table 2.2, Experimental Design 4). Four preparative and six analytical 2-D DIGE gels were generated (Fig. 2.6). Each preparative gel

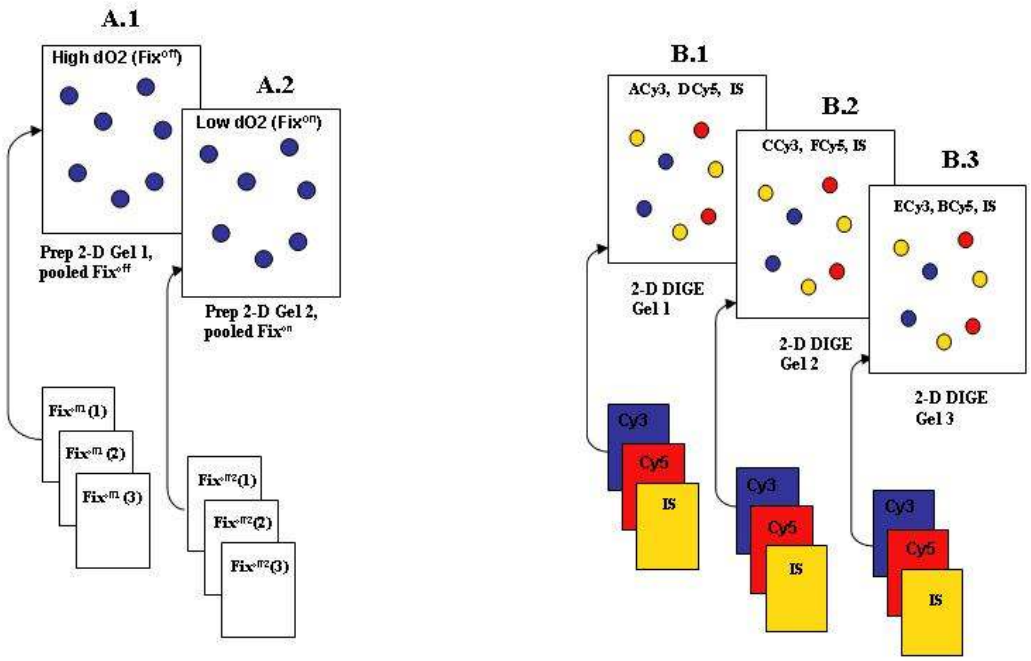


Figure 2.5. 2-D gel experimental design 3

(A) Preparative 2-D gels representing pooled replicate protein samples for Y88^{T} cells grown at DO^{high} at an OD_{600}=2.0, pH not controlled (Fix^{off}) (A.1) and DO^{low} at an OD_{600}=2.0, pH not controlled (Fix^{on}) (A.2); (B) 2-D DIGE gels representing combinations of protein samples comprising, per gel, one sample from DO^{high} labeled with CyDye 3 and one from DO^{low} with CyDye 5. An internal standard representing equal quantities of each of the two samples was labeled with CyDye 2. Thus, for an experiment with two different conditions, 6 biological replicates were generated (3 for each condition) resulting in 3 (B.1-B.3) 2-D DIGE analytical gels and 2 preparative 2-D gels (A.1, A.2).

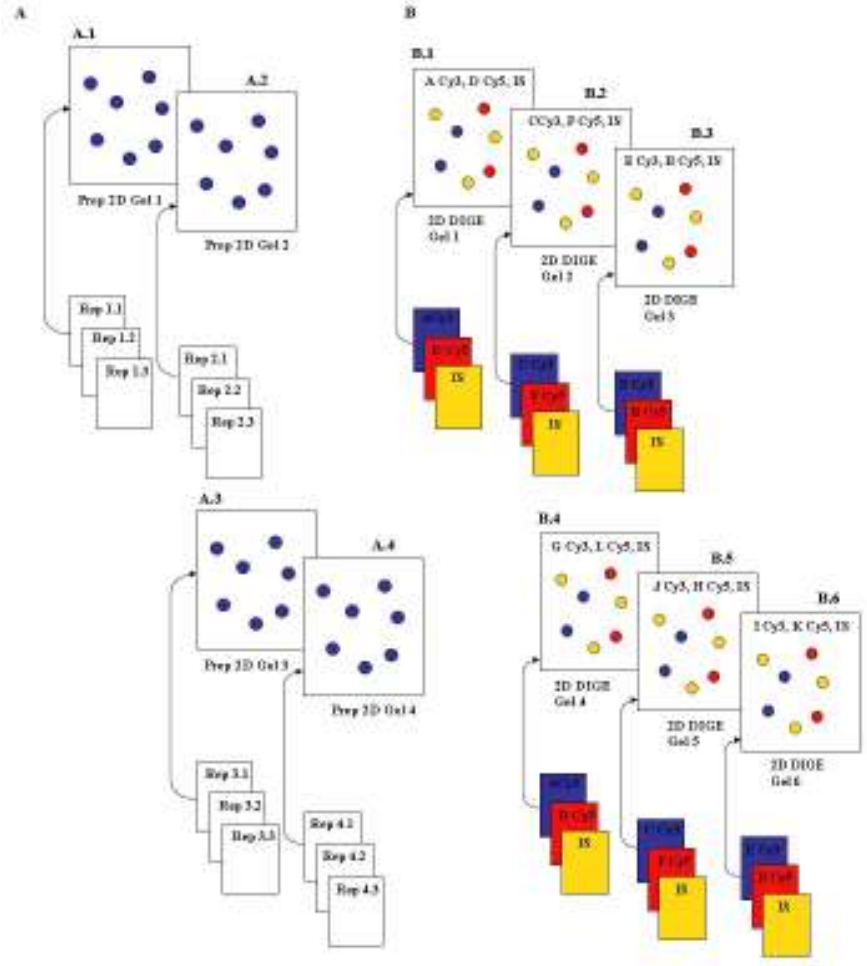


Figure 2.6. 2-D gel experimental design 4

(A) Preparative 2-D gels representing pooled replicate protein samples for Y88^{T} cells grown at DO^{high} at an OD_{600}=0.5 and 2.0, pH controlled (Fix^{off} state of the cells) (A.1) and DO^{low} at an OD_{600}=0.5 and 2.0, pH controlled (Fix^{on} state of the cells) (A.1) and a second gel representing pooled replicate protein samples for the second state of the cells (A.2); (B) 2-D DIGE gels representing combinations of protein samples comprising, per gel, one sample (cell lysate) labeled with CyDye 3 and the other with CyDye 5 for each of the two states of the cell and an internal standard, labeled with CyDye 2. Thus, for an experiment with two different conditions, 6 biological replicates were generated (3 for each condition) resulting in 3 (B.1-B.3) 2-D DIGE gels and 2 preparative 2-D gels (A.1, A.2). For an experiment with 4 different conditions, 12 biological replicates were generated (3 for each condition) resulting in 6 (B.1-B.6) 2-D DIGE analytical gels and 4 (A.1-A.4) preparative 2-D gels.

comprised pooled samples for the three replicates representing i) the Fix^{off} state of the cell at low growth, DO^{high} with pH controlled and ii) the Fix^{off} state of the cell at high growth, DO^{high} with pH controlled, iii) the Fix^{off} state of the cell at low growth, DO^{low} with pH controlled and iv) the Fix^{on} state of the cell state at high growth, DO^{low} with pH controlled. Each 2-D DIGE gel (Fig.2.6, B.1-B.6) contained three samples, two of them randomly selected and representing a different Fix^{on} or Fix^{off} state of the cell (Table 2.2, Experimental Design 4) with the third being a pooled sample of equal quantities of Fix^{on} and Fix^{off} samples representing the internal standard. The combination of samples for each of the two states of the cell plus the internal standard included per 2-D DIGE gel, allowed a comparison at the protein level between the physiological state of the cell at low (ammonium-supplemented) and high (ammonium-depleted) growth at DO^{high} and the physiological state of the cell at low (ammonium-supplemented) and high (ammonium-depleted) growth at DO^{low} respectively with pH controlled. Protein spots excised from the 2-D gels were processed as before (Section 2.3.1.1).

2.3.2 Sample preparation

Care was taken to minimise differences between samples due to handling and preparation, ensuring consistency across all experiments for all conditions tested. For all proteomic analyses, $Y88^T$ cell cultures were grown to a predetermined OD_{600} according to the experimental design in either batch flasks or bioreactors and harvested, using a 60 cm^3 syringe attached to a sampling port. Particular care was taken to minimise the formation of air bubbles that could potentially introduce a bias for the DO^{low} or DO^{high} oxygen results. Following harvesting, all samples were immediately incubated on ice and processed directly. At least 120 mL of culture for each protein sample was centrifuged at 1814 × g, 4 °C, for 30 min. Most of the spent medium was decanted, leaving behind approximately 1 mL of medium to resuspend the cells and facilitate transfer to 1.5 mL microcentrifuge tubes. Microcentrifuge tubes were tared and weighed before transferring the cells to the tubes. The cells were again centrifuged, this time in a microcentrifuge at 15,682 × g and any

remaining spent medium was removed with a pipetter, leaving only the cell pellet. Cell pellets were stored at -20 °C until further processing for protein extraction.

2.3.3 Protein extraction

Total cellular protein was extracted using a CelLytic B Plus Kit (Sigma), known to be particularly useful in lysing Gram-negative bacteria. The kit contained a proprietary bacterial lysis reagent plus lysozyme, benzoase (an exogenous nuclease to minimise nucleic acid contamination in the protein extract) and a protease inhibitor cocktail (Sigma) to prevent the proteolytic breakdown of proteins. A minimum of 0.5 g (wet weight) per sample, equivalent to approximately 0.2-0.3 g (dry weight) of Y88^T cells, was required when using the protein extraction CelLytic B Plus Kit to ensure sufficient protein for downstream analyses. This criterion was achieved by growing a 500 mL culture to an OD_{600} of greater than 0.5, then centrifuging 60 mL of the culture per sample to obtain a cell pellet of approximately 0.5 g. Since three replicates for each experimental condition were generated from growing cells in triplicate flasks, three independent protein samples were obtained for each condition being tested. A working solution comprising the bacterial lysis reagent (5 mL), lysozyme (0.1 mL), protease inhibitors (0.05 mL) and benzoase (0.76 µL = 250 units) was used to resuspend the cell pellet (0.4 mL of working solution per 0.5 g cell pellet). The suspension was briefly vortexed (approximately 10 sec) and then mixed by repeatedly inverting for 10 min to ensure full extraction of the soluble proteins. The cell lysate was centrifuged at 15,682 × g, 4 °C, for 5 min to pellet any insoluble material. The soluble proteins were transferred to a new microcentrifuge tube and the protein concentration for each extract was determined. Protein extracts were stored at -20 °C until required for further protein analysis.

2.3.4 Protein quantitation

2.3.4.1 BCA protein assay

All protein quantitation was carried out using the Bicinchoninic acid (BCA) kit (Pierce, Rockford). BCA is a stable and sensitive reagent that is highly specific for the detection of cuprous ion (Cu^{+}). When incubated with protein, a Cu^{2+} - protein complex is formed which is reduced to Cu^{+} under alkaline conditions. The reduction of Cu^{2+} to Cu^{+} depends on the amount of protein present with more protein resulting in more Cu^{+} being produced. A purple colour develops as a result of the chelation of one Cu^{+} ion with two molecules of BCA. This Cu^{+}/BCA complex is water soluble and exhibits a strong absorbance maximum at 562 nm.

To quantitate the protein, a bovine serum albumin (BSA) standard curve was created for comparison to the experimental samples. BSA was diluted according to the manufacturer's instructions to give nine standards of varying BSA concentrations. Ten µL of each BSA standard was mixed with 190 µL of BCA working solution made up as a 50:1 ratio of BCA to copper solution and incubated at 30 °C for 1 hr alongside the experimental protein samples. The experimental protein samples were measured by diluting an aliquot of extracted protein five-fold, and then mixing 10 µL of this diluted sample with 190 µL BCA working solution. Following incubation, the absorbances were measured at 562 nm using a PC-controlled GBC spectrophotometer (model UV/VIS918). BSA concentrations were plotted with A_{562} values to generate the standard curve and concentrations of the experimental samples were derived from this.

2.3.5 Gel electrophoresis

2.3.5.1 Assessment of protein by 1-D SDS-PAGE

An XCell Surelock™ Mini-Cell NuPAGE gel apparatus (Invitrogen, Carlsbad, CA) was used for 1-D SDS-PAGE of protein extracts to assess both the quality

and quantity of the protein, as well as to gain a preliminary insight into any potential altered abundance in protein between growth conditions. Protein extracts were diluted four-fold (5 µL of protein in 20 µL of MQ H_2O) and 20 µL of the diluted extract was mixed with 20 µL of 2× concentration reduced sample running buffer (100 mM DTT, 80 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, bromophenol blue solution, made up to 25 mL MQ H_2O) to give a final concentration of 1× reduced sample running buffer. Aliquots of 0.5 mL were stored at -20 ° C. Five µL aliquots were electrophoresed at 200 V and 400 mA for 35 min. Protein separation was carried out using 10% Novex pre-cast gels (1.00 mm × 15 wells) in 800 mL of 1× MES (diluted from a 20× stock solution (50 mM MES, 50 mM Tris base, 0.1% SDS, 1 mM EDTA in 1000 mL MQ H_2O) running buffer. An unstained standard protein marker, Mark12™ (Invitrogen), was used to estimate the molecular weights of the separated proteins.

2.3.5.2 Gel washing, staining and destaining of 1-D gels

Following electrophoresis, gels were removed from the gel cassettes and briefly rinsed in MQ H_2O. Gels were then fixed in a solution of 40% acetic acid and 10% methanol in MQ H_2O for at least 15 min. After rinsing twice in MQ H_2O for 2 min per rinse, gels were stained in colloidal Coomassie Brilliant Blue stain (Fluka; Sigma-Aldrich Group) which is an anionic triphenylmethane dye that binds noncovalently to the lysyl residues of proteins and has a detection limit of about one microgram of protein per spot. Any dye that is not bound to protein diffuses out of the gel during the destaining process. A 5% w/v stock solution of Coomassie Brilliant Blue G-250 (Fluka; Sigma-Aldrich Group) was used to make up a staining solution containing 10% w/v $(NH_4)_2SO_4$, 1% v/v phosphoric acid and 2% v/v Coomassie Blue stock solution. All gels were incubated at room temperature in the stain solution on an orbital shaker at 40 rpm overnight. Gels were destained in MQ H_2O to remove excess stain and then photographed using a Canon EOS SLR digital camera. Gel images were saved as JPG files.

2.3.5.3 2-DE

2.3.5.3.1 Preparative gel sample preparation

Protein in microbial extracts was precipitated with nine volumes of cold 100% acetone at -20 °C for a minimum of 2 hr. Precipitated protein was collected by centrifugation at 15,682 × g for 20 min at 4 °C. The acetone was removed by decanting carefully and the pellet left to air dry for approximately 10 min at room temperature.

2.3.5.3.2 Sample rehydration and loading onto IPG strip

Protein (50 µg per sample) from three pooled technical replicates was resolubilised in rehydration buffer (2 M thiourea, 7 M urea, 2% (v/v) IPG buffer, 2% (w/v) DTT, 4% (w/v) CHAPS and a trace of bromophenol blue for pH 4-7 IPG strips; 2 M thiourea, 7 M urea, 1% (v/v) IPG buffer (v/v), 2.5% (w/v) DTT, 2% (w/v) CHAPS, 5% (v/v) glycerol, 10 % (v/v) isopropanol in MQ H_2O for pH 6-11 IPG strips) to a final volume of 125 µL and incubated for 15 min at room temperature. Each sample was then centrifuged for 1 min at 15,682 × g to remove the insoluble material, and loaded onto either an Immobiline™ DryStrip (linear gradient) 7 cm pH 3-10, pH 4-7 or pH 6-11 strip (GE Healthcare, Uppsala, Sweden) according to the experimental design. The sample was distributed across the length of the gel strip. Gel strips were placed gel side down in an Immobiline™ Drystrip Reswelling Tray (GE Healthcare, Uppsala, Sweden) with each gel strip covered with 3 mL of PlusOne DryStrip Cover Oil (GE Healthcare, Uppsala, Sweden). The gel strips were incubated overnight at room temperature.

2.3.5.4 2-D DIGE

2.3.5.4.1 Sample preparation

Protein in microbial extracts was precipitated with nine volumes of cold 100% acetone at -20 °C for a minimum of 2 hr. Precipitated protein was collected by centrifugation at 15,682 × g for 20 min at 4 °C. The acetone was removed by decanting carefully and the pellet left to air dry for approximately 10 min at

room temperature. The pellet was then solubilised in 2 µL cell lysis buffer (30 mM Tris-Cl, 7 M urea, 2 M thiourea, 4% (w/v) CHAPS in MQ H_2O) to a final volume of 5 µg protein /µL, providing 10 µg total protein per sample required for each 2-D DIGE gel. The pH of each sample was tested to ensure a final pH after addition of lysis buffer of 7.5-8.5.

2.3.5.4.2 Sample labeling

Protein extracts were labeled for DIGE analysis using an Ettan DIGE minimal labeling kit (CyDye™ DIGE Fluor minimal dye) according to the manufacturer instructions (GE Healthcare, Uppsala, Sweden). Minimal labeling is stated to react with a single lysine per protein molecule (Marouga et al., 2005). CyDyes were reconstituted in dimethylformamide to a final concentration of 1 mM working CyDye Fluor solution. Ten µg protein per sample was labeled with 80 pmol of working CyDye Fluor solution, briefly vortexed (10 sec) and centrifuged for 5 sec in a microcentrifuge to ensure collection of the solution at the bottom of the tube. Samples were incubated on ice for 30 min in the dark to minimise degradation and photo-bleaching of fluorescence dyes. Following incubation on ice, 1 µL of 10 mM lysine was added to each sample to stop the reaction. The solution was mixed well, centrifuged briefly (5 sec) and incubated on ice for 10 min in the dark.

2.3.5.4.3 Sample mixing and loading onto IPG strip

Each set of samples was loaded onto an IPG strip as determined by the experimental design (Fig. 2.1) following resolubilisation in rehydration buffer (see section 2.3.5.3.2). A total of 30 µg protein (10 µg for each of two cell states and 10 µg for the internal standard) was loaded per strip. To estimate the variability in sample handling, sample loading as well as the inherent variability in carrying out the 2-D DIGE technique, one sample was labeled with all three CyDyes (Cy2, Cy3 and Cy 5) separately. These were loaded onto one IPG strip. Initially, the pH 3-10 range was chosen to obtain an overview of total protein distribution. Subsequently, narrower 4-7 and 6-11 pH range strips were used to increase resolution of the protein spots on the gels, providing a

more detailed overview of acidic (pH 4-7) and basic (pH 6-11) protein distribution.

2.3.5.4.4 Sample rehydration

Rehydration of protein samples were carried out as before (see section 2.3.5.3.2) and samples loaded onto an Immobiline™ DryStrip (linear gradient) 7 cm pH 3-10, pH 4-7 or pH 6-11 strip (GE Healthcare, Uppsala, Sweden) as before according to the experimental design. The gel strips were incubated overnight at room temperature in the dark to minimise photo-bleaching of the fluorescence dyes.

2.3.5.4.5 First dimension IEF

After gel strip rehydration, each sample was subjected to IEF using either an IPGPhor™ IEF Unit (pH 3-10, pH 4-7) or a Multiphor™ II Flatbed Electrophoresis Unit (pH 6-11) according to manufacturer's (GE Healthcare, Uppsala, Sweden) recommended protocols (Table 2.3). During focusing of the IPG strips, the gel strips were protected from light by placing aluminium foil on the outside of the apparatus to minimise photo-bleaching of the fluorescence dyes. The same practice was applied to the buffer tank during SDS-PAGE.

2.3.5.4.6 Protein Reduction

On completion of IEF, each gel strip was placed in a reduction buffer (50 mM Tris, 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS, 10 % DTT in MQ H_2O) for 15 min to reduce protein disulfides to dithiols. Fifteen mL Falcon tubes (covered in aluminium foil to protect the gel strips from light and to minimise photo-bleaching of the fluorescence dyes) containing 10 mL of reduction buffer were used per gel strip and the tubes were placed on a rocker at room temperature to undergo reduction for 15 min. Following reduction, the gel strips were removed

pH 3-10, pH 4-7				pH 6-11			
Step	**Type**	**Voltage**	**Time (min)**	**Step**	**Type**	**Voltage**	**Time (min)**
1	Step and hold	300V	30	1	Gradient	200	0.01
2	Gradient	1000V	30	2	Gradient	3500	90
3	Gradient	5000V	90	3	Gradient	3500	65
4	Step and hold	5000V	25				

Table 2.3. Protocols for first dimension IEF focusing of protein

Protocols for first dimension IEF focusing of protein using IPG strips in the linear gradient pH 3-10 and pH 4-7 range on the IPGPhor™ IEF Unit and IPG strips in the linear gradient pH 6-11 range on the Multiphor™ II Flatbed Electrophoresis Unit (temperature=10 °C, current=2 mA, power=5 W).

from the tubes and excess buffer was drained off each gel strip by turning each one on its side for all sides of the strip, blotting the edges on a lint-free laboratory wipe.

2.3.5.4.7 Protein Alkylation

After reduction, each gel strip was transferred to an alkylation buffer (50 mM Tris, 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS, 2% (w/v) iodoacetamide in MQ H_2O) for 15 min. Alkylation of the free cysteine thiols in the protein prevents disulfide bridges from reforming. Fifteen mL Falcon tubes were covered in aluminium foil to protect the gel strips from light so as to minimise photo-bleaching of the fluorescent dyes as well as to prevent degradation of the light-sensitive IAA. Each tube contained 10 mL of alkylation buffer per gel strip and the tubes were placed on a rocker to incubate at room temperature during alkylation.

2.3.5.4.8 Second dimension protein separation

The protein on each strip was subjected to second dimension separation in an XCell Surelock™ Mini-Cell NuPAGE gel apparatus (Invitrogen, Carlsbad, CA). Electrophoresis was carried out on each focused strip using 10% Bis-Tris NuPAGE gels (Invitrogen, Carlsbad, CA), at 200 V, 400 mA constant for 1 hr in 800 mL of running buffer (1× MOPS). Five hundred µL antioxidant (Invitrogen) was added to the upper chamber of the gel apparatus to maintain the protein in the reduced state during electrophoresis.

2.3.6 Gel image acquisition

Electrophoresed gels were rinsed in MQ H_2O and scanned using a Fuji FLA-5000 Imaging System (FUJIFILM, Life Science). Each fluorescence dye was consecutively excited to avoid fluorescence "crosstalk" and scanned with the appropriate filter (Cy2, 473 nm; Cy3, 520 nm; Cy5, 600 nm). After scanning, each gel image was analysed using the ImageQuant™ image analysis software (ImageQuant v5.2, Amersham Biosciences) to determine the limit of detection and the linear dynamic range. The volume review tool was used to determine

whether the image exceeded or was below the linear dynamic range (Max. Val >50,000). A lower or higher photomultiplier tube setting was selected to bring all the protein spots within the linear range where required.

2.3.6.1 Gel Image analysis

DeCyder 2-D image analysis software (ImageQuant v5.2, Amersham Biosciences) was used for all gel analyses. Gel images were aligned and cut to the region of interest using ImageQuant v5.2 Tools. Thereafter, each gel image was analysed using the DeCyder DIGE Differential In-gel Analysis (DIA) interface software (ImageQuant v5.2, Amersham Biosciences) with the optimal settings for spot detection and exclusion (estimated spot number set to 2500 spots). The selected parameters for spot detection and exclusion were also applied to the DeCyder DIGE Batch Processor (ImageQuant v5.2, Amersham Biosciences). All gels in a batch were processed and analysed under the same parameters and automatically matched. Individual protein patterns were assigned to groups dependent on the conditions they were representing as dictated by the experimental design.

Following batch processing of the gels, the complete batch was loaded into the DeCyder Biological Variation Analysis (BVA) software (ImageQuant v5.2, Amersham Biosciences). This software matches the different gel images across groups and provides statistical data on differential protein abundance levels between groups. Mismatches within the analysis were corrected manually by setting landmarks and adding spot matches. The groups of gels that had been assigned in the batch processor were compared and statistical *t*-test values were calculated. The obtained data were filtered according to the parameter set, with a threshold factor of two between the groups, a *t*-test value i) $p<0.01$ and the spot detected in at least 99% of the analysed protein patterns and (ii) $p<0.05$ and the spot detected in at least 95% of the analysed protein patterns. Despite this critical selection, each difference that was detected was further manually checked and where required, spots were rematched. All confirmed spots were marked and the gel image was captured in a screen shot. Analytical spot data,

e.g. *t*-test values and the average ratios of the differences computed, were exported to a Microsoft Excel file together with the master numbers of the spots.

2.3.7 Staining 2-D DIGE gels

Electrophoresed preparative gels were rinsed briefly with MQ H_2O and soaked separately in fixative (5% (v/v) ethanol, 3% (v/v) phosphoric acid) for at least one hr, following which they were stained with Coomassie Brilliant Blue G-250 (Fluka; Sigma-Aldrich Group). Each gel was then separately washed three times for 30 min periods in MQ H_2O. Following washing, each gel was separately stained by saturating with dye solution (17% (w/v) NH_4SO_4, 34% (v/v) MeOH, 3% (v/v) phosphoric acid) and soaking the gel for at least one hr in this solution. After one hr of soaking, 0.06% (w/v) Coomassie Brilliant Blue G-250 (Fluka; Sigma-Aldrich Group) was added and the gels were stained for 3-4 days.

2.3.8 Protein spot picking

Coomassie-stained protein spots were manually excised from the gels using a manual One Touch Plus Spot Picker (1.5 mm diameter, The Gel Company, San Francisco, California) and transferred to a sterile 96 Well V Bottom Polypropylene plate (Greiner Bio-One, Germany). Each gel spot corresponding to the particular spot number on a master gel (that was automatically assigned to the spot map with the highest number of detected spots) was recorded. Spot numbers were determined in the batch processor during batch analysis of the gel images.

2.3.9 Protein preparation for LC-MS/MS analysis

The protein in each spot was prepared for mass spectrometric analysis using the procedure as detailed in the remainder of this section.

2.3.9.1 Destaining of excised gel spots

All excised gel spots were washed twice in 100 µL of MQ H_2O and destained by washing three times in 200 µL of 50 mM NH_4HCO_3/ACN (1:1) at 2 hr intervals.

Following dehydration, the gel pieces were rehydrated twice in 100 µL ACN for 10 min at a time and then air dried in the laminar flow hood for 1 hr.

2.3.9.2 In-gel reduction and alkylation of protein prior to LC-MS/MS analysis

Once the gel pieces were destained, the protein in the gel pieces was reduced to ensure that all monothiols were maintained in a reduced state for tandem mass spectrometric analysis. To achieve reduction, approximately 10 µL of 10 mM DTT (enough to cover the gel piece containing a protein spot) was made up to a final concentration of 0.1 M NH_4HCO_3 and added to each well containing one gel spot. The gel pieces were then incubated at 56 °C for 30 min. Following incubation, the samples were briefly centrifuged and excess liquid was removed by aspirating with a pipetter. Ten µL of ACN was then added to each well to shrink the gel piece. The gel pieces were centrifuged and the liquid was removed by aspirating with a pipetter. The gel pieces were dried in a vacuum centrifuge (Heraeus, Hanau, Germany) following which 50 µL of 55 mM iodoacetamide in a final concentration of 0.1 M NH_4HCO_3 was added to reswell the gel pieces. The gel pieces were incubated in the dark at room temperature for 20 min. Following alkylation, the iodoacetamide solution was removed by aspirating with a pipetter and the gel pieces were then washed in 150-200 µL of 0.1 M of NH_4HCO_3 for 15 min and centrifuged. The liquid was removed before adding 50 µL of ACN to the gel pieces to shrink them. They were then centrifuged again, the liquid was removed by aspirating with a pipetter and the gel pieces dried in a vacuum centrifuge (Heraeus, Hanau, Germany).

2.3.9.3 In-gel enzymatic digestion of protein

The dried gel pieces were rehydrated in 2 µL (12.5 ng/µL) trypsin (2.5 µg of trypsin Modified Sequencing grade (Roche, Mannheim, Germany) in 50 µL of 50mM NH_4HCO_3). An additional 10 µL of 50mM NH_4HCO_3 without trypsin was then added to the gel pieces and incubated on ice for approximately 1 hr. Once all the trypsin solution had been absorbed by the gel pieces, an additional 15 µL of the same buffer (50mM NH_4HCO_3 without trypsin) was added to

maintain swelling of the gel during enzymatic digestion so that the trypsin could have maximum access to the protein. Digestion was carried out overnight in a 37 °C water bath.

2.3.9.4 Extraction of peptides following protein digestion

Following overnight enzymatic digestion of the protein, 30 µL of 25 mM NH_4HCO_3 was added to the gel pieces and incubated in a 37 °C water bath for 2 hr with shaking. The liquid was then transferred to a microcentrifuge tube (1 microfuge tube per sample well) to pool with all liquid subsequently collected from the corresponding well. To ensure maximum recovery of the digested peptides, enough ACN (1-2 times the volumes of the gel pieces) was added to the gel pieces and again incubated in a 37 °C water bath for 1 hr with shaking. All liquid was again transferred from the well to the corresponding microcentrifuge tube (1 per sample well) to pool with the previously collected liquid. Thereafter, approximately 40-50 µL of 5 % (v/v) formic acid was added to each well and was incubated for 2-3 hr with shaking in a 37 °C water bath. The liquid was transferred as before from the well to the corresponding microcentrifuge tube, pooling with the liquid previously collected. An additional quantity of ACN (1-2 times the volumes of the gel pieces) was again added to the gel pieces in the wells and incubated in a 37°C water bath for 30 min to 1 hr with shaking. The liquid was then collected and transferred to the corresponding microcentrifuge tube (1 per sample well) and pooled with all previously collected liquid. All pooled extracts were dried to approximately 5 µL in a vacuum centrifuge.

2.3.9.5 C_{18} membrane binding of peptides for reversed phase LC-MS/MS

ZipTip C_{18} pipette tips (Millipore), appropriate for peptides and low molecular weight proteins, contain C_{18} beads that hydrophobically bind peptides via hydrophobic interactions. Maximum binding occurs in the presence of TFA (0.1% (v/v) or other ion-pairing agents. The binding of peptides to the membrane and subsequent elution of the peptides allows maximal recovery of peptides for MS analysis. The pooled, lyophilised peptide extracts for each

sample (section 2.5.4) were resuspended in 10 µL of 0.25% (v/v) TFA. Each Zip Tip C_{18} was then pre-wet with 0.25% (v/v) TFA/50% ACN by gently aspirating twice. After pre-wetting, each Zip Tip C_{18} was equilibrated with 0.25% TFA by aspirating gently three times. To effect binding of peptides to the C_{18} beads, the pooled peptide extracts were gently aspirated repeatedly (30 times per sample). Once the peptides were bound, the Zip Tip C_{18} was rinsed with 0.25% TFA by again gently aspirating three times. To ensure efficient elution of peptides from the membrane, the peptides for each sample were eluted in 0.1% (v/v) formic acid in 70% ACN by aspirating 25 times. The eluted peptides were either stored at -20 °C for later analysis or they were directly prepared for MS analysis. In either case, just prior to LC-MS/MS analysis, the peptides in each sample were diluted with 2% (v/v) ACN, 0.1% (v/v) formic acid to a final volume of 60 µL for the 5 µL sample loop. Each 60 µL sample was transferred to a sample vial and loaded onto the carousel of the tandem MS for MS data acquisition.

2.3.10 LC-MS/MS

2.3.10.1 Instrumentation

A Dionex UltiMate™ 3000 LC system and a Thermo Finnigan LTQ mass spectrometer equipped with a nanospray ion source were used for all peptide analyses. The analytical column used was a 75 µm × 15 cm PepMap C18 (3 µm, 300 Å, Dionex) with a constant flow rate of 200 nL/min. The following Merck (Merck Ltd.USA) HPLC grade solvents were used: 0.1% (v/v) formic acid (solvent A) ; 0.1% (v/v) formic acid in 80% (v/v) ACN (solvent B): 0% - 20% solvent B for 20 min; 20% - 55% solvent B for 40 min; 55% - 100% solvent B for 5 min; 100% solvent B for 5 min; 100% to 0% solvent B for 2 min. The LTQ was operated in data-dependent tandem MS mode where the three most abundant precursor ions detected in a single MS scan from m/z 400 to m/z 2000 were dynamically selected for subsequent MS/MS scans with the collisional energy set to 35%, simultaneously incorporating the dynamic exclusion option to prevent re-acquisition of MS/MS spectra of the same peptides.

2.3.10.2 Data analysis and protein identification

All acquired MS and MS/MS data were interpreted using MASCOT (Matrix Science, London, UK). The public database initially utilised was downloaded from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/, 26/03/2007, Bacteria (Eubacteria), 4761919 sequences, 164309855 residues). Once the Y88^T genome was sequenced, a Y88^T database was created in-house using ORF translated protein sequences. Since at the time of writing approximately 99% of the Y88^T genome had been sequenced, all spectra were first searched against this ORF translated protein sequences database (Y88_ORFs 130208 (3912 sequences; 1254784 residues) and then repeated against the NCBI database to maximise the opportunity of peptide matches to some proteins homologous to those that might not yet be found in the Y88^T database. Searches were also repeated for all spectra previously searched against NCBI prior to establishing the Y88^T database.

All searches were carried out using a peptide mass tolerance of 1.2 Da and a fragment ion tolerance of 0.6 Da, allowing for a fixed modification of cysteine carbamidomethylation and a variable modification of methionine oxidation. A maximum of two missed internal full trypsin cleavages was allowed (Elias et al., 2005). A peptide sequence that was uniquely matched to a protein was classified as a unique peptide. Proteins with two or more unique peptides having an individual Mascot score for each peptide MS/MS spectrum of at least 25 (although in most cases peptides were high-scoring with the lowest score of at least 40), were considered to be positive protein identifications. Any protein identified by a single unique peptide, had its corresponding MS/MS spectrum further manually inspected and validated (Chen et al., 2005).

Several filtered peptide sequences matched to multiple proteins. In some cases, this was due to the sequence similarities of the proteins resulting in proteins matched with the same peptide sequences into a protein group. Consequently, the protein group with greater than three peptide matches was also considered

a positive identification and only a single protein from the group with either the largest sequence coverage or with representative annotation was reported (Bae et al., 2004).

In some cases, a high-scoring peptide did not have all peaks accounted for. In these instances, an error-tolerant search was carried out to determine if the mass difference was due to a possible post-translational amino-acid modification.

A decoy database consisting of reverse Y88^{T} ORF translated protein sequences was created. All spectra were searched against this decoy database to exclude the possibility of false positive protein identifications. The false discovery rate (false positive / (false positive plus true positive)) was set at 5 %.

Chapter 3: Microbiology of Y88^T

3.1 Introduction

Preliminary work has shown that Y88^T accumulated PHA to 41±16% of its dry weight when grown in the presence of glucose (Addison et al., 2007). Since Y88^T was isolated from an environment where nitrogen is naturally limiting, but carbon is found in excess, it is potentially an ideal organism for the study of PHA production that is not constrained by nitrogen supply. The aim of this study was to investigate nitrogen and carbon utilisation as well as PHA production by Y88^T under conditions where the cells were either fixing nitrogen or not fixing nitrogen to determine whether PHA production had any dependence on nitrogen fixation. The growth profiles of Y88^T under conditions of carbon enrichment, nitrogen supplementation and nitrogen depletion were examined. This study has shown that the type of PHA produced by Y88^T, PHB, is accumulated intracellularly throughout the growth cycle to high levels of 80.5±7.4% (1.12 g/L) of the culture dry weight after 72 hr in glucose-enriched minimal nitrogen medium. Predictable growth and lower variability in PHB concentrations was achieved. Also, PHB synthesis occurred independently of nitrogen fixation. A baseline was established for further investigation into the physiology of Y88^T at the protein level during PHB accumulation and nitrogen fixation.

3.2 Results

3.2.1 Y88^T cultivation in batch flasks

3.2.1.1 Growth in a glucose-enriched minimal nitrogen medium

3.2.1.1.1 Modeling Y88^T growth to the Gompertz model

Y88^T growth in a glucose-enriched (5 g/L, approximately 27 mM) minimal nitrogen medium (0.1 g/L, approximately 2.8 mM NH_4Cl) followed the predicted successive lag, growth and asymptotic phases of bacterial growth that

are described by the Gompertz sigmoidal function (Zwietering et al., 1990). $Y88^T$ growth curves were fitted with the Gompertz model using the Gompertz equation $Y=N0+C*\exp(-\exp((2.718*\mu/C)*(\lambda-X)+1))$ (Zwietering et al., 1990). $Y88^T$ growth, as determined by OD_{600} measurements at time points over 72 hr, showed a sigmoidal curve (Fig. 3.1), R^2 =0.99 for all three replicates (Table 3.1). All three replicates showed a very good fit of the model to the observed data, indicating that the parameter values are nearly precise and will have little variation. Although maximum growth was achieved by approximately 30 hr, the maximum (exponential) growth rate occurred between approximately 12-20 hr of growth (Fig. 3.1). Between 20-24 hr, growth was near linear. Early stationary phase was reached between 30-40 hr post-inoculation.

3.2.1.1.2 Biomass produced by $Y88^T$

Biomass as an indicator of growth was determined for $Y88^T$ cells grown in glucose-enriched (5 g/L), minimal nitrogen (0.1 g/L) medium. $Y88^T$ produced a maximum dry biomass of approximately 1.3 g/L by 52 hr with a linear increase in biomass between 20 to 44 hr (Fig. 3.2). The greatest change in biomass between sampling intervals as well as the greatest rate of biomass production occurred between 20-28 hr with an increase of 0.423 g/L at a rate of 0.05 g/hr (Table 3.2). No measurable change in biomass occurred between 52-72 hr (Table 3.2).

The yield of biomass (grams of biomass produced per gram of glucose consumed) was more or less constant throughout the growth period with the exception of the 0-8 hr growth interval. Although OD_{600} measurements showed a lag period of growth between 0-8 hr (Fig. 3.2), the biomass measurements showed one of the highest rates of increase during this growth interval (Table 3.2). With the exception of the lag phase of growth (coinciding with the first two time points), a direct correlation exists between biomass and OD_{600} (Fig. 3.3). Both OD_{600} and biomass are therefore good predictors of $Y88^T$ growth, after the initial lag phase.

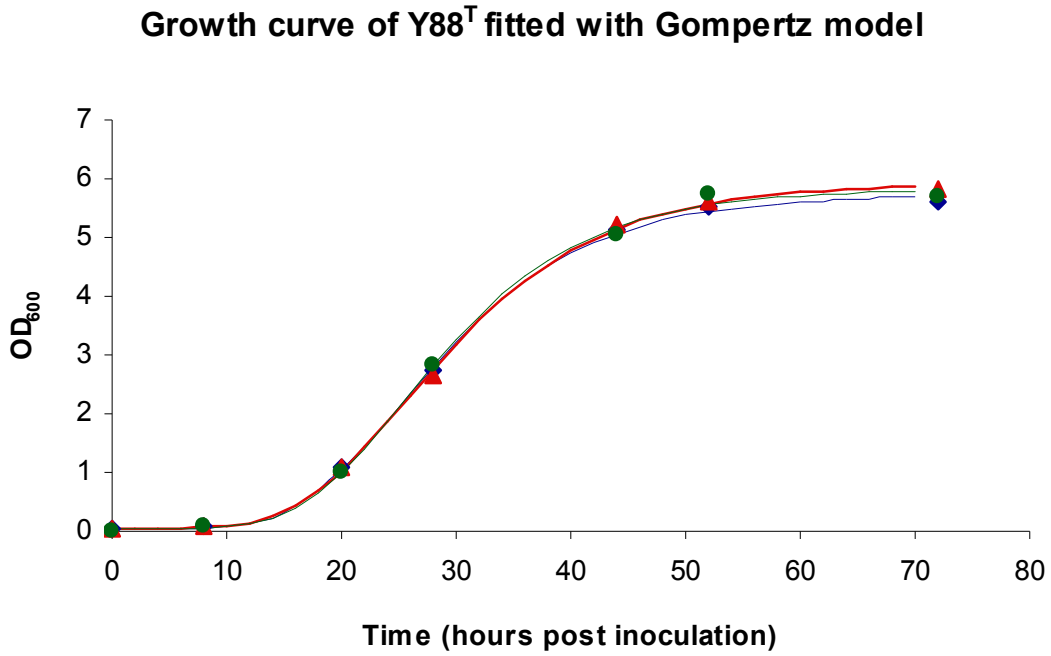


Figure 3.1 Y88^T growth data fitted to the Gompertz model

The Gompertz model used for bacterial growth shows the goodness-of-fit of the observed values (individual observed data points presented as coloured markers) to the predicted values (continuous coloured lines) for each replicate. Note that for each replicate, marker colour of observed values matches line colour of predicted values, (n=3).

$Y88^T$	Gompertz equation	R^2
Replicate 1	y=0.05585+5.667*exp(-exp((2.718*0.2312/ 5.667)*(16.12-x)+1))	0.99
Replicate 2	y=0.062148+5.858226*exp(-exp((2.718*0.228818/ 5.858226)*(16.18539-x))+1))	0.99
Replicate 3	y=0.05384+5.774*exp(-exp((2.718*0.2394/ 5.774)*(16.42-x))+1))	0.99

Table 3.1 Gompertz equations for $Y88^T$ replicate growth curves

Gompertz equation Y=N0+C*exp(-exp((2.718*μ/C)*(λ-X)+1)) describing the growth curve for each $Y88^T$ replicate. N0=log (initial number of cells), C=difference between initial and final number of cells, λ=delay before growth (same units as "x"), μ=maximum specific growth rate, "x"=time, "y"=log (cell number).

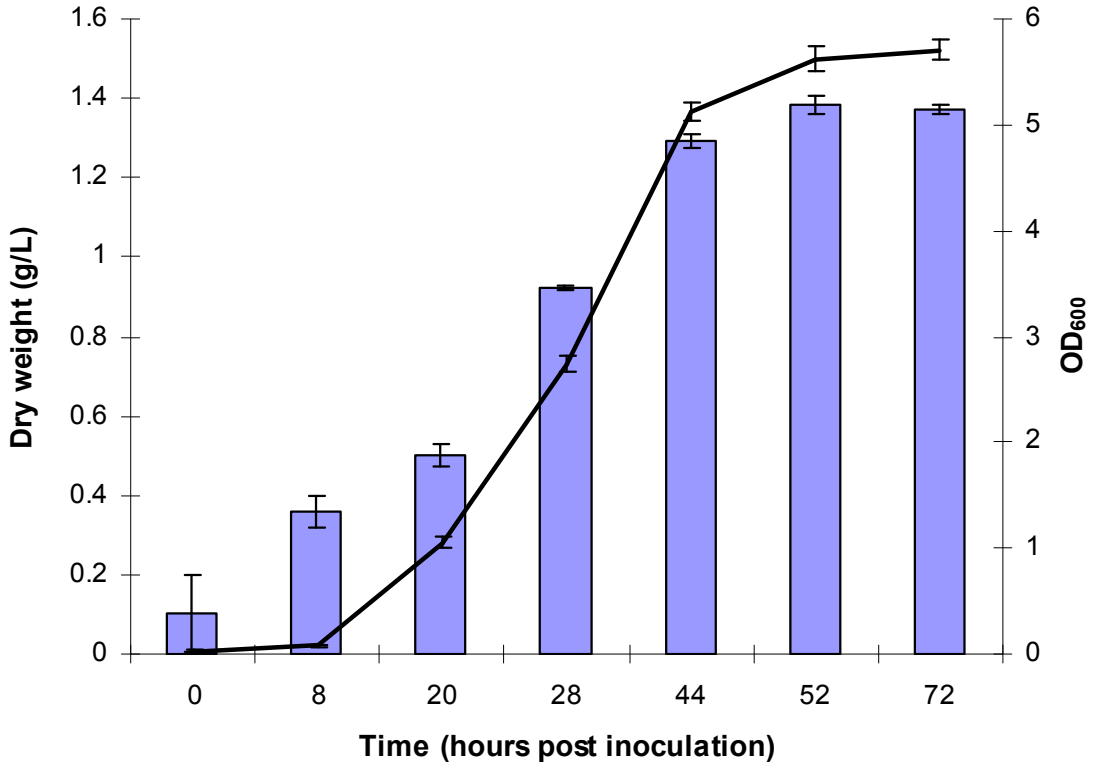


Figure 3.2 OD_{600} and dry biomass produced by $Y88^T$ in a glucose-enriched minimal nitrogen medium over 72 hr

OD_{600} (data points joined by lines; right y axis) and dry biomass (bars; left y axis) produced by $Y88^T$ in a glucose-enriched minimal nitrogen medium over 72 hr shows near linear growth and biomass production between 20-44 hr. Highest rate of biomass produced was between 20-28 hr. Error bars represent standard deviations, (n=3).

Time (hours post inoculation)	Dry biomass (g/ L)	Change in dry biomass (g/ L)	Rate of change in dried biomass (g/ L/ hr)	Yield in biomass (g biomass/ g glucose consumed)
0	0.103±0.09	0.000±0.00	0.000±0.00	0.000±0.00
8	0.357±0.04	0.254±0.12	0.03±0.02	3.31±0.93
20	0.500±0.03	0.143±0.17	0.01±0.01	0.25±0.01
28	0.923±0.01	0.423±0.04	0.05±0.01	0.27±0.00
44	1.29±0.017	0.367±0.15	0.02±0.01	0.27±0.00
52	1.383±0.02	0.093±0.03	0.01±0.01	0.28±0.00
72	1.373±0.01	0.000±0.00	0.000±0.00	0.27±0.00

Table 3.2 Dry biomass (cumulative over 72 hrs), actual change in dry biomass (for each time interval separately), rate of change in dry biomass produced (for each time interval separately) and yield of biomass by Y88^T in a glucose-enriched minimal nitrogen medium.

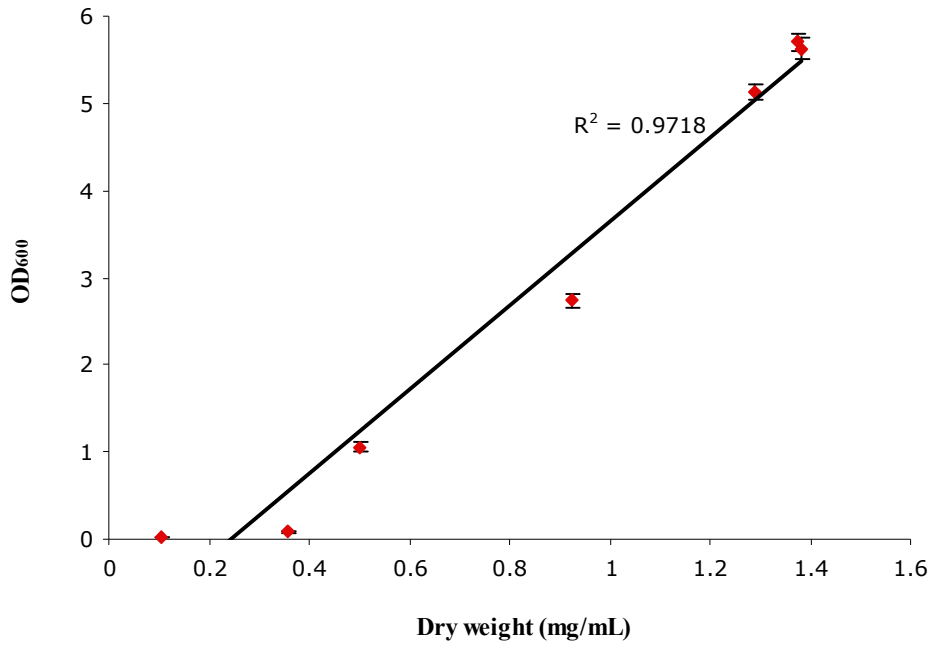


Figure 3.3 Plot of Y88^T biomass versus OD_{600}

Error bars represent standard deviations, (n=3). Note that the linear relationship breaks down in the initial stages of growth, as the cells are in lag phase to approximately eight hours of growth. See also Fig. 3.2.

3.2.1.2 Y88^T growth and carbon substrate utilisation

3.2.1.2.1 Single substrate utilisation: glucose vs. acetate

The ability of Y88^T to utilise either glucose or acetate as its sole carbon substrate in a minimal nitrogen medium was investigated to determine whether there was a difference in the amount of biomass and PHA produced from each substrate and which of these carbon sources provided the best growth substrate. Concentrations of glucose and acetate supplemented in the medium contained the same amount of carbon on a molar basis. Different concentrations of each substrate ranging from low (1.9 mM) to high (27 mM) were used. Since OD_{600} and biomass were equally good predictors of Y88^T growth (section 3.2.1.1.2), biomass was used as the measure of Y88^T growth for this section.

Y88^T cells cultivated for 24 hr on glucose as the sole carbon source achieved an overall higher growth for a given concentration of input carbon than those cultivated primarily on acetate (Fig. 3.4, see 20 mM A, acetate; 6 mM G, glucose). Growth on glucose resulted in acidification of the medium from an initial pH of 7.2, to a final pH of 4.5. In contrast, during Y88^T growth on acetate as the sole carbon source, the medium became more basic, from an initial pH of 7.2 to a final pH of 8.5. A dry biomass of approximately 1.125 g/L was achieved for the highest concentration of glucose (27 mM) compared to the highest biomass of approximately 0.2 g/L achieved for Y88^T cells grown on 20 mM acetate (Fig. 3.4). Biomass produced by Y88^T after 24 hr on 20 mM acetate was therefore approximately one-fifth that of the biomass produced on 27 mM glucose after the same time. Since 20 mM acetate contains approximately 25% of the carbon that 27 mM glucose contains and the amount of biomass produced on acetate is approximately 20% of that produced on glucose, Y88^T uses slightly less carbon if grown on glucose to achieve a similar amount of biomass on a carbon for carbon basis compared to growth on acetate. The amount of carbon converted to PHA differs significantly, however, depending on whether glucose or acetate is used as the carbon substrate. More carbon is used for PHA

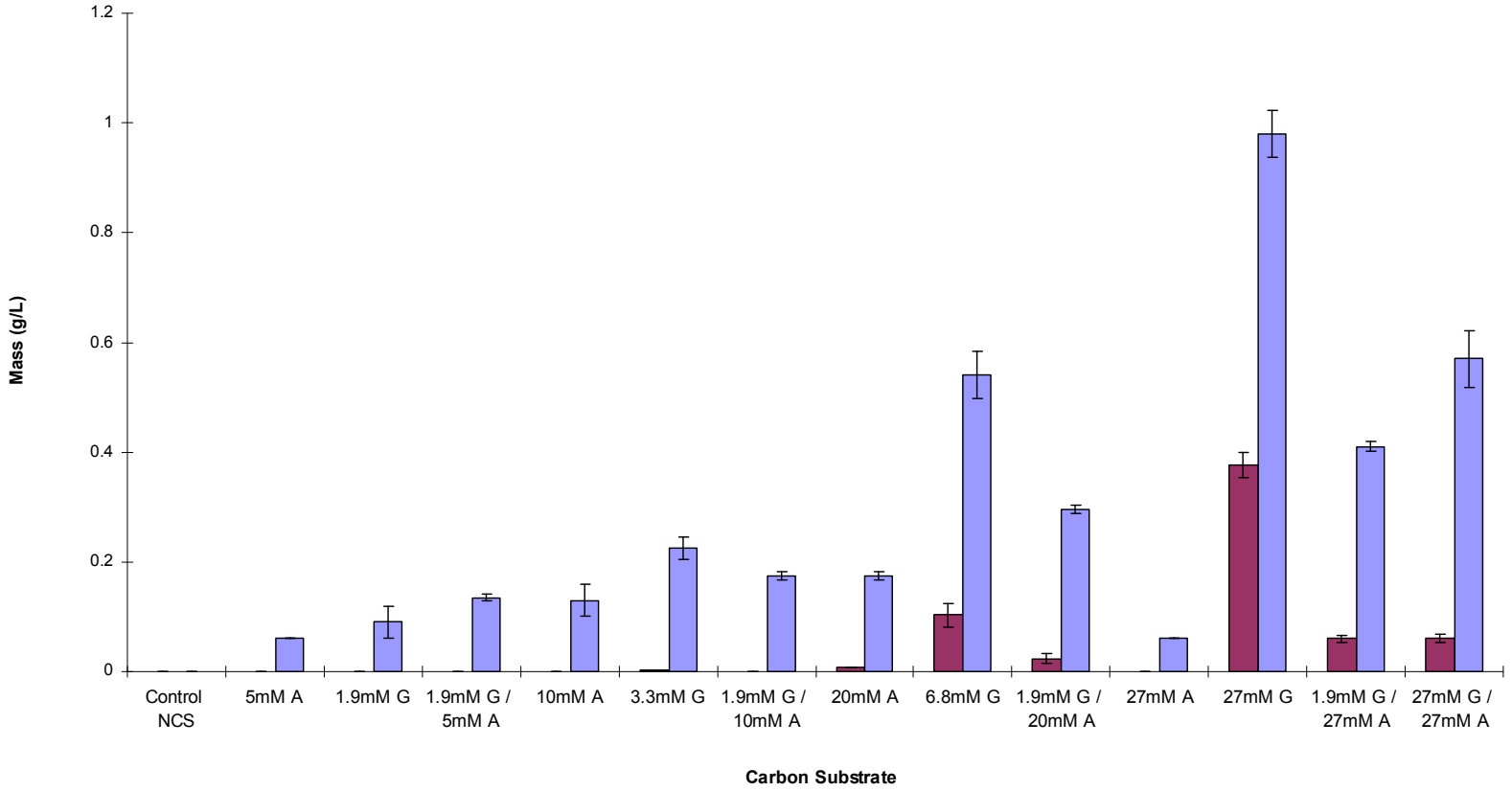


Figure 3.4. Total dry biomass and PHB mass produced by Y88^T after 24 hr growth in a minimal nitrogen medium with various carbon substrates

Total dry biomass (blue bars) and PHB mass (red bars) produced by Y88^T after 24 hr growth (left y axis) and 24 hr OD_{600} measurements (bars outlined by dotted lines; right y axis) achieved in a minimal nitrogen medium enriched with a single carbon (C) substrate (glucose (G), acetate (A)) or a mixed carbon substrate (G and A) . NCS=no carbon substrate in growth medium. Error bars represent standard deviations, (n=2).

production from glucose than that used from acetate since approximately 40% of the biomass produced by Y88^T in the glucose-enriched medium consisted of PHA whereas only 1.2% of the biomass produced by Y88^T in the acetate-enriched medium (0.01 g/L) consisted of PHA.

3.2.1.2.2 Mixed substrate utilisation: glucose and acetate

The ability of Y88^T to utilise glucose and acetate as a mixed carbon source in a minimal nitrogen medium was investigated to determine the amount of biomass and PHA produced by Y88^T when cultivated in the presence of a combination of these two substrates. Comparisons were made between the amounts of biomass and PHA produced by Y88^T on the combined substrates and that produced under the same growth condition from a single carbon source. The aim was to determine whether the presence of a second carbon source in the growth medium would increase the amount of biomass and PHA produced by Y88^T. Different concentrations of each substrate were used in combinations ranging from low (1.9 mM) to high (27 mM).

Y88^T cells cultivated on low amounts of glucose (1.9mM) in combination with 5 mM, 10 mM or 20 mM acetate as a mixed carbon source, achieved overall higher growth than Y88^T cells cultivated on the same concentrations of acetate alone (Fig. 3.4). Thus the addition of small amounts of glucose promoted Y88^T growth in acetate-enriched media, as was particularly evident in the 1.9 mM glucose/27 mM acetate combination. This combination showed superior biomass and PHB production to 27 mM acetate alone, which itself was unable to achieve even the low biomass and PHB production of the 20 mM acetate alone. An acetate concentration higher than 20 mM as a sole substrate therefore appears to inhibit Y88^T cell growth. This conclusion was confirmed by demonstrating that Y88^T cells grown on 81 mM acetate as a sole carbon substrate control showed no measurable growth.

Although a higher biomass (approximately 0.57 g/L) was obtained from growth on 27 mM glucose/27 mM acetate compared to that on 1.9 mM

glucose/27 mM acetate (approximately 0.4 g/L), minimal amounts of PHB (0.06 g/L) were produced in each case. These results demonstrate that the presence of 27 mM acetate in the growth medium, irrespective of the high or low amounts of glucose present, inhibits PHB production (Fig. 3.4). Thus it appears that the additional glucose in these cultures was used by Y88^T as an energy source for growth rather than conversion into PHB. Therefore no benefit in terms of PHB production was derived from increasing the concentration of glucose in combination with 27 mM acetate.

Growth on 1.9 mM glucose/20 mM acetate resulted in PHB constituting approximately 10% of the cell dry weight, which was twice the PHB produced during growth on the 20 mM acetate alone. However, the combination of 1.9 mM glucose/27 mM acetate resulted in approximately 15% of the biomass constituting PHB, which was approximately 33% more PHB produced from 1.9 mM glucose/20 mM acetate. This translates into approximately 33% more PHB produced from roughly 26% more acetate thus the additional acetate resulted in a higher PHB production as long as a small amount of 1.9 mM glucose was present in the culture medium. However, the overall amount of PHB produced was still low compared to that on 27 mM glucose alone.

Y88^T cells cultivated on a mixed carbon substrate of 27 mM glucose/27 mM acetate grew to a much lower biomass than that achieved on a single carbon source of 27 mM glucose with approximately 40% less biomass than that produced on 27 mM glucose alone. Dry biomass measurements determined for growth on 27 mM glucose and acetate were approximately 0.6 g/L with approximately 10% of the biomass composed of PHB (0.06 g/L). Concomitant with the decreased total biomass, the amount of PHB produced during growth on the 27mM glucose/acetate combination was approximately 75% less than that produced during growth on 27 mM glucose alone. These results suggest that acetate has a limiting effect on growth and PHB production at higher concentrations of combined glucose and acetate compared with growth and

PHB produced from high concentrations of glucose alone. Since biomass and PHB production were consistently superior with glucose as the sole carbon source, all subsequent experiments were carried out with glucose as the sole carbon substrate.

3.2.1.2.3 Y88^T growth in high and low glucose-enriched minimal nitrogen medium

Y88^T growth was investigated in glucose-enriched media with either 5 g/L of glucose (high glucose; HG) or 1 g/L glucose (low glucose; LG) supplemented in the medium. The aim was to determine if carbon in the form of glucose is required in excess to produce high yields of PHB or whether similar yields can be achieved using a lower concentration of glucose. Also investigated was the question of whether a glucose concentration of 1 g/L can be considered limiting to growth and PHB production in Y88^T.

During the first 8 hr of culture, growth kinetics on LG were similar to cells grown in HG (Fig. 3.5). After 8 hr, it was evident that growth on LG was becoming limited by low carbon supply with a biomass of only 0.04 g/L attained after 24 hr. In comparison, Y88^T cells on HG grew exponentially between 8-24 hr and attained a biomass of 0.9 g/L by 24 hr (Figs. 3.5 and 3.6). Approximately 40% (0.34 g/L) of this biomass consisted of PHB, whereas no PHB was produced during growth on LG after 24 hr (Fig. 3.5). After 48 hr, Y88^T cells on HG attained a biomass of 1.5 g/L (Fig. 3.6). Approximately 66% (1.1 g/L) of this biomass consisted of PHB (Fig. 3.6) whereas the cells grown on LG attained a biomass of 0.16 g/L over the same period and still did not produce any PHB (Fig. 3.6). By 48 hr, Y88^T therefore achieved ten-fold less biomass on LG than on HG but consumed five-fold less glucose.

Final dry biomass measurements after 72 hr showed that Y88^T cells grown on LG attained about 1/5 the biomass (0.3 g/L) of those grown on HG, which reached a final concentration of 1.5 g/L (Fig. 3.6). Since five times more glucose was used at HG than at LG and five times more biomass was produced after 72

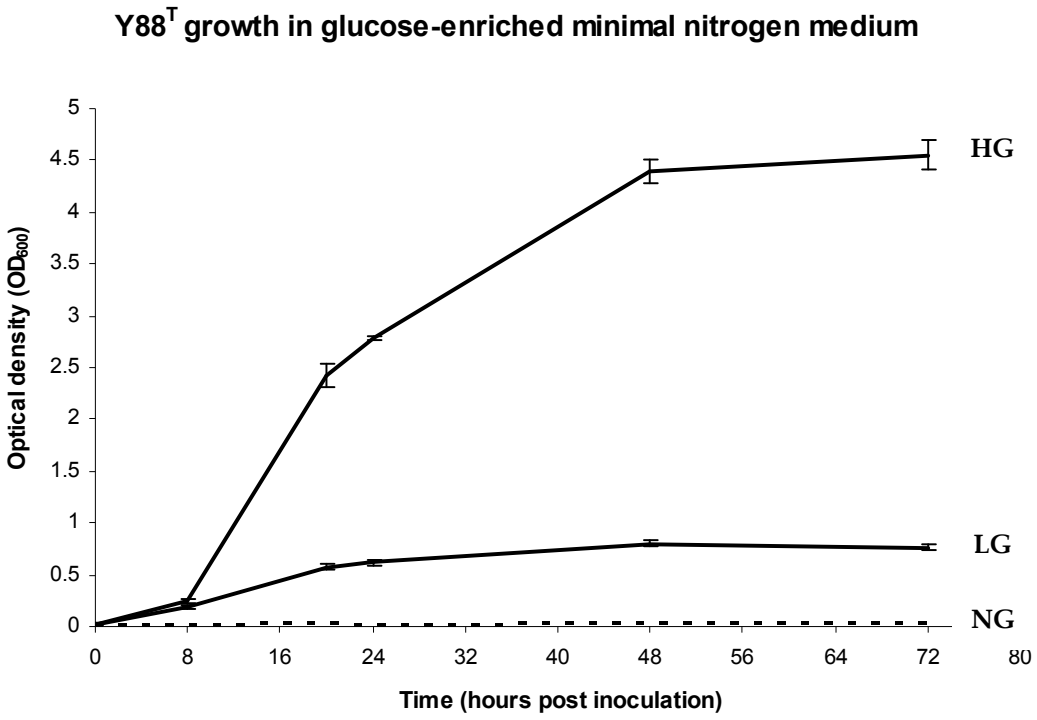


Figure 3.5 Comparative growth of Y88^T in high and low glucose minimal nitrogen media.
Y88^T growth over 72 hr in a glucose-enriched (5 g/L) (HG) and low glucose (1 g/L) (LG), minimal nitrogen (0.1 g/L NH_4Cl) medium, (n=3). A control (NG) containing Y88^T cells but no glucose in the medium was included. Error bars represent standard deviations, (n=2).

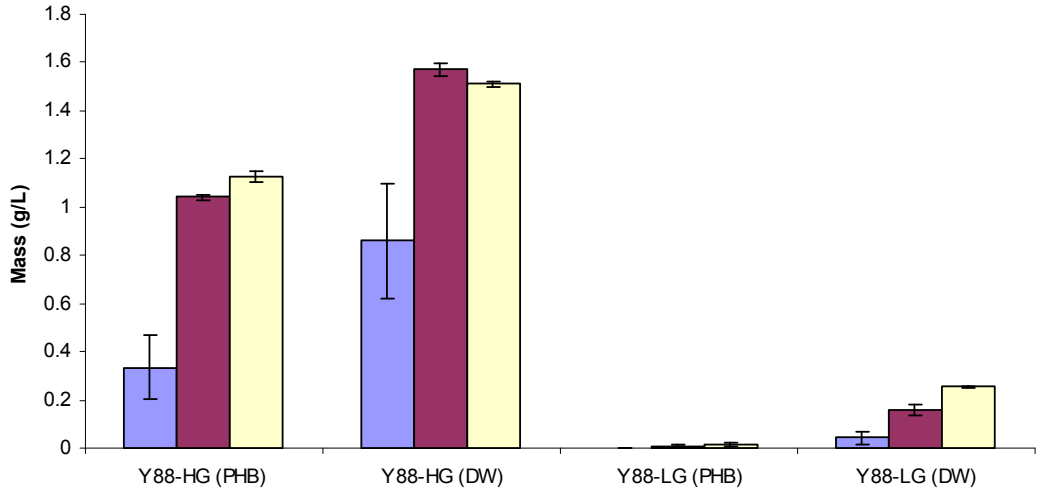


Figure 3.6 PHB and dry biomass produced by Y88^T in high and low glucose minimal nitrogen medium

PHB and dry biomass (DW, g/L) produced by Y88^T cultivated for 72 hr in glucose-enriched (HG, 27mM) and low glucose (LG, 5.6mM) minimal nitrogen (2.8 mM NH_4Cl) medium. Bars represent time that cells were cultivated as follows: 24 hr = blue, 48 hr = maroon, 72 hr = cream, (n=3).

hr, a correlation exists between the amount of glucose supplied and the biomass produced. No such correlation was evident between glucose supplied and PHB produced since by 72 hr on LG, Y88^T had not produced any measurable PHB. This suggests that the glucose supplied was used for energy requirements during growth and that the supply at 1 g/L was insufficient to sustain both growth and PHB production.

Y88^T cells therefore require glucose in excess at the beginning of the growth phase to achieve high levels of PHB accumulation in the cells throughout the growth phase when grown under our laboratory conditions. Since approximately 1 g/L glucose as a starting carbon substrate constrains biomass to low levels and is insufficient for PHB production during growth relative to the higher starting carbon substrate of 5 g/L, glucose at a starting concentration of 1 g/L can be considered limiting to Y88^T growth and PHB production.

Additional data (not shown) for Y88^T growth on 2 g/L glucose (selected as an intermediate concentration between 1 and 5 g/L) confirmed that Y88^T produced 0.5 g/L dry biomass after 24 hr. Of this, only 0.1 g/L constituted PHB compared to the 0.34 g/L PHB produced from 5 g/L glucose. Thus 3.4-fold more PHB was produced from 2.5-fold more glucose at 5 g/L than 2 g/L. Since 5 g/L glucose gave a better PHB yield than either 1 or 2 g/L, the former concentration was selected for all subsequent growth experiments with glucose as carbon substrate.

3.2.1.2.4 Composition of the PHB polymer produced by Y88^T from glucose

The monomer composition of PHB produced by Y88^T grown on a glucose-enriched (5 g/L), nitrogen-limiting (0.1 g/L NH_4Cl) medium was confirmed by infrared spectroscopy and gas chromatography analysis to be 3-hydroxybutyrate. ^{13}C-NMR analysis determined the subunit structure of the PHB extract from Y88^T cells (Fig. 3.7). The cell biomass of Y88^T constitutes approximately 90% poly-(β-hydroxybutyrate) of the extracted product by weight. Transmission electron microscopy of Y88^T cells confirmed that this

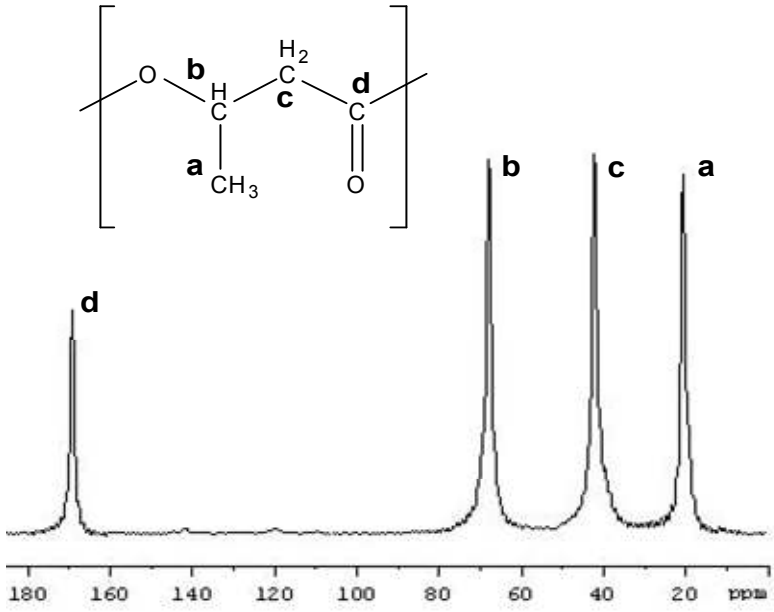


Figure 3.7 ^{13}C-NMR spectrum of PHB produced by Y88^{T}.

Inset: PHA monomer structure with carbon atoms labeled corresponding to the NMR peaks they produce in the spectrum.

bacterium accumulates PHB granules to a high percentage (80%) in its cells (Fig. 3.8). Analysis of the extracted polymer by gel permeation chromatography (GPC) showed that Y88^T is able to synthesise a high molecular weight polymer of approximately 1.4±0.2 MDa when grown in our standard minimal nitrogen medium with excess glucose as the sole carbon source.

3.2.2 Comparison between Y88^T, *A. vinelandii* and *P. putida* mt2

3.2.2.1 Growth and nitrogen fixation in NH_4^+-free and NH_4^+-supplemented glucose-enriched media

Y88^T growth in a minimal nitrogen medium enriched with glucose was investigated in parallel with two species from the Gammaproteobacteria: *A. vinelandii* (capable of fixing molecular nitrogen, positive control for nitrogen fixation) and *P. putida* mt2 (incapable of fixing molecular nitrogen, negative control for nitrogen fixation). To determine the effect that ammonium supplementation or the lack thereof has on Y88^T growth and nitrogen fixation during 24 hr in a glucose-enriched medium, Y88^T growth was investigated under ammonium-free conditions and under conditions where ammonium initially supplemented in the medium (0.1 g/L NH_4Cl) was depleted and compared to that of *A. vinelandii* and *P. putida* mt2. C_2H_2 reduction assays, to determine the activity of the nitrogenase enzyme required for nitrogen fixation, were carried out on cells from all three bacteria grown in ammonium-free and ammonium-supplemented media. Y88^T cells grew more poorly over a 24 hr period in ammonium-free medium (Fig. 3.9) reflecting similar lower growth to that of *A. vinelandii* (OD_{600}<1.0), although the latter reached a higher OD_{600} than that for *P. putida* mt2 OD_{600}<0.3) under the same growth conditions. Y88^T reached a maximum OD_{600} of approximately 0.7 by 24 hr post-inoculation in ammonium-free medium (Fig. 3.9). Y88^T gave a negative C_2H_2 reduction assay for those cells grown in ammonium-free medium although *A. vinelandii* gave a positive C_2H_2 reduction assay under the same growth conditions (see section 3.2.2.2.1 for a more detailed description of the C_2H_2 reduction assay and results obtained). Under the same growth conditions, but with ammonium chloride

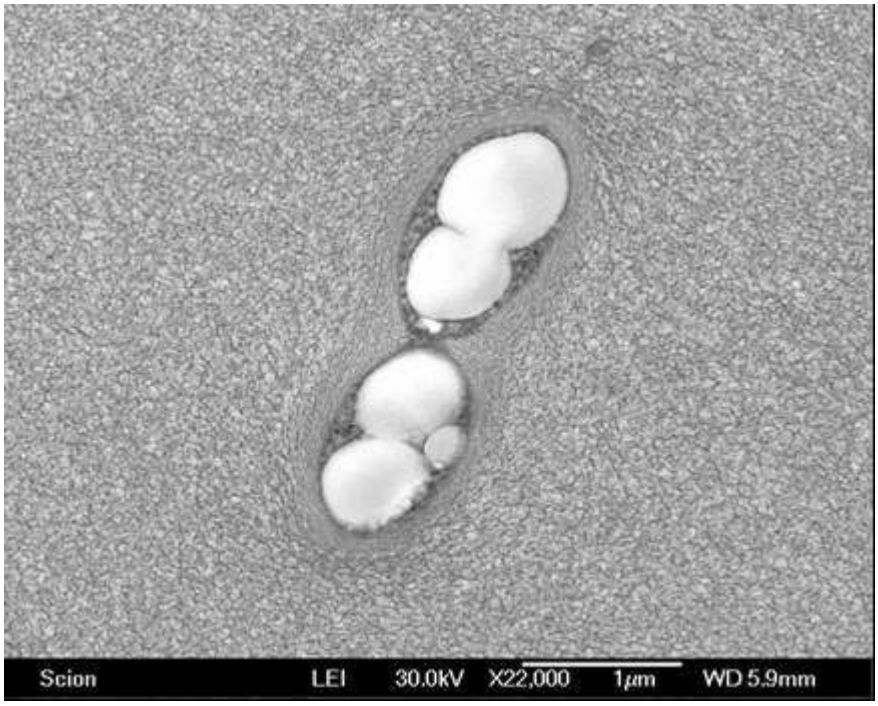


Figure 3.8 Transmission electron micrograph (TEM) of Y88^{T} cells containing PHB accumulated up to 48 hr growth.

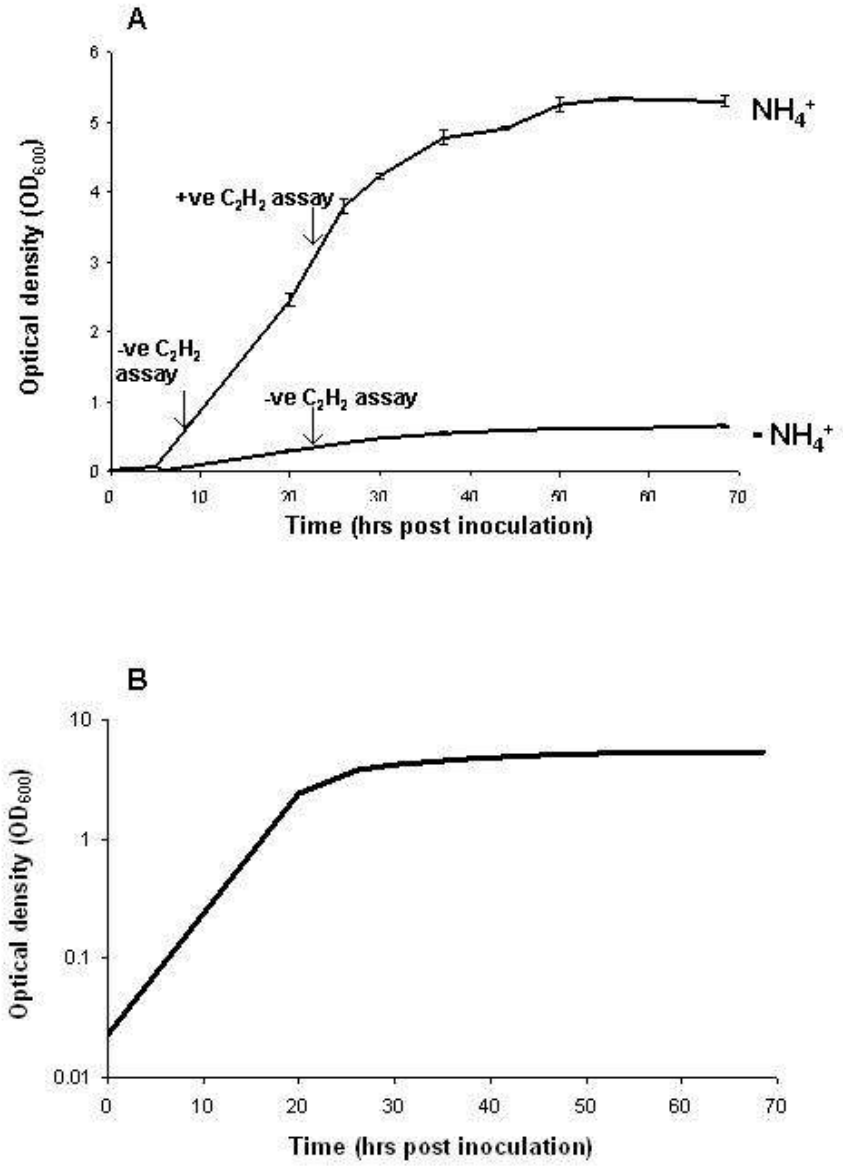


Figure 3.9 $Y88^T$ growth in minimal nitrogen and nitrogen-free media.

(A) $Y88^T$ growth (OD_{600}) measured over 72 hr in a nitrogen -free (-NH_4^+) and in a minimal nitrogen (0.1 g/L NH_4Cl) (NH_4^+) medium. An C_2H_2 reduction assay to determine nitrogenase activity gave a negative result at OD_{600}= 0.5 (-ve C_2H_2 assay) for $Y88^T$ cells grown under $NH_4^+$and -NH_4^+- conditions and a positive result at OD_{600}= 2 (+ve C_2H_2 assay) for $Y88^T$ cells grown under NH_4^+ conditions. (B) Growth curve of cells grown in NH_4^+ medium fitted to a semi-log scale. Error bars represent standard deviations, (n=3).

supplemented in the medium at 0.1 g/L (2.2mM NH_4Cl comprising 0.034 g/L NH_4^+), Y88^T showed significantly higher growth over 24 hr to either *A. vinelandii* or *P. putida* mt2. Y88^T showed a slow increase in OD_{600} during the first 8 hr, which is consistent with the lag phase of growth (Fig. 3.9). During this time, an ammonium assay determined that the ammonium initially supplemented in the medium showed no measurable change (Fig. 3.10). After the first 8 hr of growth, ammonium was steadily depleted until no measurable ammonium remained in the medium by approximately 20 hr of growth. Further monitoring of the ammonium in the medium showed no measurable change in concentration between 20-72 hr. Therefore, all the ammonium had been removed by Y88^T from the medium by 20 hr (Fig. 3.10). Despite this, Y88^T OD_{600} increased steadily between 20-30 hr corresponding to mid to late log phase (Figs. 3.9 and 3.11). An C_2H_2 reduction assay gave a positive result for nitrogenase activity in cells harvested at 24 hr following ammonium depletion in the cultures initially supplemented with ammonium, indicative of the cells fixing nitrogen (see chromatogram E in Fig. 3.14 and section 3.2.2.2.1 for a more detailed description of the C_2H_2 reduction assay results). Thus the only nitrogen available to the cells during that phase of growth was in the form of molecular nitrogen. In contrast, a negative result for nitrogenase activity in cells harvested at the same time from ammonium-free cultures was indicative of the cells not fixing nitrogen in these cultures. Y88^T OD_{600} measurements in ammonium-supplemented medium reached a maximum OD_{600} >5.0 for all replicates (n=3) by 35-40 hr post-inoculation. At this stage, the cells had entered the stationary phase of growth (Figs. 3.9 and 3.11). The comparison between ammonium-free and ammonium-supplemented growth showed that Y88^T cells attain high optical density and can fix nitrogen when ammonium is supplemented in the medium. In contrast, Y88^T cells grown in ammonium-free medium, only achieved a low relative OD_{600} (approximately 0.7) with no measurable nitrogenase activity, suggesting the cells were not able to give a positive acetylene reduction assay without ammonium added to the medium. Thus it appears that Y88^T cells require small amounts of ammonium to promote initial

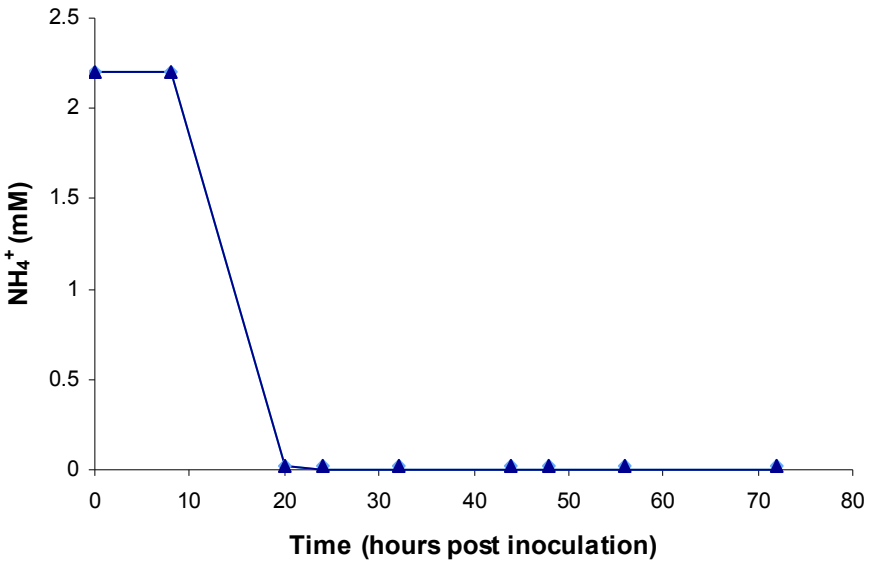


Figure 3.10 Ammonium utilisation by Y88^T

Ammonium assimilation by Y88^T shows depletion of ammonium after 20 hrs to almost undetectable levels with no change in ammonium concentrations thereafter, (n=2).

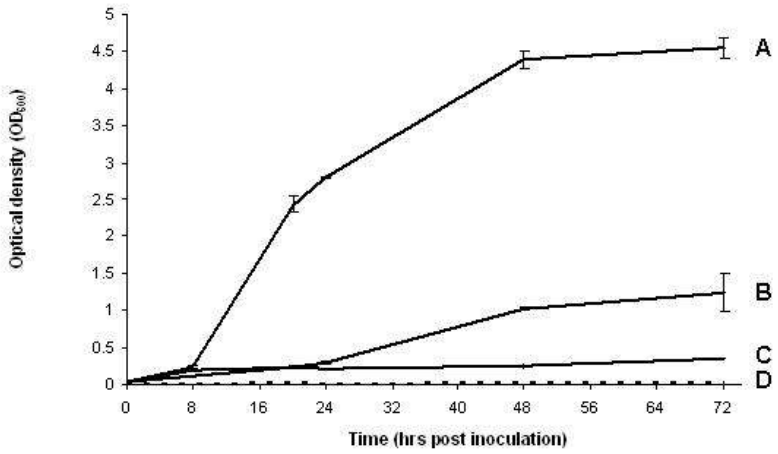


Figure 3.11 Comparative growth of Y88^T, *A. vinelandii*, and *P. putida* mt2

Comparative growth of Y88^T (A), *A. vinelandii* (B) and *P. putida* mt2 (C) during 72 hr in a glucose-enriched (5 g/L), minimal nitrogen NH_4Cl (0.1 g/L) medium, (n=3). A control (D) containing Y88^T cells but no glucose in the medium was included. Error bars represent standard deviations.

cell growth and nitrogen fixation in batch culture.

3.2.2.2 Growth, biomass and PHB production in NH_4^+-supplemented media

Y88^T and *P. putida* mt2 showed superior growth after 24 hr in an ammonium-supplemented compared to their growth in an ammonium-free, glucose-enriched medium (as discussed above in section 3.2.2.1). Since *A. vinelandii* reached an OD_{600} of approximately 1.0 in the same ammonium-supplemented medium, this medium was used as the preferred medium to compare PHB accumulation, nitrogen fixation and biomass in these three species. Y88^T, *A. vinelandii* (positive control for PHB accumulation) and *P. putida* mt2 (negative control for PHB accumulation which synthesises PHAs but not of the PHB type) were grown in a glucose-enriched minimal nitrogen (0.1 g/L NH_4Cl) medium to determine how much biomass and PHB each strain could produce under standardised growth conditions. OD_{600}, biomass and PHB produced by Y88^T and *A. vinelandii* under ammonium assimilatory and nitrogen-limiting conditions demonstrated the ability of these nitrogen-fixing bacteria to grow on glucose as a carbon substrate and convert it to PHB under these growth conditions (Figs. 3.12, 3.13). Although both nitrogen fixers displayed the ability to fix nitrogen under the tested growth conditions (Fig. 3.14), Y88^T's ability to produce PHB was superior to that of *A. vinelandii* in a nitrogen-limiting, glucose-enriched growth medium as is evident from the higher PHB produced by Y88^T (Figs. 3.12, Fig. 3.13).

During the early lag phase of growth, little difference was evident among the cell densities of the three species with an OD_{600} of approximately 0.3 measured for all (Fig. 3.11). This lag phase coincided with the period during Y88^T growth in which supplemented ammonium (0.1 g/L NH_4Cl equivalent to 2.8 mM) showed no detectable depletion in the medium (see Fig. 3.10, Section 3.2.2.1). Following the lag phase of growth and by 24 hr post inoculation, Y88^T attained a dry biomass of approximately 0.9 g/L (Fig. 3.12). Approximately 40% (0.34

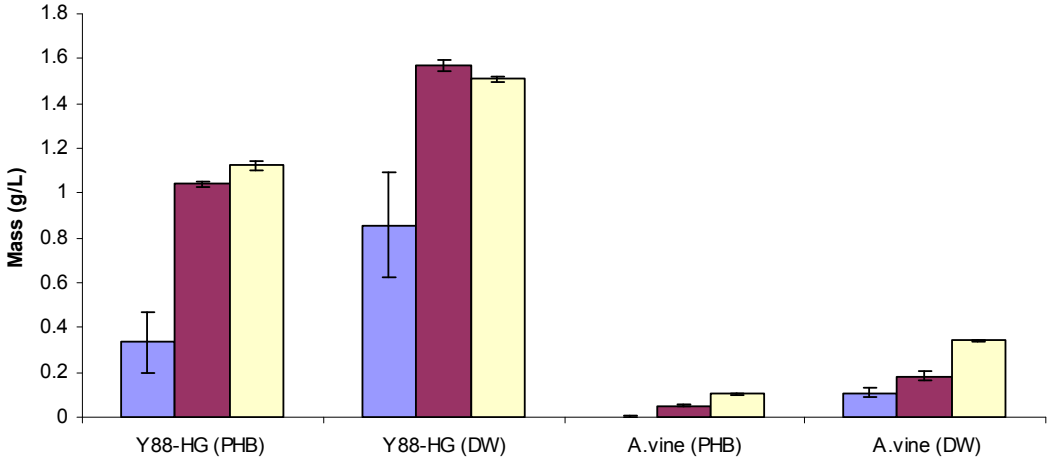


Figure 3.12 Comparative PHB mass and total dry biomass production by Y88^T and *A. vinelandii* PHB and dry biomass (dry weight (DW) produced by the nitrogen-fixing Y88^T and *A. vinelandii* (A. vine) when cultivated for 72 hr in glucose-enriched (27mM) minimal nitrogen-supplemented (2.8 mM NH_4Cl) medium. Bars represent time: 24 hr = blue, 48 hr = maroon, 72 hr = cream. Error bars represent standard deviations, (n=3).

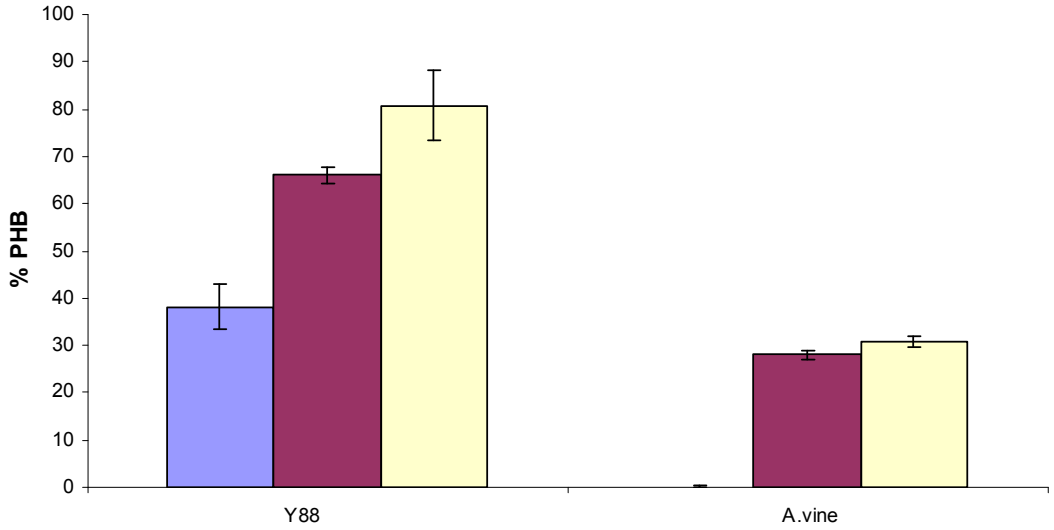


Figure 3.13 Comparative PHB production by Y88^T and *A. vinelandii* at 24, 48 and 72 hr.
Inter-specific comparison of PHB expressed as a percentage of dry biomass produced by the nitrogen-fixing Y88^T and *A. vinelandii* cultivated for 72 hr in a glucose-enriched (27mM) minimal nitrogen-supplemented (2.8 mM NH_4Cl) medium. Bars represent time: 24 hr = blue, 48 hr = maroon, 72 hr = cream. Error bars represent standard deviations, (n=3).

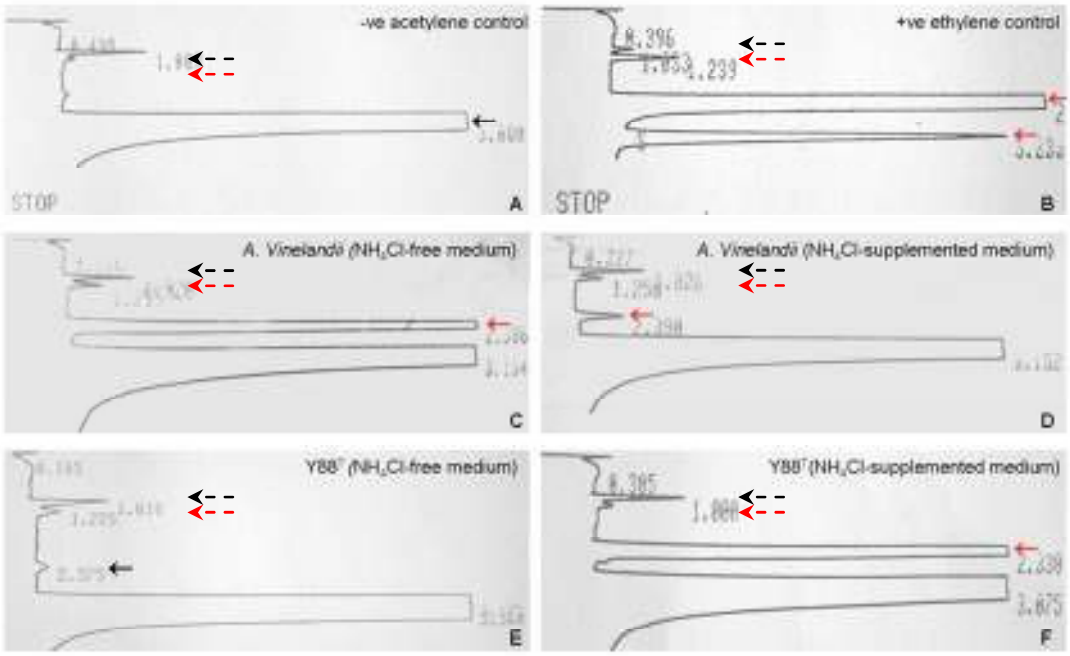


Figure 3.14 Acetylene reduction assays of Y88^T and *A. vinelandii* under various growth conditions

Representative chromatograms of C_2H_2 and ethylene peaks depicting inactive and active nitrogenase detected in C_2H_2 reduction assays. Red arrows (with solid lines) indicate positive ethylene peaks in positive ethylene control (B), Y88^T grown in ammonium-supplemented medium (F) and *A. vinelandii* grown in ammonium-free (C) and ammonium-supplemented (D) medium. Black arrows indicate C_2H_2 peak in negative C_2H_2 control (A) and the lack of an ethylene peak for Y88^T grown in ammonium-free (E) glucose-enriched medium. The representative chromatogram in each assay also has an argon peak (black arrows with broken lines) and an oxygen peak (red arrows with broken lines); n=2 for A to F but only one representative chromatogram of each pair is shown.

g/L) of the dry biomass consisted of PHB (Fig. 3.13). In comparison, *A. vinelandii* attained a dry biomass of 0.1 g/L with 0.14% (0.0035 g/L) consisting of PHB (Fig. 3.13). For the same period of growth, *P. putida* mt2 produced 0.04 g/L dry biomass but no PHB (data not shown).

After 48 hr, Y88^T attained a dry biomass of 1.5 g/L compared to a dry biomass of 0.2 g/L for *A. vinelandii* (Fig. 3.12). At this time, *P. putida* mt2 attained a dry biomass of 0.2 g/L, demonstrating how little growth 0.1 g/L NH_4Cl can support for a confirmed non-nitrogen fixer. The amount of PHB produced by Y88^T after 48 hr constituted 66% (1.04 g/L) of its dry biomass, whereas *A. vinelandii* produced 28% (0.05 g/L) (Fig. 3.12) under these standardised growth conditions.

Between 48-72 hr post inoculation, Y88^T produced no additional biomass growing to a final biomass of 1.5 g/L. Despite this lack of increase in biomass, 80% (1.125 g/L) of the dry biomass now consisted of PHB. In comparison for this same time period, dry biomass measurements for *A. vinelandii* showed an increase of more than 50% to 0.34 g/L with 31% (0.11 g/L) of this biomass comprising PHB. *P. putida* mt2 reached a final dry biomass of 0.2 g/L and produced no PHB.

These results confirm that Y88^T produced a greater biomass and PHB than the nitrogen-fixing, PHB-accumulating *A. vinelandii* under identical growth conditions. Y88^T produced up to 97% more PHB after 72 hrs than *A. vinelandii.* Compared to *A. vinelandii,* which other authors have shown to produce up to 40% PHB (including 3-hydroxyvalerate) on glucose but with valerate added to the glucose as a precursor (Page et al., 1992; Pettinari et al., 2001), Y88^T has demonstrated a superior ability to produce PHB to high levels of its biomass under our standard growth conditions compared to another nitrogen fixer such as *A. vinelandii* under similar growth conditions.

3.2.2.2.1 Acetylene reduction assay

Y88^T, *A. vinelandii* and *P. putida* mt2 cells were assayed for nitrogenase activity in cells cultivated in glucose-enriched medium for 24 hrs either with or without supplemented ammonium. Y88^T cells cultivated in ammonium-free medium attained an OD_{600} equal to 0.7 by 24 hrs. A negative C_2H_2 reduction assay was obtained (Fig. 3.14E) despite its ability to fix molecular nitrogen, demonstrating the lack of assayable nitrogenase activity in Y88^T under growth conditions with no initial supplemented ammonium. In contrast, *A. vinelandii* cells grown in the ammonium-free medium to an OD_{600} of 1.0 at 24 hrs gave a positive C_2H_2 reduction assay (Fig. 3.14C). As expected, *P. putida* mt2, incapable of fixing molecular nitrogen, gave a negative C_2H_2 reduction assay at 24 hrs and did not show any measurable growth during this period (chromatogram not shown).

Y88^T cells cultivated in ammonium-supplemented medium gave a negative assay for nitrogenase activity at an OD_{600}=0.5 but gave a positive assay for nitrogenase activity at 24 hr when the cells reached a high OD_{600} (> 2.0) (Fig. 3.14F). Under these same growth conditions, *A. vinelandii* cells gave a positive C_2H_2 reduction assay for nitrogenase activity but the peak was much smaller (Fig. 3.14D) compared to that obtained for *A. vinelandii* cells grown in the ammonium-free medium (Fig. 3.14C). *P. putida* mt2 cells grown in the ammonium-supplemented medium gave a negative C_2H_2 reduction assay as expected (chromatogram not shown).

The difference between Y88^T and *A. vinelandii* in their abilities to produce a positive ethylene peak when grown in ammonium-free medium is surprising given that both are capable of fixing nitrogen. Two points are noteworthy in this regard. First, Y88^T cells gave a positive ethylene peak only when their OD_{600} was greater than approximately 1.4 in the ammonium-supplemented medium. At OD_{600} measurements less than 1.4, despite growing at a high rate, the cells always gave a negative C_2H_2 reduction assay, even when an equivalent cell biomass was used in the C_2H_2 reduction assays to compensate for the difference

in the number of cells in the assays (data not shown). The switch to nitrogen fixation coincided with mid-log-phase growth, usually between 20 and 22 hr post-inoculation and an OD_{600} greater than 1.4. Since $Y88^T$ never reached an OD_{600} of greater than 0.7 in the ammonium-free medium, this could have been a factor in producing the negative result, since any active nitrogenase may not have been present at detectable levels in the assay when measured at low OD_{600}. Secondly, the results in this study suggest that the ability of $Y88^T$ to fix nitrogen relies on the presence of minimal amounts of available ammonium ion at low OD_{600} to initially promote growth. Thus in the absence of sufficient initial quantities of this supply, growth is restricted, as is the ability to initiate nitrogen fixation.

The ethylene peak produced by $Y88^T$ in ammonium-supplemented medium (Fig. 3.14F) was much larger relative to that produced by *A. vinelandii* cells grown in parallel in the same medium (Fig. 3.14C). This result is not surprising considering that *A. vinelandii* experiences variable inhibition of nitrogenase activity by the addition of ammonium chloride (Laane *et al*., (1980). Thus, growth in our glucose-enriched medium supplemented with 0.1 g/L of ammonium chloride, while favourable for $Y88^T$ growth and nitrogen fixation, does not promote growth in *A. vinelandii* during the first 24 hrs post-inoculation in the same way that it does in $Y88^T$. The addition of ammonium chloride to the growth medium therefore does not appear to have the same extended inhibitory effect on$Y88^T$ growth and nitrogenase activity as it does on *A. vinelandii* growth in our standard ammonium supplemented medium.

3.2.2.3 Comparison between $Y88^T$ and WP01

3.2.2.3.1 Growth and PHB production

The growth and PHB-producing abilities of $Y88^T$ and WP01 were compared to determine which of these two PHB producers accumulated the most PHB when cultivated on glucose or acetate. This work was done primarily to examine the potential of these bacteria as PHB production strains. Although it was

established in section 3.2.1.2.2 that Y88^T preferred glucose over acetate as a carbon substrate and produced more PHB when grown on glucose, it was of interest to determine whether another member within the same family produced similar amounts of PHB to Y88^T under the same glucose or acetate-enriched growth conditions and to directly compare the two species side-by-side. Also of interest was whether the ability of Y88^T to fix molecular nitrogen conferred an advantage to produce higher levels of PHB on glucose and acetate than that produced by a phylogenetically similar PHB producer, WP01, lacking any nitrogen-fixing genes in its genome and therefore incapable of fixing molecular nitrogen (Strabala, Macdonald & Liu, unpublished results).

Y88^T and WP01 were cultivated in parallel in either a glucose- or acetate-enriched minimal nitrogen medium. For Y88^T acetate growth, two different minimal nitrogen media were used: the same minimal nitrogen (0.1 g/L NH_4Cl-supplemented) medium used for all subsequent Y88^T experiments in this study and the basal salts medium (1× BSM) used in our laboratory as the standard WP01 growth medium. For WP01, three different media were used: the Y88^T minimal nitrogen medium described above, basal salts medium (1× BSM) supplemented with 0.1 g/L NH_4Cl and a variant of the 1× BSM containing the same ingredients as the 1× BSM but with no added nitrogen source.

Cells cultivated in the glucose-enriched nitrogen-fixing medium (supplemented with 0.1 g/L NH_4Cl) attained a higher OD_{600} for Y88^T after 24 hr than WP01 (Fig. 3.15). The dry biomass of 0.8 g/L obtained for Y88^T consisted of approximately 60% PHB (0.47 g/L) whereas a lower dry biomass of 0.5 g/L for WP01 consisted of 54% (0.27 g/L) PHB (Fig. 3.16). The ability of Y88^T to fix nitrogen therefore confers a growth advantage in a glucose-enriched minimal nitrogen medium, enabling it to attain a higher amount of biomass and PHB production than WP01 under the same growth conditions.

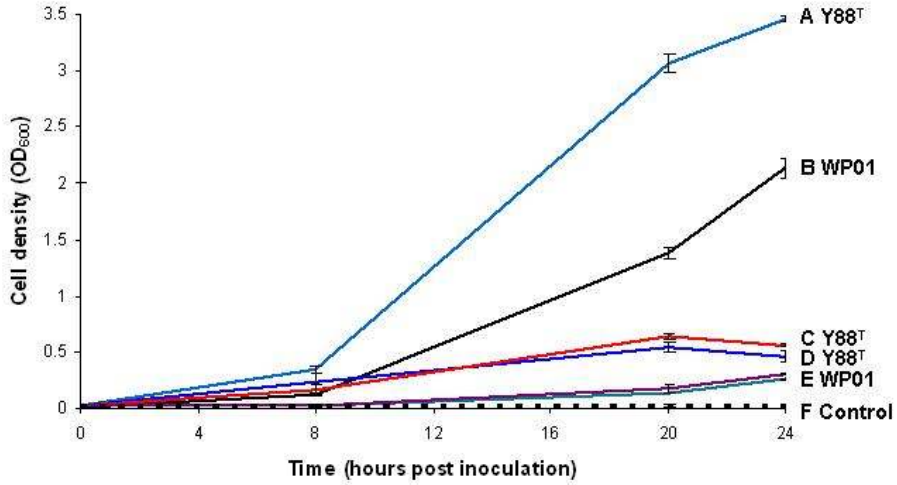


Figure 3.15 Growth of Y88^T and WP01 in a glucose-enriched, acetate-enriched and carbon-free medium

Comparative growth of Y88^T on glucose-enriched nitrogen-fixing medium (A), acetate-enriched nitrogen-fixing medium (C) and acetate-enriched 1x BSM (D) with WP01 on glucose-enriched nitrogen-fixing medium (B), acetate-enriched variant medium without supplemented nitrogen (E purple line) and acetate-enriched 1x BSM (E green line). Three negative controls included were mock inoculations in acetate-enriched 1x BSM medium and mock inoculations in acetate-enriched nitrogen-fixing medium and WP01 cells in acetate-free 1x BSM medium. The three control growth curves are superimposed (F) due to no measurable growth for all. Error bars represents standard deviations, n=2.

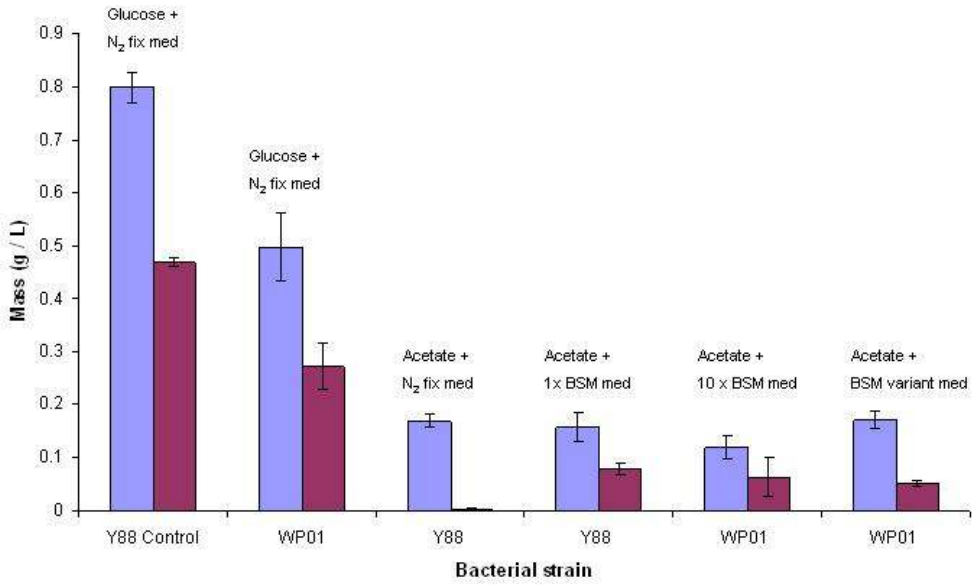


Figure 3.16 PHB and biomass produced by Y88^T and WP01 grown for 24 hr in glucose-enriched nitrogen-fixing medium, acetate-enriched nitrogen-fixing medium, acetate-enriched 1x BSM and acetate-enriched variant 1x BSM with no nitrogen source

Blue bars=dry biomass produced after 24 hr, maroon bars=PHB produced after 24 hr. Error bars represent standard deviations, (n=2).

Y88^T cells cultivated in an acetate-enriched nitrogen-fixing medium and acetate-enriched 1x BSM demonstrated that Y88^T achieved a similar OD_{600} of approximately 0.5 after 24 hr and a similar dry biomass (0.16 g/L) in both media (Fig. 3.16). Despite producing a similar biomass, 50% of the dry biomass produced from Y88^T cells grown in 1× BSM consisted of PHB (0.08 g/L) compared to only 1.3% (0.002 g/L) of the dry biomass produced from Y88^T cells grown in the nitrogen-fixing medium (Figs. 3.15 & 3.16). Thus for acetate growth, it would be beneficial to substitute the nitrogen-fixing medium I generally used for Y88^T PHB production, for 1× BSM. Although use of this medium could boost Y88^T PHB production in an acetate-enriched environment, Y88^T cells produce a higher amount of PHB (0.5 g/L) in glucose-enriched nitrogen-fixing medium than in acetate-enriched 1× BSM (0.08 g/L) after 24 hr.

Like Y88^T, WP01 produced significantly less dry biomass and PHB (Fig. 3.16) in both acetate-enriched media compared with that produced in the glucose-enriched media. The amount of dry biomass produced by WP01 was similar to that produced by Y88^T in 1× BSM enriched with acetate. The amount of PHB produced by these sphingomonads in the acetate-enriched media did not differ significantly from each other.

For both Y88^T and WP01, a glucose-enriched minimal nitrogen medium promotes higher biomass production compared to acetate-enriched minimal nitrogen medium. Glucose is therefore the preferred carbon substrate for both Y88^T and WP01 rather than acetate, if achieving high amounts of PHB is desirable, since greater PHB production correlates with higher biomass in these bacteria under these experimental conditions. Furthermore, Y88^T produces a higher amount of biomass and PHB than WP01 on glucose-enriched and acetate-enriched 1x BSM media, although in the latter medium the difference in PHB produced between the two strains is less pronounced.

3.2.2.4 Carbon and nitrogen utilisation by Y88^T

3.2.2.4.1 Glucose assay results

A glucose assay was carried out on Y88^T cells grown in a glucose-enriched (5 g/L) minimal nitrogen (0.1 g/L) medium over 72 hr (Fig. 3.17) with sampling times coinciding with different phases of growth during which i) ammonium supplied in the growth medium was being assimilated, ii) nitrogen fixation occurred as determined by a positive acetylene reduction assay and iii) the cells had reached stationary phase. The assay was carried out to characterise the rate at which Y88^T consumes glucose for each growth interval between sampling, as well as at which stage of the growth process most of the glucose is consumed. This assay was required to determine when the highest rate of glucose consumption occurred and if the rate at it was consumed correlated with the rate at which PHB was produced (see Fig. 3.18, blue bars, below).

Glucose assay results showed that Y88^T consumed glucose almost linearly between 10 and 30 hr post inoculation and that most of the initial glucose supply of 27mM (5 g/L), had been depleted from the medium by 50 hr (Fig. 3.17). The depletion of glucose from the medium correlated with the production of PHB being produced (Fig. 3.18). All the glucose was depleted by 72 hr.

3.2.2.4.2 Total dissolved organic carbon (DOC) in spent medium

The amount of total inorganic carbon (TIC) and DOC in the spent medium was determined during Y88^T growth on a glucose-enriched minimal nitrogen medium. This was to establish how much DOC remained in the growth medium following glucose depletion and PHB accumulation throughout the growth process.

The DOC results confirmed that more than 1 g/L carbon, which equals 50% of the initial 2 g/L of carbon in 5 g/L of glucose, was removed from the medium

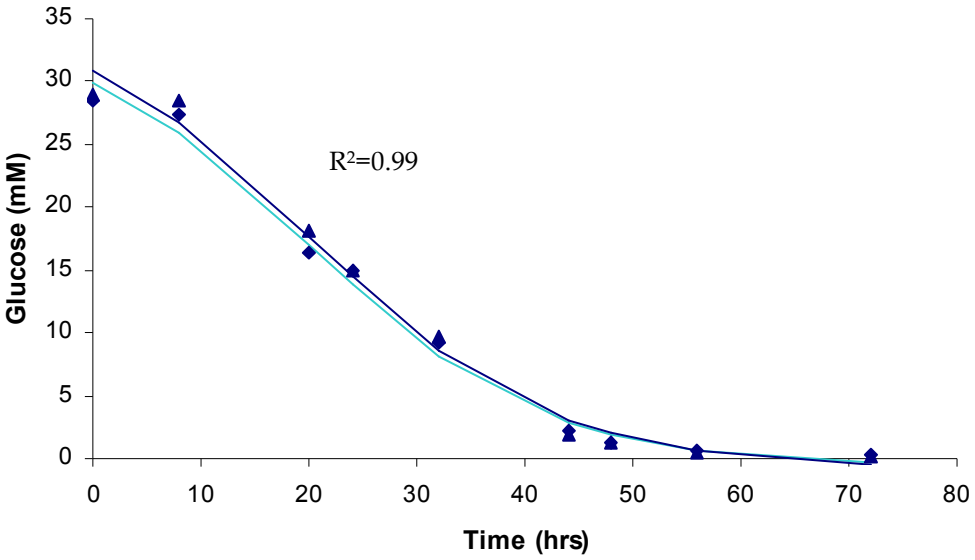


Figure 3.17 Glucose depletion by Y88^T

Glucose remaining in the spent medium during Y88^T growth over 72 hrs in a carbon-enriched, minimal nitrogen medium, (n=2).

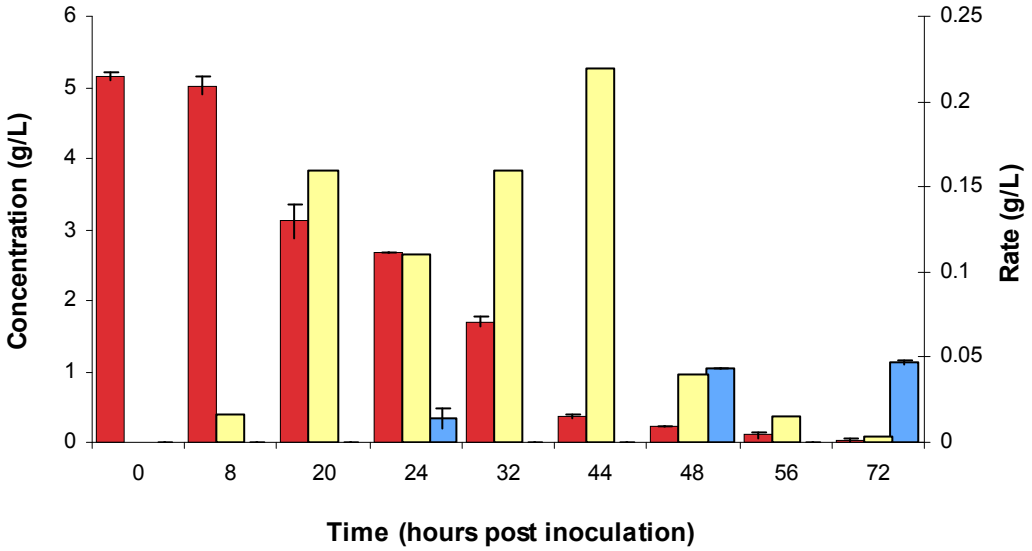


Figure 3.18 PHA production by Y88^T during glucose depletion

Glucose remaining in spent medium (red bars) during Y88^T consumption of glucose showing rate of glucose consumption (grams of glucose consumed per hr (g/L/hr) calculated for each time period between sampling (yellow bars) and the amount of PHB (blue bars) converted from glucose. Error bars represent standard deviations, n=2. Some error bars are not evident on graph due to small sampling errors.

by 44 hr (Fig. 3.19). The rate of carbon consumption by Y88^T during the first 8 hr of growth was 0.02 g/L/hr. Between 8 and 20 hrs, the rate of carbon consumed increased slightly to 0.03 g/L/hr. Between 20 and 28 hr, carbon was depleted at a rate of 0.08 g/L/hr, corresponding to the mid to high log phase of growth when Y88^T cells are actively fixing nitrogen (as shown by a positive C_2H_2 reduction assay for nitrogenase activity). During the late log phase of growth and early stationary phase, carbon consumption by Y88^T occurred at a rate of 0.07 g/L/hr. The highest rate of carbon consumption occurred between 44 and 48 hr at a rate of 0.3 g/L/hr. This coincided with the early- to mid- stationary phase, suggesting that although the cells were well past the highest rate of growth by then, an increased demand for carbon was presumably required to meet underlying metabolic demands of more cells after the growth spurt, in addition to other activities such as PHB production. Thus significantly more carbon is required during this period of the stationary phase than during any preceding stage of growth, including the mid to high log phase of growth when the cells were shown to fix nitrogen, a metabolically demanding activity, both for ATP production and reducing equivalents, which are produced via respiration.

3.2.2.4.3 Ammonium assimilation

Ammonium assimilation by Y88^T during growth over 72 hr on a glucose-enriched minimal nitrogen-supplemented (0.1 g/L) medium was determined. The aim was to determine when ammonium becomes depleted in the medium as a result of Y88^T utilisation as well as identifying when the probable switch from ammonium assimilation to fixing nitrogen occurs. An ammonium assay was carried out on the spent medium at the same time intervals as that for which glucose and carbon were assayed. Ammonium assay results show that the ammonium supplemented in the medium at 0.1 g/L, was completely removed by 20 hr growth with only minimal amounts at the limits of detection extant in the medium thereafter (Fig. 3.20). By 20 hr, Y88^T cells had reached the log phase of growth and with no ammonium nitrogen available, Y88^T cells had to fix nitrogen to maintain an increase in OD_{600}. A positive C_2H_2 reduction assay

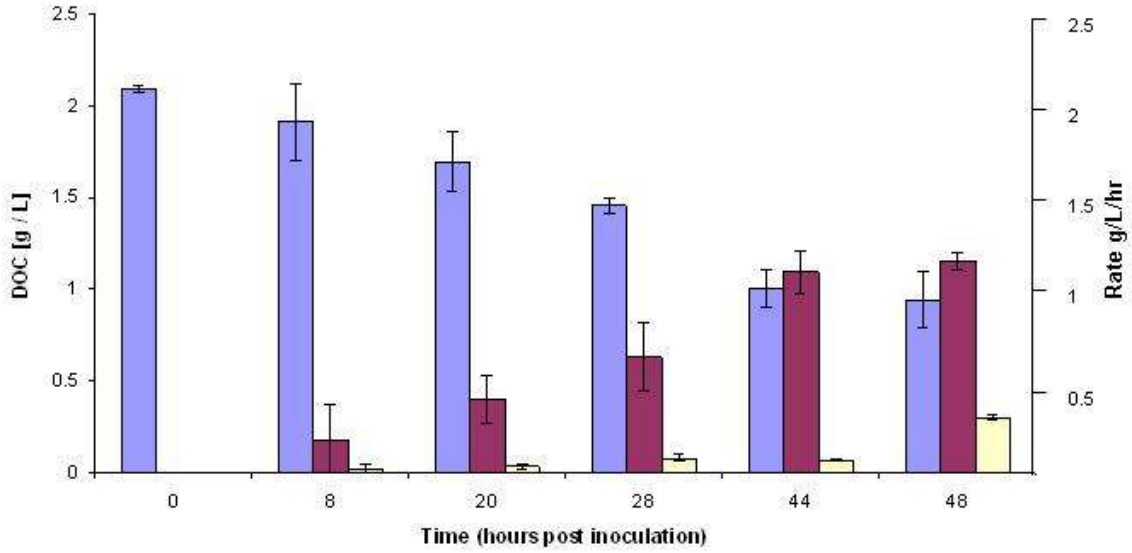


Figure 3.19 Total dissolved organic carbon (DOC) change during Y88^T growth

DOC removed (red bars) by Y88^T from the glucose-enriched minimal nitrogen medium during 48 hr growth. Blue bars=carbon remaining in medium, red bars=carbon utilised, yellow bars=the rate of carbon consumption. Error bars represent standard deviations, (n=3).

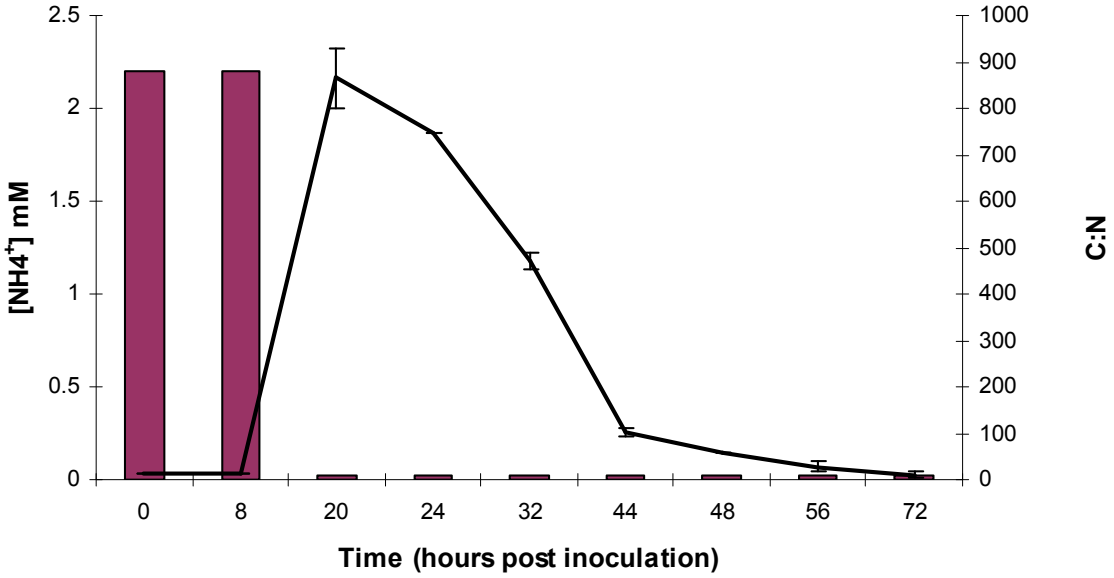


Figure 3.20 Change in C:N ratio during ammonium assimilation by Y88^T

Ammonium remaining in medium (red bars, left y axis) during Y88^T utilisation of glucose and the change in the C:N ratio (line graph, right y axis) during 72 hr of growth. Error bars represent standard deviations, (n=2).

at 24 hr confirmed active nitrogenase in Y88^T cells, indicative of nitrogen fixation by these cells at that stage of growth. The switch from ammonium assimilation to nitrogen fixation therefore occurs between 8-20 hr of growth although a positive C_2H_2 reduction assay is only obtained during 22-24 hr despite the cells growing exponentially at 20 hr.

3.2.2.4.4 Change in carbon to nitrogen (C:N) ratio.

The consumption of glucose and ammonium assimilation by Y88^T over 72 hr changes the C:N ratio in the medium during growth, which is known to affect the amount of PHB that is accumulated. The C:N ratio was therefore calculated from the amount of glucose and ammonium consumed.

During the first 8 hr post-inoculation, very little change in the C:N ratio was evident (from approximately 13:1 to 12.5:1) (Figs. 3.20 and 3.21). From 8 to 20 hr, a significant change was observed with a high rate of glucose consumption (Fig. 3.21) and ammonium assimilation (Fig. 3.20). Ammonium nitrogen levels decreased to minimal levels during that time (Fig. 3.20) concomitant with a decrease in glucose by at least one-third of its original concentration. This change in glucose coincided with a dramatic change in ammonium levels (Fig. 3.20) to barely detectable levels which changed the C:N ratio from approximately 12.5:1 to 820:1 (Figs. 3.20 and 3.21). This change coincided with the early- to mid- log phase of growth. At 24 hr, when an C_2H_2 reduction assay gave a positive result for nitrogenase activity, the C:N ratio showed a slight decline to 745:1. At 30 hr, the C:N ratio was 460:1, which was a decrease in the C:N ratio approaching half that observed at 24 hr. Thereafter, the C:N ratio continued to decrease to 60:1 by 48 hr, coinciding with the highest rate of glucose removal (sections 3.2.1.7.1 and 3.2.1.7.2, Figs. 3.17 and 3.21). Although a 60:1 C:N ratio may still be considered high (Kostyál et al., 1997), relative to the C:N ratios obtained between 20 and 32 hr, it is much lower at 48 hr which likely indicates a high rate of metabolic activity. Indeed, at this point, the decrease in the C:N ratio in the medium correlates with the highest rate of PHB accumulation for Y88^T (see section 3.2.2.3.1). After 72 hr, the C:N ratio was

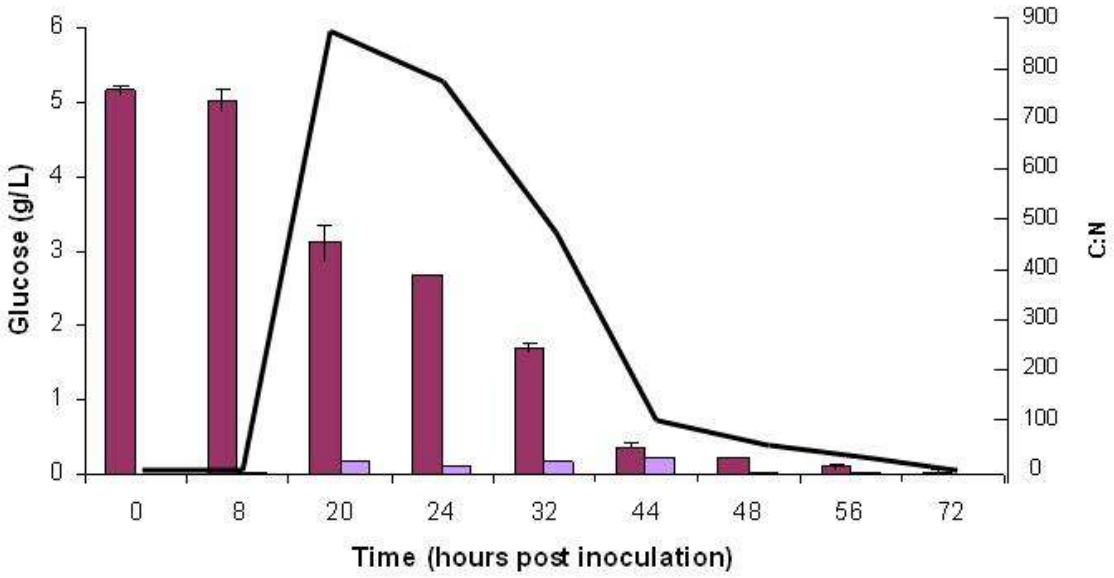


Figure 3.21 Y88^T glucose depletion: change in C:N ratio and rate of glucose depletion

Glucose remaining in spent medium (maroon bars; left y axis) during Y88^T consumption of glucose, showing the rate of glucose consumption (grams of glucose consumed per hr calculated for each time period between sampling (blue bars) and the change in the C:N ratio during 72 hr of growth (line graph; right y axis). Error bars represent the standard deviation, (n=2).

approximately 10:1, at which point approximately all the glucose had been utilised (Fig. 3.21).

3.2.2.4.5 Acetic acid assay

Since the pH of the growth medium acidifies to approximately 4.5 during Y88^T growth in a glucose-enriched minimal nitrogen medium, the possibility of any acetic acid buildup from glucose utilisation was considered. Growth on excess glucose under aerobic conditions causes the formation of acidic byproducts of which acetate is the most predominant (Luli & Strohl, 1990). An acetic acid assay (n=2), carried out to determine whether any acetic acid was present in the growth medium during Y88^T growth, showed that there are no detectable levels of acetic acid in the spent growth medium. An independent in-house analysis confirmed this result (M. Robinson, unpublished results). The decrease in pH could therefore not be attributed to the accumulation of acetic acid due to glucose-mediated aerobic acidogenesis.

Analysis of the extracellular biomass using gas chromatography revealed that during glucose-enriched growth, Y88^T produced an extracellular carbohydrate. Subsequent ^{13}C NMR ascertained that the carbohydrate consisted of two components likely to contain α- and β-carbon subunits of glucose (data not shown). Therefore, this carbohydrate is not likely to be the cause for the acidification of the media. Further ^{13}C NMR and ^{1}H NMR analysis is required to identify the type of carbohydrate produced by Y88^T in the extracellular environment likely to contribute to the acidification.

3.2.2.4.6 Conversion of glucose to PHB

Accumulated PHB was detected in Y88^T cells under all growth conditions tested in batch culture. Since the amount of glucose consumed by Y88^T and the amount of PHB produced was determined, it was possible to calculate the yield of PHB as well as the rate that glucose was being consumed and converted to PHB (Fig. 3.22, Table 3.3). The amount of PHB converted from glucose was determined at 20, 24, 48 and 72 hr with a three-fold increase in the amount of

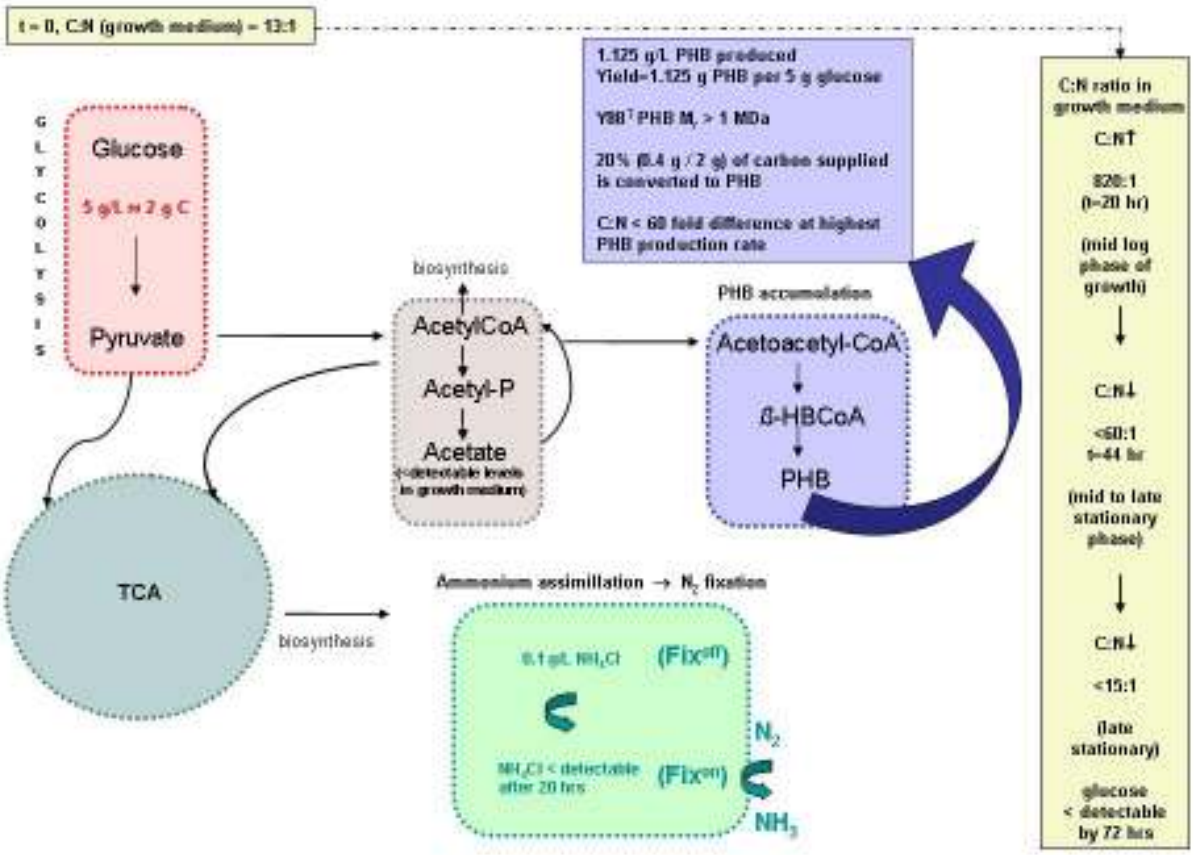


Figure 3.22 Schematic representation of glucose conversion to PHB and ammonium assimilation by Y88^T

Maximum amount of PHB accumulated, yield of PHB and percentage of carbon conversion to PHB calculated, are shown. Changes in the carbon to nitrogen ratio (C:N) during glucose utilisation, PHB accumulation and nitrogen fixation are indicated

Time (hours post inoculation)	Amount of glucose remaining in medium (g/ L)	Amount of glucose removed from medium (g/ L)	Amount of glucose consumed between each sampling period g/ L	Rate of glucose consumed g/ hr	Amount of PHA produced g/ L	Average Rate of PHA produced g/ hr	Yield of PHA (w/ w)	Yield of PHA %
0	5.157±0.06	0	0	0	–	–	–	–
8	5.022±0.13	0.135±0.06	0.14	0.017	–	–	–	–
20	3.114±0.23	2.043±0.17	1.9	0.16	–	–	–	–
24	2.682±0.00	2.475±0.06	0.43	0.11	0.34±0.13	0.014	0.14±0.00	14
32	1.701±0.06	3.456±0.0	0.98	0.16	–	–	–	–
44	0.369±0.04	4.788±0.1	1.33	0.22	–	–	–	–
48	0.216±0.00	4.941±0.06	0.15	0.04	1.04±0.01	0.03	0.21±0.00	21
56	0.099±0.04	5.058±0.06	0.12	0.015	–	–	–	–
72	0.036±0.03	5.121±0.08	0.06	0.004	1.125±0.02	0.004	0.22±0.03	22

Table 3.3. Glucose consumed by Y88^T during 72 hr growth in a glucose-enriched minimal nitrogen medium

The total amount of glucose removed and remaining in the medium at the end of each sampling period as well as the amount of glucose consumed between each sampling period was calculated to determine the average rate of glucose consumed was calculated for each sampling period. The amount of glucose converted by Y88^T into PHB was determined at 24-hr intervals (24 hr, 48 hr and 72 hr) as was the yield (g PHB produced / g glucose utilised) of PHB. From this yield, it could be determined that Y88^T is capable of producing PHB at a cost of $2.11US per kg with glucose as a substrate (see text over the page for more detail on yield). Standard deviations are given, (n=2).

PHB produced between 24 and 48 hr. A ten-fold increase in glucose consumption occurred between 20 and 44 hr with the highest rate of glucose consumption between 32-44 hr (Table 3.3). Importantly, this peak glucose consumption rate did not correlate with growth rate or nitrogen fixation. Growth rate was maximal in the 22-24 time interval, when Y88^T cells were also fixing nitrogen as indicated by a positive C_2H_2 reduction assay. A positive acetylene reduction assay was never obtained after 24 hr of growth, indicating that the cells fixed nitrogen briefly, but ceased this activity in favour of PHB accumulation in later stages of growth. Notably, PHB production doubled from approximately 40% to 80% of the dry biomass between 24-72 hr (Fig. 3.22) during which time nitrogen fixation had ceased, confirming that PHB production was not directly dependent on nitrogen fixation. Since the highest rate of PHB production occurred in the time interval following that during which the highest rate of glucose consumption occurred, this suggests a correlation between the highest rate of glucose consumption and its subsequent conversion to PHB. A final concentration of 1.125 g/L PHB was produced by 72 hr from 5 grams of glucose initially supplemented in the growth medium. During this time, the C:N ratio in the medium decreased from 820:1 (after 20 hr) to 15:1 (after 72 hr). A final yield of 0.22 g PHB/g glucose was achieved. Therefore approximately 20% of the supplemented glucose is converted to PHB (Table 3.3, Fig. 3.22). The yield of PHB at 0.22 g/g glucose therefore consists of 0.132 g carbon/g glucose used to make PHB. At a market price of US $0.44/kg of glucose, the cost of using glucose as a substrate for PHB production in Y88^T can be calculated as: (market price of glucose per kg) / (attainable yield of PHB) / 0.95 (unknown cost of adjustment to scale up from bench- to production-scale) (Webb et al., 2004). Thus factoring glucose in as a substrate, Y88^T can be used to produce PHB at a cost of US $2.11 per kg. Utilising glucose from lignocellulosic biomass, the cost is estimated to be less than US 50¢ per kg. (G. Lloyd-Jones, personal communication). These results should be taken into account in planning large-scale production of PHB from Y88^T.

Chapter 4: "Proteomic signatures" for the Fix^{on}/Fix^{off} physiology of $Y88^T$ during PHA synthesis

4.1 Introduction

"Proteomic signatures" are defined as a group or subset of proteins in a proteomic experiment that are characteristic of a specific physiological state or response (VanBogelen et al., 1999b). Proteins are identified as signatures by comparing protein profiles to identify proteins with altered abundance and mapping the altered state of the protein to a particular response. Proteomic signatures can be used as indicators of the growth phase of a cell population to predict the physiological state of cells (VanBogelen et al., 1999a; Voigt et al., 2004). Additionally, predictions can be made of other aspects of metabolism such as relating a protein to a specific pathway or function. For physiological studies, the examination of approximately 300 cytosolic proteins is considered to be a reasonable starting point to identify proteins that can be mapped as signatures to a particular state of the cell under specified growth conditions (Voigt et al., 2004).

The $Y88^T$ proteome was examined using 2-D DIGE and MS to gain a direct assessment of proteins in terms of their presence and relative abundance in actively growing $Y88^T$ cells. 2-D DIGE allowed the visualisation of the response of the $Y88^T$ proteome to conditions where no ammonium was supplemented in the growth medium, ammonium was supplemented at 0.1 g/L and the supplemented ammonium was depleted. Proteomic signatures of nitrogen-fixing and PHA-synthesising $Y88^T$ cells showed that the abundance of the $Y88^T$ nitrogenase enzyme was significantly altered at high OD_{600} (defined here as OD_{600} =2.0) in response to a switch from ammonium assimilation to nitrogen fixation, but that PHA synthesis proteins undergo more subtle changes and occur in high abundance regardless of the growth conditions. Several proteins not usually associated with nitrogen fixation were coordinately upregulated

during nitrogen fixation. Furthermore, this study supported a hypothesis that PHA synthesis in Y88^T is not dependent on nitrogen fixation.

4.2 Results

4.2.1 1DE

Based on examination of the growth and metabolism of Y88^T (Chapter 3), it was expected that the Fix^{on} and Fix^{off} states (Table 2.2) would show differences in protein signatures reflecting the physiological state of the cell. Protein was extracted from Y88^T cells grown for 24 hr in ammonium-free or ammonium-depleted media, representing the Fix^{off} and Fix^{on} physiologies of cell growth, respectively. To assess the quality and quantity of Y88^T protein extracts for subsequent 2-D DIGE analysis, Y88^T extracts were first subjected to 1-DE to determine whether protein concentrations were appropriate for 2-D DIGE analysis. Also, 1-DE was used to determine if differences in protein expression could be observed in 1-DE protein bands between extracts representing the Fix^{off} or Fix^{on} states. Since protein was extracted from whole cell pellets obtained by centrifugation of cell cultures, total protein extracts comprised cytoplasmic as well as membrane-bound/associated proteins. Protein extracts were also obtained from the same cultures after 72 hr, both at this time point corresponding to the Fix^{off} state when the cells were in stationary phase. These were compared with the 24 hr samples to see if any differences in protein abundance and/or band patterns could be detected on 1-D gels. Total protein extracts were diluted four-fold and 5 µL of soluble protein were subjected to 1-DE under reducing conditions (Fig. 4.1).

Additionally, protein extracts were obtained at different optical densities representing the same two physiological states of the cells as before. Since growth to OD_{600} of 0.5 and 1.0 did not produce a positive C_2H_2 reduction assay, protein extracts from cells grown to these optical densities were considered to represent the Fix^{off} state of the cells. Growth to an OD_{600} of 2.0 produced a positive C_2H_2 reduction assay, therefore protein extracts from cells grown to this OD_{600} were considered to represent the Fix^{on} state of the cells. Total protein

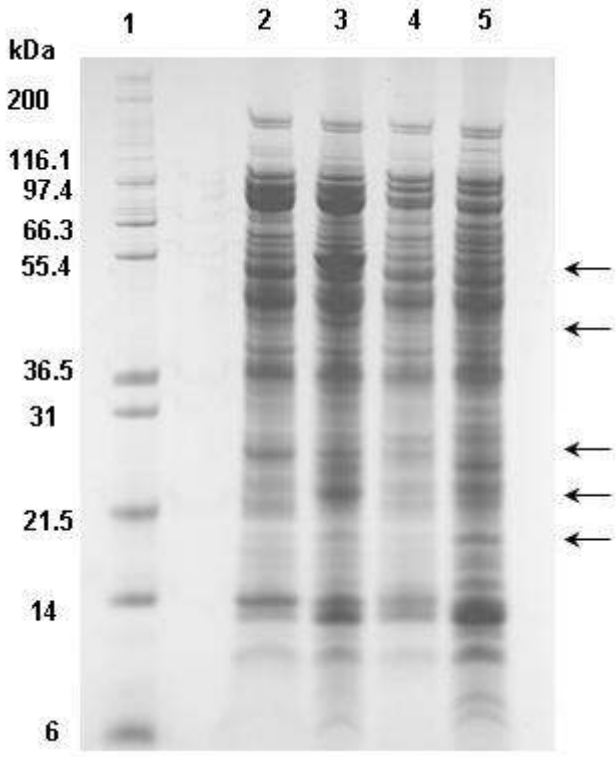


Figure 4.1 Coomassie-stained 1-DE gel showing comparison of Y88^{T} total protein extracts
Lanes 2 and 4: total protein extracts representing the Fix^{off} after 24 hr (lane 2) and 72 hr (lane 4) growth in ammonium-free media. Lane 3: total protein extract representing the Fix^{on} state after 24 hr growth in ammonium-supplemented medium that had become depleted of ammonium ion. Lane 5: total protein extract from stationary phase cells representing Fix^{off} after 72 hr. Differences in protein band intensities for extracts representing different growth conditions are indicated by arrows.

extracts were diluted four-fold, and 5 µL of soluble protein were subjected to 1-DE under reducing conditions (Fig. 4.2).

Differences were observed from 1-DE bands particularly at approximately 23 and 55.4 kDa in the Fix^{on} extracts representing 24 hr growth (Fig. 4.1, lane 3). Differences in several other highly abundant bands were evident across similar extracts in the 10-100 kDa range. For example, differences in protein abundance were evident for bands at approximately 23 and 55.4 kDa for cells grown to high OD_{600} (Fig. 4.2, lanes 2, 3, 8, 9, OD_{600}=2.0) similar to those at the same M_r in Fig. 4.1, lane 3, when compared to the corresponding band at 55.4 kDa in cells grown to lower OD_{600} (Fig. 4.2, lanes 4 and 5 (OD_{600}=1.0), lanes 6 and 7 (OD_{600}=0.5). The bands at 23 and 55.4 kDa that appeared in the protein extracts from cultures grown to high OD_{600} (Fig. 4.2, lanes 2, 3, 8, 9) as well as in the 24 hr extract (Fig. 4.1, lane 3) appear to be typical of the Fix^{on} state of the cells since these extracts were from the same cultures that gave positive C_2H_2 reduction assays (Chapter 3). Proteins extracted from cultures grown to lower OD_{600} (Fig. 4.2, lanes 4 to 7) showed similar band patterns and abundance to those obtained for cells grown in ammonium-free cultures that gave negative C_2H_2 reduction assays (Chapter 3). These 1-D protein profiles therefore appeared to be typical of the Fix^{off} state of the cells.

4.2.2 2-D/2-D DIGE

4.2.2.1 Differential protein abundance between the Fix^{on} and Fix^{off} physiologies of $Y88^T$ during PHA accumulation

To visualise differences in protein abundance between $Y88^T$ cells grown in ammonium-free and ammonium-supplemented conditions, representative samples of $Y88^T$ protein from each of the growth conditions were subjected to 2-D DIGE analysis. The selected samples were similar in total protein concentration and 1DE band patterns for a particular growth condition but showed differences in band profiles across growth conditions.

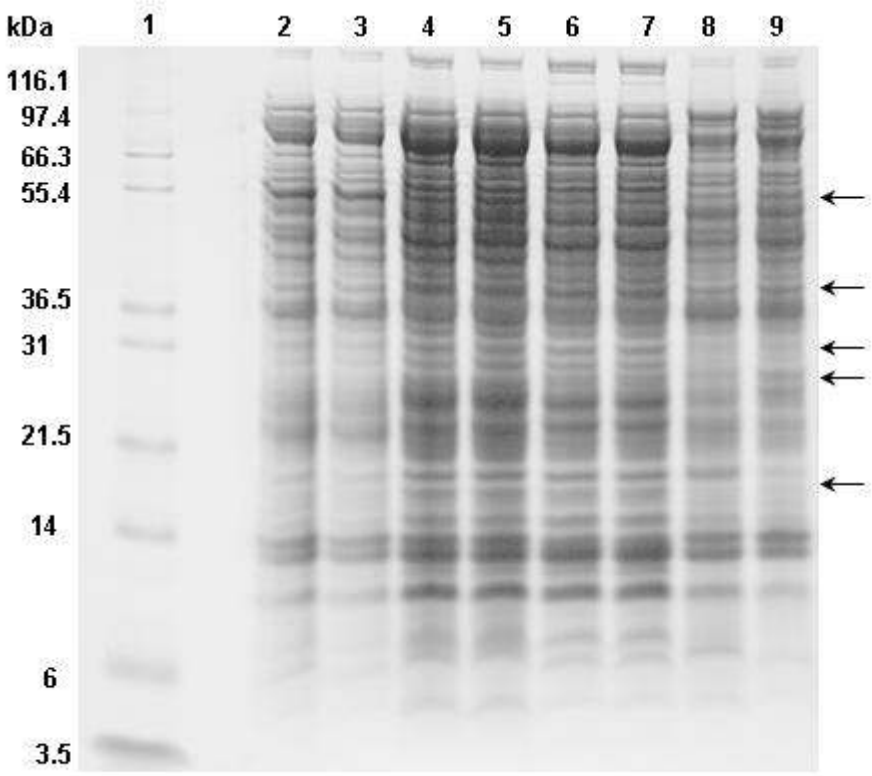


Figure 4.2.Total protein gel of Y88^{T} at various stages of culture growth

Coomassie-stained 1-DE gel showing comparison of Y88^{T} total protein extracts representing the Fix^{off} state at different optical densities (OD_{600}=0.5 (lanes 6, 7); OD_{600}= 1.0 (lanes 4, 5) and the Fix^{on} state of the cells at OD_{600}=2.0 (lanes 2, 3, 9); Fix^{off} state (no NH_4Cl) after 24 hrs (lane 8). Observed differences in protein band intensities between the different physiological states are indicated by arrows.

Since this was the first analysis of Y88^{T} proteins using 2-D DIGE, a master gel consisting of protein extracted from cells under ammonium-free (negative control) and ammonium-supplemented conditions (during which the ammonium was depleted and the cells switched from assimilating ammonium to fixing nitrogen) was generated using linear gradient pH 3-10 strips for first dimension IEF. This pH range was chosen to obtain an overview of total protein distribution for the standard Fix^{on} versus Fix^{off} physiology to gauge whether i) Y88^{T} protein extracts were amenable to 2-D DIGE analysis, ii) the Y88^{T} proteome consisted of predominantly basic or acidic proteins and how well these proteins separated within the pH range selected and iii) whether protein differences for the selected growth conditions could be readily detected for Y88^{T} using this type of analysis.

2-D DIGE of Y88^{T} extracts in the pI 3-10 range revealed in excess of 300 abundant proteins in the range 10-120 kDa (Fig. 4.3). These proteins represent only a small fraction of the Y88^{T} proteome as predicted from the Y88^{T} genome sequence. The presence of many proteins could therefore be masked by these abundant proteins or were below the set spot detection limit. Also proteins could be inadvertently excluded if they are outside the analytical window of pI 3-10 and 10-120 kDa. Several proteins showed altered abundance under both ammonium-free and ammonium-supplemented conditions. Several spots unique to each condition were detected. These included highly abundant proteins in particularly the 10-30 kDa range as well as higher molecular weight proteins in the 50-100 kDa range (Fig. 4.3). Although proteins were observed across the pI range, at least two-thirds of the Y88^{T} proteins were found within the neutral to acidic region of the gel (Fig. 4.3), corresponding to a pI range of between 4 and 7. Therefore, it was more appropriate to analyse subsequent samples using a narrower pI range of 4-7 that would allow greater resolution of these proteins, particularly some of those possibly masked beneath the more abundant proteins.

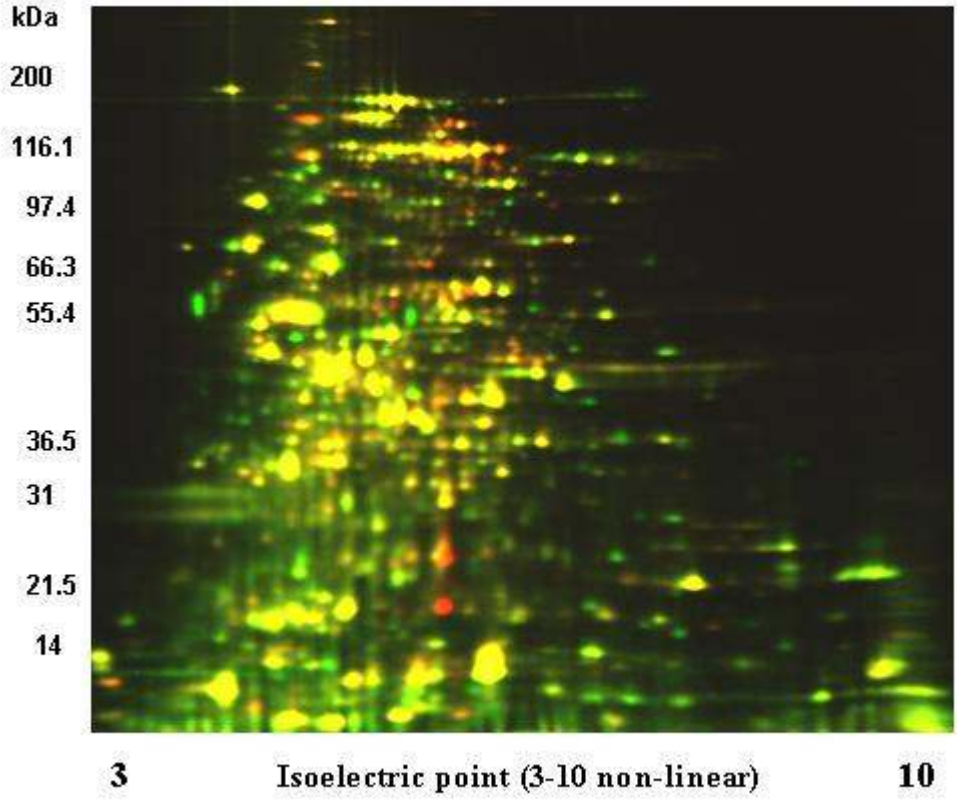


Figure 4.3 Overlay image of Y88^{T} protein profiles generated from minimally labeled pooled Fix^{on} and Fix^{off} samples

(Cy2 - yellow), Fix^{on} (Cy3 - green) and Fix^{off} (Cy5 - red) samples showing differentially altered proteins in the linear gradient pH 3-10 range. Yellow=proteins unaltered for both the Fix^{on} and Fix^{off} states of the cells; green=proteins with increased abundance for the Fix^{on} state of the cells; red=proteins with increased abundance for the Fix^{off} state of the cells.

2-D DIGE using a linear gradient of pH 3-10 confirmed that Y88^T total protein extracts were amenable to such analysis and that differences in protein abundance were readily detected using this method. Given these results and that the majority of Y88^T proteins appeared to be more acidic, subsequent samples were analysed using linear gradient pH 4-7 strips to increase the resolution of proteins over this narrower acidic range, unless otherwise stated.

4.2.2.2 Differential protein abundance at different optical densities for the Fix^{on} and Fix^{off} physiologies of Y88^T during PHA accumulation

2-D DIGE was carried out to examine the comparative response of the Y88^T proteome to conditions of ammonium supplementation (Fix^{off}) at OD_{600} of 0.5 and 1.0 relative to an OD_{600} of 2.0 under ammonium-depleted (Fix^{on}) conditions. Samples were similar in protein concentration and 1-DE band patterns for a given OD_{600} but different across optical densities. Notably, all samples clearly showed PHA granules in the discarded cell debris following protein extraction from cells at low or high OD_{600}, although differences in the amount of PHA were evident with a greater amount of PHA from cells grown to high OD_{600} (Chapter 3).

Comparison of the proteins in the pI 4-7 range for Y88^T cells for the Fix^{off} and Fix^{on} states showed approximately a total of 200 protein spots irrespective of growth conditions. Approximately 80 proteins were observed to be highly abundant on each of the three gels (Figs. 4.4A, 4.4B and 4.5B). Of these, forty-eight were selected for initial analysis. Selection criteria included specific targeting of proteins predicted to have altered abundance between physiological states. Predicted pI and M_r of proteins in public databases known to be involved in nitrogen fixation and PHA synthesis were used to target the most likely candidates, since at that time the Y88^T genome had not yet been sequenced. Additionally, several highly abundant proteins that appeared to be unaltered between physiological states were selected.

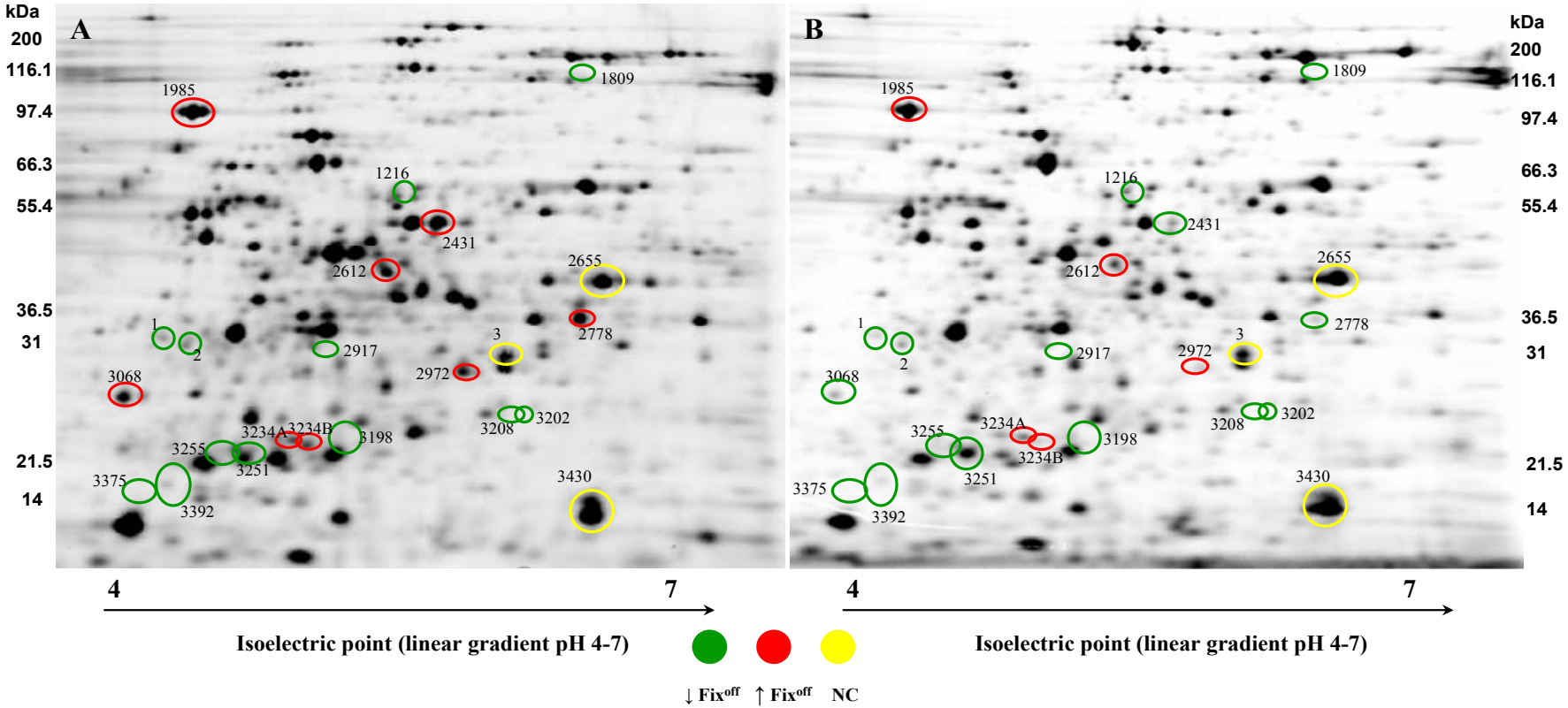


Figure 4.4 2-D DIGE (pI 4-7) profiles of Y88^{T} cells for the Fix^{off} physiological state of the cells (A and B)

2-D DIGE images (pI 4-7) showing protein profiles of Y88^{T} cells representing the Fix^{off} state of the cells at OD_{600}=0.5 (A) and OD_{600}= 1.0 (B). Red circles denote protein spots with increased abundance for the Fix^{off} state of the cells; green circles denote protein spots with decreased abundance for the Fix^{off} state of the cell; yellow circles denote protein spots with no significant change in abundance between the Fix^{off} and Fix^{on} physiological states of the cells. Spots 2655, 3 and 3430 were later identified to be acetyl-CoA C-acetyltransferase and 3-hydroxybutyryl-CoA dehydrogenase believed to be involved in PHA synthesis. Proteins identified corresponding to spot numbers in A and B that showed significant altered abundance are discussed in the text and are listed in Table 4.1, (n=3).

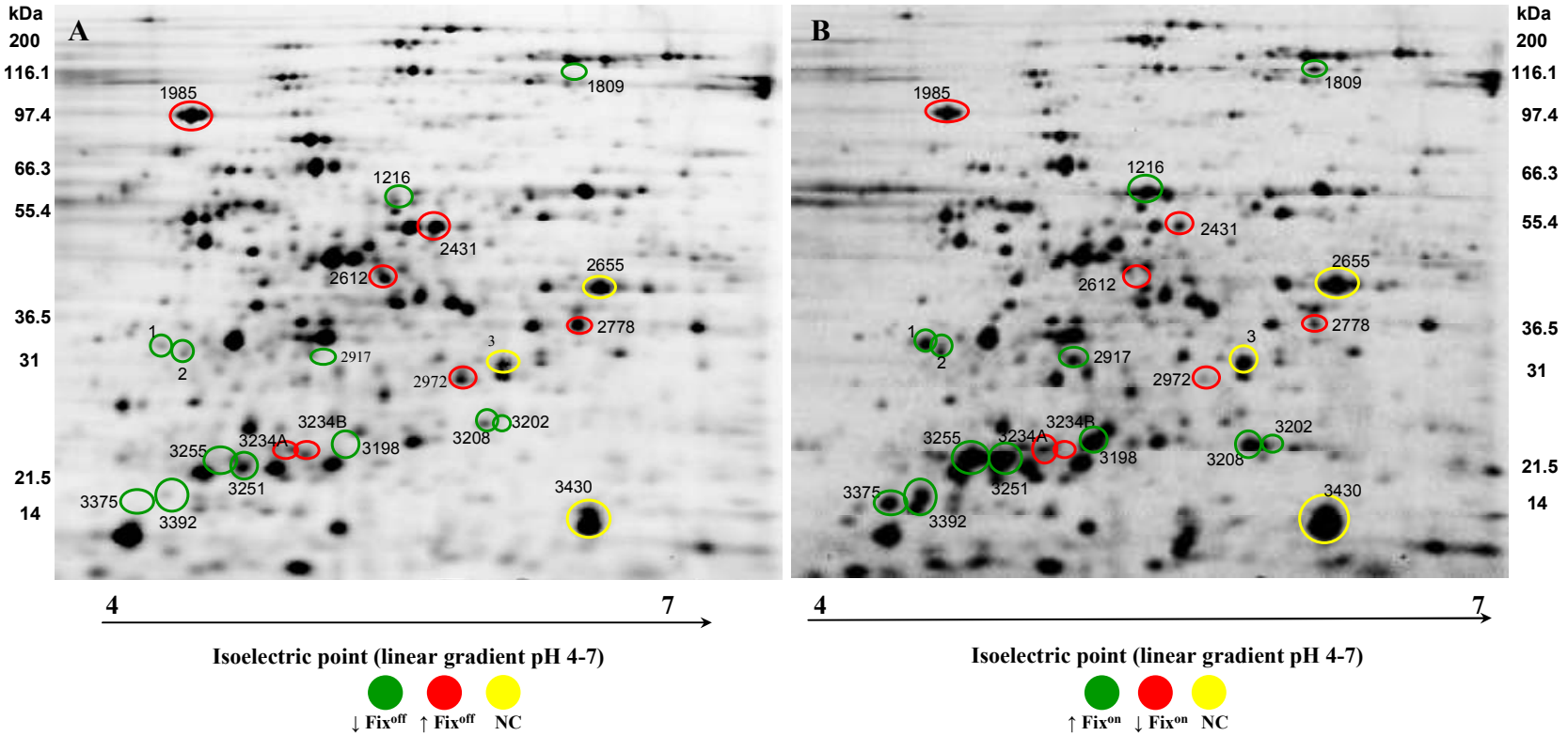


Figure 4.5. 2-D DIGE (pI 4-7) profiles of Y88^{T} cells for the Fix^{off} (A) and Fix^{on} (B) physiological state of the cells

2-D DIGE images (pI 4-7) showing protein profiles of Y88^{T} cells representing the Fix^{off} state of the cells at OD_{600}=0.5 (A, repeated from previous page for ease of comparison with the Fix^{on} state) and the Fix^{on} state of the cells at OD_{600}= 2.0 (B). In A, red circles denote protein spots with increased abundance for the Fix^{off} state of the cells and green circles denote protein spots with decreased abundance for the Fix^{off} state of the cells. In B, green circles denote protein spots with increased abundance for the Fix^{on} state of the cells and red circles denote protein spots with decreased abundance for the Fix^{on} state of the cells. Yellow circles in both A and B denote protein spots that show no change in abundance between the two physiological states. Spots 2655, 3 and 3430 were later identified to be acetyl-CoA C-acetyltransferase and 3-hydroxybutyryl-CoA dehydrogenase believed to be involved in PHA synthesis. Proteins identified corresponding to spot numbers in A and B that showed significant altered abundance are discussed in the text and are listed in Table 4.1, (n=3).

Generally, there was good congruency between the experimental pI and M_r and the theoretical pI and M_r based on the Y88^T translated ORF sequences for most of the proteins selected for analysis, with a few exceptions. Furthermore, analysis of the 2-D DIGE gels showed very little change for the majority of proteins observed at low or high OD_{600}, suggesting that the changes in extracellular nutrient availability and culture conditions had no visible effect on their abundance throughout growth, irrespective of whether the cells were assimilating ammonium or fixing nitrogen. This is not unexpected since many proteins would be involved in maintaining core metabolic activities within the cell and would not be expected to show significant changes.

Despite high similarity among protein profiles representing Fix^{on} and Fix^{off} physiological states, several proteins were either unique to a specific physiological state or showed altered abundance depending on growth conditions. For instance, proteins involved in nitrogen fixation were identified exclusively in extracts from cells grown to high OD_{600}. Some of these are discussed in the following sections. Also, proteins usually associated with other physiological functions were coordinately upregulated with the nitrogen-fixing proteins, suggesting putative functional roles (discussed in the following sections) for these proteins in nitrogen recycling under nitrogen-depleted conditions when the cells have to fix nitrogen.

4.2.2.2.1 Nitrogen-fixing and associated proteins

Two spots (spots 3255 and 3198) corresponding to the homodimeric dinitrogenase reductase (NifH), which together form component II of the hetero-oligomeric nitrogenase complex and are required for nitrogen fixation, were identified at approximately 23 kDa. One subunit had a pI of 4.7 (Fig. 4.5B, spot 3255), while the other had a pI of 5.2 (Fig. 4.5B, spot 3198). The protein spots of both subunits differed from the 31.8 kDa and pI 4.9 predicted from the Y88^T genome sequence, suggesting possible post-translational modification of this protein (see Chapter 6). NifH is known to be post-translationally modified by ADP-ribosylation in response to ammonium which would explain a

difference in observed and predicted M_r and pI. However, since these subunits are upregulated at high OD_{600} representative of the nitrogen-fixing physiology of $Y88^T$ at which point no ammonium is detected in the growth medium, this is unlikely. More likely is that the ADP-ribosylation response may be reversed under these conditions due to the depletion of ammonium in the medium. The differences between observed and predicted M_r and pI may therefore be a response to some post-translational modification other than that associated with ADP-ribosylation. Both subunits showed altered abundance under ammonium-depleted conditions at high OD_{600}: the former was upregulated by approximately 16.7-fold (spot 3255, *t*-test=0.01, Table 4.1, Fig. 4.5B) and the latter was upregulated 10.5-fold (spot 3198, *t*-test=0.019) (Table 4.1). These subunits were among the most abundant spots on the gels representing nitrogen-fixing physiology. Note that the LC-MS/MS data supporting this and all subsequent protein identifications in this chapter are given in Table 4.2.

An abundant protein spot was evident at approximately 58 kDa and pI 5.4 on the gel at high OD_{600} of 2.0 (Fig. 4.5B, spot 1216). This protein was later confirmed to be the NifK protein at its predicted M_r and pI. NifK (see following chapters) and constitutes part of the $\alpha_2\beta_2$-heterotetrameric component I (molybdenum iron β-subunit) of the nitrogenase complex and was more abundant for the Fix^{on} state relative to the Fix^{off} state. NifK was also identified in spot 2917 on the same gel and showed a 12-fold increase in abundance for the Fix^{on} state (*t*-test=0.015). This spot is likely to be a NifK fragment as it occurred at a much lower M_r and more acidic pI than expected from the $Y88^T$ genome sequence. Notably, the two NifH as well as the two NifK spots were absent from the profile of proteins representing the Fix^{off} state at OD_{600} of 0.5 and 1.0 (Figs. 4.4A, 4.4B). Given that these components are essential for nitrogenase activity, their absence on the gels representing growth to low OD_{600} suggests that at low growth, the presence of ammonium in the medium inhibits nitrogenase expression and thereby, its activity, explaining why $Y88^T$ gives a negative C_2H_2 reduction assay in the presence of ammonium at low OD_{600} (see Chapter 3, Fig. 3.8). This effect is physiologically consistent with ammonium-

Spot number	Protein ID	OD_{600}=0.5/ OD_{600}=0.1		OD_{600}=2.0/ OD_{600}=1.0		OD_{600}=0.5/ OD_{600}=2.0	
		Average Ratio	T-test value	Average Ratio	T-test value	Average Ratio	T-test value
1985	Chaperone DnaK (DnaK)	2.25	0.003*	–	–	–	–
2778	Ketol acid reductoisomerase[a]	2.78	0.004*	-1.18	0.5	3.27	0.036**
	tryptophanyl-tRNA synthetase[a]	2.78	0.004*	–	–	–	–
	Glyceraldehyde-3-phosphate dehydrogenase[a]	2.78	0.004*	–	–	–	–
2972	Translation elongation factor (Tuf)[b]	2.31	0.007*	–	–	3.62	0.006*
	Malate dehydrogenase[b]	2.31	0.007*	–	–	3.62	0.006*
2431	S-adenosylhomocysteine hydrolase (AdoHcy)	4.16	0.02**	–	–	6.76	0.024**
3208	hypothetical protein Saro_2615	-1.42	0.81	30.69	0.014**	-26.82	0.012*
3375	OmpA/MotB[c]	–	–	18.56	0.019**	-12.1	0.02**
	Ribosomal protein L7/L12[c]	–	–	18.56	0.019**	-12.1	0.02**
3392	OmpA/MotB[d]	–	–	15.55	0.02**	-12.24	0.019**
	Ribosomal protein L7/L12[d]	–	–	15.55	0.02**	-12.24	0.019**
3202	Translation elongation factor EF-1α[e]	-1.42	0.81	7.43	0.081	-10.57	0.007*
	nitrogenase molybdenum-iron protein beta chain (NifK)[e] (fragment)	-1.42	0.81	7.43	0.081	-10.57	0.007*
2612	S- adenosylmethionine (SAM) synthetase	–	–	–	–	6.21	0.009*
3255	Dinitrogenase reductase (NifH)	–	–	–	–	-16.7	0.01*

Table 4.1 Y88^{T} proteins showing altered abundance in for the Fix^{off} state (OD_{600} of 0.5 and 1.0) compared to the Fix^{on} state (OD_{600} of 2.0), identified using LC-MS/MS

Proteins identified in the same spots are marked with the same superscripts (a, b, c, d e). *t*-tests show significantly altered protein abundance at 99% confidence level (*) and at 95% confidence level (**) for a given OD_{600} pair. Unmarked *t*-test values represent altered proteins that did not attain the minimum >2-fold difference at 95% level of confidence. Positive average ratios denote enhanced protein abundance for 1st OD_{600} relative to 2nd OD_{600} in each OD_{600} pair; negative average ratios denote enhanced protein abundance for 2nd OD_{600} relative to 1st OD_{600} in each OD_{600} pair. Refer to Table 4.2 for supporting LC-MS/MS data for peptide matches to protein identifications that correspond to gel spot numbers in Table 4.1 and that are discussed in the body of the text throughout this chapter. Table 4.1 continued on next page.

Spot number	Protein ID	OD_{600}=0.5/ OD_{600}=0.1		OD_{600}=2.0/ OD_{600}=1.0		OD_{600}=0.5/ OD_{600}=2.0	
		Average Ratio	T-test value	Average Ratio	T-test value	Average Ratio	T-test value
3198	Dinitrogenase reductase (NifH)	–	–	–	–	-10.5	0.02**
1809	NI	–	–	–	–	-8.47	0.013*
2917	Nitrogenase molybdenum beta chain (NifK) fragment	1.84	0.53	6.57	0.09	-12.11	0.015**
3234A	ATP-dependent Clp protease, proteolytic subunit (ClpP)	2.42	0.037**	-1.36	0.095	3.28	0.015**
3234B	isochorismatase	2.42	0.037**	-1.36	0.095	3.28	0.015**
3430	phasin	-1.95	0.19	1.61	0.36	-3.12	0.056
3251	hypothetical protein BRADO5426	-2.39	0.05**	-4.64	0.05**	-1.94	0.05**
1216	Nitrogenase molybdenum beta chain (NifK)	-1.05	0.92	1.21	0.63	3.4	0.04**
1	NifU	–	–	–	–	–	–
2	Electron-transfer flavoprotein α-subunit (FixB)	–	–	–	–	–	–
	OmpA/MotB	–	–	–	–	–	–
3430	Phasin	–	–	–	–	–	–

Table 4.1 (cont.) $Y88^T$ proteins showing altered abundance in for the Fix^{off} state (OD_{600} of 0.5 and 1.0) compared to the Fix^{on} state (OD_{600} of 2.0), identified using LC-MS/MS

Proteins identified in the same spots are marked with the same superscripts (a, b, c, d e). *t*-tests show significantly altered protein abundance at 99% confidence level (*) and at 95% confidence level (**) for a given OD_{600} pair. Unmarked *t*-test values represent altered proteins that did not attain the minimum >2-fold difference at 95% level of confidence. Positive average ratios denote enhanced protein abundance for 1st OD_{600} relative to 2nd OD_{600} in each OD_{600} pair; negative average ratios denote enhanced protein abundance for 2nd OD_{600} relative to 1st OD_{600} in each OD_{600} pair. Refer to Table 4.2 for supporting LC-MS/MS data for peptide matches to protein identifications that correspond to gel spot numbers in Table 4.1 and that are discussed in the body of the text throughout this chapter.

Spot number	Protein ID	Unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
1985	Chaperone DnaK (DnaK)	R.QAVTNPDNTIFAVK.R(93) K.LGQAIYEK.E(81)	67507	2706	70	4.77	22
2778	Ketol acid reductoisomerase[a]	R.ADVDVIMIAPK.G(94) R.VLEDIQAGR.F(69) K.NFVLDNR.A(47)	36252	988	29	5.57	15
	tryptophanyl-tRNA synthetase[a]	R.INLTDDADAIMQK.V(89) R.SALDAILAK.G(66)	36986	468	11	5.76	6
	Glyceraldehyde-3-phosphate dehydrogenase[a]	R.AVGEVLPELK.G(55) K.VAINGFGR.I(47)	36036	148	10	6	7
2972	Translation elongation factor (Tuf)[b]	K.LLDQGEAGDNVGALIR.G(112) K.LIAPIAMDEGLR.F(71) K.TTVTGVEMFR.K(72) K.VGDEVEIIGLKPTAK.T(69)	43045	2263	117	5.16	32
	Malate dehydrogenase[b]	R.GGGGEIVALLK.T(93) K.VVGMAGVLDSAR.F(73) K.ANLQVSVDAVK.E(71) R.DDLLGINLK.V(58)	33461	755	23	4.92	16
2431	S-adenosylhomocysteine hydrolase (AdoHcy)	R.TANLILDDGGDATMFALWGAR.V(138) R.VEAGETLPEPANAEEIEFQR.A(97) R.VVVTEIDPICALQAAMEGYEVVTMEEAVKR.A(92) K.VACVAGFGDVGK.G(88) R.AEIAIAETEMPGLMALR.D(87)	50877	580	19	5.33	66

Table 4.2 LC-MS/MS peptide analysis of Y88^T proteins showing differential abundance for the Fix^{off} state (OD_{600} of 0.5 and 1.0) compared to the Fix^{on} state (OD_{600} of 2.0)

Highest scoring peptides (individual peptide scores in brackets) matched to Y88^T proteins digested from differentially abundant 2-D DIGE spots are given. Protein spots numbers correspond to those in Table 4.1 that are discussed in the body of the text throughout this chapter. Supporting peptide data confirming positive protein identifications is also given.

Continued on next page.

Spot number	Protein ID	Unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
3208	hypothetical protein Saro_2615	K.AAYLQPNFSSTK.D(61) K.TDFGEPFIK.V(46)	29433	322	22	5.92	4
3375	OmpA/MotB[c]	R.LEGQVGYLDAGNK.T(92) K.TLNVDNQTITAVGGR.T(89)	25899	697	12	7.03	11
	Ribosomal protein L7/L12[c]	K.LTVLEAADLAK.A(88) R.AITSLGLTEAK.A(79)	12683	552	11	4.65	43
3392	OmpA/MotB[d]	R.LEGQVGYLDAGNK.T(80) K.TLNVDNQTITAVGGR.T(91) R.QTIPNDDYMIYQSR.T(68)	25899	651	12	7.03	17
	Ribosomal protein L7/L12[d]	K.LTVLEAADLAK.A(88) M.ADIAALVEELSK.L(94) K.KIEEAGGTVEIK.X(78) K.IEEAGGTVEIK.X(76)	12683	1556	32	4.65	47
3202	Translation elongation factor EF-1α[e]	K.LLDQGEAGDNVGALIR.G(112) K.LIAPIAMDEGLR.F(71) K.TTVTGVEMFR.K(72) K.VGDEVEIIGLKPTAK.T(69)	43045	2263	117	5.16	32
	nitrogenase molybdenum-iron protein beta chain (NifK)[e] (fragment)	R.LVDAIADSNAHIHGK.K(37) R.VGFPIFDR.H(43) R.DTGVPLIR.V(44)	58437	335	29	5.37	7

Table 4.2 (cont.) LC-MS/MS peptide analysis of Y88^T proteins showing differential abundance for the Fix^{off} state (OD_{600} of 0.5 and 1.0) compared to the Fix^{on} state (OD_{600} of 2.0)
Highest scoring peptides (individual peptide scores in brackets) matched to Y88^T proteins digested from differentially abundant 2-D DIGE spots are given. Protein spots numbers correspond to those in Table 4.1 that are discussed in the body of the text throughout this chapter. Supporting peptide data confirming positive protein identifications is also given.
Continued on next page.

Spot number	Protein ID	Unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
2612	S- adenosylmethionine (SAM) synthetase	K.GIMDTDGNWAPGVPEEVER.V(98) K.ILQQLAADR.H(75) K.NIVAAGLAHR.C(70) K.IIVDTYGGASPHGGGAFSGK.D(66)	43503	2729	119	5.27	28
3255	Dinitrogenase reductase (NifH)	K.STTSQNTLAALADLGQR.I(119)	31.859	563	9	4.94	5
3198	Dinitrogenase reductase (NifH)	R.CVESGGPEPGVGCAGR(118) R.LGGLICNER.K(82) R.DNIVQHAELR.R(83) K.AQEIYIVMSGEMMAMYAANNISK.G(74)	31.859	9714	490	4.94	38
1809	NI	–	–	–	–	–	–
2917	Nitrogenase molybdenum beta chain (NifK) fragment	R.MYDGGTTLEEAGQAVHAR.A(54) R.LAGVEIPDDLAKER.G(46)	58437	192	13	5.37	7
3234A	ATP-dependent Clp protease, proteolytic subunit (ClpP)	R.MNDLYVK.Y(47) R.SFDIYSR.L(43) R.GMASDIEIQAR.E(77)	25292	730	37	4.92	13
3234B	isochorismatase	K.SIDATLLR.N(53) K.LGGAYGLGITYAK.T(82) R.TETYEHTTGIAR.K(72)	23573	775	29	5.06	15
3430	phasin	K.SQASFGELGEFAK.G(106) K.LANEAFQPISNR.V(89) K.ILSTGLQELGK.G(93) K.HFDAAVAASSK.N(73) K.SAFETLTAEFK.D(76)	33.963	12142	288	9.56	27

Table 4.2 (cont.) LC-MS/MS peptide analysis of Y88^T proteins showing differential abundance for the Fix^{off} state (OD_{600} of 0.5 and 1.0) compared to the Fix^{on} state (OD_{600} of 2.0)

Highest scoring peptides (individual peptide scores in brackets) matched to Y88^T proteins digested from differentially abundant 2-D DIGE spots are given. Protein spots numbers correspond to those in Table 4.1 that are discussed in the body of the text throughout this chapter. Supporting peptide data confirming positive protein identifications is also given. Continued on next page.

Spot number	Protein ID	Unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
3251	hypothetical protein BRADO5426	R.ISDTAFGTLEAAR.R(96) M.SATTSPSLETVR.I(64)	22861	450	11	4.93	15
1216	Nitrogenase molybdenum beta chain (NifK)	no high-scoring peptide data obtained*	58437	–	–	5.37	–
1	NifU	R.TNGLTSIDEVTNYTK.A(107) K.AGGGCSTCAEGIEGVLER.V(93) K.NSGILEDADGVGDVGAISCGDALR.L(89)	33542	4140	13	4.72	59
2	Electron-transfer flavoprotein α-subunit (FixB)	R.GVELGAVVMGDDR.A(99) R.IVEYAPAIVEADIVTK.V(83) K.LAEVLGAEYGGSRPLVQK.G(81)	39443	3555	256	5.95	37
	OmpA/MotB	K.TLNVDNQTITAVGGR.T(105) R.LEGQVGYLDAGNK.T(98) K.FAWQLIAGVR.A(90) R.QTIPNDDYMIYQSR.T(86)	25899	1655	44	7.03	30
3430	Phasin	K.VALVFGQMNEPPGAR.A(116) R.VALSGLTMAEYFR.D(112) K.APEFIEQSTEAAILVTGIK.V(113)	36948	5250	222	4.78	51

Table 4.2 (cont.) LC-MS/MS peptide analysis of $Y88^T$ proteins showing differential abundance for the Fix^{off} state (OD_{600} of 0.5 and 1.0) compared to the Fix^{on} state (OD_{600} of 2.0)
Highest scoring peptides (individual peptide scores in brackets) matched to $Y88^T$ proteins digested from differentially abundant 2-D DIGE spots are given. Protein spots numbers correspond to those in Table 4.1 that are discussed in the body of the text throughout this chapter. Supporting peptide data confirming positive protein identifications is also given. * denotes NifK peptide scores did not attain high enough scores to verify positive identification in this experiment. All subsequent experiments had high-scoring peptide data for positive verification of NifK.

mediated control of nitrogenase as previously demonstrated in *A. vinelandii* (Laane et al., 1980) in which ammonium chloride represses nitrogen fixation due to inhibition of the electron transport system to nitrogenase. However, the results presented here differ from those of Laane *et al.,* (1980) in that these authors investigated the activity of the nitrogenase enzyme whereas our study investigated changes in the expression of the nitrogenase enzyme. In $Y88^T$, nitrogen fixation appears to be low due to lack of expression of the proteins themselves. Whether this lack of expression of the nitrogenase proteins in $Y88^T$ is a consequence of the inhibition of the electron transport system to nitrogenase, is unknown. What is apparent is that whatever the constraint on the $Y88^T$ nitrogenase is, it is only removed at higher OD_{600} of greater than 1.0 once the ammonium has been depleted since at an OD_{600} of 1.0, NifHK are also absent from the gels (Fig. 4.4B).

Several other proteins known to be involved in nitrogen fixation displayed altered abundance at high OD_{600} under ammonium-depleted conditions relative to lower OD_{600} under ammonium-supplemented conditions. A highly abundant spot at approximately 23 kDa on the high OD_{600} gel (Fix^{on}) was identified as the hypothetical protein BRADO5426 (Fig. 4.5B, spot 3251). This protein showed a 2-fold difference in abundance between high (OD_{600}=2.0) and low OD_{600} (OD_{600}=0.5) and a 4.6-fold difference between OD_{600}=2.0 and OD_{600}=1.0 and may be implicated in some aspect of nitrogen fixation, since it is abundantly present for the Fix^{on} state. Its exact function has yet to be elucidated. It is located in the $Y88^T$ genome just upstream of *nif*U and *nif*W and the two subunits of the α/β-heterodimeric electron transfer flavoprotein (ETF), the electron transfer flavoprotein β chain (FixA) and electron transfer flavoprotein α chain (FixB) as well as several other genes involved in nitrogen fixation. This gene cluster likely forms an operon that co-coordinates the switch from the Fix^{off} to the Fix^{on} state. The NifU protein was upregulated 3.7-fold (*t*-test=0.09) in the Fix^{on} state (Fig. 4.5B, spot 1) relative to the Fix^{off} states (Figs. 4.4A, 4.4B, spot 1). NifU is an assembly scaffold protein that plays a role (together with NifS) in the activation of the Fe and MoFe proteins required for maturation of the nitrogenase enzyme

(Johnson et al., 2005). Similarly, FixB was upregulated for the Fix^{on} state at the high OD_{600} of 2.0 (Fig. 4.5B, spot 2) relative to the Fix^{off} states at lower OD_{600} (0.5 and 1.0). This protein is considered to sustain nitrogen fixation by providing reducing equivalents to the nitrogenase complex (Finocchiaro et al., 1988; O'Neill et al., 1998; Edgren & Nordlund, 2004; Scott & Ludwig, 2004). At low OD_{600}, the downregulation of this protein correlates with the absence of the nitrogenase metalloproteins for the Fix^{off} state and is consistent with the ammonium depression hypothesis of nitrogenase as a consequence of ammonium uptake inhibiting the flow of reducing equivalents to nitrogenase, thereby inhibiting nitrogen fixation (Scott & Ludwig, 2004). At a high OD_{600}, the enhanced abundance of these proteins suggests a coordinated response whereby ammonium depletion results in a derepression of the operon transcribing these proteins, concomitantly promoting expression of the genes transcribing the nitrogenase metalloproteins and allowing nitrogen fixation to proceed.

4.2.2.2.2 PHA synthesis and granule-associated proteins

One of the three major enzymes involved in PHB synthesis, β-Ketothiolase (acetyl-CoA C-acetyltransferase, PhbA) was identified at 41 kDa (Figs. 4.4, 4.5; spot 2655) and was common to all physiologies. This protein showed very little change in abundance (1.3-fold) at OD_{600} of 2.0 (Fig. 4.5B) from that at lower OD_{600} although it possible to see this change in abundance on the gel. At a 1.3-fold difference, PhbA was found to not be significantly different relative to its abundance at the lower OD_{600} of 0.5 (*t*-test=0.37, Table 4.1, Fig. 4.5A) and an OD_{600} of 1.0 (*t*-test=0.34, Table 4.1, Fig. 4.4B) despite the presence of higher PHB content in the cells. Additionally, the PHB granule-associated protein, phasin (PhaP) was identified at approximately 13-14 kDa as the most abundant spot on the gel under all growth conditions (Figs. 4.4, 4.5, spot 3430). PhaP was most highly expressed under high OD_{600} of 2.0 for the Fix^{on} state (Fig. 4.5B). BVA results confirmed an increase in PhaP abundance of 3.12-fold at a high OD_{600} of 2.0 compared to the low OD_{600} of 0.5 (*t*-test=0.05, Table 4.1), and a 2-fold (*t*-test= 0.19) increase in abundance from an OD_{600} of 0.5 to 1.0 (Table 4.1). An increase in abundance of PhaP therefore occurred from low to high OD_{600} consistent

with growth and PHB accumulation (Figs. 4.4, 4.5) but did not attain statistical significance at the 99% level of confidence.

Three points are notable in terms of increased PhbA and PhaP abundance throughout growth. First, the increase in PhbA, PhaP and OD_{600} occur concomitantly suggesting an interdependence of these two proteins from the early stages of growth and throughout changing growth conditions. Secondly, although changes in abundance occur with growth, both PhbA and PhaP are expressed at high levels even at low OD_{600}. It would therefore take a huge increase in absolute abundance to achieve at least the two-fold difference in abundance required to attain statistical significance at a 99% level of confidence. Thirdly, since PhaP was also highly expressed at the lower OD_{600} of 0.5 and 1.0, at which points no detectable nitrogen fixation occurred, these proteomic results augment the microbiology data by demonstrating that neither the presence nor the abundance of both PhbA and PhaP are dependent on nitrogen fixation. By inference, PHA synthesis is independent of nitrogen fixation.

4.2.2.2.3 Stress-related proteins

Several stress-related proteins were identified, of which two were upregulated for the Fix^{off} state. Chaperone DnaK (spot 1985) was upregulated 2.3-fold (*t*-test=0.0032) at an OD_{600} of 0.5 (Fig. 4.4A, Table 4.1) relative to an OD_{600} of 1.0 (Fig. 4.4B). DnaK is the catalytic component of the DnaK chaperone system involved in a wide variety of biological processes such as protein folding, disassembly, disaggregation and translocation as well as targeting proteins for degradation (Bukau & Horwich, 1998; Fink, 1999). Its higher abundance for the Fix^{off} state of $Y88^{T}$ cells coupled to its roles in disassembly and targeting proteins for degradation suggests that it may be upregulated in response to nitrogen turnover at lower OD_{600} when the cells are establishing themselves and starting to undergo higher rates of cell division. Additionally, DnaK has also been identified as a granule-associated protein in an *R. eutropha* phasin mutant strain (Pötter et al., 2004), as well as in *E. coli* expressing the phbCAB operon (Han et al., 2001). DnaK has been shown to enhance the specific activity of PHA synthase (Jossek et al., 1998) during granule formation and has also been shown

to be upregulated as a result of stress to the cells caused by the accumulating PHA granules increasing in size and displacing the intracellular contents during PHA production (Hecker & Volker, 2004; Tessmer et al., 2007). Given this latter role for DnaK, it could be expected that the abundance of this protein would be enhanced at higher PHA production. Its greater abundance at low OD_{600} when ammonium ion is available in the growth medium coupled to lower PHA production during the early stage of growth, therefore suggests that DnaK may play an increased role in protein turnover to accommodate an increase in the cells "scavenging" nitrogen and could be a link between nitrogen cycling and PHA accumulation in $Y88^T$. Considering its role in cell stress, one other possibility for the enhanced abundance of DnaK at a lower OD_{600} of 0.5 when the cells would be experiencing a high rate of cell division, could be a cellular response to stress resulting from rapid metabolism.

Another stress response protein, ATP-dependent Clp protease, proteolytic subunit (ClpP) was identified together with a second protein, isochorismatase. These proteins were picked as one spot since they did not completely resolve as individual spots on the 7 cm gels. Based on the observed pIs and M_r of both spots compared with the pIs and M_r predicted from the $Y88^T$ genome sequence, it was possible to distinguish ClpP as spot 3234A and isochorismatase as spot 3234B (Figs. 4.5A, 4.5B). Isochorismatase was upregulated 3.3-fold (*t*-test=0.015) for the Fix^{off} state at low OD_{600} of 0.5 (Fig. 4.5A, spot 3234B) relative to a high OD_{600} of 2.0 (Fig. 4.5B). Isochorismatase belongs to the family of hydrolase enzymes and catalyses the second step in the overall enterobactin biosynthetic pathway as well as the first part of the pathway converting isochorismate, in the presence of water, to 2,3-dihydroxybenzoate and pyruvate. Isochorismatase is produced in *E. coli* during the biosynthesis of enterobactin, an iron-chelating product derived from chorismic acid and involved in the transport of iron from the bacterial environment into the cell cytoplasm (Hamano et al., 2006). Isochorismatase is also involved in the biosynthesis of the antimicrobial compound phenazine in pseudomonads and is considered to contribute to the persistence of pseudomonad infections by affording a competitive advantage to

pseudomonads over other bacteria in its environment (Parsons et al., 2003). Its enhanced abundance in Y88^{T} at low OD_{600} of 0.5 (Fig. 4.4A) relative to the later stages of growth at an OD_{600} of 1.0 (Fig. 4.4B) or high OD_{600} of 2.0 (Fig. 4.5B), may be related to iron ion transport of the supplemented iron in the growth medium into the cells during early growth. Its downregulation at high OD_{600} may reflect the depletion of iron in the growth medium during nitrogen fixation since a high turnover rate of nitrogenase requires iron and increased demand for iron would result in depletion of the extracellular iron supply. Since isochorismatase is produced during enterobactin biosynthesis, its increased abundance in Y88^{T} cells during early growth was considered to possibly be related to enterobactin biosynthesis. However, this possibility was discounted in Y88^{T} since enterobactin biosynthesis is associated with limiting iron availability and during early growth of Y88^{T}, more iron is available than at the later stages of growth. Consistent with this notion, only one enterobactin biosynthesis gene, ferric enterobactin esterase-related protein (Fes), was identified from the Y88^{T} ORF translated protein sequences.

ClpP showed enhanced abundance for the Fix^{on} state of the cells at high OD_{600} of 2.0 (Fig. 4.5B, spot 3243A) relative to the Fix^{off} state of the cells at lower OD_{600} of 0.5 and 1.0 (Figs. 4.5A, 4.4B). This protein forms part of the Clp protease complex that can function either as a proteolysis regulator or as a molecular chaperone. As a subunit of this complex, ClpP is important for the degradation of misfolded proteins generated when cells are exposed to stress. During stress, abnormal or misfolded proteins will accumulate in the cell due to denaturation and errors in biosynthesis (Gottesman, 1996). Cellular response to the accumulation of these proteins involves increasing the synthesis of molecular chaperones and proteases, the former to assist with the correct folding/re-folding of proteins targeted for degradation and the latter to degrade the proteins that can not be folded (Schelin et al., 2002). In addition to eliminating abnormal proteins, proteolysis controls the level of naturally short-lived regulatory proteins and is therefore essential for cell homeostasis and optimal metabolic activity (Gottesman, 1996). The enhanced abundance of ClpP at a

high OD_{600} of 2.0 in $Y88^T$ cells fixing nitrogen may therefore be due to increased proteolysis of non-functional, short-lived regulatory proteins that may function in nitrogen cycling during nitrogen fixation. In $Y88^T$, accumulated non-essential or non-functional proteins would place additional stress on cells already stressed by the competing energy intensive metabolic processes of nitrogen fixation and PHA accumulation, if these were not removed. A higher abundance of ClpP may therefore be required to ensure optimal nitrogen cycling and nitrogen-fixing activities as well as optimal functioning of PHA regulatory proteins.

4.2.2.2.4 Other proteins

Ketol acid reductoisomerase, tryptophanyl-tRNA synthetase and glyceraldehyde-3-phosphate dehydrogenase (NAD-dependent) were identified in the same spot (Figs. 4.4, 4.5; spot 2778). These three proteins co-migrated as a result of having similar M_r and pIs (Table 4.1). Spot 2778 was more abundant (2.8-fold upregulated, *t*-test=0.004) at high OD_{600} of 2.0 (Fig. 4.5B) relative to the lower OD_{600} of 1.0 (Fig. 4.4B). It is unclear which of these proteins, if not all three, are responsible for this increase. What is clear is that this spot is present at OD_{600}=0.5 (Figs. 4.4A, 4.5A), disappears from the gel at OD_{600}=1.0 (Fig. 4.4B) and is again evident at OD_{600}=2.0 (Fig. 4.5B). Ketol acid reductoisomerase is implicated in the synthesis of branched-chain amino acids, and such branched-chain amino acids as valine and isoleucine have been reported to induce the synthesis of phosphotransbutyrylase (YqiS), an enzyme which enhances PHB accumulation in *Bacillus megaterium* (Vazquez et al., 2003). Enhanced abundance of ketol acid reductoisomerase at high OD_{600} could therefore result from an increased demand for biosynthesis of these amino acids due to increased PHB synthesis. Tryptophanyl-tRNA synthetase is involved in protein biosynthesis and its upregulation at high OD_{600} could be related to increased protein synthesis of proteins required for nitrogen fixation and increased PHA production. NAD-dependent glyceraldehyde-3-phosphate dehydrogenase catalyses the conversion of glyceraldehyde-3-phosphate, the first 3-carbon metabolite of glycolysis, that is eventually used for PHB synthesis (Han & Lee, 2003). An increase in abundance of glyceraldehyde-3-phosphate dehydrogenase

at higher OD_{600} could be due to the greater demand for glyceraldehyde-3-phosphate conversion to meet the increased demand of acetyl-coA and NADPH required for PHB synthesis (Han & Lee, 2003).

Several proteins involved in protein synthesis and translation also showed altered abundance. Translation elongation factor and NAD-dependent malate dehydrogenase (comigrating in spot 2972, Figs. 4.4, 4.5) were upregulated 2.3-fold (*t*-test=0.007) at an OD_{600} of 0.5 (Fig. 4.4A) relative to that at an OD_{600} of 1.0 (Fig. 4.4B) and 3.6-fold (*t*-test=0.006) relative to that at the highest OD_{600} of 2.0 (Fig. 4.5B, Table 4.1). Spot 3202, identified as elongation factor EF-1α, was upregulated 10.6-fold (*t*-test=0.0076) at high OD_{600} (Fig. 4.5B) relative to that at the low OD_{600} of 0.5 (Fig. 4.5A). OmpA/MotB was identified in two highly abundant spots (Fig. 4.5B, spots 3392 and 3375) together with ribosomal protein L7/L12 at approximately 12kDa and pIs of 4.5 and 4.7 on the high OD_{600} gel (Fig. 4.5B). OmpA is a porin outer membrane protein present in the outer membrane of many Gram-negative bacteria whereas MotB is a flagellar torque-generating protein and shares a conserved C-terminal domain with OmpA described as a peptidoglycan-anchoring domain (De Mot & Vanderleyden, 1994). Baev *et al*. (2006) demonstrated that nitrogen limitation activates the nitrogen regulatory response (Ntr) which induces the expression of several operons including genes coding for outer membrane proteins involved in peptide transport. Expression of the porin ompA gene, amongst others, that encodes the protein OmpA, was enhanced under nitrogen-limiting conditions. These authors further demonstrated that OmpA, in conjunction with two other porins, plays an important role in peptide transport during nitrogen limitation. Spot 3392 was upregulated 15.6-fold (*t*-test=0.02) at high OD_{600} (Fig. 4.5B) relative to the lower OD_{600} of 1.0 (Fig. 4.4B) and 12.3-fold (*t*-test=0.019) relative to the OD_{600} of 0.5. Spot 3375 was upregulated 18.6-fold (*t*-test=0.019) at high OD_{600} of 2.0 (Fig. 4.5B) relative to that at an OD_{600} of 1.0 (Fig. 4.4B) and 12-fold (*t*-test=0.02) relative to an OD_{600} of 0.5 (Fig. 4.5A, Table 4.1). High scoring peptide matches to OmpA/MotB in both spots made it difficult to ascertain which of the two OmpA/MotB proteins matched which spot. $Y88^T$ has seven

OmpA/MotB proteins. The upregulation of OmpA/MotB in Y88^T cells following ammonium depletion and during nitrogen fixation may be a response to low nitrogen levels in the cell or possibly increased cell membrane synthesis during nitrogen fixation.

Since Y88^T has only one gene encoding L7/L12 (based on the Y88^T genome, Scion in-house proprietary database) and the same high-scoring peptide matched each spot, this protein could have undergone post-translational modification, resulting in two charged isoforms of the same protein. Had the presence of the two spots been due to proteolysis, different peptides could have been expected to match uniquely to each spot and the proteins would have been observed at different molecular weights. This protein is absent from both the lower OD_{600} gels at corresponding M_r and pI (Figs. 4.4A, 4.4B). Since the ribosomal protein L7/L12 plays an important role in mRNA-ribosome interactions and in the binding of several factors essential for accurate translation (Rosen et al., 2001), its enhancement and possible post-translational modification as observed at high OD_{600} may reflect an adaptive requirement of the cell for the rapid synthesis of a specific set of proteins required for increased growth rate during nitrogen fixation. Its induction from low to high OD_{600} reflects the shift from the Fix^{off} to the Fix^{on} state of the cell and supports the notion that ribosomes play a role in sensing environmental changes as suggested by Rosen *et al*. (2001).

A hypothetical protein, Saro_2615 (spot 3208), showed highly altered abundance (26.8-fold) for the Fix^{on} state at high OD_{600} of 2.0 relative to the Fix^{off} state at lower OD_{600} of 0.5 (*t*-test=0.01). This protein is considered to be a putative outer membrane protein but its function is yet to be elucidated. Its coordinated upregulation, together with the aforementioned OmpA/MotB, suggested in this study to possibly be a response to low nitrogen levels in Y88^T, may reflect an enhanced nitrogen-sensing regulatory response that promotes nitrogen fixation.

Spot 2612, identified as S-Adenosylmethionine (AdoMet) synthetase, which is known to play a central metabolic role in all organisms, was downregulated 6.21-fold (*t*-test=0.009) at high OD_{600} (Fig. 4.5B) relative to that at the low OD_{600} of 0.5 (Fig. 4.5A). AdoMet synthetase is implicated in an intrinsic role in microbial secondary metabolism and the possible regulatory control of morphological differentiation (Okamoto et al., 2003). It catalyses the formation of AdoMet, which is the primary methyl group donor in biological systems and known to be involved in polyamine biosynthesis (Takusagawa et al., 1996). After donation of the methyl group from AdoMet to a methyl acceptor, S-Adenosylhomocysteine (AdoHcy) is formed which is hydrolysed by AdoHcy hydrolase. The latter protein (identified in spot 2431) was shown to have a 4.2-fold higher abundance at OD_{600} 0.5 relative to 1.0 and a 6.8-fold decrease in abundance for the Fix^{on} state at OD_{600} 2.0 relative to the Fix^{off} state at OD_{600} 0.5. Elevated levels of AdoMet synthetase are associated with increased levels of intracellular AdoMet (Okamoto et al., 2003). Another important function of AdoMet is its role as the sole donor of an aminopropyl group that is conjugated with putrescine to form the polyamine spermidine and then spermine (Chiang et al., 1996). This is of particular interest in the proteomic signature of $Y88^T$, since one of the distinguishing characteristics of sphingomonads from other Gram-negative bacteria is that they utilise homospermidine as the major polyamine (Busse et al., 2005). Although the function of homospermidine is unclear (Tholl et al., 1996), is has been implicated to have a functional role in nitrogen fixation and a role in nitrogen metabolism (Hamana et al., 1983). The coordinated upregulation of AdoMet synthetase and AdoHcy hydrolase in $Y88^T$ cells for the Fix^{off} state is more likely to be a response either to nitrogen scavenging or the stress related to scavenging nitrogen since methylation of heat shock proteins by AdoMet is known to occur under stress (Chiang et al., 1996). Interestingly, AdoMet produces the autoinducer Al-2 that is implicated in a unique "universal" signal suggested to be used in bacterial communication (Schauder et al., 2001).

Translation elongation factor and malate dehydrogenase were identified in the same protein spot (spot 2972), which was upregulated at OD_{600} 0.5 (Fig. 4.4A) relative to OD_{600} of 1.0 (2.31-fold, *t*-test 0071, Fig. 4.4B, Table 4.1) and 2.0 (3.62-fold, 0064, Fig. 4.5B, Table 4.1). These two proteins co-migrated so it is unclear which of these proteins, if not both, are responsible for this increase.

Three spots, two of which have yet to be identified, also show altered protein abundance. Spot 3208, formerly identified as hypothetical protein Saro_2615 which on closer inspection appeared to be two closely migrating spots rather than one, was later also identified as enoyl-CoA hydratase (see Appendix A, Fig. A.1, spot D11) an enzyme involved in fatty acid β-oxidation. It was highly upregulated 30.7-fold (*t*-test=0.014, Appendix A, Fig. A.1) irrespective of growth conditions at high OD_{600}. Spot 2431 was upregulated 4.16-fold (*t*-test=0.02) at an OD_{600} of 0.5 (Fig. 4.4A) relative to that at an OD_{600} of 1.0 (Fig. 4.4B) and 6.8-fold (*t*-test=0.024) relative to that at the highest OD_{600} of 2.0 (Fig. 4.5B). Spot 1809, yet to be identified, was upregulated 12-fold (*t*-test=0.013) at a high OD_{600} of 2.0 (Fig. 4.5B) for the Fix^{on} state relative to the low OD_{600} of 0.5 (Fig. 4.5A).

4.3 Summary

An examination of the $Y88^T$ proteome identified proteomic signatures for the Fix^{off} and Fix^{on} states of $Y88^T$ growth. The proteomic signature for the $Y88^T$ Fix^{on} state included a number of nitrogen-fixing proteins which showed a clear upregulation in their abundance with a switch from assimilating extracellular ammonium to fixing nitrogen. Thus the presence of ammonium in the growth medium appeared to inhibit the expression of the genes encoding the $Y88^T$ nitrogen-fixing proteins. The OmpA/MotB proteins included in this signature were similarly affected. The proteomic signature obtained for the $Y88^T$ Fix^{on} state contrasted with that for $Y88^T$ PHA synthesis physiology, which showed no obvious effect of the presence or absence of ammonium. Therefore, PHA production appears to be independent of nitrogen fixation in $Y88^T$.

The coordinated upregulation of several proteins observed in the proteomic signature for Y88^T nitrogen-fixing physiology that would not usually be associated with nitrogen fixation, suggests a functional role for these proteins initiated by ammonium-depleted conditions. Based on their general functions, these proteins could play a role in optimising metabolic activities as part of a coordinated response to changing nitrogen fixation and PHA activities, ensuring that the activity of specific regulatory proteins required for these functions is maintained to allow nitrogen fixation to proceed while maintaining high levels of PHA production. The nitrogenase enzyme is a case in point. The high turnover rate of nitrogenase during nitrogen fixation may result in the upregulation of proteins involved in protein degradation. The degraded protein products can be used to replenish the amino acids required for nitrogenase synthesis. Without the coordinated regulatory response of proteins that target other proteins for degradation, the increased demand for nitrogenase during nitrogen fixation could not be maintained and nitrogen fixation would not occur.

Chapter 5: The paradox of Y88^T growth at high oxygen concentration

5.1 Introduction

The MoFe nitrogenase enzyme is rapidly and irreversibly inactivated by oxygen (Staal et al., 2007). As an obligate aerobe, Y88^T requires oxygen to support aerobic respiration and ATP synthesis to meet the high energy demands of nitrogen fixation and other cellular activities, yet it must prevent irreversible damage to its nitrogenase when growing under nitrogen-limiting conditions. To assess the possible effects of DO on the Y88^T nitrogenase enzyme, Y88^T cells were grown in parallel in bioreactors in which the DO concentrations were maintained at low (10% DO) or high (70% DO) levels. Proteomic analysis showed that the Y88^T nitrogenase enzyme was present at barely detectable levels under DO^{high} conditions. Despite this fact and regardless of the vast differences in DO concentrations, the ability of Y88^T to increase its biomass and produce high amounts of PHB was not affected. The yields of PHB produced under both oxygen concentrations were also similar to that produced by Y88^T grown in batch flasks (Chapter 3). These results indicate the versatility of Y88^T as a PHB-producing nitrogen fixer that appears to have the ability to protect its nitrogenase (possibly through a conformational change) under adverse growth conditions.

5.2 Results

5.2.1 Cultivation of Y88^T in a bioreactor at DO^{low} or DO^{high} concentrations

5.2.1.1 Growth and nitrogen utilisation

The microbiology of Y88^T was investigated during batch cultivation for 24 hr in parallel bioreactors with DO concentrations held constant at 10% and 70% DO in a glucose-enriched (5 g/L) minimal nitrogen medium (0.1 g/L NH_4Cl) (see Chapter 2, section 2.1.4.2, Table 2.1 for experimental design and parameter details). Y88^T attained similar high OD_{600} of approximately 2.0 at both DO^{low}

and DO^{high} concentrations after approximately 20 hr (Fig. 5.1A). The amount of biomass produced was also similar (Fig. 5.1B) at both DO concentrations. Acetylene reduction assays carried out at low (0.5) or high (2.0) OD_{600} at both DO concentrations to determine whether the nitrogenase enzyme was active under these conditions, showed that $Y88^T$ cells grown in DO^{low} produced substantial amounts of ethylene, indicative of nitrogen fixation (Fig. 5.1C). In contrast, $Y88^T$ cells grown at DO^{high} concentrations showed a much smaller ethylene peak, indicative of decreased nitrogenase activity (Fig. 5.1D). Despite the presumable decrease in nitrogen fixation suggested by the smaller ethylene peak in the acetylene reduction assays for the cells grown at DO^{high}, the cells grew similarly to those at DO^{low} with no difference in the amount of biomass produced by $Y88^T$ between DO^{low} and DO^{high} conditions. At the time of the assay, only traces of nitrogen were measured in the spent medium for both DO^{low} and DO^{high} conditions using the Kjeldahl method (Fig. 5.2). Independent verification of the removal of nitrogen from the growth medium by $Y88^T$ was also demonstrated by an ammonium assay. Also, the C:N ratios of total biomass of $Y88^T$ (Fig. 5.2, maroon triangles, blue squares on graph; Table 5.1) were similar for the cells grown at both concentrations of DO, suggesting that nitrogen (and carbon) changes in the cells were occurring similarly irrespective of DO concentrations. The nitrogen content from the total biomass produced during growth at both DO concentrations increased from 0.01 mg/mL to 0.03 mg/mL after approximately 8 hr growth (Fig. 5.3, Table 5.1), after which it was maintained at 0.03 mg/mL until the last sampling at 24 hr. Therefore, despite the ammonium being depleted from the medium during growth at both DO concentrations, the nitrogen content of the biomass was maintained at the same level for DO^{low} and DO^{high} conditions.

Despite a much smaller ethylene peak for the DO^{high} conditions (Fig. 5.1D) and the fact that the ammonium had been depleted from the growth medium (Fig. 5.2, Table 5.1), $Y88^T$ cells at DO^{high} showed no difference in total nitrogen levels or biomass to that at DO^{low} (Fig. 5.3, Table 5.1). Since a two-fold increase in biomass occurred between 16 and 24 hr for $Y88^T$ cells grown under both DO

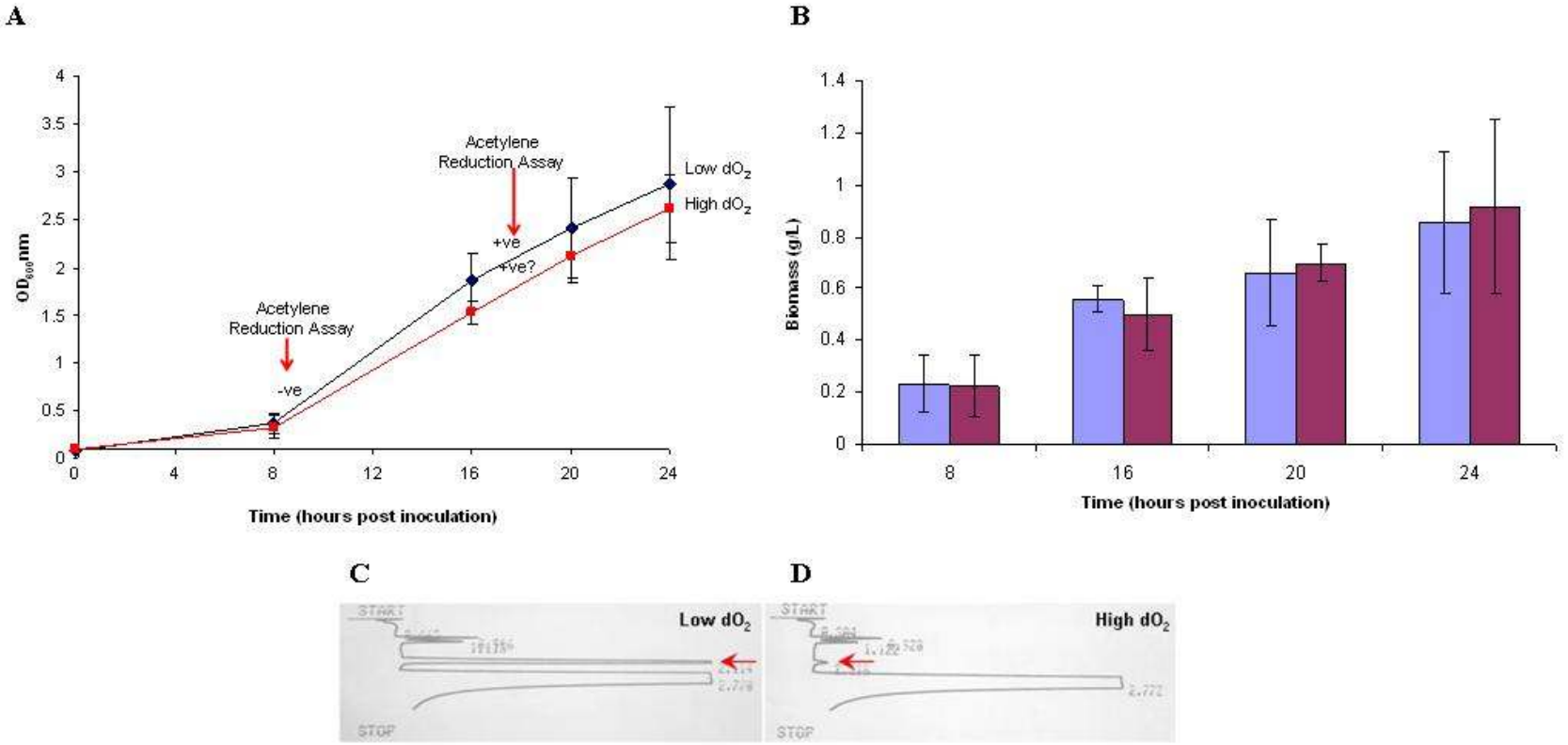


Figure 5.1 OD_{600}, biomass production and nitrogenase assays of $Y88^T$ grown at DO^{low} or DO^{high} for 24 hrs in a glucose-enriched, minimal nitrogen medium (A), OD_{600} of the $Y88^T$ culture media vs. time with cells grown under low (*black diamonds*) or high (*red squares*) DO. Acetylene reduction assays carried out at OD_{600} equal to 0.5 and 2.0 are indicated by the red arrows. (B) $Y88^T$ biomass yield vs. time with cells grown under DO^{low} (*blue bars*) or DO^{high} (*maroon bars*). Ethylene peaks obtained for cells of equal density at an OD_{600} of 2.0 at DO^{low} (C) and DO^{high} (D) are indicated by red arrows. Error bars (A, B) represent the standard deviations of the measurements, (n=3).

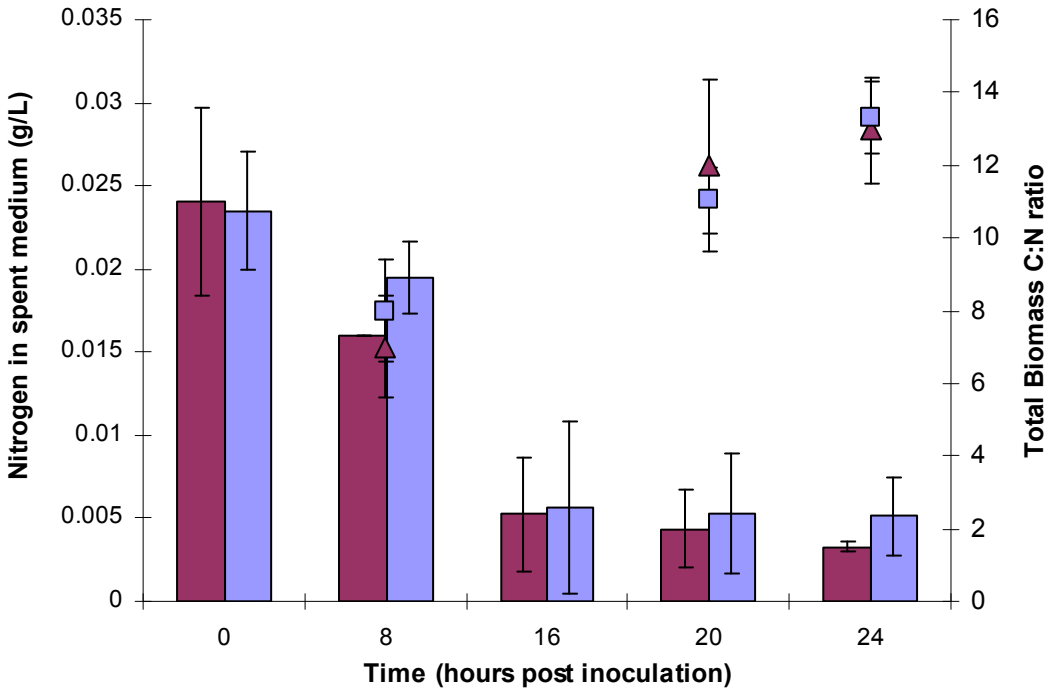


Figure 5.2 Total nitrogen (Kjeldahl) determined in the spent medium after removal of Y88^{T} cells grown at DO^{low} or DO^{high}

Total nitrogen content in spent medium from growth at DO^{low} (*maroon bars*) and DO^{high} (*blue bars*) during 24 hr growth. Biomass C:N ratios determined at 8, 20 and 24 hr are represented by maroon triangles (DO^{low}) and blue squares (DO^{high}). Error bars represent standard deviations, (n=2).

Time (hrs)	Total biomass C (g/L)		Total biomass N (g/L)		Biomass C:N		NH_4^+ in spent medium		Kjeldahl N (total) in spent medium (g/L)		PHB (g/L)	
	Do^{low}	Do^{high}	Do^{low}	Do^{high}	Do^{low}	Do^{high}	Do^{low}	Do^{high}	Do^{low}	Do^{high}	Do^{low}	Do^{high}
0	0.0302±0.006	-	0.002±0.001	-	-	-	0.0195±0.012	0.0200±0.008	0.024±0.006	0.021±0.000	-	-
8	0.065±0.022	0.051±0.017	0.01±0.005	0.01±0.002	7±1.414	8±1.414	0.013±0.007	0.0155±0.004	0.016±0.000	0.001±0.016	0.023±0.010	0.0241±0.013
12	0.382±0.168	0.361±0.118	0.04±0.014	0.04±0.011	-	-	-	-	0.005±0.003	0.006±0.004	0.334±0.024	0.231±0.071
20	0.393±0.157	0.375±0.107	0.034±0.012	0.03±0.008	11.985±2.382	12.31±0.898	0.003±0.001	0.003±0.000	0.004±0.002	0.007±0.003	0.418±0.035	0.312±0.140
24	0.452±0.059	0.525±0.017	0.04±0.001	0.04±0.002	12.956±1.476	15.14±0.989	0.002±0.001	0.002±0.001	0.003±0.000	0.003±0.000	0.555±0.052	0.562±0.106

Table 5.1 Carbon and nitrogen content of $Y88^T$ cells grown in glucose-enriched, nitrogen-limiting medium

The total carbon and nitrogen (determined by the modified Dumas method) at each sampling interval, the total carbon to nitrogen ratio of dry biomass and the total amount of nitrogen remaining in the spent medium (determined using an ammonium assay and verified by total Kjeldahl nitrogen values) at each sampling were confirmed for total dry biomass obtained from $Y88^T$ growth in a glucose-enriched, nitrogen-limiting medium. PHB accumulated during growth is also shown, (n=2, errors are standard deviations).

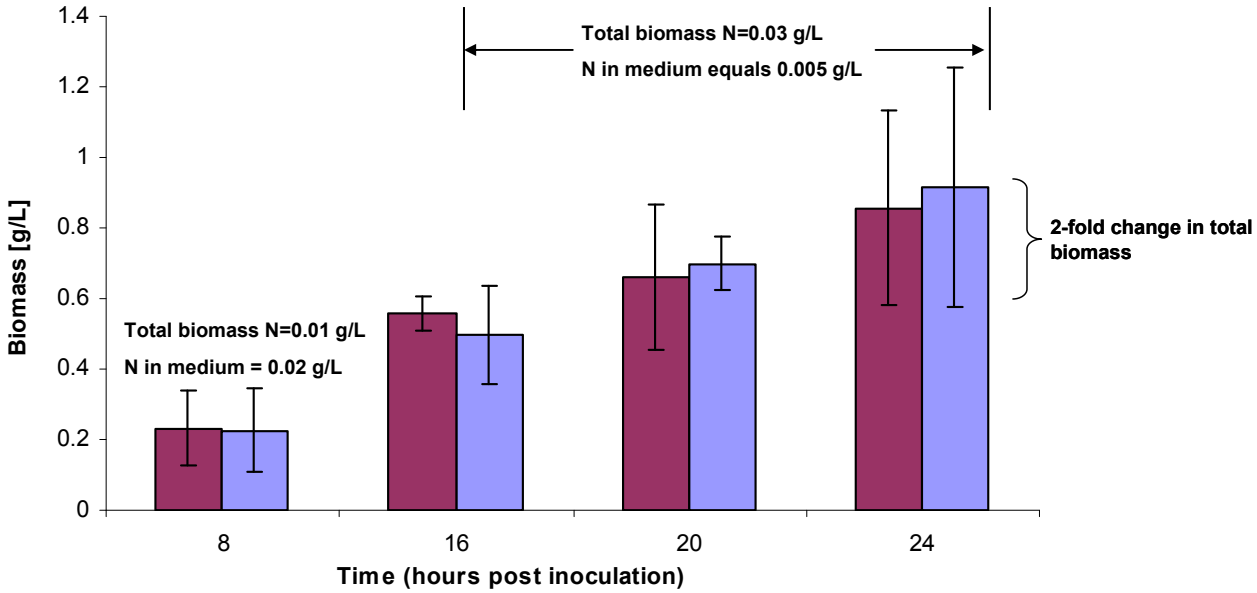


Figure 5.3 Change in total biomass nitrogen content with change in biomass during growth in nitrogen-depleted medium DO^{low} (maroon bars) and DO^{high} (blue bars) during 24 hr growth. Error bars represent standard deviations, (n=2).

conditions, during which time the nitrogen levels in the medium appeared to be too low to sustain growth without fixation, the cells at DO^{high} appear to have fixed nitrogen similar to that at DO^{low} conditions.

The paradox of an increase in biomass and intracellular nitrogen content at DO^{high} under nitrogen-depleted growth conditions, despite very little measurable nitrogenase activity to promote nitrogen fixation, suggests either that (i) the nitrogenase of $Y88^{T}$ is sensitive to DO^{high} and that an alternative mechanism is used by $Y88^{T}$ at DO^{high} to compensate for the loss of nitrogenase activity; (ii) an alternative mechanism is used by $Y88^{T}$ at DO^{high} to prevent inactivation of its nitrogenase and allow nitrogenase activity to proceed albeit at a reduced rate; (iii) biomass increase is due to the synthesis and accumulation of an alternative byproduct such as an exopolysaccharide that requires more carbon but does not require much nitrogen, thereby still altering the C:N ratio. The latter explanation, however, does not fully account for the data as it does not explain the similar nitrogen content of $Y88^{T}$ cells at both DO^{high} and DO^{low}. A ^{15}N isotope labeling experiment, during which $Y88^{T}$ cells could be grown under both DO conditions while sparging the headspace of the reactor with ^{15}N gas and then determining the total cellular ^{15}N content, would be one way of irrefutably confirming that nitrogen is fixed under both DO^{high} and DO^{low} conditions. This would confirm the similar nitrogen content of $Y88^{T}$ cells determined in these assays at both DO concentrations.

5.2.1.2 PHB production and glucose utilisation

The amount of PHB produced by $Y88^{T}$ in DO^{low} after 24 hr was 0.6±0.05 mg/ml, (n=2). The same amount of PHB (0.6±0.1 mg/ml, n=2) was produced by $Y88^{T}$ after 24 hr in DO^{high}. Therefore, DO^{high} had no effect on the total amount of PHB produced by $Y88^{T}$ after 24 hr when compared to that produced at DO^{low} for the same period of growth (Fig. 5.4). Glucose utilisation was similar in the early stages of growth at DO^{low} or DO^{high}, although at 20 hr, this was not the case. $Y88^{T}$ consumed more glucose at 20 hr in DO^{high} than in DO^{low} (Figs. 5.5, 5.6). In fact, the amount of glucose consumed by 24 hr at DO^{high} was approximately 1 g/L whereas 1.6 g/L was consumed for the same period of growth at DO^{high}

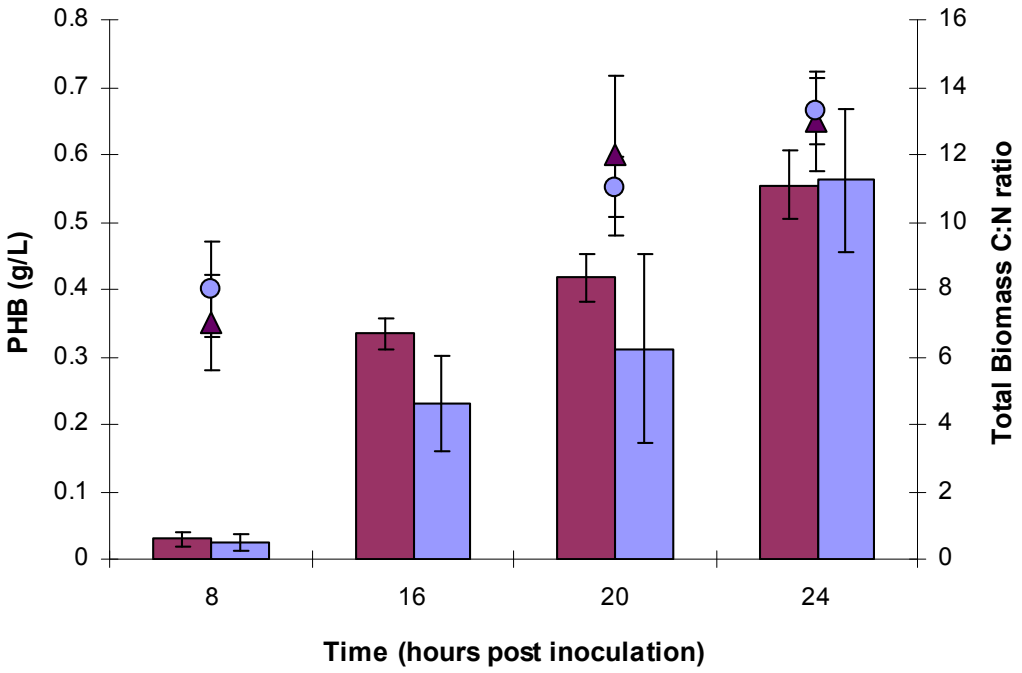


Figure 5.4 PHB produced by Y88^{T} at DO^{low} and DO^{high} during changing C:N ratios

PHB produced during growth under DO^{low} (*maroon bars*) or DO^{high} (*blue bars*) conditions over 24 hr in a glucose-enriched, minimal nitrogen medium. Biomass C:N ratios determined at time-points throughout the growth period from cells grown at DO^{low} (represented by maroon triangles) and at DO^{high} (represented by blue circles), (n=2).

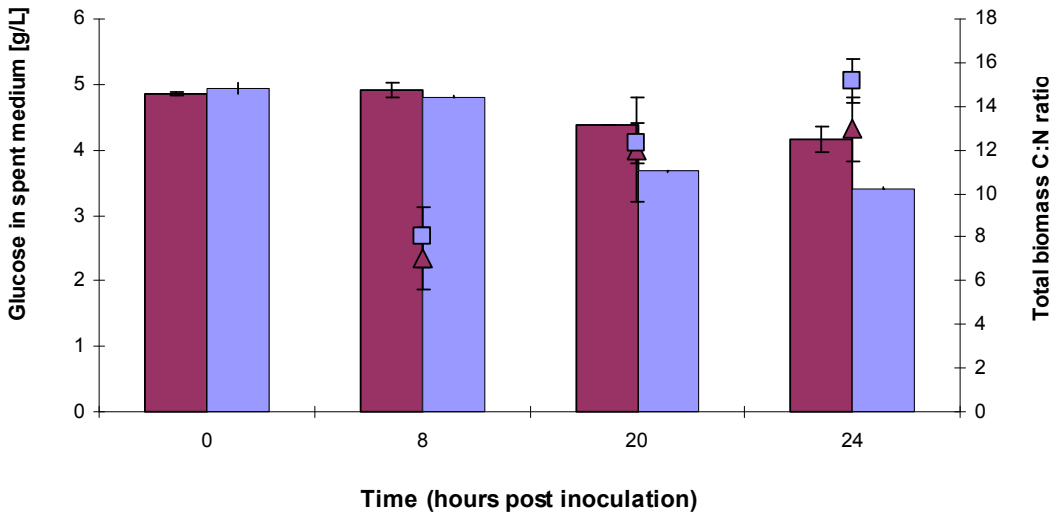


Figure 5.5 Glucose utilisation by Y88^{T} and change in C:N ratio

Glucose remaining in the spent medium following removal of Y88^{T} cells grown at DO^{low} (*maroon bars*) or DO^{high} *(blue bars*) after 24 hr of growth. Biomass C:N ratios determined at time-points throughout the growth period are represented by maroon triangles (DO^{low}) and blue squares (DO^{high}), (n=2).

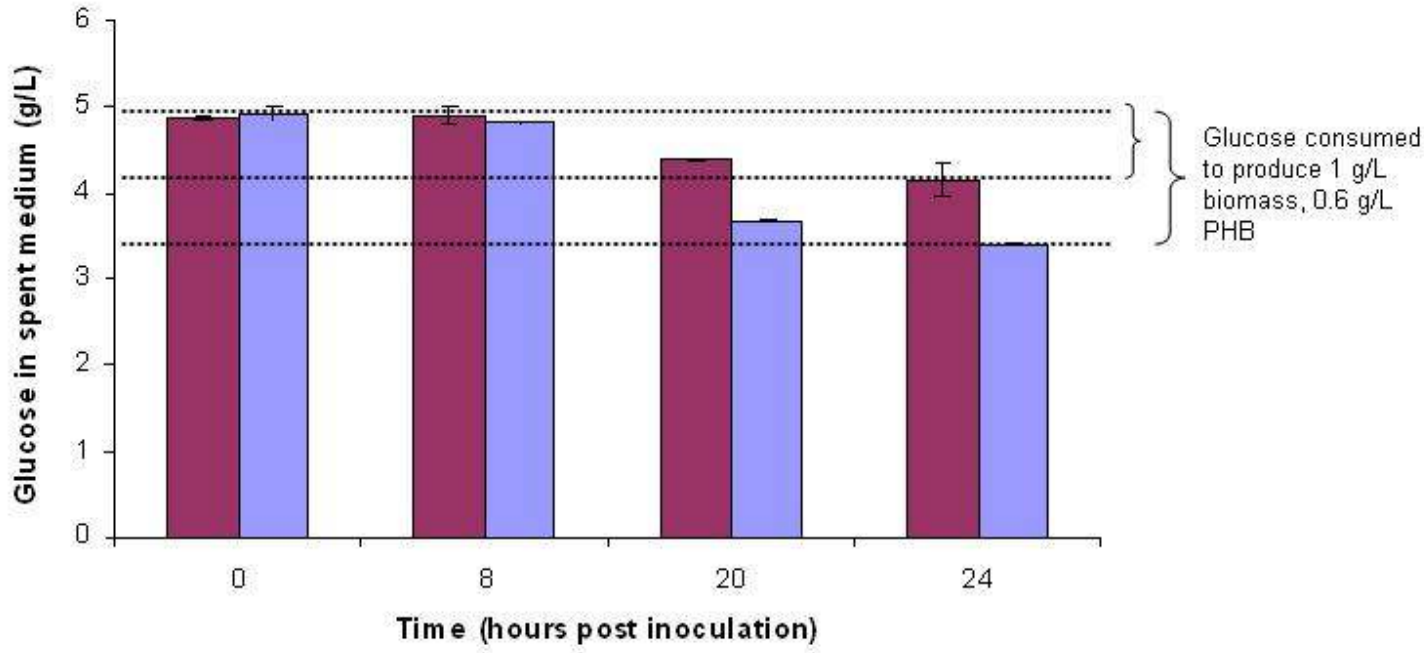


Figure 5.6 Glucose consumption by Y88^{T} showing inefficiency of glucose utilisation at DO^{high}

Glucose remaining in the spent medium following removal of Y88^{T} cells grown at DO^{low} (*maroon bars*) or DO^{high} *(blue bars*) after 24 hr of growth, (n=2). The amount of glucose removed after 24 hr at DO^{high} was twice as much as that consumed for the same period at DO^{low}. Despite double the amount of glucose consumed at DO^{high}, the same biomass (1 g/L) in PHB (0.6 g/L) was produced.

(Fig. 5.6). Since biomass and PHB production were similar (1 g/L biomass, 0.6 g/L PHB) at both DO^{high} and DO^{low}, it appears that at DO^{high}, glucose utilisation was less efficient than at DO^{low}. Since the increased glucose consumption at DO^{high} was not reflected in an increase in the amount of PHB produced, relative to that at DO^{low} at that stage of growth (Fig. 5.4), these results suggest that more glucose was likely being used to meet energy demands such as, for instance, higher respiration or the conversion to an additional byproduct other than PHB. Oxygen uptake rate at this stage of growth increased rapidly at DO^{high} (consistent across triplicate experiments) relative to DO^{low} supporting this notion (Table 5.2). The similar amount of PHB produced irrespective of DO concentrations and efficiency of glucose utilisation, suggest that PHB synthesis in $Y88^T$ is not limited by a changing supply of carbon or nitrogen during 24 hr in a bioreactor, nor by DO^{low} or DO^{high} concentrations at high growth.

5.2.1.3 Comparison of protein profiles obtained for DO^{low} or DO^{high} conditions at high cell densities (OD_{600}=2.0)

5.2.1.3.1 Acidic proteins

2D DIGE was used to visualise the $Y88^T$ proteins and differences in their abundance resulting from growth at DO^{low} or DO^{high}. Examination of the proteins in the pI 4-7 range for $Y88^T$ cells grown at DO^{low} or DO^{high} showed approximately 200 protein spots for both states of the cells with approximately 80 of these occurring in high abundance (Fig. 5.7). Ninety-six protein spots were selected from these gels for analysis, most of which were not differentially expressed. Thirty-two protein spots showed statistically significant difference in abundance between the different DO states in the pI 4-7 range. Since $Y88^T$ continued to increase in biomass (refer back to Fig. 5.1B) despite producing a very small ethylene peak at DO^{high} (Fig. 5.1D), emphasis in this section is placed on the examination of the statistically significant differentially expressed proteins between DO^{high} and DO^{low} growth conditions in an attempt to explain this apparent paradox.

DO condition	Sparging oxygen level	
	Equation for the line fitted to oxygen demand data	**Total (accumulative) Input oxygen over 24 hrs (L)**
Maintain DO^{low} (low OD_{600})	y=21.59x-1.71	1857.63±683.98
Maintain DO^{high} (low OD_{600})	y=30.53x-5.21	2928.36±924.31
Maintain DO^{low} (high OD_{600})	y=44.2x-10.03	18278.96±1734.41
Maintain DO^{high} (high OD_{600})	y=108.22x-40.05	28303.39±1102.66

Table 5.2 Oxygen input to meet the demand of $Y88^T$ cells and maintain set DO levels constant during $Y88^T$ growth in a glucose-enriched, nitrogen-fixing medium in bioreactors

The amount of oxygen that was required to maintain DO levels constant and to meet the respiratory requirements of $Y88^T$ cells during growth in the bioreactors at DO^{low} and DO^{high} is shown. During the first 9 hours of growth, there was little difference in the maximum amount of oxygen required to maintain set levels at the two DO levels. Between 9 to 24 hrs, a much higher amount of input oxygen was required to meet the respiratory requirements of $Y88^T$ cells grown at DO^{high} compared to $Y88^T$ cells grown at DO^{low}. This is confirmed by the slope of the line for a trendline fitted through the input oxygen data under the different DO conditions.

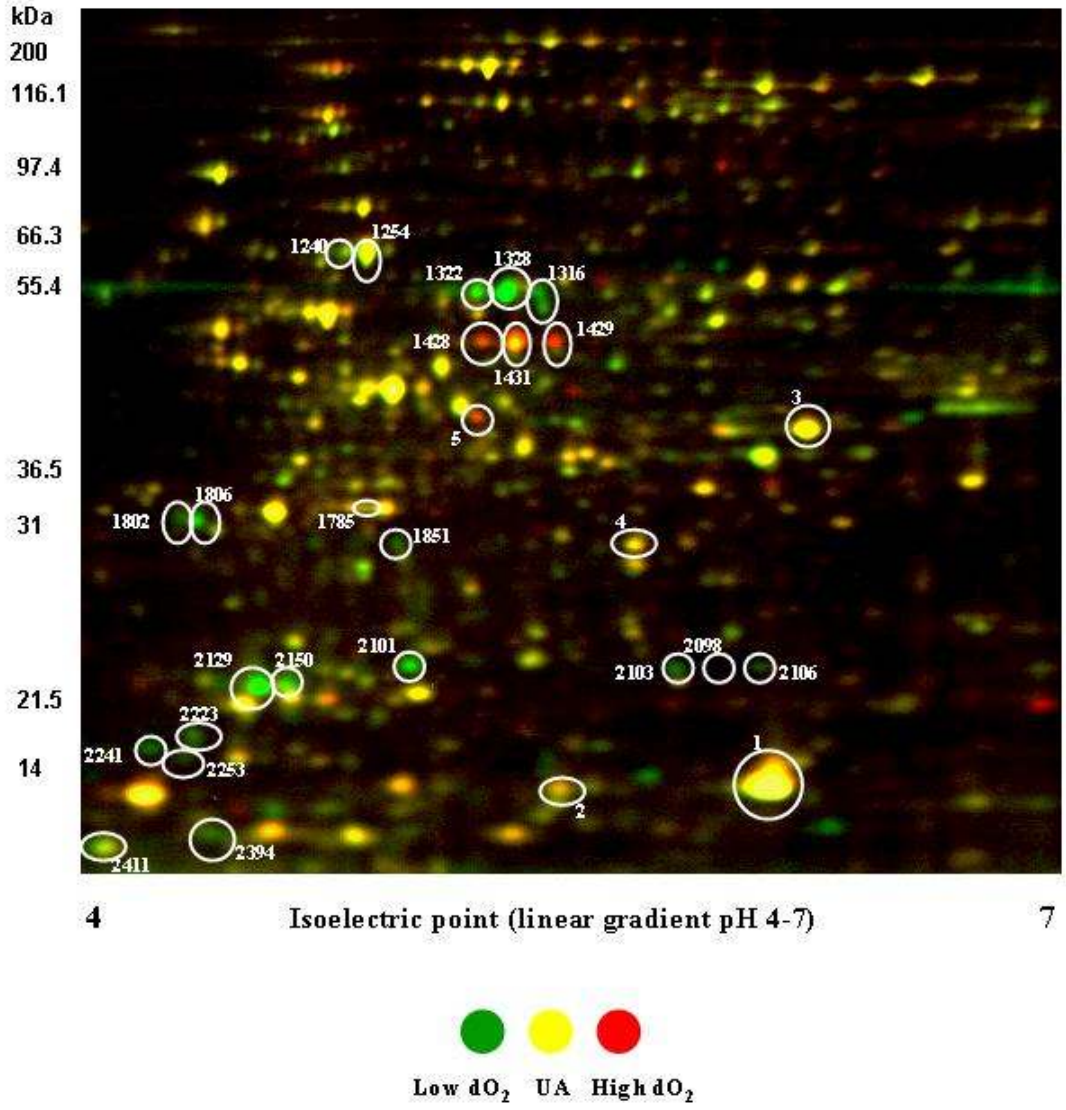


Figure 5.7 Overlay image of Y88^{T} protein profiles generated from minimally labeled samples showing differential protein abundance under DO^{low} or DO^{high}, uncontrolled pH conditions in the linear gradient pH 4-7 range

Pooled (Cy2 - yellow), DO^{low} (Cy3 - green) or DO^{high} (Cy5 - red) samples are overlaid. UA (yellow) =unaltered abundance: proteins expressed similarly in both states of the cells at DO^{low} or DO^{high}; DO^{High} (red): proteins either unique to or more abundant in the cells at DO^{High}; DO^{low} (green): proteins either unique to or more abundant in the cells at DO^{low}. Spot numbers represent proteins determined by BVA analysis to have significantly altered abundance between the two DO concentrations, with the exception of spots 1 through 5 which did not show statistically different altered abundance (n=3). Spots selected for analysis were those likely to be involved in the nitrogen-fixing and PHB pathways based on predicted pI and molecular weight, as well as some spots with altered abundance that are not usually associated with these pathways. Spot numbers correspond to protein identifications in Table 5.3.

5.2.1.3.1.1 Nitrogen-fixing proteins

The β-subunit of nitrogenase component I, NifK, was identified in a spot at approximately 56kDa (spot 1328) that had 6.8-fold enhanced abundance (*t*-test=0.01) at DO^{low} relative to DO^{high} (Figs. 5.7, 5.8A, Table 5.3). Its position on the gel was consistent with the M_r and pI 5.4 predicted from the $Y88^T$ translated ORFs. NifK was barely detectable at the corresponding M_r and pI on the gel at DO^{high} (see corresponding spot 1328, Fig. 5.8B).

Two forms of the homodimeric dinitrogenase reductase (NifH) subunits were seen on the gels at approximately 23 and 24 kDa and were more abundant at DO^{low}. One subunit had a pI of 4.7 (Figs. 5.7, 5.8A, spot 2129) and showed a 3-fold increase *(t*-test=0.026) in abundance at DO^{low}. The second subunit had a pI of 5.2 (Figs. 5.7, 5.8A, spot 2101 and showed a 4.5-fold increase (*t*-test=0.01) in abundance at DO^{low}. These two forms were present at different abundance under DO^{low} conditions. Their observed M_r differed from that predicted (31.8 kDa) as did the pI 4.9 predicted from the $Y88^T$ NifH conceptual translation, suggesting possible post-translational modification of this protein as mentioned previously (Chapter 4, section 4.2.2.2.1). The difference in abundance of each subunit is consistent with the notion that each may be differently modified post-translationally for the nitrogen-fixing state, or that at least one is modified even if the other is not. It is known that when conditions are unfavourable for nitrogen fixation, such as in the presence of ammonium, one subunit of the homodimeric NifH protein is inactivated by the addition of an ADP-ribose group from NAD^+ to an arginine residue on the subunit, thereby preventing nitrogen fixation (Nordlund & Ludden, 2004; Huergo et al., 2006). Once the ammonium is exhausted from the medium, the ADP-ribose group is removed from the NifH subunit. Although ADP-ribosylation of dinitrogenase reductase plays a significant role in the regulation of nitrogenase activity, it is not the only mechanism of such regulation (Zhang et al., 1996). The post-translational modification of NifH in *Anabaena* sp. (Durner et al., 1994; Durner & Böger, 1995) was found to not be ADP-ribosylated as in *Rhodospirillum rubrum* (Ludden et al., 1989; Martin & Reinhold-Hurek, 2002). The fact that the two homodimeric

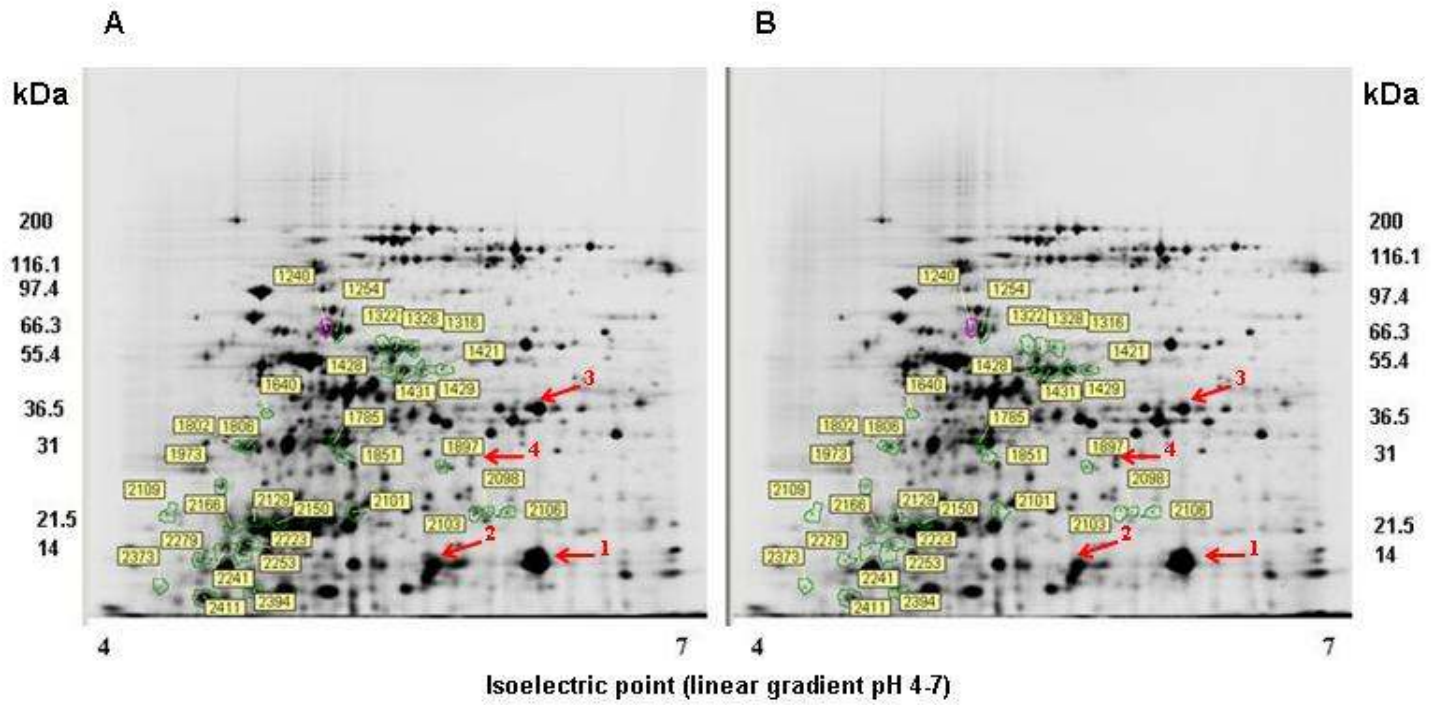


Figure 5.8 2-D DIGE images (pI 4-7) showing protein profiles of Y88^{T} cells grown to high OD_{600} (2.0) under ammonium-depleted conditions at DO^{low} or DO^{high} concentrations, pH not controlled.

(A) DO^{low} gel, (B) DO^{high} gel. Spot numbers (excluding those marked by red arrows) represent proteins determined by BVA analysis to have significantly altered abundance between the two DO concentrations. Spot numbers correspond to protein identifications in Table 5.3. Spots marked by red arrows represent PHB synthesis proteins: 1=PhaP, 2=PhaP, 3=PhbB, 4=PhbA, (n=3).

Spot number	Protein ID	Do^{low}/Do^{high}		Highest scoring unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
1240	Chaperonin GroEL	3.42	0.030**	K.ALAGLTGANEDQTR.G(87)	57484	571	29	5.09	12
				K.EGVITVEEAK.G(61)	–	–	–	–	–
				K.SVAAGINPMDLK.R(57)	–	–	–	–	–
1254*	Chaperonin GroEL	3.36	0.035**	K.ALAGLTGANEDQTR.G(85)	57484	3666	170	5.09	20
				K.SVAAGINPMDLK.R(63)	–	–	–	–	–
				K.VIEDLKGR.S(38)	–	–	–	–	–
				K.LAGGVAVIK.V(64)	–	–	–	–	–
1316	GMP synthase	3.26	0.003*	R.IVAVSDGAPFAVIADDTR.K (101)	57278	243	15	5.43	12
				R.LGEAEQVVGLFR.N (71)	–	–	–	–	–
				R.ELGLPDIFVGR.H (48)	–	–	–	–	–
				K.LIANFVR.H (43)	–	–	–	–	–
	Glycine dehydrogenase (decarboxylating)/ ribosomal protein L9	3.26	0.003*	R.IDLAALLER.L(79)	55278	186	4	5.34	6
				K.AALDQFIAAMR.D(68)	–	–	–	–	–
	prolyl-tRNA synthetase			R.IDLAALLER.L(79)	58055	125	4	5.39	4
				K.AALDQFIAAMR.D(68)	–	–	–	–	–
1322	D-3-phosphoglycerate dehydrogenase (PGDH)	7.52	0.023**	R.VTVGTDEGEK.A (45)	56092	93	6	5.44	3
1328	Nitrogenase molybdenum beta-chain (NifK)	6.75	0.011*	R.MYDGGTTLEEAGQAVHAR.A(109)	58437	17167	983	5.37	49
				K.VEYTILGDNSDVWDTPTDGEFR.M(84)	–	–	–	–	–

Table 5.3 2-D DIGE protein spot analysis (pH 4-7) and LC-MS/MS analysis of protein from cells grown at DO^{low} or DO^{high}

2-D DIGE (pH 4-7) protein spot analysis showing statistical values and LC-MS/MS identification of proteins with altered abundance at DO^{low} or DO^{high} concentrations or high OD_{600} (2.0). * denotes $p<0.01$ excepting for those spots denoted by ** where $p<0.05$. Positive values represent increased abundance of proteins at DO^{low} and negative values represent increased abundance of proteins at DO^{high}. NI=protein not identified. See body of the text throughout section 5.2.1.3 for discussion of these proteins. Table 5.3 continued on next page.

Spot number	Protein ID	Do^{low}/Do^{high}		Highest scoring unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
1328	Nitrogenase molybdenum beta-chain (NifK)	6.75	0.011*	K.MQDLFDSSPFGAGCK.A(95)	58437	17167	983	5.37	49
				R.ATISMQEFCTEK.T(72)	–	–	–	–	–
1421	NI	-3.02	0.034**	–	–	–	–	–	–
1428	S-adenosylhomocysteine (AdoHcy) hydrolase	-2.97	0.007*	K.VACVAGFGDVGK.G(80)	50877	788	38	5.33	20
				R.DEFGPTQPLK.G(47)	–	–	–	–	–
1429	NI	-2.86	0.008*	–	–	–	–	–	–
1431	S-adenosylhomocysteine (AdoHcy) hydrolase	-2.65	0.010*	R.TANLILDDGGDATMFALWGAR.V(138)	50877	14979	631	5.33	66
				R.VVVTEIDPICALQAAMEGYEVVTMEEAVKR.A(92)	–	–	–	–	–
				R.VEAGETLPEPANAEEIEFQR.A(97)	–	–	–	–	–
				R.AEIAIAETEMPGLMALR.D(87)	–	–	–	–	–
	histidyl-tRNA synthetase			R.IDDFLTAEAQDFFGR.V(100)	–	–	–	–	–
				K.VLAGLSLADGAPR.W(86)	–	–	–	–	–
				R.VTSGLDAAGVAWTR.S(97)	–	–	–	–	–
				R.SLGETTDVVSK.E(85)	–	–	–	–	–
1640	NI	2.01	0.043**	–	–	–	–	–	–
1785	fructose-biphosphate aldolase	2.14	0.010*	K.NNGIIASFSR.A(64)	32291	132	17	5.06	5
				R.ALLEDLR.H(32)	–	–	–	–	–
	translation elongation factor TS			K.TPIAQVVEAAGK.A(54)	32125	121	26	5.09	3
				R.IVAVSDGAPFAVIADDTR.K(89)	–	–	–	–	–

Table 5.3 (cont.) 2-D DIGE (pH 4-7) protein spot analysis showing statistical values and LC-MS/MS identification of proteins with altered abundance at DO^{low} or DO^{high} concentrations or high OD_{600} (2.0). * denotes $p<0.01$ excepting for those spots denoted by * where $p<0.05$. Positive values represent increased abundance of proteins at DO^{low} and negative values represent increased abundance of proteins at DO^{high}. NI=protein not identified. See body of the text throughout section 5.2.1.3 for discussion of these proteins. Table 5.3 continued on next page.

Spot number	Protein ID	Do^{low}/Do^{high}		Highest scoring unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
1802	NifU	3.08	0.018**	R.TNGLTSIDEVTNYTK.A(107)	33542	864	123	4.72	59
				K.AGGGCSTCAEGIEGVLER.V(93)	–	–	–	–	–
				R.VNAEMVAEGALAPER.A(91)	–	–	–	–	–
				R.DGGDCELVDVEGNR.V(85)	–	–	–	–	–
				R.APAKPLTNLQK.I(65)	–	–	–	–	–
				K.VDEPSQTISDAR.F(72)	–	–	–	–	–
				K.CFGIDEGMIER.T(59)	–	–	–	–	–
1806	Electron transfer flavoprotein beta-subunit	3.94	0.009*	R.GVELGAVVMGDDR.A(99)	39443	3555	256	5.95	37
				R.DLAGAVATTLR.T(96)	–	–	–	–	–
				R.IVEYAPAIVEADIVTK.V(83)	–	–	–	–	–
				K.LAEVLGAEYGGSRPLVQK.G(81)	–	–	–	–	–
1851	Chaperonin GroEL	11.52	0.009	K.EGVITVEEAK.G(45)	57484	202	10	5.09	5
				K.SVAAGINPMDLK.R(50)					
1897	NI	2.04	0.045**	–	–	–	–	–	–
2101	dinitrogenase reductase	4.47	0.013*	R.CVESGGPEPGVGCAGR.G(118)	31859	380	59	4.94	38
				K.STTSQNTLAALADLGQR.I(112)	–	–	–	–	–
				K.LGTQLIHFVPR.D(66)	–	–	–	–	–
				R.DNIVQHAELR.R(61)	–	–	–	–	–
				R.LGGLICNER.Q(60)	–	–	–	–	–
2103	NI	12.46	0.005*	–	–	–	–	–	–

Table 5.3 (cont.) 2-D DIGE (pH 4-7) protein spot analysis showing statistical values and LC-MS/MS identification of proteins with altered abundance at DO^{low} or DO^{high} concentrations or high OD_{600} (2.0). $p<0.01$ excepting for those spots denoted by * where $p<0.05$. Positive values represent increased abundance of proteins at DO^{low} and negative values represent increased abundance of proteins at DO^{high}. NI=protein not identified. See body of the text throughout section 5.2.1.3 for discussion of these proteins. Table 5.3 continued on next page.

Spot number	Protein ID	Do^{low}/Do^{high}		Highest scoring unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
2106	NI	6.46	0.011*	–	–	–	–	–	–
2109	NI	5.38	0.011*	–	–	–	–	–	–
2129	dinitrogenase reductase	2.8	0.026**	R.CVESGGPEPGVGCAGR.G(120)	31859	4864	151	4.94	32
				K.STTSQNTLAALADLGQR.I(118)	–	–	–	–	–
				R.DNIVQHAELR.R(66)	–	–	–	–	–
				R.LGGLICNER.Q(65)	–	–	–	–	–
2150	hypothetical protein BRAD05426	7.85	0.002*	R.ISDTAFGTLEAAR.R(100)	22861	4004	108	4.93	58
				K.ETLELVDVDKNDVK.K(92)	–	–	–	–	–
				K.VAAFTAAYPELSYEDGVAAMGPVR.N(68)	–	–	–	–	–
				M.SATTSPSLETVR.I(64)	–	–	–	–	–
				K.IPVLVGYLHNFAWLK.D(50)	–	–	–	–	–
2166*	alkyl hydroperoxide reductase	2.03	0.034**	K.FVDLTDADIAGK.W(97)	20864	10135	373	4.79	66
				K.WEEGEATLAPSIDLVGK.I(93)	–	–	–	–	–
				R.GTFVVDPDGVIQLVEITPEGVGR.N(88)	–	–	–	–	–
				R.EGVGLADR.G(67)	–	–	–	–	–
				K.AWHDSSAAIAK.I(64)	–	–	–	–	–
2223	NI	5.78	0.012*	–	–	–	–	–	–
2241	OmpA/MotB[d]	5.61	0.01*	K.YNLGLSAR.R(50)	11983	963	81	5.76	21
				K.AFGEANPR.V(56)	–	–	–	–	–
				R.VPTADGVR.E(46)	–	–	–	–	–

Table 5.3 (cont.) 2-D DIGE (pH 4-7) protein spot analysis showing statistical values and LC-MS/MS identification of proteins with altered abundance at DO^{low} or DO^{high} concentrations or high OD_{600} (2.0). * denotes $p<0.01$ excepting for those spots denoted by ** where $p<0.05$. Positive values represent increased abundance of proteins at DO^{low} and negative values represent increased abundance of proteins at DO^{high}. NI=protein not identified. See body of the text throughout section 5.2.1.3 for discussion of these proteins. Table 5.3 continued on next page

.

Spot number	Protein ID	Do^{low}/Do^{high}		Highest scoring unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
2241	OmpA/MotB[d]	5.61	0.01*	K.TLNVDNQTITAVGGR.T(89)	25899	674	12	7.03	11
				R.LEGQVGYLDAGNK.T(92)	–	–	–	–	–
2241	Ribosomal protein L7/L12[d]			R.AITSLGLTEAK.A(79)	12683	533	11	4.65	28
				K.LTVLEAADLAK.A	–	–	–	–	–
2253	OmpA/MotB[e]	10.32	0.013*	K.TLNVDNQTITAVGGR.T(89)	–	–	–	–	–
				R.LEGQVGYLDAGNK.T(92)	–	–	–	–	–
	Ribosomal protein L7/L12[e]			R.AITSLGLTEAK.A(79)	–	–	–	–	–
				K.LTVLEAADLAK.A	–	–	–	–	–
1	Phasin	–	–	K.ILSTGLQELGKGYATEGK.S(95)	33963	8163	295	9.56	40
				K.TPTELLEKQSALLR.K(90)	–	–	–	–	–
				K.SQASFGELGEFAK.G(87)	–	–	–	–	–
				K.GNVEALVESTK.I(85)	–	–	–	–	–
				K.ILSTGLQELGK.G(84)	–	–	–	–	–
				K.LANEAFQPISNR.V(78)	–	–	–	–	–
2279	NI	5.31	0.014*	–	–	–	–	–	–
2373	NI	3.64	0.001*	–	–	–	–	–	–
2394	NI	3.02	0.005*	–	–	–	–	–	–
2411	NI	3.04	0.027**	–	–	–	–	–	–
2	Phasin	–	–	K.SQASFGELGEFAK.G(85)	33963	3862	131	9.56	34
				K.LANEAFQPISNR.V(71)	–	–	–	–	–
				K.GNVEALVESTK.I(81)	–	–	–	–	–
				K.SAFETLTAEFK.D(63)	–	–	–	–	–

Table 5.3 (cont.) 2-D DIGE (pH 4-7) protein spot analysis showing statistical values and LC-MS/MS identification of proteins with altered abundance at DO^{low} or DO^{high} concentrations or high OD_{600} (2.0). * denotes $p<0.01$ excepting for those spots denoted by ** where $p<0.05$. Positive values represent increased abundance of proteins at DO^{low} and negative values represent increased abundance of proteins at DO^{high}. NI=protein not identified. See body of the text throughout section 5.2.1.3 for discussion of these proteins. Continued on next page.

Spot number	Protein ID	Do^{low}/Do^{high}		Highest scoring unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
3	Acetyl-CoA C-acetyltransferase (β-Ketothiolase)			R.SGVAAEKFDR.V(79)	41038	530	15	5.18	14
				R.AMGTFAQDTANAYQLTR.E(73)	–	–	–	–	–
4	3-Hydroxybutyryl-CoA dehydrogenase			K.IVAADAEAALSR.I(86)	30939	175	10	5.42	22
				K.VLAPQAILASNTSSIPITR.M(51)	–	–	–	–	–
5	S-Adenosylmethionine synthetase (AdoMet)			K.GIMDTDGNWAPGVPEEVER.V(86)	43392	210	5	5.27	10
				K.NIVAAGLAHR.C(49)	–	–	–	–	–

Table 5.3 (cont.) 2-D DIGE (pH 4-7) protein spot analysis showing statistical values and LC-MS/MS identification of proteins with altered abundance at DO^{low} or DO^{high} concentrations or high OD_{600} (2.0).

* denotes $p<0.01$ excepting for those spots denoted by ** where $p<0.05$. Positive values represent increased abundance of proteins at DO^{low} and negative values represent increased abundance of proteins at DO^{high}. NI=protein not identified. Spot numbers 3, 4 and 5 (Fig. 5.7) showed altered abundance but were determined by BVA analysis not to be significantly altered. See body of the text throughout section 5.2.1.3 for discussion of these proteins.

NifH subunits of Y88^{T} are observed at different pIs for the nitrogen-fixing state (when the ADP-ribose group would have already been removed if either of them had been ADP-ribosylated to begin with), suggests a possible different post-translational modification in this instance. Gallon *et al*., (2000) proposed that *Gloeothece* nitrogenase is subject to modification by palmitoylation rather than by ADP-ribosylation. Palmitoylation is an important reversible post-translational lipid modification of proteins that provides a mechanism for binding cytosolic proteins to membranes, segregating proteins to microdomains or mediating protein-protein interaction (Gallon et al., 2000; Basu, 2004). Since this mechanism is reversible, de-palmitoylation and re-palmitoylation could provide a mechanism whereby these processes are regulated (Basu, 2004). Notably, the NifH subunits were barely visible on the gel at DO^{high} suggesting a different regulatory mechanism to that at DO^{low} that results in a decrease in abundance at DO^{high}.

The NifU protein (spot 1806) showed a 3.1-fold increase (*t*-test=0.018) in abundance at DO^{low} relative to DO^{high} (Figs. 5.7, 5.8A, Table 5.3). At DO^{high}, NifU was coordinately downregulated with both the NifK and NifH components of nitrogenase. It is possible that this protein plays a regulatory role in negative feedback of Y88^{T} nitrogenase enzyme expression at DO^{high}, since inactivation of NifU has been shown to affect the activity of both nitrogenase components (Dos Santos et al., 2004). The proteomic signature obtained at DO^{low} with respect to the NifK, NifH and NifU proteins is consistent with that obtained for the Fix^{on} state of Y88^{T} cells at high OD_{600} (Chapter 4). The decreased abundance of these proteins at DO^{high} is consistent with the decreased nitrogenase activity shown in the acetylene reduction assays for the Fix^{off} state of Y88^{T} cells at lower OD_{600} (Chapter 4).

Electron transfer flavoprotein β-subunit (ETF-FixB subunit of the ETF-FixAB protein) (spot 1802) showed a 3.9-fold increase (*t*-test=0.001) in abundance at DO^{low} (Figs. 5.7, 5.8). Functionally, in bacteria (and eukaryotes), ETF is a soluble one- or two-electron carrier that couples various flavoprotein dehydrogenases

(one acting oxidatively, the other reductively) via transient formation of alternately ETF-oxidized and ETF-reduced binary complexes (Scott & Ludwig, 2004). ETF-FixAB is specifically required for dinitrogenase complex activity (Tsai & Saier, 1995) and is only expressed when nitrogenase is also expressed (Edgren & Nordlund, 2004; Sarma & Emerich, 2006).

Hypothetical protein BRADO5426 (spot 2150, Figs. 5.7, 5.8) was 8-fold more abundant (*t*-test=0.002) at DO^{low} relative to DO^{high}. At DO^{high}, this protein was barely visible on the gel at the corresponding pI and M_r. Its coordinated enhanced abundance for the Fix^{on} state of $Y88^T$ cells at DO^{low} suggests a possible regulatory role in some aspect of nitrogen fixation (suggested in Chapter 4).

5.2.1.3.1.2 PHA-accumulating proteins

Several PHB synthesis proteins were abundantly expressed at DO^{low} and DO^{high}. PhbB was identified in spot 4 (Fig. 5.7) and like PhbA (spot 3, Fig. 5.7) was not significantly altered between the different cell states representing $Y88^T$ physiology at DO^{low} and DO^{high}. A PHB granule-associated protein, PhaP, was identified in spots 1 and 2 at approximately 14 kDa and pIs of approximately 5.8 and 6.1 respectively (Figs. 5.7, 5.8). PhaP was observed to be the most highly abundant protein on all gels for all growth conditions. The M_r and pIs of both spots differed significantly from the predicted M_r and pI of the $Y88^T$ phasin. Unlike *R. eutropha,* which is reported to have four phasins (Neumann et al., 2008), $Y88^T$ has only one phasin predicted from the $Y88^T$ genome to be 33.9 kDa and have a pI of 9.56. Since the same high-scoring peptides were identified in each case (Table 5.3), the spots were considered not to be degradation products of the same protein. Spots identified to be the same protein with similar M_r but different pI, are suggested to represent isoforms with potential post-translational modifications (Büttner et al., 2001; Rosen et al., 2004; Zhu et al., 2005). These findings will be discussed in greater detail in Chapter 6.

5.2.1.3.1.3 Other proteins

GroEL was upregulated 3.4-fold at DO^{low} compared to DO^{high} in two adjacent spots with the same M_r but slightly different pIs of 5.07 and 5.09, spot 1240 and 1254 (*t*-test=0.03 and 0.034 respectively) and 11.52-fold (*t*-test=0.009) in spot 1851 (Fig. 5.7). The enhanced abundance of two adjacent spots at the same M_r is consistent with one of the spots possibly representing a phosphorylated form of the same protein. As a molecular chaperone, GroEL (together with DnaK) is mainly known for its essential role in core physiological responses (Lemos et al., 2007), participating in, amongst other things, a variety of cellular processes such as protein folding and preventing misfolded proteins aggregating, protein translocation as well as the assembly and disassembly of protein complexes. A lesser known role of GroEL is its involvement in the nitrogen fixation system where it fulfills both regulatory and structural functions and is required for the correct folding of NifA (Govezensky et al., 1991). It has also been shown that upon heat shock, a fraction of GroEL is phosphorylated with this modification enhancing the capacity of GroEL to bind to certain unfolded proteins (Sherman & Goldberg, 1992; Sherman & Goldberg, 1994). The capacity of GroEL to be phosphorylated is consistent with an additional, possibly phosphorylated form induced in $Y88^T$ cells (spot 1240, Fig. 5.7) at DO^{low} in addition to the enhanced abundance of the native form of the $Y88^T$ GroEL (spot 1254) at DO^{low} during nitrogen fixation. The lack of this potentially phosphorylated form on the gels at DO^{high} is consistent with the low to undetectable levels of all nitrogenase components compared to that at DO^{low}. It is therefore conceivable that a phosphorylated form of GroEL may be induced during nitrogen fixation to increase the binding affinity of GroEL for one or more of the nitrogen-fixing proteins, ensuring correct folding of these proteins as part of the nitrogenase assembly system and thereby regulation of nitrogen fixation in $Y88^T$.

GMP synthase (spot 1316) showed a 3.3-fold (*t*-test 0.003) increase in abundance at DO^{low} compared to DO^{high}. GMP synthase is involved in nucleotide biosynthesis and is an amidotransferase that catalyses the amination of xanthosine 5′-monophosphate to form GMP in the presence of glutamine and

ATP (Nakamura et al., 1995). Glutamine hydrolysis produces the necessary amino group, while ATP hydrolysis provides the energy to drive the reaction. Ammonia can also serve as an amino group donor (Nakamura et al., 1995). Additionally, glycine dehydrogenase, ribosomal protein L9 and prolyl-tRNA synthetase were identified in what appeared to be the same spot as GMP synthase (spot 1316, Fig. 5.7), although on magnification of the gel image it was evident that three separate spots existed (Fig. 5.9, see *). Due to their close co-migration, it was impossible to excise them as individual spots, so they were excised as one spot. Ribosomal protein L9 is considered to serve as a "molecular strut" most likely playing a role in ribosome assembly and/or maintaining the catalytically active conformation of ribosomal RNA (Lillemoen et al., 1997). Prolyl-tRNA synthetase is one of 20 aminoacyl-tRNA synthetases that catalyse the attachment of an amino acid to its cognate transfer RNA molecule. All four proteins in spot 1316 play some role in core metabolic processes of nucleotide, amino acid or ribosome (and thus protein) synthesis. High-scoring peptide matches to each of them ensured their inclusion as positive identifications since it was probable that due to their similar M_r and pIs (Table 5.3), they had co-migrated. Due to their mutual proximity, it was unclear which, if not all, of these proteins contributed to the increased abundance of this spot. Since the increase in abundance at DO^{low} coincided with nitrogen fixation, any or all of them may be implicated in some aspect of RNA/DNA repair or synthesis or amino acid synthesis to replenish requisite proteins for specialised activities such as nitrogen fixation.

Fructose-bisphosphate aldolase (Fba1p) and translation elongation factor Ts (EF-Ts) were both identified in spot 1785 and showed a 2.1-fold increase (*t*-test=0.01) in abundance at DO^{low}. These two spots migrated as adjacent spots in a "string" formation with four other spots. Due to their proximity on the 7 cm gels, it was difficult to excise the spot detected by BVA analysis to be altered from the more abundant protein spot that showed no altered abundance between the two DO conditions. Based on pI and the previous identification of the larger spot as EF-Ts, it was possible to determine that the upregulated spot

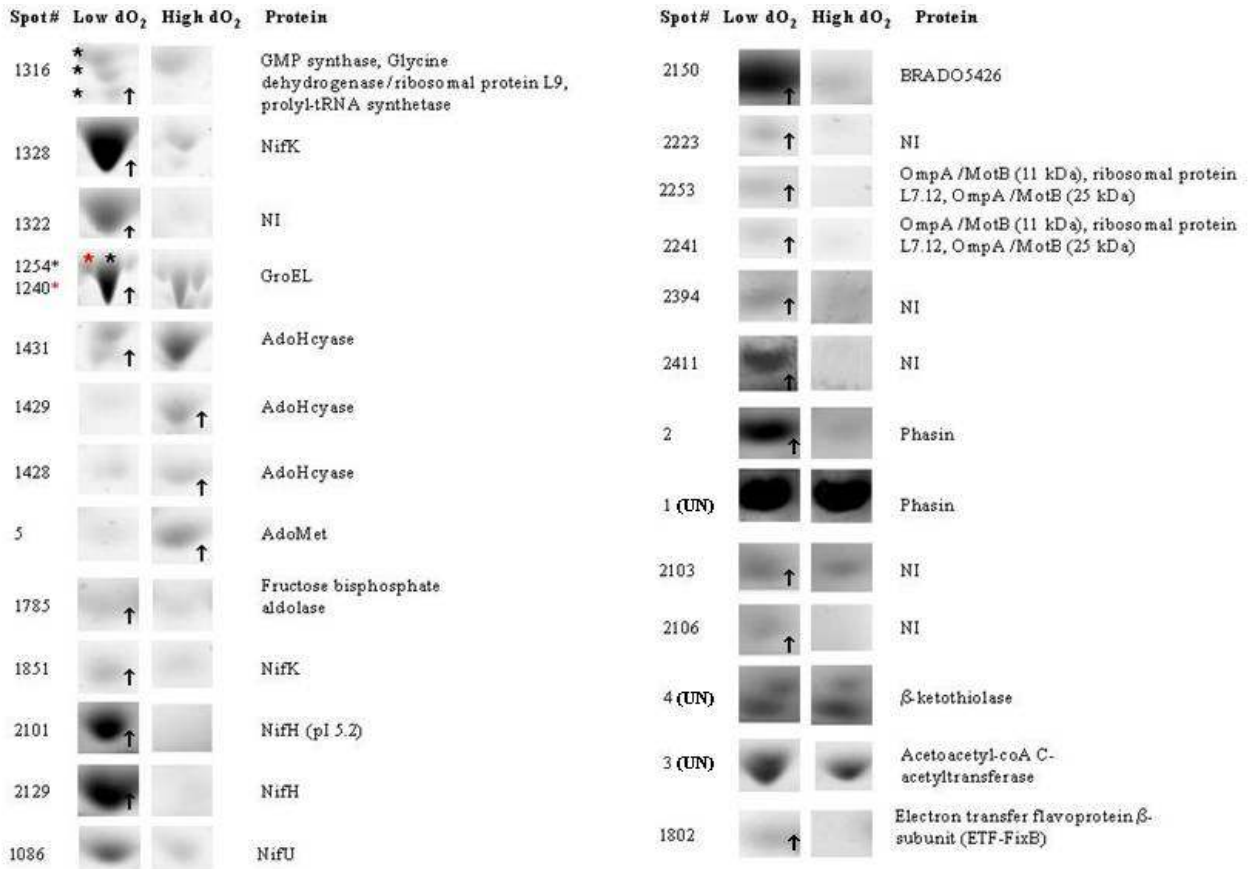


Figure 5.9 Magnification of the protein spots from 2-D DIGE pH 4-7 gels that showed altered abundance between DO^{low} and DO^{high} under constant pH 7 conditions in the growth medium.

Arrows denote those proteins with increased abundance at DO^{low} (left spot in each pair of spots) or DO^{high} (right spot in each pair of spots). Spots with no arrows were determined by BVA analysis to be unaltered (UN) between DO^{low} and DO^{high}. See body of the text throughout section 5.2.1.3 for discussion of these proteins.

was indeed Fba1p. This protein is required for glycolysis and gluconeogenesis and catalyses the cleavage of fructose-1,6-bisphosphate into glyceraldehyde-3-phosphate and dihydroxyacetone phosphate (Fitzgerald-Hughes et al., 2007). Enhanced abundance of this protein in $Y88^T$ may signify increased energy requirements during nitrogen fixation. Notably, KDPG aldolase (also known as 2-dehydro-3-deoxyphosphogluocnate aldolase which is unique to the ED pathway, see spot number C2, Chapter 8, Appendix A.1, Fig. A.1) and 6-phosphogluconate dehydratase, two diagnostic enzymes of the ED pathway were also abundant on the gels irrespective of DO conditions. It therefore appears that $Y88^T$ may utilise more than one pathway to catabolise glucose.

S-Adenosylhomocysteine hydrolase (AdoHcyase) was identified in three adjacent spots: 1428, 1429 and 1431 (Figs. 5.7, 5.8, 5.9). All showed a 3-fold increase in abundance at DO^{high} (*t*-test=0.006, 0.008, 0.01 respectively). Based on their similar M_r but different pIs (Figs. 5.7, 5.8), spots 1428 and 1429 may possibly be phosphorylated forms of the native AdoHcyase in spot 1429. AdoHcyase is a ubiquitous enzyme that plays a pivotal role in methylation-based processes by maintaining the intracellular balance between AdoHcy and S-Adenosyl-L-methionine (AdoMet). AdoHcy is the product of all AdoMet-dependent biological transmethylations. About 95% of AdoMet is reported to be used for transmethylation reactions that result in the formation of AdoHcy and is subsequently hydrolysed by AdoHcyase to form adenosine and homocysteine (Merali et al., 2000). AdoHcyase is the only known enzyme to catalyse the breakdown of AdoHcy (Turner et al., 2000) which if accumulated, can potentially be toxic to the cell (James et al., 2002). Since the accumulation of AdoMet is known to be significantly affected by DO as well as nitrogen source and pH (Yu et al., 2003), this would cause an imbalance between AdoMet and AdoHcy if not effectively regulated. It is therefore possible that the potentially phosphorylated forms of AdoHcyase on the gels at DO^{high} are induced to cope with increased activity of AdoHcy at DO^{high} levels. This is supported by the fact that AdoHcyase was also upregulated at the same M_r but different pI (from that considered to be the pI of the native form of AdoHcyase) at low OD_{600} for the

Fix^{off} state, where DO was not controlled (Chapter 4). Phosphorylation of AdoHcyase may therefore serve to increase the activity of this protein under DO^{high} to modulate AdoHcy and maintain the balance between AdoHcy and AdoMet. Since AdoMet is known to be significantly affected by DO, nitrogen source or pH as mentioned above (Yu et al., 2003), in Y88^{T}, nitrogen availability, DO^{high}, pH, or a combination of all three during growth under nitrogen-limiting and DO^{high} conditions may affect the balance between AdoHcy and AdoMet. Consequently, a heightened response by AdoHcyase to maintain this balance may be the reason for the observed increase in abundance as well as its appearance in three different spots at similar M_r but different pI.

In addition to the positive identification of AdoHcyase in spots 1428 and 1429, the third spot (spot 1431) matched to histidyl-tRNA synthetase as well as AdoHcyase. Since histidyl-tRNA synthetase may have co-migrated with AdoHcyase, it is unclear which of these two proteins, if not both, is responsible for the increase in abundance in spot 1431. Similar to prolyl-tRNA synthetase (spot 1316 mentioned above), histidyl-tRNA is one of the 20 aminoacyl-tRNA synthetases that catalyse the attachment of an amino acid to its cognate transfer RNA molecule and its upregulation in Y88^{T} is likely to be associated with increased protein synthesis.

Alkyl hydroperoxide reductase (AhpC) was identified in spot 2166 with a 2-fold increase in abundance (*t*-test=0.03) at DO^{low}. This protein migrated at a similar M_r but slightly lower pI than the unaltered AhpC protein identified in spot C6 (see Appendix A, Fig. A.1), which showed similar abundance at both high and DO^{low} concentrations. AhpC confers resistance to alkyl hydroperoxides that are produced in response to oxidative stress, by reducing these compounds to alcohols (Storz et al., 1990). Reactive oxygen intermediates are continuously produced in any aerobically metabolising cell and need to be dealt with continuously to avoid damage to DNA, RNA, proteins, and lipids. AhpC is also iron-regulated and proposed to be expressed at a significantly higher level in cells grown under iron-limiting conditions than in iron-sufficient conditions

(Baillon et al., 1999). Since iron is known to participate in the production of reactive oxygen intermediates via the Fenton reaction (Storz et al., 1990), the increased abundance of AhpC under iron-limiting conditions appears to be counterintuitive given that it is required to protect against iron-generated free radical production. With respect to Y88^{T} growth under nitrogen-depleted conditions, AhpC may be enhanced in Y88^{T} cells at DO^{low} due to the increased iron requirements for the highly abundant nitrogen fixation apoenzymes. Consequently, AhpC may be induced to protect the cells against the Fenton reaction under these growth conditions.

Additionally, the production of AhpC has been coupled to the activation of iron acquisition mechanisms such as those in *E. coli* that control superoxide dismutase (SOD) by means of the iron-responsive gene regulator Fur (Niederhoffer et al., 1990). Notably, under nitrogen-depleted growth at DO^{low} and DO^{high}, SOD was always abundant on the gels and showed no altered state between nitrogen-depleted growth conditions. Again, because of the iron requirements of nitrogen fixation proteins such as NifH, NifK and NifU (involved in the synthesis of the Fe-S cluster of nitrogenase (Schilke et al., 1999)), it is likely that the increased abundance of the isoform of AhpC is related to the activation of iron acquisition mechanisms required to maintain nitrogenase activity. Although enhanced abundance of AhpC is observed during nitrogen fixation at DO^{low}, a basal level of AhpC was evident for both DO^{low} or DO^{high} conditions, since the 20 kDa native AhpC protein and a 16 kDa isoform (confirmed to be two separate gene products based on the Y88^{T} genome) of this protein were present in equal abundance for both DO conditions at their deduced M_r and pIs. The similar abundance of these proteins irrespective of DO concentration, suggests that a background level of oxidative stress proteins may be required for Y88^{T} growth under conditions requiring increased amounts of iron.

Two highly abundant spots (spots 2241 and 2253, Figs. 5.7, 5.8) were evident at acidic pIs of 4.7 and 5.0 at approximately 12 kDa. Spot 2241 showed a 5.6-fold

(*t*-test=0.01) increase in abundance at DO^{low}, whereas spot 2253 showed a 10.3-fold (*t*-test=0.01) increase in abundance at DO^{low}. Two OmpA/MotB proteins encoded by two separate genes based on the $Y88^T$ genome, and ribosomal protein L7/L12 appear to have co-migrated in each of these spots. The M_r of the two OmpA/MotB proteins was 11.9 and 25.9. Their respective pIs were 5.76 and 7.03. Ribosomal protein L7/L12 is predicted to be 12.7 kDa and have a pI of 4.7. The same set of OmpA/MotB and ribosomal L7/L12 proteins was identified with unique peptide matches to both spots. Manual validation of all peptide matches suggested that any of the three proteins were equally likely to be the correct identification in both spots. Despite the 25 kDa OmpA/MotB protein having high-scoring peptides, the predicted pI of 7.03 and M_r for this protein were both higher than that for the observed spot. Therefore, the 11.9 kDa OmpA/MotB protein and the 12 kDa ribosomal protein L7/L12 were the more likely candidates represented in these spots. It was also considered that the 25 kDa OmpA/MotB protein may have been degraded and a fragment of this may have co-migrated with the other two proteins but this was discounted since the same unique, high-scoring peptides for this protein matched both spots. Alternatively, the 25 kDa protein could be processed post-translationally by cleavage, which would explain its migration at a much lower M_r and pI than predicted. Since the same migration pattern was evident in $Y88^T$ cells for all protein samples representing the Fix^{on} state (for both batch flasks and bioreactors), this suggests directed processing of this protein post-translationally during nitrogen fixation.

5.2.1.3.2 Basic proteins

Based on the pH 4-7 2-D gels, a major portion of the $Y88^T$ proteome seemed to comprise more acidic proteins. A more detailed profile of the $Y88^T$ basic proteome was obtained to determine whether any additional proteins could be mapped to differences in the response of $Y88^T$ to different DO concentrations. 2-D DIGE was used to visualise the more basic proteins in the linear gradient pI 6-11 range. Comparison of the proteins in the linear gradient pI 6-11 range for $Y88^T$ cells grown at DO^{low} or DO^{high} showed approximately 70 protein spots with about 30 of these present in high abundance (Fig. 5.10). Of these, only six

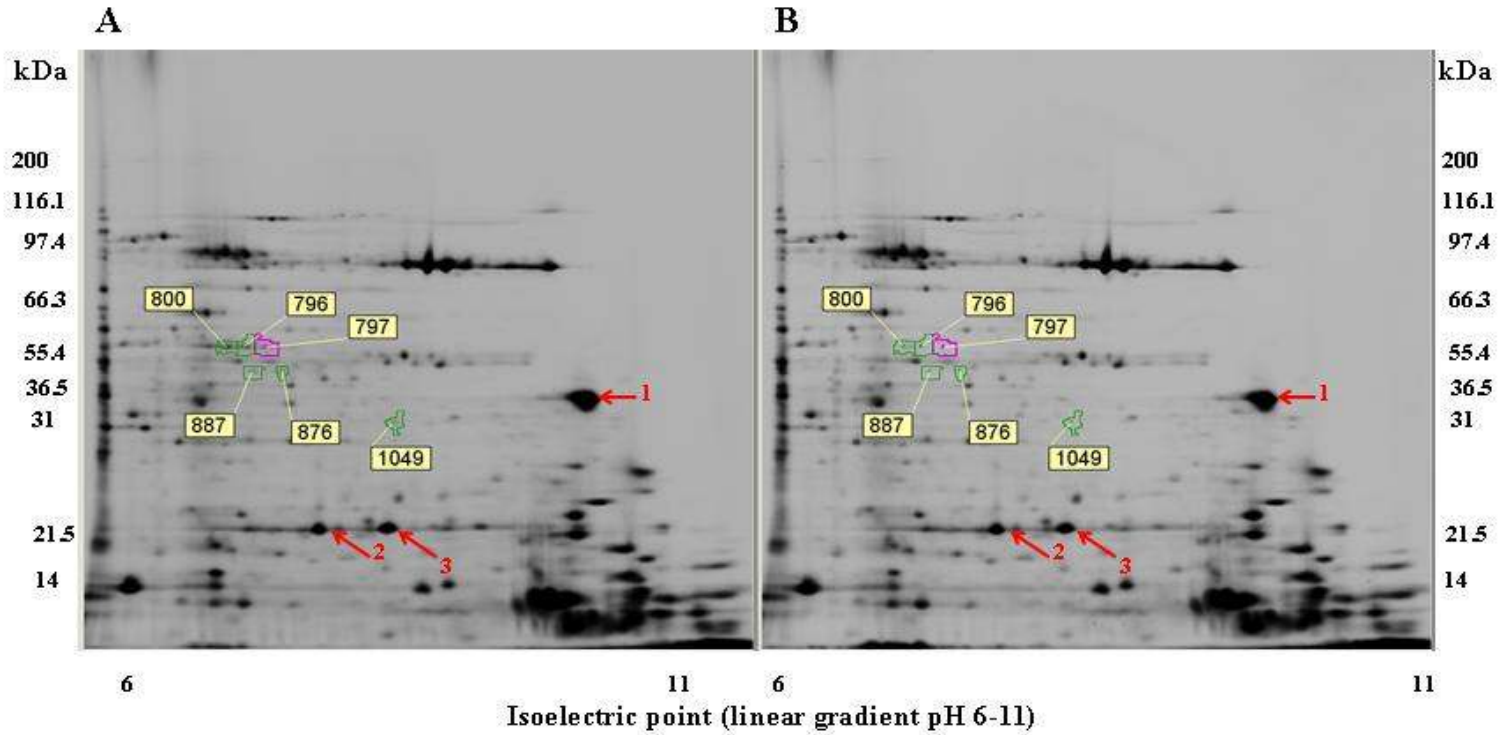


Figure 5.10 2-D DIGE images (pI 6-11) showing protein profiles of Y88^{T} cells grown to high OD_{600} (2.0) at DO^{low} or DO^{high} concentrations, pH not controlled (A) DO^{low}; (B) DO^{high}. Spot numbers (excluding those marked by red arrows) represent proteins determined by BVA analysis to have significantly altered abundance between the two DO concentrations. Spots marked by red arrows represent PHB synthesis proteins PhaP (spot 1) and two potential forms of PhbB with different pIs (spots 2, 3), (n=3)

(discussed in the following sections, see also Table 5.4) showed statistically significant altered abundance. In addition to those displaying altered abundance, 10 of the most abundant proteins that showed no change were selected for further analysis.

5.2.1.3.2.1 Nitrogen fixation proteins

The α-subunit of dinitrogenase (component I), NifD (spot 800, Fig. 5.10), showed a 14-fold increase in abundance (*t*-test=0.01) relative to DO^{high}. This 56 kDa protein was identified on the gel for the pH 6-11 range (spot 800) and had a predicted pI of 6.3. Despite this predicted pI, NifD was never observed on any gels in the pH 4-7 range. It was only ever identified on pH 6-11 gels for the Fix^{on} state at DO^{low} and migrated at a more basic pI (approximately 7.2) than that predicted from the $Y88^{T}$ genome sequence. Although it was observed at its predicted M_r, a distinct pI shift from the conceptual protein sequence suggests that this protein may be modified post-translationally during nitrogen fixation. It is well established that nitrogenase activity is regulated by post-translational modification of the NifH Fe component of nitrogenase (Ludden & Roberts, 1989; Martin & Reinhold-Hurek, 2002; Huergo et al., 2006). However, there is no apparent precedent in the literature that NifD is subject to post-translational modification. This provides the first evidence that a NifD protein may be post-translationally modified.

5.2.1.3.2.2 PHB synthesis proteins

Two PHB synthesis associated proteins were identified in the pI 6-11 range: PhaP (spot 1(UN), Fig. 5.9) and PhbB (spots 2 and 3, Fig. 5.9, Table 5.4). PhbB catalyses the second step in the PHB biosynthetic pathway. Both PhaP and PhbB were expressed at similar abundance irrespective of DO concentrations. PhaP was observed at its expected M_r and pI as predicted from the $Y88^{T}$ genome but the appearance of PhbB in two separate spots differing in pI but at the same M_r, suggests a pI shift for this protein indicative of an additional, potentially phosphorylated, isoform.

Spot number	Protein ID	Do^{low}/DO^{high}		Highest scoring unique peptides matched (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
800	Nitrogenase molybdenum-iron protein α-chain (NifD)	14.12	0.010	M.SVSTPTTIQEVK.D(60) R.SMNYISR.H(38)	56185	171	7	6.3	3
1	Phasin	–	–	K.AATAQTPTTAGVTETTK.E(101) K.TVAAAETLAPVAPK.A(92) K.SQASFGELGEFAK.G(87) K.GNVEALVESTK.I(84) K.TTKPAAAPLPR.T(73)	33963	5802	214	9.56	57
2	Acetoacetyl-CoA reductase (PhbB)	–	–	R.QVVANYAGNEEK.A(71) K.SGIHGFTK.A(65) R.VAIVTGGTR.G(60) M.ARVAIVTGGTR.G(50)	25370	3241	145	6.92	21
	hypothetical protein Saro_1378	–	–	R.TQAAQIQQIQEAGER.E(86) K.NAVQSQIPTTYK.A(85) K.VTLVLDQR.V(66) R.KVTLVLDQR.V(62) K.SPKPDQAALR.T(39)	24314	3264	132	9.1	33
3	Acetoacetyl-CoA reductase (PhbB)	–	–	R.QVVANYAGNEEK.A(70) R.VAIVTGGTR.G(60)	25370	1276	55	6.92	17

Table 5.4 2-D DIGE protein spot analysis showing statistical values and LC-MS/MS identification of proteins with altered abundance at high OD^{600} (2.0), and pH not controlled at DO^{low} or DO^{high}. * denotes $p<0.01$ excepting for those spots denoted by ** where $p<0.05$. Positive values represent increased abundance of proteins at DO^{low} and negative values represent increased abundance of proteins at DO^{high}. NI=protein not identified. Spot numbers 3, 4 and 5 (Fig. 5.7) showed altered abundance but were determined by BVA analysis not to be significantly altered. See body of the text throughout section 5.2.1.3 for discussion of these proteins.

Notably, PhaP was previously identified in two spots at approximately 14 kDa, well away from its predicted 33.96 kDa, but at different pIs (see previous section 5.2.1.3.3), suggesting that two different isoforms of the protein exist, at least one of which is additionally post-translationally modified. Numerous high-scoring peptide matches to PhaP in this spot on the pH 6-11 gel found at the predicted M_r and pI for this protein, not only confirmed the genome prediction of a non-canonical phasin protein with an N-terminus rich in alanine, proline and lysine residues, but also confirmed additional unique peptide matches to the N-terminus that had not previously been matched by any of the former peptides for the two phasin isoforms identified on the pH 4-7 gels (n=8 independent gels). The additional peptide matches to the N-terminus increased the sequence coverage to account for almost the entire protein sequence for PhaP. This suggests that the form of PhaP observed on the pH 6-11 gels may be the precursor of the other two forms observed on the pH 4-7 gels, the amino terminal segment of which is cleaved to produce the 13 kDa phasin isoforms. The implications of these findings are discussed in Chapter 6.

5.2.1.3.2.3 Other

Several additional spots were determined to be significantly altered between DO^{low} or DO^{high} conditions in the pI 6-11 range. None of these spots yielded positive identifications due to insufficient quantities of protein.

5.2.1.4 Comparison of protein profiles obtained for DO^{low} or DO^{high} conditions at high cell densities (OD_{600}=2.0), pH controlled

5.2.1.4.1 Acidic proteins

The batch flask and bioreactor experiments (Chapter 4, Chapter 5, Section 5.1.2.5.2) carried out to compare protein profiles for the physiological states of $Y88^T$ cells during growth at DO^{low} or DO^{high} conditions did not have the pH of the medium controlled. Consequently, $Y88^T$ glucose-mediated growth resulted in acidification of the medium. To determine the effect, if any, of pH on the abundance of those proteins identified in the proteomic signatures for the Fix^{on} and Fix^{off} states, the growth experiments were repeated in the bioreactors at DO^{low} and DO^{high} but with pH controlled at 7.0 for the duration of the

experiment. 2-D DIGE was carried out for different combinations of low or high OD_{600} and DO^{low} or DO^{high} concentrations under pH-controlled conditions as determined by the experimental design (see Chapter 2, Tables 2.1, 2.2) to take into account all possible effects of growth and DO concentrations within the existing growth parameters. Ninety-six protein spots were selected for analysis, which included those proteins determined to be statistically altered under DO^{low} or DO^{high} conditions with pH controlled as well as many of those that showed no altered abundance. The goal was to attempt to identify as many proteins as possible in underlying metabolic pathways to obtain a more complete profile of the $Y88^{T}$ proteome and more representative protein signatures for the specified growth conditions.

The number of proteins showing statistically significant differences in abundance between different DO concentrations with pH not controlled compared to different DO concentrations with pH controlled at 7.02±0.2, was dramatically reduced. Previously, 38 proteins showed altered abundance between DO^{low} and DO^{high} conditions with pH not controlled with 32 of these in the pH 4-7 range (Fig. 5.8) and six in the pH 6-11 range (Fig. 5.10). Under the same DO conditions but with pH controlled, this number was reduced to seven for pH 4-7 DIGE (Figs. 5.11, 5.12) and from six (Fig. 5.10) to three for pH 6-11 DIGE (see section 5.2.1.5.4 below), suggesting that changes of many of these proteins may be regulated or affected by pH, not DO. Despite these changes, the proteins signatures for nitrogen fixation and PHB synthesis showed similar profiles to those detected in previous experiments (sections 5.2.1.3, 5.2.1.4). Since the proteomic signatures for the different DO concentrations were highly altered once the medium was prevented from naturally acidifying, pH clearly had an effect on many proteins with the exception of the core nitrogen-fixing proteins and PHB synthesis/associated proteins.

5.2.1.4.1.1 Linear gradient pH 4-7 range

Several proteins showed altered abundance for the two DO states, with seven proteins that showed statistically significant changes. These included three proteins (spots 1249, 2038 and 2136) that had not previously shown any change

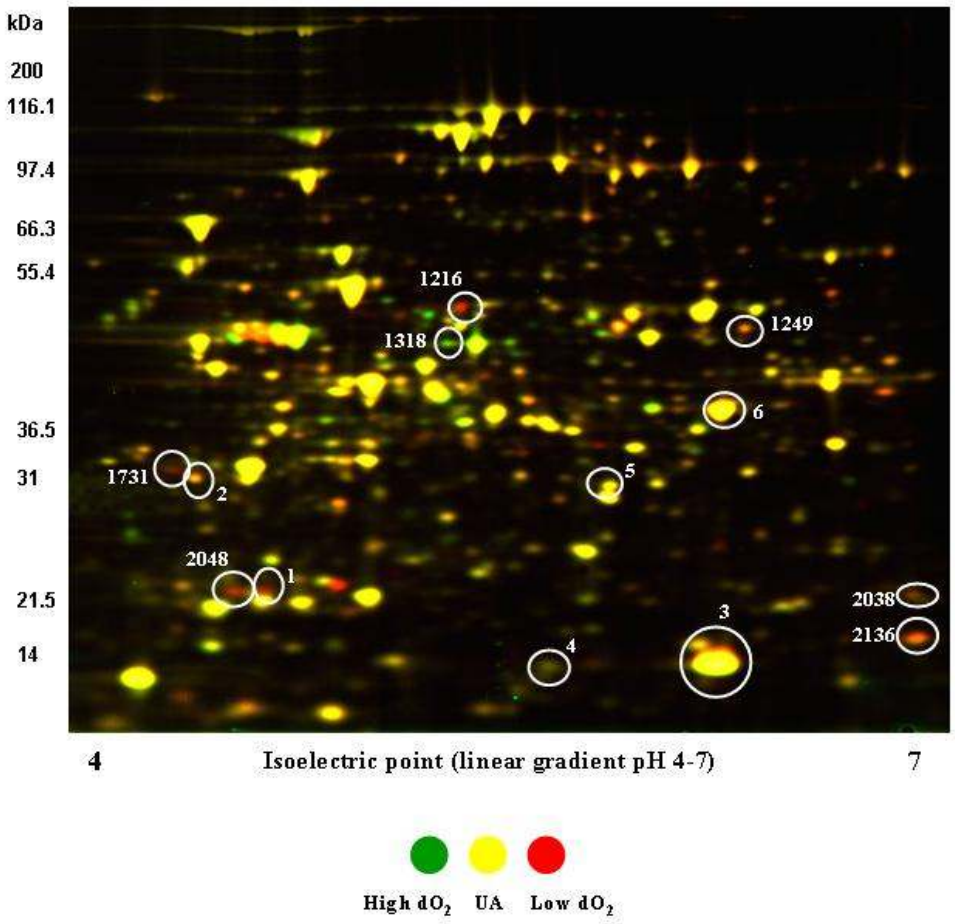


Figure 5.11 Overlay image of Y88^{T} protein profiles generated from minimally labeled samples showing differential protein abundance under DO^{low} or DO^{high}, constant pH 7 conditions in the linear gradient pH 4-7 range

Pooled (Cy2 - yellow), DO^{low} (Cy3 - green) or DO^{high} (Cy5 - red) samples are overlaid. Overlay image of Y88^{T} protein profiles generated from minimally labeled pooled (Cy2 - yellow), DO^{low} (Cy3 - green) or DO^{high} (Cy5 - red) samples showing differential protein abundance in the linear gradient pH 4-7 range. Yellow (UE)=proteins expressed similarly in both states of the cells at DO^{low} or DO^{high}, red (AE DO^{low})=proteins unique to the state of the cells at DO^{low}, green (AE DO^{high})=proteins unique to the state of the cells at DO^{high}. Spot numbers (excluding spots 1, 2, 3, 4, 5 and 6) represent proteins determined by BVA analysis to have significantly altered abundance between the two DO concentrations. Spots 1 and 2 showed altered abundance but were outside the BVA filter thresholds. Spots 3, 4, 5 and 6 are involved in PHB synthesis. Spots 1249, 2038 and 2136 that showed significant change in abundance under pH controlled conditions, did not show altered abundance when pH was not controlled. Spots selected for analysis were those likely to be involved in the nitrogen-fixing and PHB pathways based on predicted pI and molecular weight, as well as some spots with altered abundance that are not usually associated with these pathways.

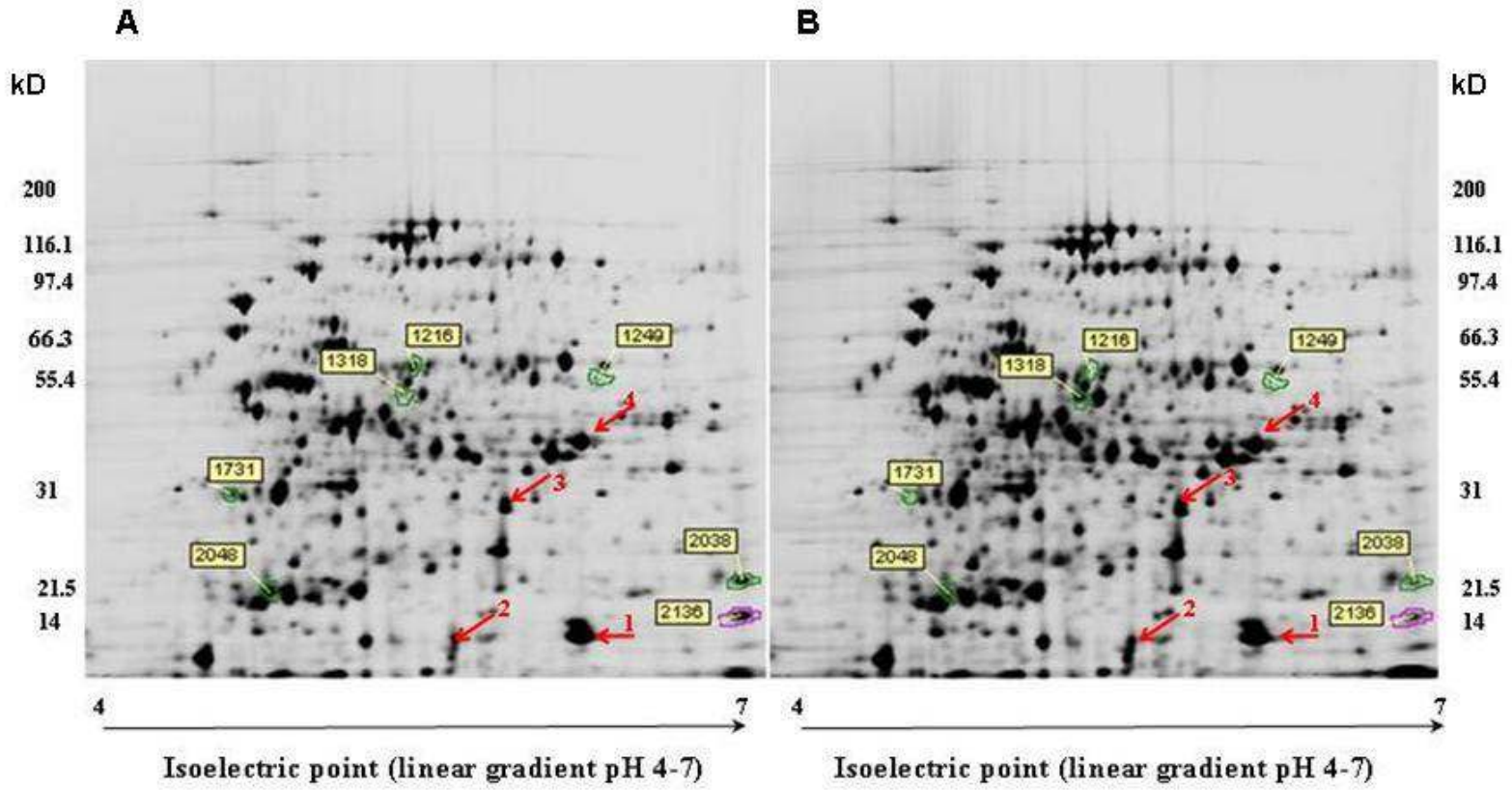


Figure 5.12. 2D DIGE images (pI 4-7) showing protein profiles of Y88^{T} cells grown to a high OD_{600} (2.0) under ammonium-depleted conditions at DO^{low} or DO^{high} at constant pH

(A) DO^{low}; (B) DO^{high}. Spot numbers represent proteins determined by BVA analysis to have significantly altered abundance between the two DO concentrations.

(Figs. 5.11, 5.12, 5.13). With both DO and pH controlled, proteomic signatures for DO^{low} or DO^{high} differed from those determined for DO^{low} or DO^{high} with pH of the medium not controlled.

5.2.1.4.1.2 Nitrogen fixation proteins

The NifK protein (spot 1216) showed significant altered abundance between the two DO conditions under controlled pH conditions. As before, NifK was identified at 58 kDa and pI 5.4 and showed a greater than 4-fold increase in abundance (*t*-test=0.019) at DO^{low} relative to DO^{high} (Table 5.5, Figs. 5.11, 5.13). This protein was once again detected at very low levels at the corresponding M_r and pI on the gel at DO^{high} (see corresponding spot 1216, Figs. 5.12B, 5.13).

The fact that DO^{high} results in decreased abundance of nitrogenase is not unexpected given the susceptibility of nitrogenase to oxygen (Staal et al., 2007). What is surprising is that this low level of NifK may represent a still-active form of the protein based on a very low level of nitrogenase activity detected in the acetylene reduction assay (see Fig. 5.1D at the beginning of this chapter). Similarly, the abundance of the NifH protein (spot 2048, Figs. 5.11, 5.12B, 5.13) was altered at DO^{high}, showing a 2.5-fold (*t*-test=0.033) decrease in abundance of this component of the nitrogenase enzyme at DO^{high} conditions. Again, this subunit of the nitrogenase enzyme was barely visible on the gel at DO^{high} compared to the highly abundant spots evident at the same M_r and pI on the gel at DO^{low}. Considering the obvious abundance of spot 2048 at DO^{low} but its huge decrease in abundance to only barely detectable levels on the gel at DO^{high}, it is surprising that BVA analysis determined only a 2.5-fold difference in abundance between the two. Nevertheless, the spot was highly downregulated as evident on the gels at DO^{high} suggesting an effect of DO^{high} on dinitrogenase. Despite its downregulation at DO^{high}, NifH, like NifK may still be present at very low levels under DO^{high} conditions.

Notably, only one form of NifH was visible at pI 4.7 on the gel at DO^{low} with pH controlled instead of the two forms previously observed at different pIs and M_r with pH not controlled. Thus at constant exposure to DO^{low}, with the

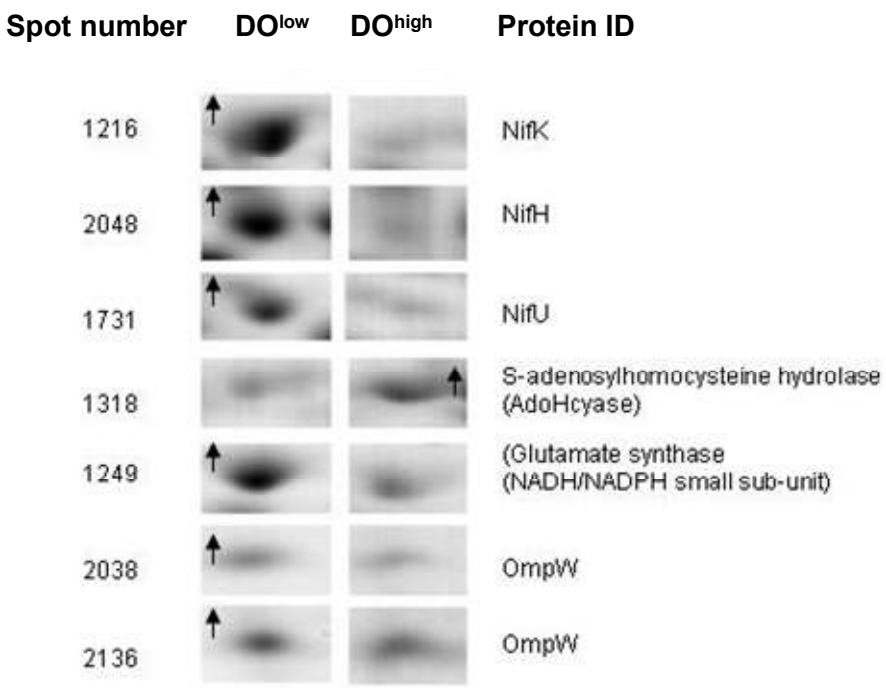


Figure 5.13 Magnification of the protein spots from 2-D DIGE pH 4-7 gels that showed altered abundance between DO^{low} and DO^{high} under constant pH 7 conditions in the growth medium. Arrows denote increased abundance for the protein under DO^{low} or DO^{high} conditions.

Spot number	Protein ID	Do^{low}/DO^{high}		Highest scoring unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
1216	Dinitrogenase (NifK)	4.3	0.019**	R.MYDGGTTLEEAGQAVHAR.A(109)	58437	13674	708	5.37	44
				K.MQDLFDSSPFGAGCK.A(95)					
				K.INFLGGFDGYTVGNTR.E(95)					
				K.VEYTILGDNSDVWDTPTDGEFR.M(84)					
				R.ATISMQEFCTEK.T(72)					
1249	Glutamate synthetase (NADP/NADPH), small subunit	2.13	0.015**	R.VTGENDVTGLECVR.V (80)	52955	150	16	5.89	12
				R.FAMEFLTQQNK.R(65)					
				QEVPGSEFTLK.A (51)					
1318	S-adenosylmethionine (SAM) synthetase	-2.13	0.045**	K.GIMDTDGNWAPGVPEEVER.V (86)	43503	130	14	5.27	10
				K.NIVAAGLAHR.C(49)					
				R.TDLVADLK.A(49)					
1731	NifU	2.52	0.040**	R.TNGLTSIDEVTNYTK.A(107)	33542	3630	98	4.72	59
				K.AGGGCSTCAEGIEGVLER.V(93)					
				K.NSGILEDADGVGDVGAISCGDALR.L(89)					
				R.DGGDCELVDVEGNR.V(85)					
				K.VDEPSQTISDAR.F(72)					

Table 5.5 2-D DIGE protein spot analysis (pH 4-7) and LC-MS/MS analysis of protein from cells grown at DO^{low} or DO^{high} at constant pH and high OD_{600} (2.0)

2-D DIGE protein spots in the linear gradient pI 4-7 range showing differential abundance between DO^{low} or DO^{high} concentrations, constant pH or high OD_{600} (2.0). For those spots denoted by * $p<0.01$ otherwise $p<0.05$ (**). Positive values represent increased abundance of proteins at DO^{low} and negative values represent increased abundance of proteins at DO^{high}. Table 5.5 continued on next page.

Spot number	Protein ID	Do^{low}/DO^{high}		Highest scoring unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
2038	Outer membrane protein W (OmpW)	2.24	0.003*	K.LLATGVLADGNIDTVR.S(94)	25096	251	6	9.1	4
				K.TPYSLTFDAK.K(71)					
2048	Dinitrogenase reductase (NifH)	2.5	0.033*	R.CVESGGPEPGVGCAGR.G(120)	31859	4200	123	4.94	32
				K.STTSQNTLAALADLGQR.I(118)	–	–	–	–	–
				R.QTDKELELAESLAK.K (85)	–	–	–	–	–
				K.ELELAESLAK.K(70)	–	–	–	–	–
				R.DNIVQHAELR.R(66)	–	–	–	–	–
2136	Outer membrane protein W (OmpW)	2.12	0.012*	K.LLATGVLADGNIDTVR.S(94)	25096	851	16	9.1	10
				K.TPYSLTFDAK.K(71)	–	–	–	–	–

Table 5.5 2-D DIGE protein spot analysis (pH 4-7) and LC-MS/MS analysis of protein from cells grown at DO^{low} or DO^{high} at constant pH, high OD_{600} (2.0)

2-D DIGE protein spots in the linear gradient pI 4-7 range showing differential abundance between DO^{low} or DO^{high} concentrations, constant pH or high OD_{600} (2.0). For those spots denoted by * $p<0.01$ otherwise $p<0.05$ (**). Positive values represent increased abundance of proteins at DO^{low} and negative values represent increased abundance of proteins at DO^{high}.

extracellular medium maintained at pH 7, the appearance of NifH was altered at DO^{low} and the two forms that were previously evident (Chapter 4 and Chapter 5, Figs. 5.7, 5.8A, spots 2129 and 2101), were no longer observed as two distinct, abundant spots on the 2-D gels (spot 2048, Figs. 5.11, 5.12A). Of the two spots previously observed, the spot considered to be the more basic NifH form (spot 2101, Figs. 5.7, 5.8A) became barely visible on the gel. This suggests that it may have been differently regulated when the growth medium was maintained at a neutral pH and it may have migrated to another position on the gel. Alternatively, the two spots could have resolved into the one, more acidic spot observed. Thus the former presence of the more basic NifH form could be attributed to the change brought about by the more acidic growth medium. Its former presence at a higher-than-predicted pI and/or higher M_r (by about 1000 Da) than that of the more acidic form (spot 2129, Figs. 5.7, 5.8A; spot 2048, Figs. 5.11, 5.12A), may have been due to some possible post-translational modification as a result of the more acidic growth conditions as previously discussed (section 5.2.1.3). Its disappearance from the corresponding pI and M_r on the gels at DO^{low} for the neutral growth conditions suggests some modification that may have caused a shift in pI or M_r. Despite the altered NifH profile for the pH controlled, DO^{low} conditions compared to the same conditions but with pH not controlled, the activity of the nitrogenase enzyme was clearly not affected by the change in this form of NifH, since a positive acetylene reduction assay was obtained for $Y88^T$ cells grown under controlled pH, DO^{low} conditions. Since NifH functions as a homodimer, it is conceivable that all the NifH was therefore reduced to a single spot under neutral growth conditions. Growth at acidic pH therefore appears to result in an intracellular response with the modification of a key component of the $Y88^T$ nitrogenase under acidic growth conditions, a finding for which there is no apparent precedent in the literature.

The NifU protein (spot 1731) showed a 2.5-fold (*t*-test=0.04) increase in abundance at DO^{low} relative to DO^{high} (Figs. 5.11, 5.12A, Table 5.5). At DO^{high}, it was coordinately downregulated with the NifK and NifH components of

nitrogenase. It is possible that this protein plays a regulatory role in negative feedback of nitrogenase enzyme activity at DO^{high} whether pH is controlled or not, since decreased abundance of NifU occurred regardless of pH. NifU has been shown to affect the activity of both nitrogenase components (Dos Santos et al., 2004). The proteomic signature obtained with respect to NifK and NifU at DO^{low} with pH controlled is consistent with that previously obtained for the nitrogen-fixing state of $Y88^T$ cells at high OD_{600} in which pH was not controlled (Chapter 4). The coordinated decrease in abundance of the nitrogen-fixing proteins at DO^{high} and controlled pH is consistent with their decrease for the Fix^{off} state of $Y88^T$ cells at lower OD_{600} (Chapter 4). Therefore, pH had no visible effect on NifK and NifU abundance, although it may have had an anomalous effect on NifH abundance, at least for the more basic form of NifH, despite apparently not having an effect on nitrogenase activity.

5.2.1.4.1.3 Nitrogen assimilation

The small subunit of NADPH-dependent glutamate synthase (also known as GOGAT), was identified on the gels at DO^{low} concentration (spot 1249, Figs. 5.11, 5.12, 5.13) and showed a >2-fold change in abundance (*t*-test=0.015) relative to DO^{high} concentrations (Table 5.5). This protein was observed at the predicted 53 kDa but at a greater pI than expected (pI=6.4 compared to the expected pI of 5.9) from the $Y88^T$ genome prediction. GOGAT (NADPH- or NADH-dependent) is a key enzyme in the early stages of ammonia assimilation in bacteria that catalyses the reductive transamidation of the amido nitrogen from glutamine and 2-oxoglutarate to form glutamate. The ammonia that enters this pathway can be supplied either by internal metabolic processes such as amino acid catabolism, or through the reduction of external nitrogen sources such as atmospheric dinitrogen. Most bacteria use NADPH-dependent GOGAT, although some are reported to show NADH-dependent activity (Dincturk & Knaff, 2000). NADPH-dependent GOGAT in bacteria consists of two different subunits, the larger of which shows considerable similarity to the ferredoxin-dependent enzymes.

The identification of NADPH-dependent GOGAT on $Y88^T$ gels at DO^{low}, suggests that $Y88^T$ presumably uses the NADPH-dependent form of GOGAT as the electron donor for the two-electron conversion of glutamine to glutamate. Since the ammonia supplemented in the medium was depleted by the time the $Y88^T$ cells were harvested for protein extraction, the abundance of NADPH-dependent GOGAT at DO^{low} concentrations must be due to the assimilation of ammonium produced intracellularly as a consequence of the reduction of external atmospheric dinitrogen. The decreased abundance of this protein at DO^{high} is consistent with the decreased ethylene peak in the acetylene reduction assays at DO^{high} as well as the decreased abundance of the nitrogen-fixing proteins on the 2-D gels. The altered abundance of NADPH-dependent GOGAT at DO^{low} or DO^{high} may reflect the ability of $Y88^T$ to establish hierarchical cascades of protein abundance in response to nitrogen limitation.

AdoHcyase (spot 1318) showed a greater than 2-fold difference (*t*-test=0.045) at DO^{high} relative to DO^{low}. Its role in balancing AdoHcy and AdoMet, an ubiquitous cofactor in protein enzymes that is mainly involved in transmethylation, transsulfuration and polyamine synthesis (Cochrane & Strobel, 2008), was previously described (Chapter 5, section 5.2.1.3.4). Since nitrogen source, pH and DO have been shown to have significant effects on accumulation of AdoMet (Yu et al., 2003), the increased abundance of a potentially phosphorylated form of AdoHcyase at DO^{high} with pH controlled is consistent with its former increased abundance for DO^{high} conditions (Chapter 5, section 5.2.1.3.4) where pH was not controlled. Thus, irrespective of pH, this protein appears to be abundant at DO^{high} as a presumably more phosphorylated form similar to that observed for the Fix^{off} state (Chapter 4).

The outer membrane porin protein OmpW was identified in two spots (spots 2136, 2038) both of which showed greater than 2-fold increase in abundance (spot 2038, *t*-test=0.003; spot 2136, *t*-test=0.01) at DO^{low} relative to DO^{high} (Table 5.5). The identification of this protein in both spots at different M_r but the same pI suggests that one isoform could be post-translationally modified, possibly by glycosylation, a modification that is does not change the pI of the protein but

does change its M_r. Bacterial outer membrane proteins (OMPs) are known to play important roles in the cell's adaptive response to environmental conditions (Nikaido, 1999; Lin et al., 2002) and OmpW has been shown to be expressed under stress conditions, such as low aeration, elevated temperature or high salt concentration (Nandi et al., 2005; Hong et al., 2006). The increased abundance of this protein at DO^{low} regardless of the growth state of the cell (see later Sections 5.2.1.5.7, 5.2.1.6.2) suggests that its altered abundance in $Y88^T$ may be a response to low aeration.

Several outer membrane proteins were identified in spots that showed no altered abundance on the gels for both DO^{low} and DO^{high} at controlled pH (see Appendix 3). A TonB-dependent siderophore receptor, additional TonB-dependent receptors and a TolB-like protein were identified in eight other spots on the gels. TonB-dependent receptors are outer membrane proteins known for, amongst other roles, the active transport of iron siderophore complexes in Gram-negative bacteria. Due to the high insolubility of Fe(III) in aerobic conditions, many Gram-negative bacteria produce small iron-chelating compounds called siderophores under iron-limiting conditions (Moeck & Coulton, 1998). TonB-dependent receptors function as receptors for these "ferrisiderophores". The increased abundance of a TonB-dependent siderophore receptor at DO^{low} is consistent with the acquisition of iron uptake that may be required for the iron-sulfur cluster of nitrogenase utilised during nitrogen fixation. At DO^{high}, the decreased abundance of this protein suggests a decreased requirement for iron, consistent with what would be expected for decreased iron requirements when the iron-dependent nitrogenase enzyme and accessory proteins are not present at such high abundance under DO^{high} conditions. The upregulation of this OMP together with the coordinated upregulation of the porin TonB-dependent siderophore receptor suggests a functional role for these OMPs in iron uptake and transport across the periplasm during nitrogen fixation.

5.2.1.4.2 Basic proteins

5.2.1.4.2.1 Nitrogen-fixing proteins

Three spots showed altered abundance in the pI 6-11 range between DO^{low} and DO^{high} in which culture pH was controlled. As in the uncontrolled pH experiments, NifD (spot 880, Fig. 5.14, Table 5.6) showed a higher abundance at DO^{low} although the 2.4-fold increase in abundance (*t*-test=0.005) under pH-controlled conditions between DO^{low} and DO^{high} was much lower than the 14-fold difference determined between these two DO levels with pH not controlled (Section 5.2.1.4.1). Thus controlling pH resulted in a smaller difference in abundance of this protein. Once again, NifD was observed only on the pH 6-11 gels at its predicted M_r but at a pI of 7.2 rather than its predicted pI of 6.3. As a result of this second set of independent, replicate experiments which confirms the presence of NifD again at a higher than predicted pI, it seems likely that the $Y88^T$ NifD may be post-translationally modified as previously suggested (Section 5.2.1.4.1).

5.2.1.4.2.2 PHB synthesis proteins

Irrespective of DO^{low} or DO^{high} conditions with pH controlled, the proteins involved in PHB synthesis (Fig. 5.14, spots 1 through to 3; Table 5.6) previously observed in the pH 6-11 range for DO^{low} or DO^{high} with pH not controlled (Fig. 5.10), were once again observed at the same pI and M_r as before. Once again there was no significant change in their abundance. The abundance of these proteins remained consistently high irrespective of growth at DO^{low} or DO^{high} or whether the growth medium was more acidic or neutral.

5.2.1.4.2.3 Other proteins

Spot 1784 (Fig. 5.14) identified as OmpW (and previously identified in spot 2136, Fig. 5.12) showed a 2.3-fold increase at DO^{low} at approximately 14 kDa and pI 6.4 (Table 5.6). Since the abundance of OmpW was unaltered between low or high OD_{600} at DO^{low}, this confirmed that its altered abundance was a response to DO^{low} conditions at constant pH. Peptide matches to this protein were the same as those in spots 2036 (at a higher M_r) and spots 2136 (the same

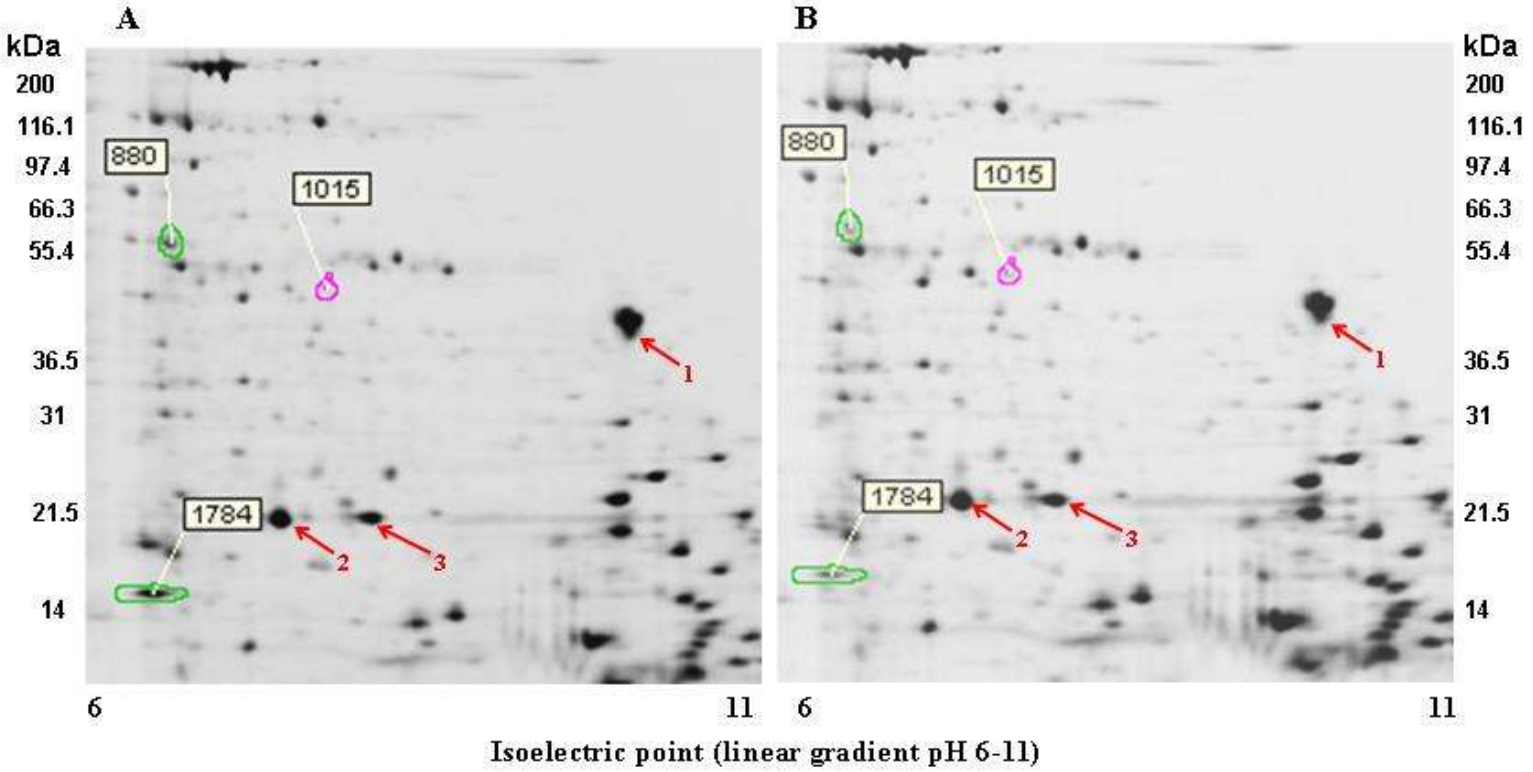


Figure 5.14 2-D DIGE images for the linear gradient pI 6-11 range showing protein profiles of Y88^{T} cells grown to a high OD_{600} at DO^{low} or DO^{high} at constant pH 7. Cells were grown to an OD_{600} of approximately 2.0 under DO^{low} (A) or DO^{high} (B) conditions. Spot numbers (excluding those marked with red arrows) represent proteins determined by BVA analysis to have significantly altered abundance between the two DO concentrations. Spots marked by red arrows represent PHB synthesis proteins PhaP (spot 1) and two potential forms of PhbB with different pIs (spots 2, 3), (n=3).

Spot number	Protein ID	Do^{low}/DO^{high}		Highest scoring unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
880	Nitrogenase molybdenum-iron protein a-chain (NifD)	2.39	0.010*	M.SVSTPTTIQEVK.D(60)	56185	171	7	6.3	3
				R.SMNYISR.H(38)	–	–	–	–	–
1015	NI	-3.49	0.016**	–	–	–	–	–	–
1784	Outer membrane protein (OmpW)	2.32	0.032**	K.LLATGVLADGNIDTVR.S(92)	25096	735	13	9.1	10
				K.TPYSLTFDAK.K(71)	–	–	–	–	–

Table 5.6 2-D DIGE protein spot analysis (pH 6-11) and LC-MS/MS analysis of protein from cells grown at DO^{low} or DO^{high} at constant pH

2-D DIGE analysis of protein spots in the linear gradient pI 6-11 range that were determined by BVA analysis to be significantly altered between DO^{low} or DO^{high} concentrations at constant pH or high OD_{600} (2.0). For those spots denoted by * $p<0.01$ otherwise $p<0.05$ (**). Positive values represent increased abundance of proteins at DO^{low} and negative values represent increased abundance of proteins at DO^{high}. NI=protein not identified.

M_r) on the pI 4-7 gels (Fig. 5, 12). Since this protein occurred at the same M_r at the extreme pH 7 end of the pH 4-7 gel (spot 2136, Fig. 5.12) and then again at the same M_r between pH 6-7 on the pH 6-11 gel (spot 1784, Fig. 5.14), this was clearly an overlap of the same protein in the corresponding section of the two different pI ranges.

Spot 1015 (Fig. 5.14) showed a 3.5-fold decrease (Table 5.6) at DO^{high} at approximately 50 kDa and pI 7.5. No identification could be obtained for this spot due to its very low abundance.

5.2.1.5 Comparison of protein profiles obtained for DO^{low} or DO^{high} conditions at low cell densities (OD_{600}=0.5), pH controlled

To determine whether the altered abundance of $Y88^T$ proteins between DO^{low} or DO^{high} conditions were due to the differences in sustained DO levels at high OD_{600} or whether they were part of the regulatory network of changes associated with changes from low to high growth, proteomic signatures were obtained for $Y88^T$ cells subjected to different DO concentrations at the low OD_{600} of 0.5.

5.2.1.5.1 Acidic proteins

As with extracts from cells grown to OD_{600}=2.0, approximately 200 protein spots were present in the linear gradient pI 4-7 range (Fig. 5.15). Of these, approximately 70 occurred in high abundance. BVA analysis determined that six proteins were significantly altered between DO^{low} and DO^{high} at low OD_{600}. Three of these were the same proteins that previously showed altered abundance between DO^{low} and DO^{high} at high OD_{600}. Since at low OD_{600}, ammonium would still have been available in the medium, it was not surprising that no nitrogen fixation proteins were expressed with either DO^{low} or DO^{high}.

AdoHcyase was identified in two adjacent spots, 1318 and 1326 (Fig. 5.15), both of which were upregulated at DO^{high} with a 2.4-fold increase in abundance in spot 1318 (*t*-test=0.01) and a 2.5-fold increase in abundance in spot 1326 (*t*-

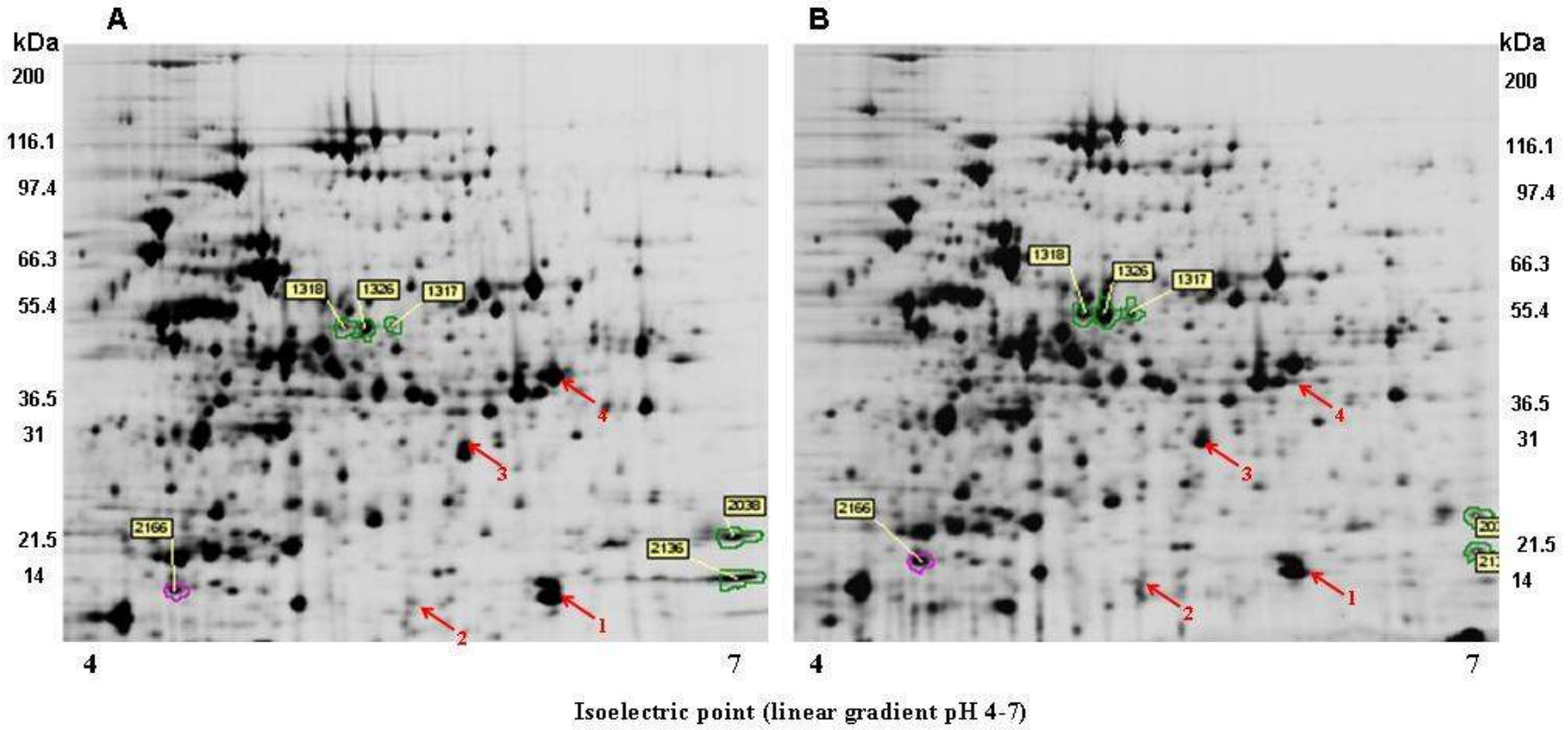


Figure 5.15 2-D DIGE images (pI 4-7) showing protein profiles of Y88^{T} cells grown to a low OD_{600} under ammonium-depleted conditions at DO^{low} or DO^{high} at constant pH 7

Cells were grown to an OD_{600} of 0.5 under DO^{low} (A) or DO^{high} (B) conditions. Spot numbers (excepting those marked with red arrows) represent proteins determined by BVA analysis to have significantly altered abundance between the two DO concentrations. Spots 1 through 4 represent PHB synthesis proteins.

test=0.003) (Table 5.7). A third spot (spot 1317) in this "train" of spots was only tentatively identified as AdoHcyase (one peptide matched in 3 separate experiments). Since this protein is upregulated at both high and low growth at DO^{high} regardless of the pH, its altered abundance can be attributed to a response to DO^{high}. The fact that it appears in two adjacent spots at the same M_r but slightly different pIs, suggests that spot 1318 may be a phosphorylated form of the protein appearing in spot 1326 as discussed before (Chapter 5 section 5.2.1.3.4). The same can be said for spot 1317. Notably, AdoHcyase was expressed at higher abundance on all gels representing the Fix^{off} state relative to those for the Fix^{on} state, regardless of extracellular pH or DO conditions. Thus the enhanced abundance of AdoHcyase generally may be related to the nitrogen source since nitrogen source is known to affect the accumulation of AdoMet (Yu et al., 2003). The induction of phosphorylated forms of AdoHcyase in $Y88^T$ may be more related to DO^{high} concentrations.

5.2.1.5.2 Basic proteins

Only one spot was determined by BVA analysis to be significantly altered in the linear gradient pI 6-11 range at low OD_{600} and at DO^{low} compared to that at low OD_{600} but at DO^{high} (Fig.5.16, Table 5.8). Spot 1784, previously identified to be OmpW, was 2.8-fold more abundant (*t*-test=0.034) at low OD_{600} for the state of the cells at DO^{low} relative to DO^{high} (Table 5.8, Fig. 5.16). Once again, peptide matches to this protein confirmed that OmpW identified in spot 1784 was the same protein identified in spot 2036 on the pI 4-7 gel (Section 5.2.1.5.7 and Fig. 5.14).

PhaP and PhbB were once again observed to be abundant at the same pI and M_r in two spots (Fig. 5.16) as in the profile of PHB synthesis proteins observed in the linear gradient pH 6-11 range for low OD_{600} or high OD_{600} at DO^{low} or DO^{high} with pH controlled (Figs. 5.14). Spot 1 had high-scoring peptide matches to PhaP. Spots 2 and 3 had high-scoring peptide matches to PhbB. The peptides matching spot 1 (Fig. 5.16) matched the entire PhaP sequence predicted from the $Y88^T$ genome which included peptide matches across the entire N-terminus

Spot number	Protein ID	Do^{low}/DO^{high}		Highest scoring unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
1317	NI	-2.88	0.023**	–	–	–	–	–	–
1318	Adenosylhomocysteinase (S-adenosylhomocysteine hydrolase, AdoHcy)	-2.39	0.011*	K.VACVAGFGDVGK.G(80)	50877	788	38	5.33	20
				R.DEFGPTQPLK.G(47)	–	–	–	–	–
1326	Adenosylhomocysteinase (S-adenosylhomocysteine hydrolase, AdoHcy)	-2.52	0.0027**	R.TANLILDDGGDATMFALWGAR.V(138)	50877	14979	631	5.33	66
				R.VVVTEIDPICALQAAMEGYEVVTMEEAVKR.A(92)	–	–	–	–	–
				R.VEAGETLPEPANAEEIEFQR.A(97)	–	–	–	–	–
2038	OmpW	2.17	0.0019*	R.AEIAIAETEMPGLMALR.D(87)	25096	851	16	9.1	10
2136	OmpW	2.14	0.017**	K.LLATGVLADGNIDTVR.S(94)	25096	251	6	9.1	4
				K.TPYSLTFDAK.K(71)					
2166	Alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen (AhpC)	-2.01	0.02*	K.ATAEGNEAVQSADYFAGK.K(83)	16726	465	21	4.84	23
				M.TIAVGDKLPDVK.L(62)	–	–	–	–	–

Table 5.7 2-D DIGE spot analysis (pI 4-7) and LC-MS/MS of $Y88^T$ protein from cells grown to a low OD_{600} under ammonium-depleted conditions at DO^{low} or DO^{high} at constant pH 7.

2-D DIGE protein spots showing differential abundance under ammonium-depleted conditions at DO^{low} or DO^{high} concentrations, constant pH and low OD_{600} (0.5). For those spots denoted by * $p<0.01$ otherwise $p<0.05$ (**). Positive values represent increased abundance of proteins under DO^{low} conditions and negative values represent increased abundance of proteins under DO^{high} conditions.

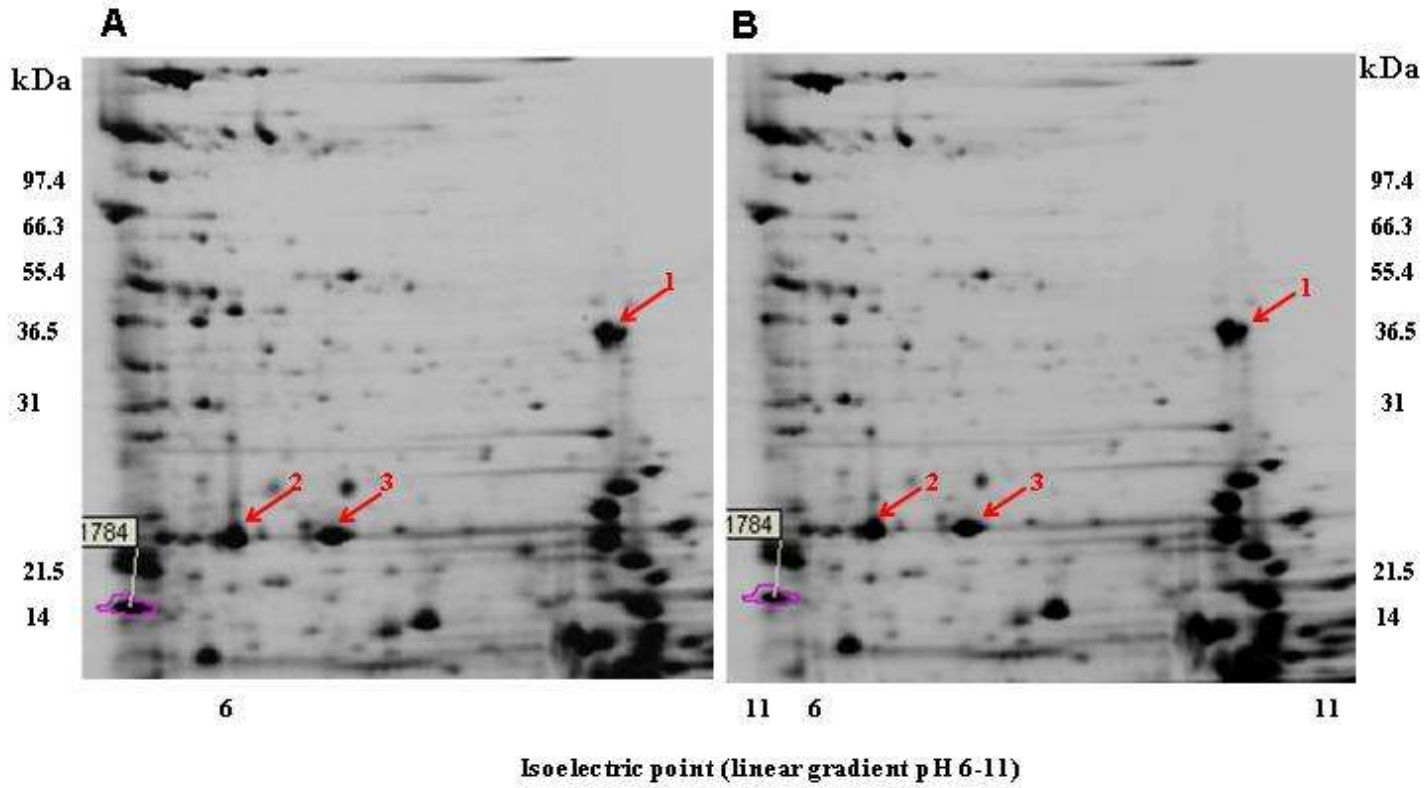


Figure 5.16 2-D DIGE images for the linear gradient pI 6-11 range showing protein profiles of Y88^T cells grown to a high OD_{600} under DO^{low} or DO^{high} conditions at constant pH 7.

Cells were grown to an OD_{600} of approximately 2.0 under DO^{low} (A) or DO^{high} (B) conditions Spot numbers (excluding those marked with red arrows) represent proteins determined by BVA analysis to have significantly altered abundance between the two DO concentrations. Spots marked by red arrows represent PHB synthesis proteins PhaP (spot 1) and two potential forms of PhbB with different pIs (spots 2, 3), (n=3).

Spot number	Protein ID	Do^{low}/DO^{high}		Highest scoring unique peptides matched (peptide score)	M_r	Score	Queries matched	pI	Sequence coverage (%)
		Average Ratio	*t*-test value						
1784	OmpW	2.18	0.034**	K.LLATGVLADGNIDTVR.S(94)	25096	851	16	9.1	10
				K.TPYSLTFDAK.K(71)	–	–	–	–	–

Table 5.8 2-D DIGE protein spot analysis (pI 6-11) and LC-MS/MS analysis of $Y88^T$ protein from cells grown to a low OD_{600} under ammonium-depleted conditions at DO^{low} or DO^{high} at constant pH 7

2-D DIGE protein spot analysis showing statistical values for proteins determined to be differentially abundant between DO^{low} or DO^{high} concentrations at constant pH and low OD_{600} (0.5). For those spots denoted by *; $p<0.01$ otherwise $p<0.05$. Positive values represent increased abundance of proteins under DO^{low} conditions and negative values represent increased abundance of proteins under DO^{high} conditions.

and C-terminus. The peptides matching the N-terminus segment of this high M_r form of PhaP were unique to this spot and were never matched in the lower M_r form at the more acidic pI. The relevance of these findings is discussed in Chapter 6.

5.3 Summary

The proteomic signatures described for $Y88^T$ nitrogen fixation and PHB accumulation confirmed that there was no coordinated regulation of PHB and nitrogen-fixing proteins in $Y88^T$. The microbiology of $Y88^T$ confirms that $Y88^T$ produces copious amounts of PHB irrespective of DO concentrations, changing nitrogen concentrations in the culture supernatant, pH conditions, and assayable acetylene reduction assay or an alternative mechanism for nitrogenase activity.

Concomitantly, PHB synthesis proteins and phasins occurred in high abundance on all gels irrespective of the growth conditions or whether cells were at low OD_{600} or high OD_{600}. Since these proteins are among the most abundant proteins on the gel at all times, core metabolic activities appear to support PHB production regardless of the growth conditions. Consistent with this, essentially all the proteins that show unaltered abundance are associated with core carbon metabolism. Those that that did show altered abundance, such as the OmpA or OmpW proteins, may be involved in specialised activities associated with membrane changes and/or iron regulation that may be heightened under conditions conducive to nitrogen fixation. Others, such as AdoHcyase, may have increased regulated activity related to stress under higher DO conditions.

The proteins involved in nitrogen fixation were all abundant at DO^{low} but showed decreased abundance at DO^{high}. Consistent with the nitrogenase enzyme's lability under high oxygen concentrations, in $Y88^T$, the different nitrogenase components were present at extremely low abundance at DO^{high}, which correlates with the decreased ethylene peak observed at DO^{high} (Fig. 5.1D). Thus, it cannot be ruled out that there is a residual level of nitrogenase

activity in these cells. Whether any possible remaining nitrogenase activity is sufficient to sustain very low levels of nitrogen fixation is not certain. Nonetheless, despite this decrease in nitrogenase activity, $Y88^T$ apparently continued to accumulate similar amounts of PHB and fixed nitrogen irrespective of DO concentrations. Interestingly, examination of the $Y88^T$ genome revealed that $Y88^T$ possesses a gene encoding the NifW protein that is considered to be a part of the oxygen-protection complex that may involve the reversible inactivation of nitrogenase via a conformational change, thereby protecting the enzyme from oxygen (Moshiri et al., 1995; Kim & Burgess, 1996; Lee et al., 1998; Ureta & Nordlund, 2002); however, this gene product was never identified proteomically. Irrespective of the mechanism that may result in a low level of nitrogenase at DO^{high} concentration, nitrogen levels are no different relative to those at DO^{low}, nor is PHB production limited. This is evident despite the fact that at DO^{high}, $Y88^T$ uses glucose less efficiently although the increased consumption of glucose may be part of a more efficient oxygen protective mechanism in promoting increased respiration (Post et al., 1983) and subsequent carbon dioxide production.

Chapter 6: Bioinformatic analysis of post-translational modifications and primary and secondary structure of selected Y88^T nitrogenase and PHB synthesis proteins

6.1 Introduction

Proteins can be post-translationally modified in a large number of ways. These modifications can be identified using various mass spectrometric techniques (Larsen & Roepstorff, 2000; Reinders et al., 2004; Rosen et al., 2004; Ekman et al., 2008; Philmus et al., 2008). Generally, a difference in a protein's observed M_r and/or pI on 2-D gels relative to its predicted M_r and/or pI is suggestive of one or more potential modifications. For instance, proteins migrating at lower than expected M_r are often the result of truncation of their C- or N-termini. Likewise, pI shifts of protein spots to more acidic isoelectric points on 2-D gels can be caused by protein phosphorylation.

The identification of the nitrogenase and phasin proteins of Y88^T on 2-D gels with unpredicted migration positions (Chapter 5) suggested that these proteins may be modified in a way that changed their isoelectric points and their molecular weights, resulting in mobility shifts. Given the differences in migration patterns of the Y88^T NifH, NifD and phasin proteins from *in silico* predictions based on the Y88^T genome sequence, it was hypothesised that there may be primary and/or secondary structure differences that may explain these anomalies. In an attempt to identify any possible modification(s) in the Y88^T nitrogenase and phasin that could explain the 2-D gel migration anomalies, their sequences were compared to orthologous proteins in a phylogenetic analysis to map known modification sites from other species that were potentially conserved in Y88^T. MS peptide data were then examined for evidence of similar modifications in these Y88^T proteins. The significance of the potential modifications in terms of the biology of Y88^T was also considered.

Analysis of the protein sequences of the Y88^T nitrogenase and phasin determined that: i) NifK (nitrogenase component I β-subunit), NifD (nitrogenase component I α-subunit) and NifH (nitrogenase component II) contain highly conserved sequences, including amino acid residues known to be integral to Fe-S cluster assembly and thought to be the site of nitrogenase reduction during nitrogen fixation; ii) a peptide match from the NifH protein observed at higher M_r and pI than the predicted Y88^T NifH protein under acidic growth conditions was a sequence known to be highly conserved and modified post-translationally in other nitrogen-fixing bacteria (Section 6.2.1.1) (Pope et al., 1985; Ekman et al., 2008); iii) the Y88^T NifH may be susceptible to a different type of modification to those generally observed in other nitrogen-fixing bacteria; iv) the Y88^T phasin has a low complexity, 199 amino acid alanine- and proline-rich N-terminal segment that is not found in phasins of most other PHA-producing bacteria, but is similar to four other sphingomonad phasins, suggesting that this may be a distinguishing feature of this family (Section 6.2.2.1); v) the Y88^T phasin may be modified post-translationally, resulting in cleavage of the 199 amino acid N-terminus as well as two potentially modified forms with lower M_r and different pIs to that of the predicted Y88^T phasin; vi) the Y88^T phasin has several amphipathic α-helices in its C-terminal half, a feature that has been identified in other phasins and is suggested here to be the reason why no common PHA-binding domain has been reported in phasins generally.

6.2 Results

6.2.1 Amino acid sequence analysis of the nitrogenase protein of Y88^T

6.2.1.1 NifH

A CLUSTALW (1.83) multiple sequence alignment of the Y88^T NifH protein sequence with 20 NifH sequences from other bacteria, (Table 6.1, Fig. 6.1), showed highly conserved regions spanning the entire sequence among all the NifH sequences examined. The most highly conserved portion of the NifH sequence, CVESGGPEPGVGCAGR (Fig. 6.1, amino acid sequence

Accession number	Species	Truncated species name as it appears in NifH sequence alignment
ref|YP_001600720.1|	*Gluconacetobacter diazotrophicus* PAl 5	Gluco
ref|NP_659736.1|	*Rhizobium etli* CFN 42	Ensifermexicanus
emb|CAD31326.1|	*Mesorhizobium* loti	Meso
gb|ACF19771.1|	*Aminobacter* sp. BA135	Aminobactersp
ref|NP_659736.1|	*Rhizobium etli* CFN 42	Rhizobiumsp
gb|ACF19777.1|	*Mesorhizobium* sp. ML105	Meso2
gb|ACM01736.1|	*Rhodobacter sphaeroides* KD131	Rhodobacterspaeroides
gb|AAB02342.1|	*Azospirillum brasilense*	Azosp
gb|ABG74606.1|	*Sinorhizobium fredii*	Sinorhizobiumfredii
gb|AAG60754.1|AF322012_59	*Bradyrhizobium japonicum* USDA 110	Bradylia
gb|ACF15255.1|	*Bradyrhizobium yuanmingense*	Bradyyuan
gb|ACF15253.1|	*Bradyrhizobium elkanii*	Bradyrhizobiumelkanii
gb|ACI23531.1|	*Bradyrhizobium canariense*	Bradcana
gb|ACI46147.1|	*Azorhizobium doebereinerae*	Azorhizobiumdoebereinerae
NA		Y88
gb|AAT06095.1|	*Burkholderi*a sp. Br3469	Burkholderiasp
dbj|BAE71134.1|	*Sphingomonas azotifigens*	Sphinogazob
gb|ABP79021.1|	*Pseudomonas stutzeri* A1501	Pseudomonasstut
gb|ABU49158.1|	*Azotobacter vinelandii*	Azoto
emb|CAD56230.1|	*Paenibacillus durus*	Paenibacillus
gb|AAX73206.2|	*Paenibacillus massiliensis*	Paenibacillusmaa

Table 6.1 Bacterial species names and the corresponding abbreviated names in the NifH alignment with Y88^T (Fig. 6.1).
No accession number (NA) yet for Y88^T. *P. durus* and *P. massiliensis* (dinitrogenase MoFe-alpha component) are outgroup sequences.

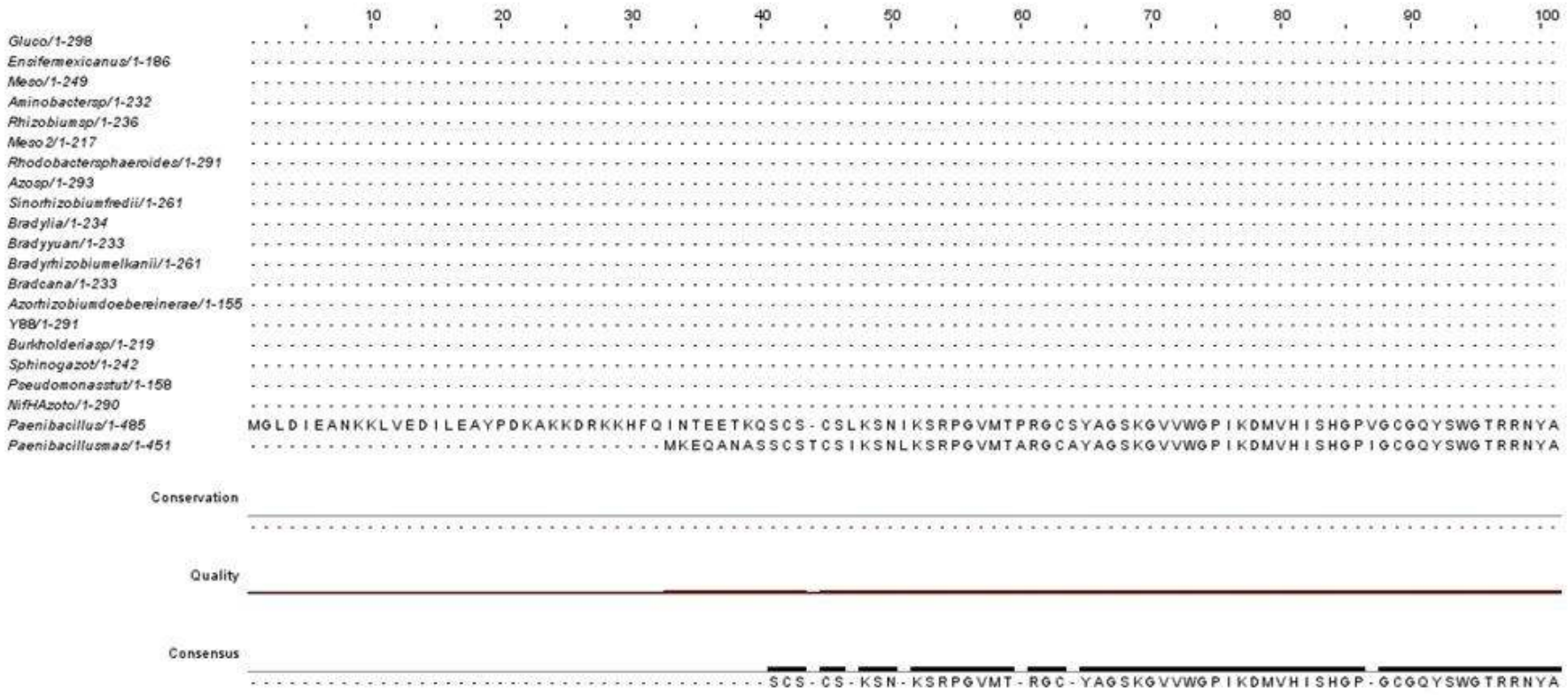


Figure 6.1 CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} NifH protein sequence with 20 NifH sequences from other bacteria (Table 6.1)

The most highly conserved region of NifH (red box, amino acids 231 to 246 on page 193) and other high-scoring peptides matching the Y88^{T} NifH protein (all other boxes on pages 192 to 194) are shown. *P. durus* and *P. massiliensis* (dinitrogenase MoFe-alpha component) are outgroup sequences. The degree of amino acid conservation across all NifH sequences are shown with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio). Figure 6.1 continued on the following page. See Table 6.1 for full species names.

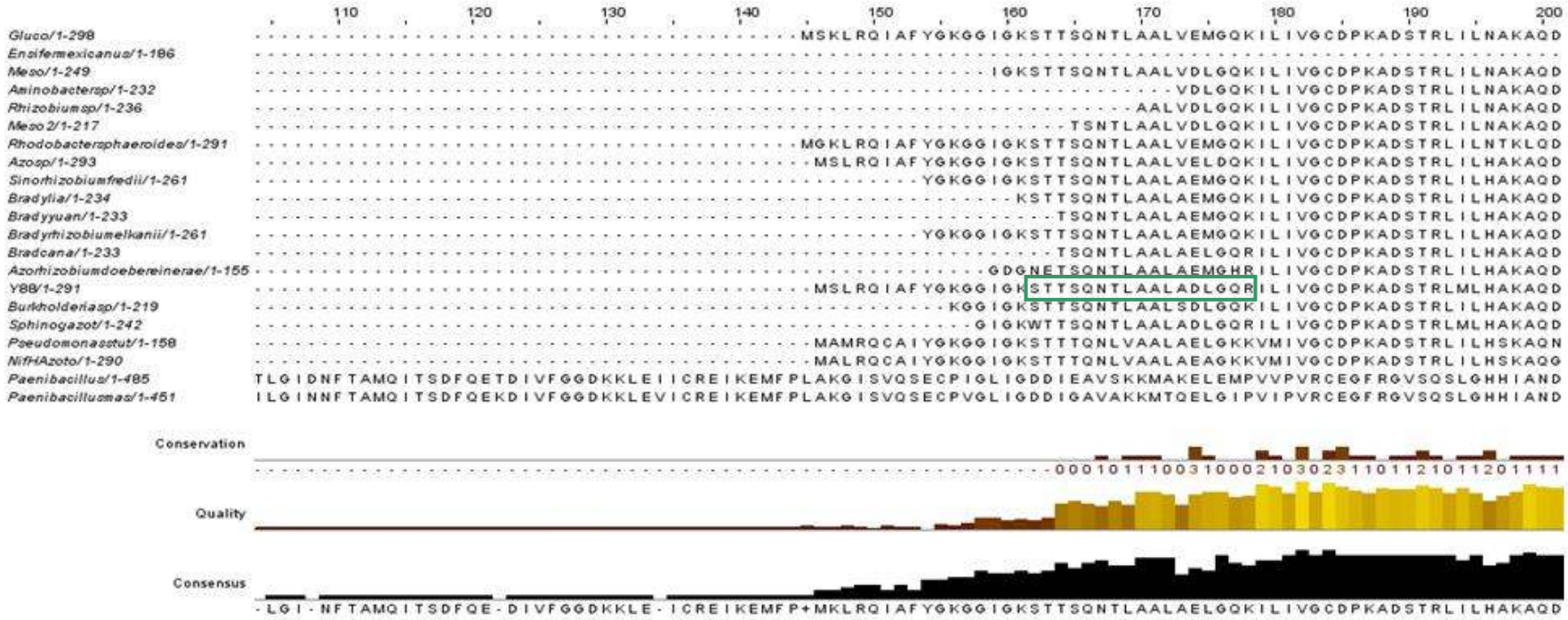


Figure 6.1 (cont.) CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} NifH protein sequence with 20 NifH sequences from other bacteria (Table 6.1)

An N-terminal peptide match to the Y88 NifH protein (green box, amino acids at positions 162-178) is shown. *P. durus* and *P. massiliensis* (dinitrogenase MoFe-alpha component) are outgroup sequences. The degree of amino acid conservation across all NifH sequences are shown with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio). Figure 6.1 continued on the following page. See Table 6.1 for full species names.

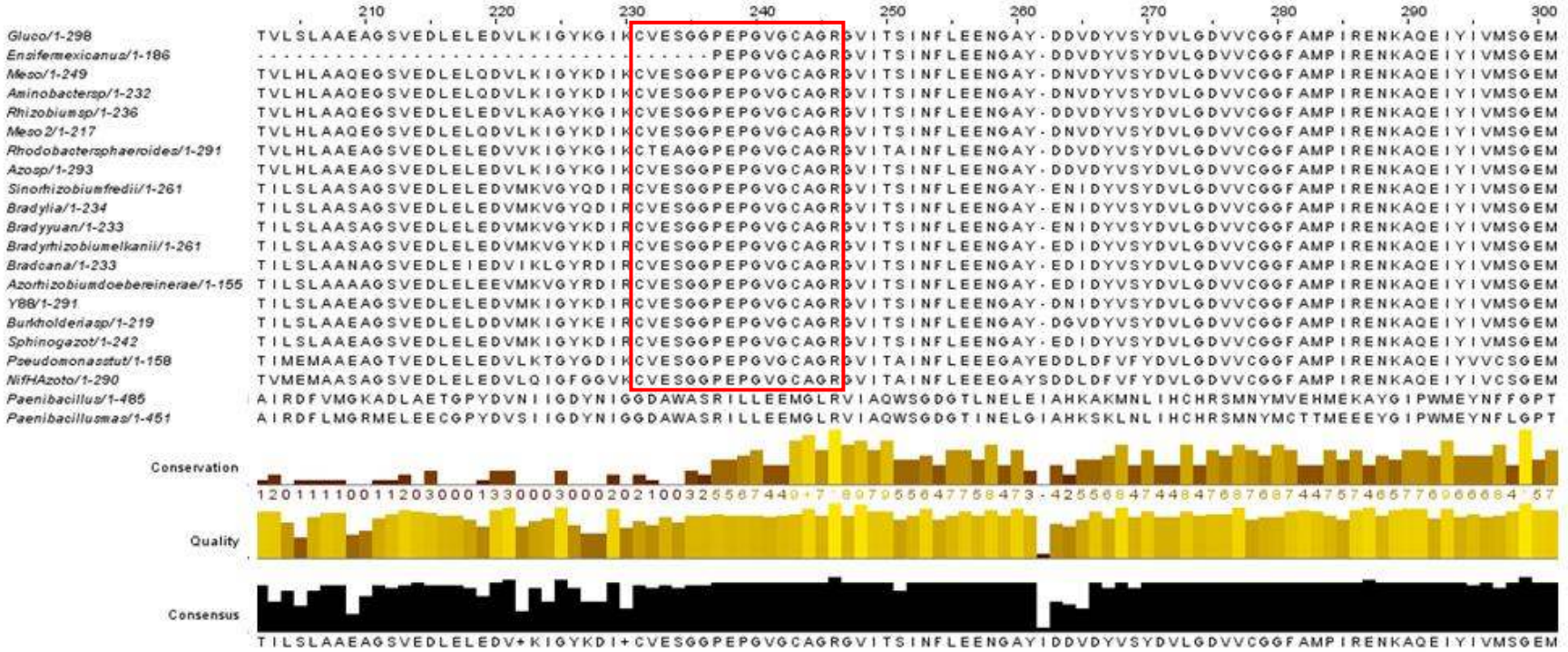


Figure 6.1 (cont.) CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} NifH protein sequence with 20 NifH sequences from other bacteria (Table 6.1)

The most highly conserved region of NifH (red box, peptide sequence corresponding to amino acid positions 231 to 246) with the amino acid at position 246 in this alignment corresponding to arginine 101 of *A. vinelandii* and arginine 101 of Y88^{T}). *P. durus* and *P. massiliensis* (dinitrogenase MoFe-alpha component) are outgroup sequences. The degree of amino acid conservation across all NifH sequences are shown with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio). Figure 6.1 continued on the following page. See Table 6.1 for full species names.

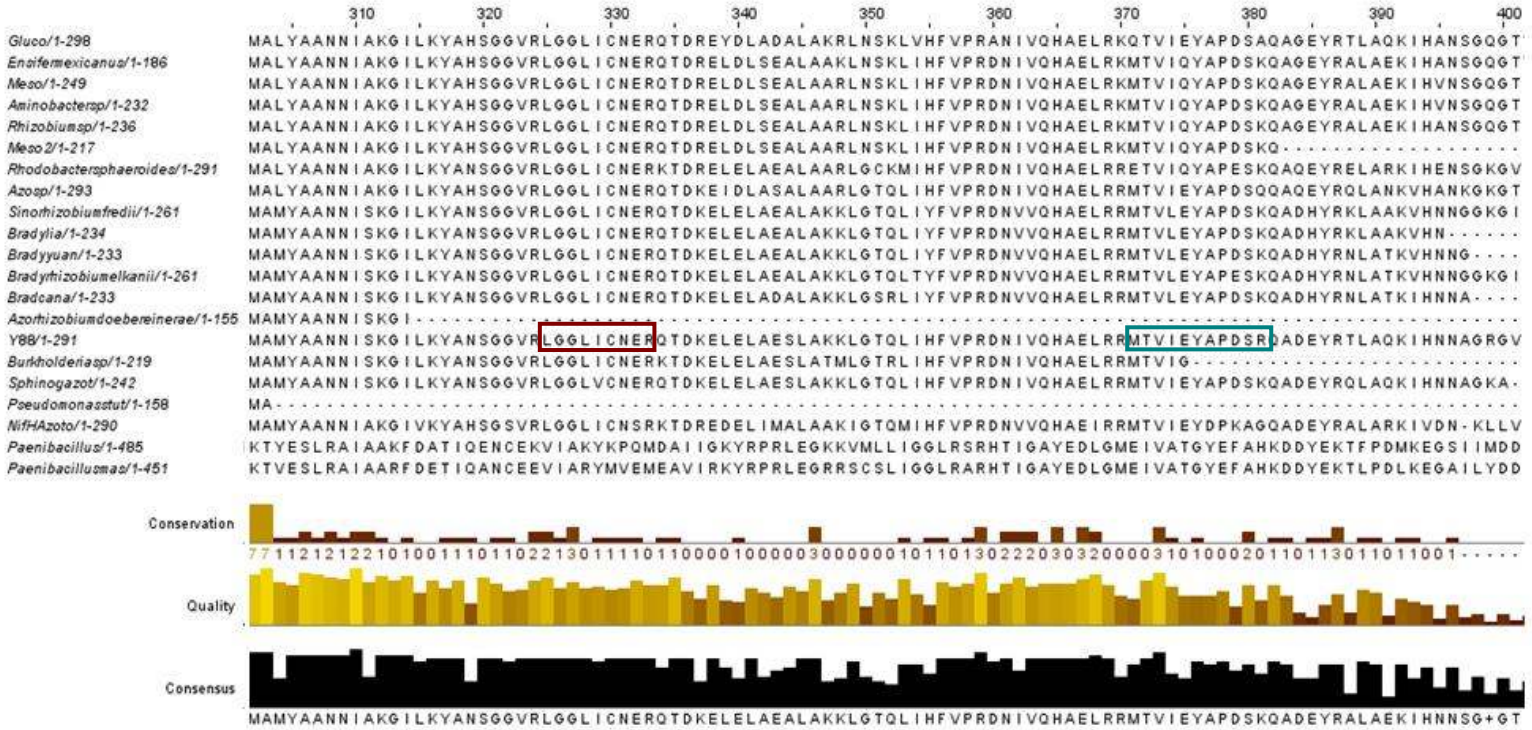


Figure 6.1 (cont.) CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} NifH protein sequence with 20 NifH sequences from other bacteria (Table 6.1)

Peptides matching to the Y88^{T} NifH protein (brown box, amino acid sequence corresponding to positions 325 to 333 and green box, amino acid sequence corresponding to positions 371 to 381). *P. durus* and *P. massiliensis* (dinitrogenase MoFe-alpha component) are outgroup sequences. The degree of amino acid conservation across all NifH sequences are shown with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio). Figure 6.1 continued on the following page. See Table 6.1 for full species names.

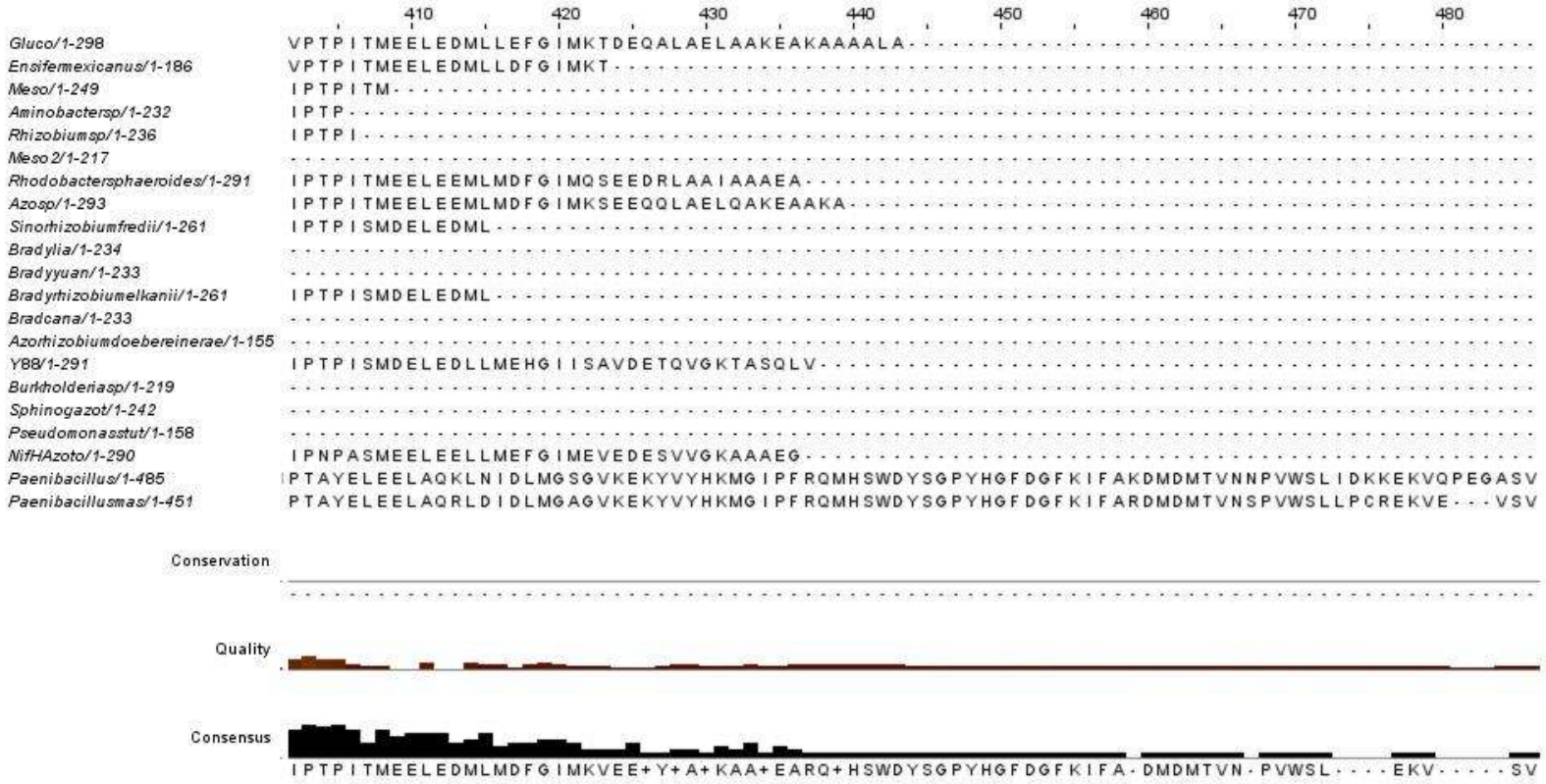


Figure 6.1 (cont.) CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} NifH protein sequence with 20 NifH sequences from other bacteria (Table 6.1)

C-terminus amino acids are shown to ensure complete sequences of the relevant species are represented in the alignment. *P. durus* and *P. massiliensis* (dinitrogenase MoFe-alpha component) are outgroup sequences. The degree of amino acid conservation across all NifH sequences are shown with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio). See Table 6.1 for full species names.

corresponding to positions 231 to 246 in this alignment, with the C-terminus arginine known to be susceptible to ADP-ribosylation, at position 246), was shown to be positioned close to the active site of nitrogenase (Georgiadis et al., 1992; Ekman et al., 2008). This conserved segment was of particular interest because coincidentally it corresponded perfectly to the highest-scoring peptide matched in the Y88^{T} NifH protein in the MS analysis (Chapter 5, Table5.1, Chapter 6, Fig. 6.1, red box at amino acid positions 231 to 246).

This peptide sequence is reported to be perfectly conserved in most sequenced nitrogenase proteins (Ekman et al., 2008), which was found to be the case for all NifH sequences examined in this study, with the exception of *Paenibacillus* sp. (Fig. 6.1, last two sequences in the alignment). In Y88^{T}, the Arg101 at the C-terminus of this peptide corresponds to Arg100 in *A. vinelandii* (Li, 2002) as well as Arg101 in *Rhodospirillum rubrum* and is the site of activity-regulating ADP-ribosylation in *R. rubrum* (Pope et al., 1985) and many other (Ekman et al., 2008), but not all (Durner et al., 1994), nitrogen-fixing bacteria.

Error-tolerant spectral analysis of unassigned candidate peptides matched to CVESGGPEPGVGCAGR in Mascot suggested several potential modifications consistent with the mobility shifts observed for NifH. One such modification biochemically identified in NifH is palmitoylation (Gallon et al., 2000). Theoretically, the peptide carrying a modification would have a higher mass in the modified protein than the peptide in the unmodified protein and the mass difference would allow identification and location of the modification (Ekman et al., 2008). Based on the molecular weight and pI differences of the two NifH forms (Chapter 4, Fig. 4.5, spots 3198, 3255; Chapter 5, Fig. 5.5, spots 2101, 2129), modification of CVESGGPEPGVGCAGR by acylation was considered. Despite the approximately 500 Da difference between the two forms corresponding to two possible palmitoylations (each equal to 238 Da), an error-tolerant search did not detect this type of modification on CVESGGPEPGVGCAGR in Y88^{T}. In bacteria, palmitoylation is usually observed only on lysine and on N-terminal cysteines following post-translational signal peptide cleavage (Ekman et al.,

2008). Since the N-terminal cysteine in CVESGGPEPGVGCAGR is the result of a trypsin digestion, not a signal peptide cleavage, the modification of the N-terminal cysteine on CVESGGPEPGVGCAGR is unlikely. The possibility that this modification does exist on another peptide or in combination with another modification (Gallon et al., 2000), also cannot be ruled out since not all peptides spanning the entire Y88^{T} NifH sequence were matched.

The potential difference in M_r between the two NifH forms does correspond to one possible ADP-ribosylation (equal to 541 Da). That Arg101 in Y88^{T} may therefore be ADP-ribosylated also cannot be ruled out, since this amino acid was never accounted for in the spectrum for this peptide. Two possible reasons for the unaccounted Arg101 were considered. First, since the enzyme trypsin was used to digest the protein during protein preparation for MS and cleaves C-terminal to arginine (or lysine), any modification of Arg101 would potentially block this cleavage, resulting in a peptide with a much higher molecular weight (Ekman et al., 2006) than that observed in the spectrum. Secondly, the mass of arginine is 174 Da and the MS parameter set limits for peak detection are between 200 and 2000 Da. Arginine may therefore not be accounted for in the spectrum (for ion trap data) since its mass is below the mass cutoff and is outside this m/z range of interest.

Despite the lack of detection of the arginine in the CVESGGPEPGVGCAGR spectrum, it is nonetheless tempting to rule out this modification for the high M_r form of NifH protein because ADP-ribosylation inhibits nitrogenase activity (Pope et al., 1985) yet nitrogen fixation was demonstrated by Y88^{T} (in an acetylene reduction assay). The higher M_r more basic form may represent a portion of the NifH pool that may be modified in an alternative way to that of ADP-ribosylation, but may have a similar effect rendering it catalytically inactive under aerobic conditions of growth (Gallon et al., 2000). This was suggested to be the case in the *Azolla* cyanobiont (Ekman et al., 2008) where approximately 25% of the NifH pool was suggested to be modified by means of ADP-ribosylation and consequently inactive, as well as possibly targeted for

degradation. An alternative type of modification of NifH would explain the presence of the higher M_r form for the nitrogen-fixing state of the cell but further experimentation is required to establish if this is the case. In $Y88^T$, any potential modification may be a consequence of the lower pH of the medium, since this higher M_r more basic form is absent from the pH 4-7 gels representing the Fix^{on} state of the cells at DO^{low} and a neutral growth medium.

An error-tolerant search of the high-scoring peptide STTSQNTLAALADLGQR (Fig. 6.1, green box) revealed a possible modification of this peptide, resulting in a higher mass. However, the mass difference between the potentially modified and unmodified peptides did not account for the mobility shift of approximately 500 Da observed between the two NifH forms on the 2-D gels.

Although the NifH protein is modified by ADP-ribosylation in *R. rubrum*, (Pope et al., 1985) more recently other modifications such as palmitoylation, acetylation and other, as yet, unidentified modifications have been suggested to regulate NifH in other bacterial species (Gallon et al., 2000; Ekman et al., 2008). Based on these precedents, one must also consider that the $Y88^T$ NifH may be modified in yet a different way in response to changes in culture growth conditions that would explain the two different forms of NifH for the Fix^{on} state of the cells at DO^{low} and at a more acidic growth medium.

6.2.1.2 NifK and NifD proteins

The $Y88^T$ NifK and NifD protein sequences were examined in a CLUSTALW (1.83) multiple sequence alignment with 16 NifK (β-subunit) and NifD (α-subunit) sequences from other Alphaproteobacteria, as well as the γ-proteobacterium *A. vinelandii* as an outgroup (Fig. 6.2). *A. vinelandii* was chosen for comparison because its nitrogenase is well studied and understood biochemically (Dean et al., 1990; Kent et al., 1990; Peters et al., 1997; Li, 2002). Sequence analysis revealed that the $Y88^T$ NifK and NifD proteins were highly conserved relative to these 16 species. The conserved amino acid residues of interest are discussed below and highlighted in sequence alignments of the NifK and NifD proteins of $Y88^T$ with those of *A. vinelandii* (Fig. 6.2).

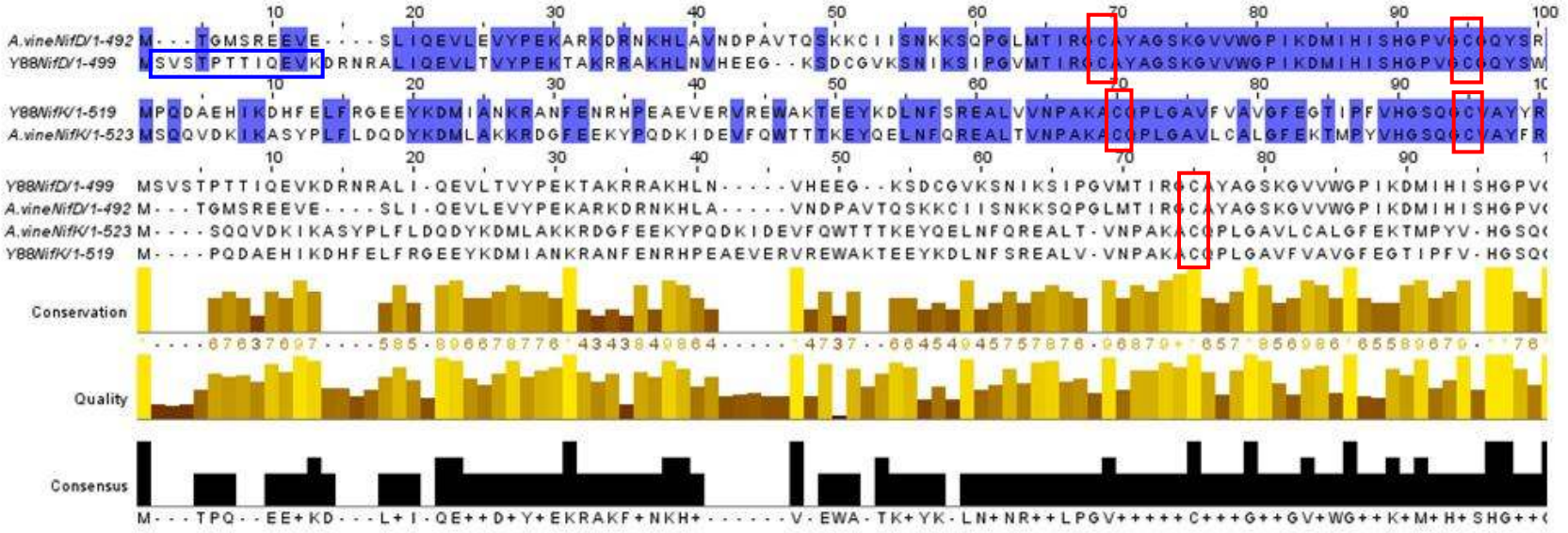


Figure 6.2 CLUSTALW (1.83) multiple sequence alignment of the Y88^T NifD and NifK protein sequences with the *A. vinelandii* NifD and NifK sequences

Highly conserved NifDK amino acids (red boxes) and peptides matching the Y88^T NifD protein (blue boxes) and the Y88^T NifK protein (green boxes) are shown. *A. vine*NifD/1-492 = *A. vinelandii* NifD protein sequence; *A.vine*NifK/1-523 = *A. vinelandii* NifK protein sequence; Y88NifD/1-499 = Y88^T NifD protein sequence; Y88NifK/1-519 = Y88^T NifK protein sequence. Blue shaded amino acids are conserved across NifD of *A. vinelandii* and Y88^T in the first two rows of sequences from top of the page and across NifK of *A. vinelandii* and Y88^T in the third and fourth rows of sequences from top of the page. The degree of amino acid conservation across NifDK for both *A. vinelandii* and Y88^T are shown for the bottom four rows of sequences with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Fig. 6.2 continued on the following page. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio).

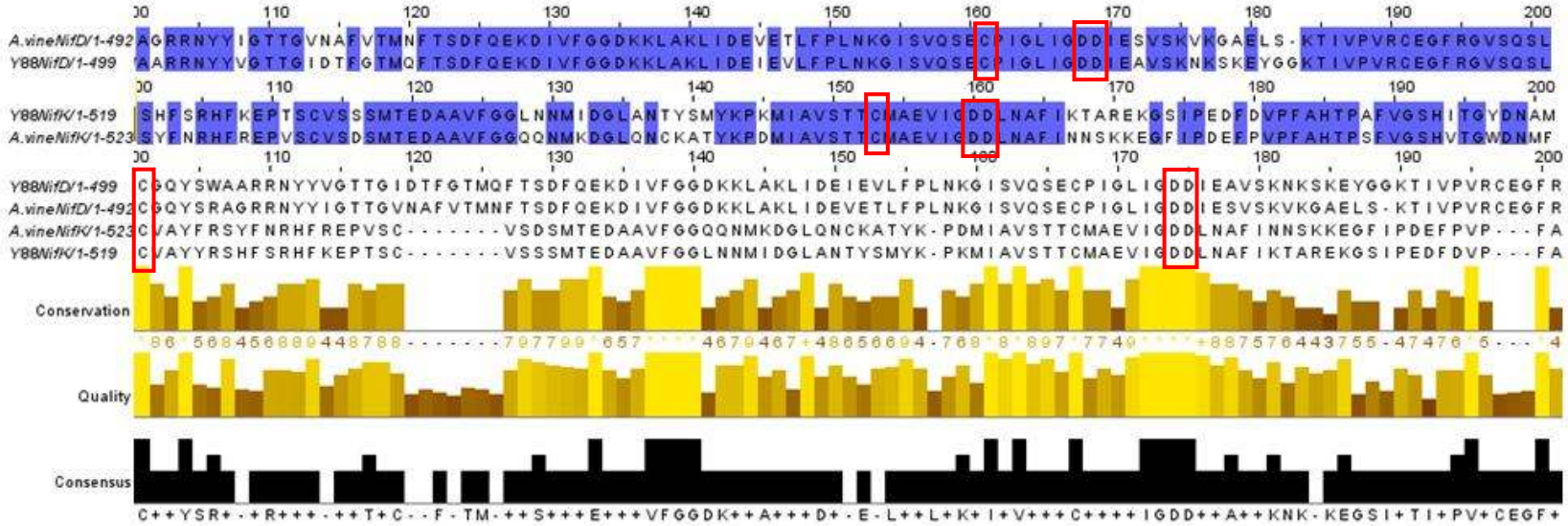


Figure 6.2 (cont.) CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} NifD and NifK protein sequences with the *A. vinelandii* NifD and NifK sequences

Highly conserved NifDK amino acids (red boxes) and peptides matching the Y88^{T} NifD protein (blue boxes) and the Y88^{T} NifK protein (green boxes) are shown. *A. vine*NifD/1-492 = *A. vinelandii* NifD protein sequence; *A.vine*NifK/1-523 = *A. vinelandii* NifK protein sequence; Y88NifD/1-499 = Y88^{T} NifD protein sequence; Y88NifK/1-519 = Y88^{T} NifK protein sequence. Blue shaded amino acids are conserved across NifD of *A. vinelandii* and Y88^{T} in the first two rows of sequences from top of the page and across NifK of *A. vinelandii* and Y88^{T} in the third and fourth rows of sequences from top of the page. The degree of amino acid conservation across NifDK for both *A. vinelandii* and Y88^{T} are shown for the bottom four rows of sequences with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Fig. 6.2 continued on the following page. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio).

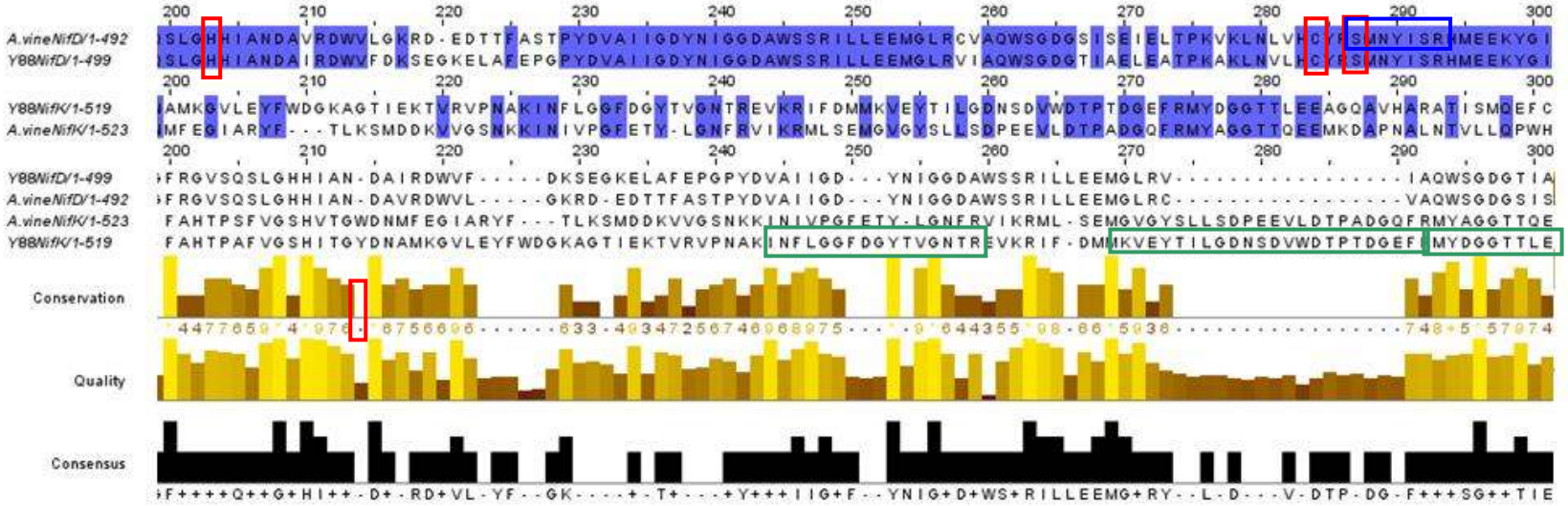


Figure 6.2 (cont.) CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} NifD and NifK protein sequences with the *A. vinelandii* NifD and NifK sequences

Highly conserved NifDK amino acids (red boxes) and peptides matching the Y88^{T} NifD protein (blue box) and the Y88^{T} NifK protein (green boxes) are shown. *A. vine*NifD/1-492 = *A. vinelandii* NifD protein sequence; *A.vine*NifK/1-523 = *A. vinelandii* NifK protein sequence; Y88NifD/1-499 = Y88^{T} NifD protein sequence; Y88NifK/1-519 = Y88^{T} NifK protein sequence. Blue shaded amino acids are conserved across NifD of *A. vinelandii* and Y88^{T} in the first two rows of sequences from top of the page and across NifK of *A. vinelandii* and Y88^{T} in the third and fourth rows of sequences from top of the page. The degree of amino acid conservation across NifDK for both *A. vinelandii* and Y88^{T} are shown for the bottom four rows of sequences with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Fig. 6.2 continued on the following page. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio).

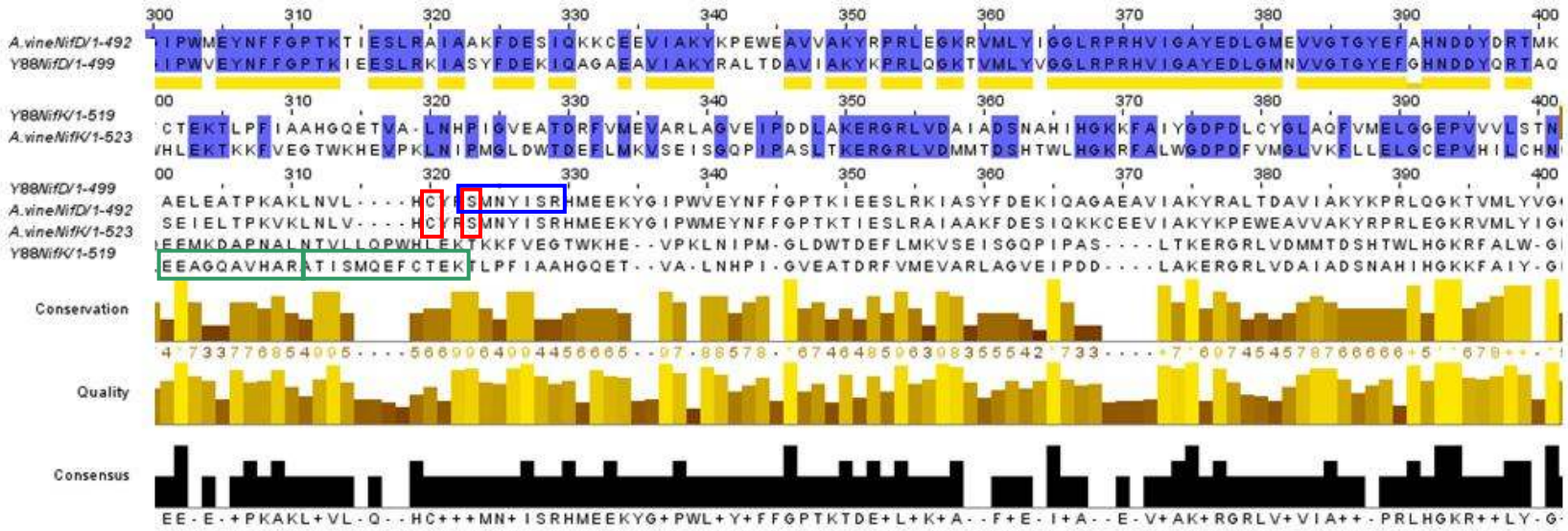


Figure 6.2 (cont.) CLUSTALW (1.83) multiple sequence alignment of the Y88^T NifD and NifK protein sequences with the *A. vinelandii* NifD and NifK sequences

Highly conserved NifDK amino acids (red boxes) and peptides matching the Y88^T NifD protein (blue box) and the Y88^T NifK protein (green boxes) are shown. *A. vine*NifD/1-492 = *A. vinelandii* NifD protein sequence; *A.vine*NifK/1-523 = *A. vinelandii* NifK protein sequence; Y88NifD/1-499 = Y88^T NifD protein sequence; Y88NifK/1-519 = Y88^T NifK protein sequence. Blue shaded amino acids are conserved across NifD of *A. vinelandii* and Y88^T in the first two rows of sequences from top of the page and across NifK of *A. vinelandii* and Y88^T in the third and fourth rows of sequences from top of the page. The degree of amino acid conservation across NifDK for both *A. vinelandii* and Y88^T are shown for the bottom four rows of sequences with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Fig. 6.2 continued on the following page. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio).

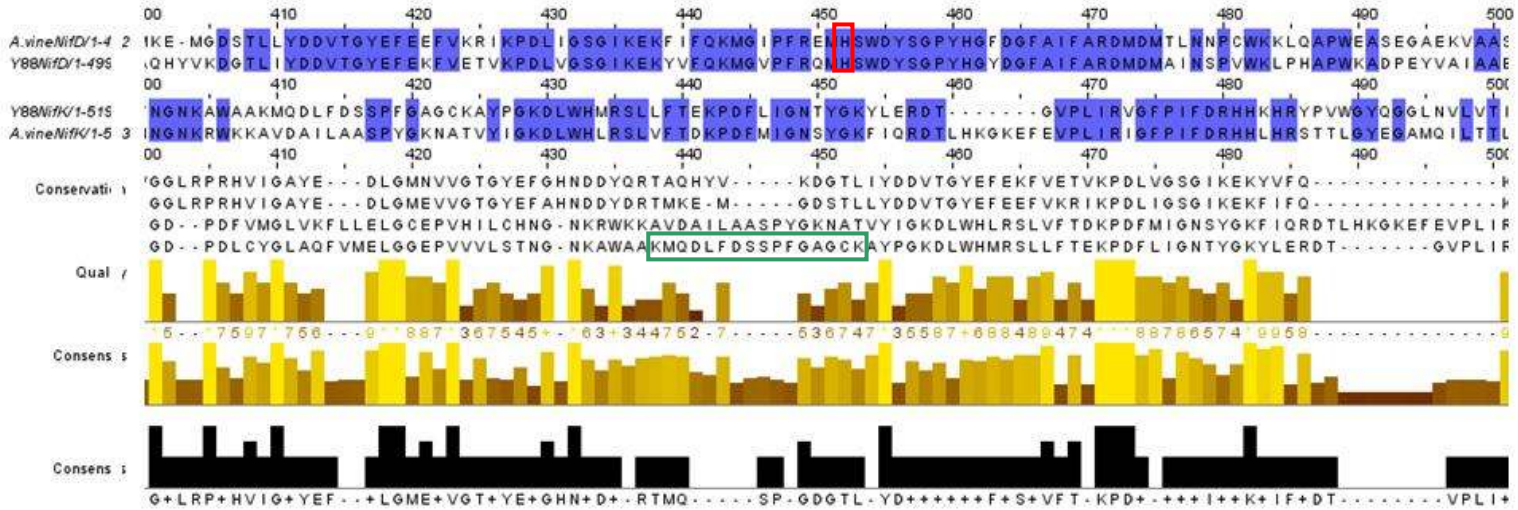


Figure 6.2 (cont.) CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} NifD and NifK protein sequences with the *A. vinelandii* NifD and NifK sequences Highly conserved NifDK amino acids (red boxes) and peptides matching the Y88^{T} NifD protein (blue boxes) and the Y88^{T} NifK protein (green boxes) are shown. *A. vine*NifD/1-492 = *A. vinelandii* NifD protein sequence; *A.vine*NifK/1-523 = *A. vinelandii* NifK protein sequence; Y88NifD/1-499 = Y88^{T} NifD protein sequence; Y88NifK/1-519 = Y88^{T} NifK protein sequence. Blue shaded amino acids are conserved across NifD of *A. vinelandii* and Y88^{T} in the first two rows of sequences from top of the page and across NifK of *A. vinelandii* and Y88^{T} in the third and fourth rows of sequences from top of the page. The degree of amino acid conservation across NifDK for both *A. vinelandii* and Y88^{T} are shown for the bottom four rows of sequences with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Fig. 6.2 continued on the following page. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio).

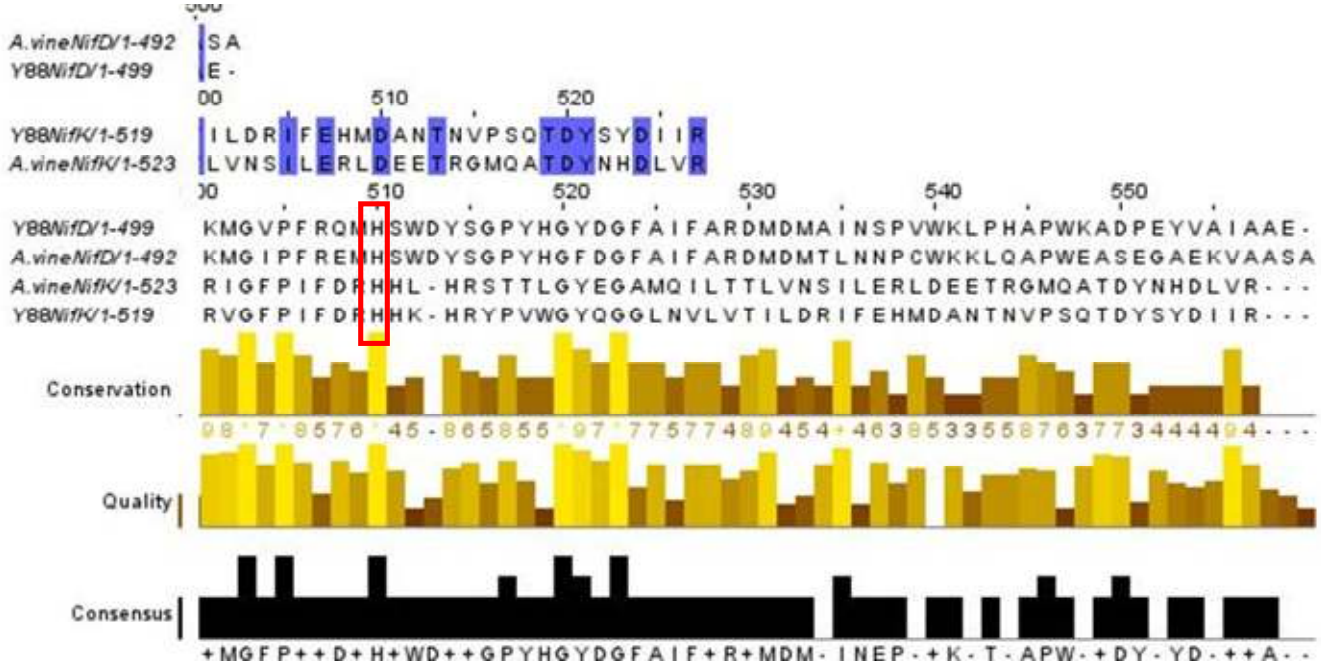


Figure 6.2 (cont.) CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} NifD and NifK protein sequences with the *A. vinelandii* NifD and NifK sequences showing the highly conserved key amino acids (red boxes) and peptides matching the Y88^{T} NifD protein (blue boxes) and the Y88^{T} NifK protein (green boxes). *A. vine*NifD/1-492 = *A. vinelandii* NifD protein sequence; *A.vine*NifK/1-523 = *A. vinelandii* NifK protein sequence; Y88NifD/1-499 = Y88^{T} NifD protein sequence; Y88NifK/1-519 = Y88^{T} NifK protein sequence. Blue shaded amino acids are conserved across NifD of *A. vinelandii* and Y88^{T} in the first two rows of sequences from top of the page and across NifK of *A. vinelandii* and Y88^{T} in the third and fourth rows of sequences from top of the page. The degree of amino acid conservation across NifDK for both *A. vinelandii* and Y88^{T} are shown for the bottom four rows of sequences with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio). Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio).

Amino acid residues corresponding to *A. vinelandii* α-subunit Cys275 and α-subunit His442 (amino acid positions 284 and 452 respectively in *A. vinelandii* NifD alignment with Y88^{T} but positions 320 (red box) and 510 (red box) respectively in alignment with both NifDK, Fig. 6.2) that are required for all FeMo protein activities, were identified as Y88^{T} αCys282 and αHis450 respectively (Fig. 6.2, red boxes at amino acid positions 284 and 510 respectively). Additionally, substitution of *A. vinelandii* αHis442 (amino acid position 452 in alignment with Y88^{T}) with Asn has been shown to completely inactivate MoFe activity and in *A. vinelandii*, αHis442 appears to play a major role in anchoring the cofactor to the NifD polypeptide (Brigle et al., 1985). *A. vinelandii* αHis195, αCys275, and αSer278 (Fig. 6.2, amino acid positions 203, 284 and 287 respectively) which are involved in proton delivery to the *A. vinelandii* FeMo cofactor (Pham & Burgess, 1993; Dilworth et al., 1998; Fisher et al., 2000), were identified as Y88^{T} αHis202, αCys282 (mentioned above) and αSer285 respectively (Fig. 6.2, red boxes at amino acid positions 203, 284 and 287 respectively). Also, *A. vinelandii* αAsp161, αAsp162, βAsp160 and βAsp161 (Fig. 6.2, amino acid positions 168, 169, 160, 161 respectively), reported to all be strictly conserved in the MoFe protein primary sequences of at least five nitrogen-fixing species (Peters et al., 1995) were conserved at sites αAsp166 and αAsp167 in the Y88^{T} NifD sequence and at βAsp160 and βAsp161 in the NifK predicted protein sequence (Fig. 6.2, red boxes at same amino acid positions 168, 169, 160, 161 respectively). Six *A. vinelandii* cysteine residues, αCys62, αCys88, αCys154 and βCys70, βCys95 and βCys153 Fig. 6.2, amino acid positions 69, 95, 161, 70, 95, 153 respectively), reported to be the sites through which the MoFe P-cluster is covalently coordinated to the MoFe protein in *A. vinelandii* (Li, 2002) were also conserved in Y88^{T} at αCys67, αCys93, αCys159, βCys70, βCys95 and βCys153, respectively (Fig. 6.2, red boxes at amino acid positions 69, 95, 161, 70, 95, 153 respectively).

Close inspection of all peptide matches to the Y88^{T} NifD and NifK protein sequences (Fig. 6.2, blue and green boxes) to determine if any peptides may

have been modified at the conserved amino acids mentioned above, which could either explain a pI shift or why the abundance of these proteins was altered at DO^{high} concentrations, showed that only one of these peptides, SMNYISR, spanned any of the conserved amino acids of interest in the NifD sequence (Fig. 6.2, blue box, amino acid positions 323 to 329). An error-tolerant spectral analysis of unmatched spectra did not account for any potential modification of this peptide that could explain a pI shift of the NifD protein to a more basic position on the 2-D gel. That NifD in particular may still be susceptible to post-translational modification cannot be ruled out, since only two peptides matched this protein (Fig.6.2; blue boxes amino acid positions 2 to 13 and 323 to 329) and only one of them, SMNYISR, as mentioned before, spanned any portion of the $Y88^T$ sequence that included any of the key amino acids identified from *A. vinelandii* to be conserved and essential to nitrogenase activity. The first serine in the peptide SMNYISR corresponds to αSer275 in *A. vinelandii* and to αSer278 in $Y88^T$ (Fig. 6.2, red box at amino acid position 323 in this sequence alignment).

The second peptide, SVSTPTTIQEVK (Fig. 6.2, amino acid positions 2 to 13), that was matched to NifD was identified in an error-tolerant search to potentially be modified either at the C-terminal lysine or the internal glutamine. Since the mass difference was 0.04 Da, this mass difference was below the set limit of the mass error tolerance, thus a modification of this amino acid could be ruled out.

6.2.2 Protein and peptide analysis of the $Y88^T$ phasin

6.2.2.1 Proteomics of the three $Y88^T$ phasin isoforms

The full-length $Y88^T$ phasin predicted from the $Y88^T$ genome sequence was observed at its predicted 34 kDa and pI of 9.6 on the basic gels (Chapter 5, spot 1, Figs. 5.10, 5.14) compared to the two "truncated" 13 kDa forms at acidic pIs (6.9 and 7.2) (Chapter 5, spots 1 and 2, Fig. 5.8). Several peptides that matched different parts of the alanine- and proline-rich N-terminal segment (Chapter 5, spot 1, Figs. 5.10, 5.14; Chapter 6, Fig. 6.3, black boxed sequences) were never

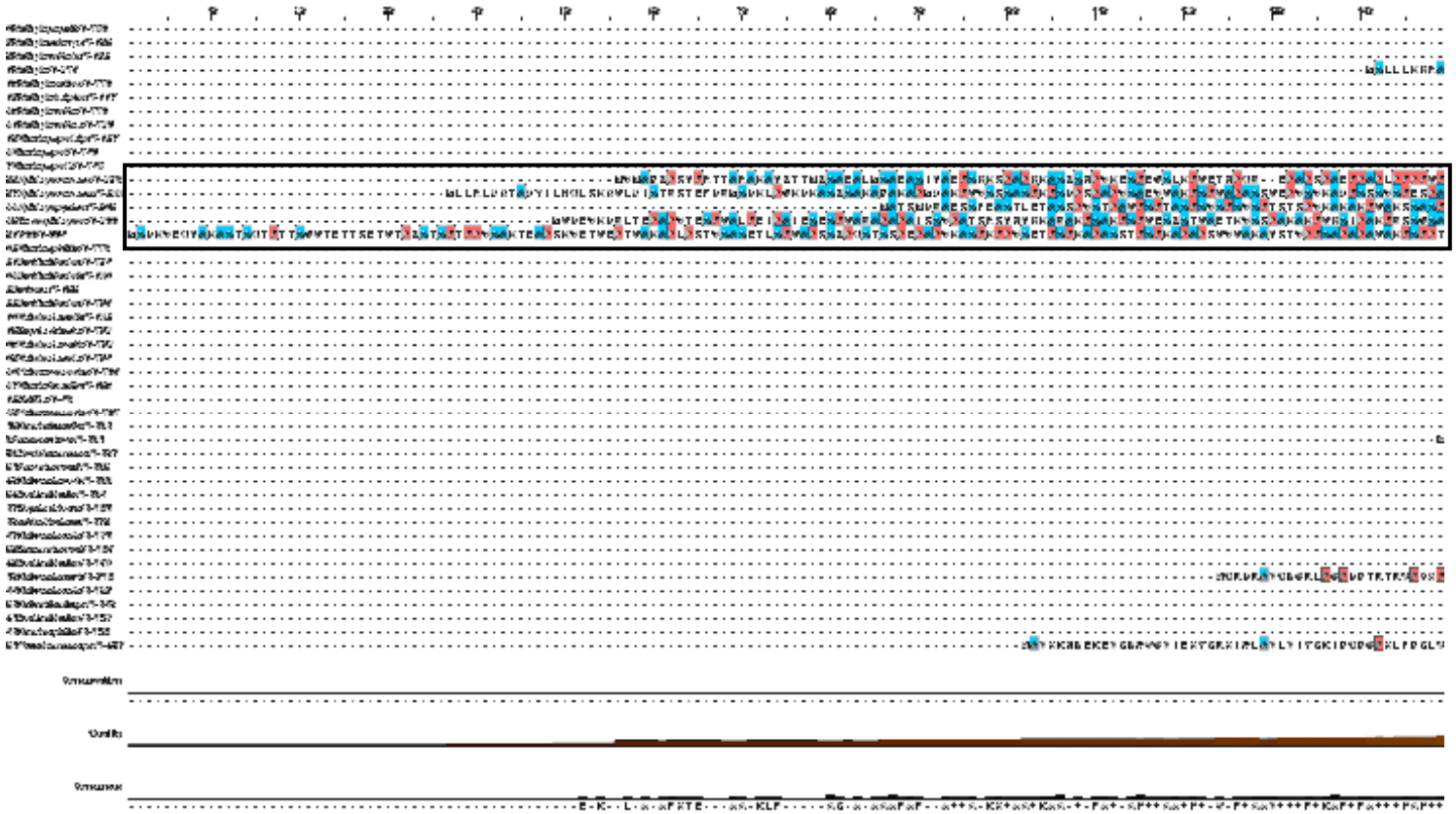


Figure 6.3 CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} phasin protein with representative phasin sequences randomly selected from Genbank (Table 6.2)

The highly conserved alanine/proline-rich N-terminal segment in the sphingomonadaceae (black box), the C-terminal segment in Y88^{T} for which 100% sequence coverage was obtained (red box) and the alanine/proline-rich C-terminus of *P. putida* (green box) are shown. Red and blue shaded colours represent alanine and proline residues. Figure 6.3 continued on the following page. The degree of amino acid conservation across all phasin sequences are shown with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio). See Table 6.2 for full species names.

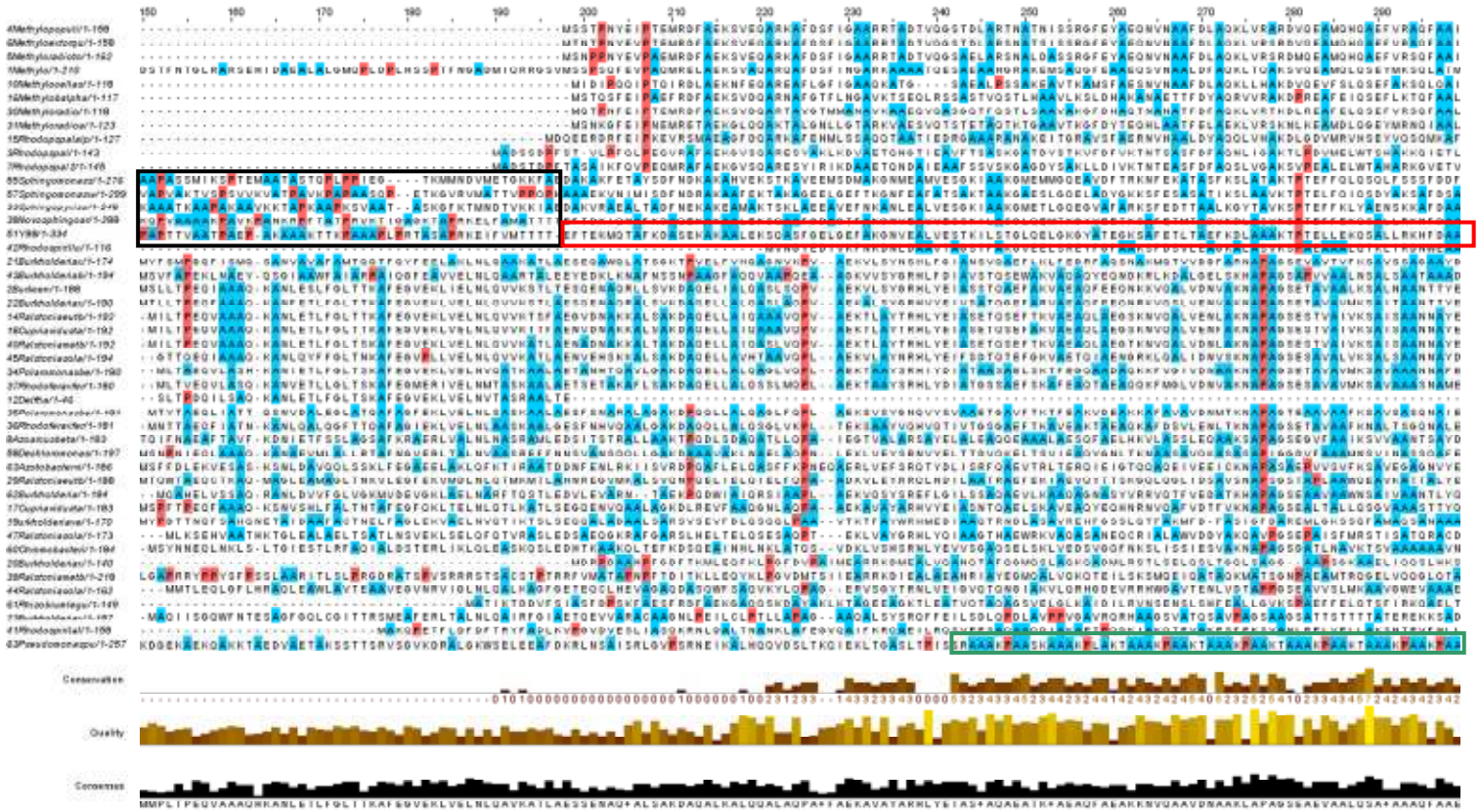


Figure 6.3 (cont.) CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} phasin protein with representative phasin sequences randomly selected from Genbank (Table 6.2)
The highly conserved alanine/proline-rich N-terminal segment in the sphingomonads (black box), the C-terminal segment in Y88^{T} for which 100% protein sequence coverage was obtained (red box) and the alanine/proline-rich C-terminus of *P. putida* (green box) are distinguishable from all other sequences. Red and blue shaded colours represent alanine and proline residues. Fig. 6.3 continued on the following page. The degree of amino acid conservation across all phasin sequences are shown with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio). See Table 6.2 for full species names.

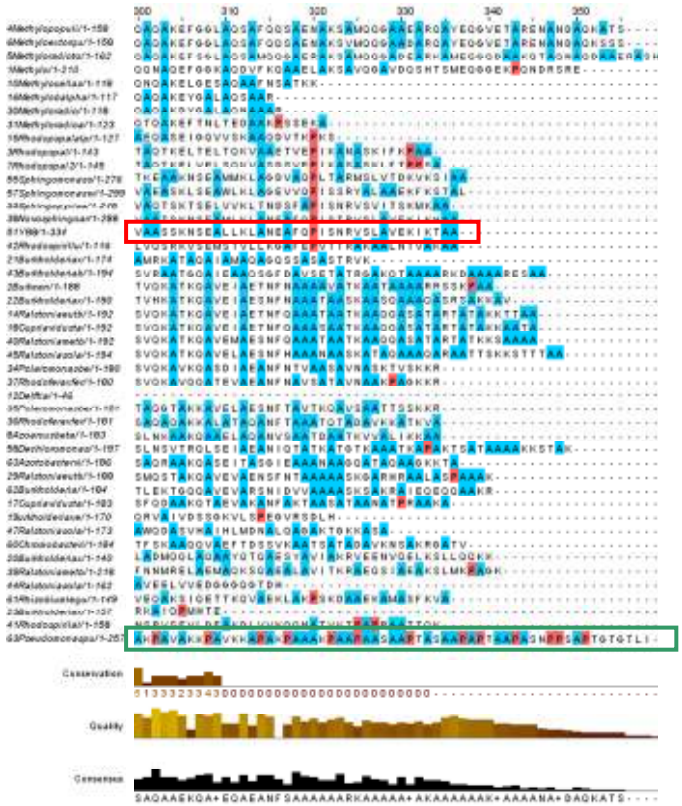


Figure 6.3 (cont.) CLUSTALW (1.83) multiple sequence alignment of the Y88^{T} phasin protein with representative phasin sequences randomly selected from Genbank (Table 6.2)
The highly conserved alanine/proline-rich N-terminal segment in the sphingomonads (black box previous page), the C-terminal segment in Y88^{T} for which 100% protein sequence coverage was obtained (red box continued from previous page to this page) and the alanine/proline-rich C-terminus of *P. putida* (green box continued from previous page to this page) are distinguishable from all other sequences. Red and blue shaded colours represent alanine and proline residues. The degree of amino acid conservation across all phasin sequences are shown with the most highly conserved (pale yellow) to least conserved (dark brown) amino acids. The degree of consensus ranging from highest consensus (tallest black bars) to lowest consensus (shortest black bars) is shown above the consensus sequence. Alignments were generated using CLC Main Workbench 4.1.2 (CLC Bio). See Table 6.2 for full species names.

observed in the two lower M_r forms (Chapter 5, Fig. 5.8, spots 1 and 2). Peptides matching the C-terminal portion of the predicted Y88^T phasin from the two lower M_r forms (Chapter 5, Figs. 5.8, spots 1 and 2; Chapter 6, Fig. 6.3, green boxed sequence) were also observed in the C-terminus of the higher M_r basic form. The occurrence of these peptides derived from corresponding protein spots on all gels supports the notion that the lower M_r forms of the Y88^T phasin are not artefacts of degradation or mobility. The difference in M_r between the high M_r form and the perceived modified forms corresponds to this "lost" segment. The observed low M_r phasins may be post-translationally modified by site-specific cleavage at the N-terminal end of Met192 or Met203 in the Y88^T phasin sequence (Fig. 6.3) or may be directly translated, either from an internal ribosome initiation site, or from an alternative phasin transcript.

6.2.2.2 Phasin protein sequence alignments

Examination of the CLUSTALW (1.83) multiple sequence alignment of the Y88^T phasin protein sequence with 45 phasin sequences taken from Genbank (Table 6.2, Fig. 6.3) showed that the Y88^T phasin has a distinctive, low-complexity, mainly alanine and proline-rich N-terminal segment, similar in composition only to corresponding N-termini in the four other known sphingomonad phasin sequences examined (Fig. 6.3, black box). This 199 amino acid segment corresponds to the difference in M_r between the expected 33.9 kDa native form and the two 13 kDa forms, so the N-terminal portion may be cleaved post-translationally. Notably, only the sphingomonad phasins had a similar N-terminal segment, although *P. putida* had an alanine/proline-rich segment at its C-terminus (Fig. 6.3, red box). Many of the phasin sequences examined had an alanine-rich segment C-terminal to the "core" phasin sequence, similar to that reported for *R. eutropha* (Neumann et al., 2008), but none of these alanines was conserved across all the sequences examined, reflecting the general sequence heterogeneity among the phasins.

To eliminate the possibility of overlooking potentially incorrectly annotated phasin sequences, a TBLASTN analysis of the Y88^T phasin protein sequence against the NCBI non-redundant database was carried out. This analysis

Accession number	Species	Truncated species name as it appears in phasin sequence alignment
gb|ACB80653.1|	*Methylobacterium populi* BJ001	Methylopopuli
gb|ABY30954.1|	*Methylobacterium extorquens* PA1	Methyloextorqu
gb|ACB27545.1|	*Methylobacterium radiotolerans* JCM 2831	Methyloradioto
gb|ACA18624.1|	*Methylobacterium* sp. 4-46	Methylo
gb|ACK51275.1|	*Methylocella silvestris* BL2	Methylocellas
gb|ACB80343.1|	*Methylobacterium populi* BJ001	Methylobalpha
gb|ACB26322.1|	*Methylobacterium radiotolerans* JCM 2831	Methyloradio
gb|ACA19562.1|	*Methylobacterium* sp. 4-46	Methyloradioa
ref|NP_949474.1|	*Rhodopseudomonas palustris* CGA009	Rhodopspalalp
gb|ACF03108.1|	*Rhodopseudomonas palustris* TIE-1	Rhodopspal
gb|ABJ07987.1|	*Rhodopseudomonas palustris* BisA53	Rhodopspal
gb|EAT09637.1|	*Sphingomonas* sp. SKA58	Sphingomonass
gb|ABQ67229.1|	*Sphingomonas wittichii* RW1	Sphingomonasw
gb|ABF52225.1|	*Sphingopyxis alaskensis* RB2256	Sphingopyxisa
gb|ABD24947.1|	*Novosphingobium aromaticivorans* DSM 12444	Novosphingoar
NA	*Novosphingobium nitrogenifigens* Y88^T	Y88^T
gb|ABC23614.1|	*Rhodospirillum rubrum* ATCC 11170	Rhodospirillu
gb|ABE33652.1|	*Burkholderia xenovorans* LB400	Burkholderiax
gb|ABB09697.1|	*Burkholderia* sp. 383	Burkholderiab
ref|YP_625781.1|	*Burkholderia cenocepacia* AU 1054	Burkcen
gb|ABE31411.1|	*Burkholderia xenovorans* LB400	Burkholderiax
gb|AAC78327.1|	*Ralstonia eutropha* H16	Ralstoniab
emb|CAQ69257.1|	*Cupriavidus taiwanensis*	Cupriavidust
gb|ABF08086.1|	*Ralstonia metallidurans* CH34	Ralstonia metb
ref|YP_002254539.1|	*Ralstonia solanacearum* MolK2	Ralstoniasola
gb|ABE45002.1|	*Polaromonas* sp. JS666	Polaromonasbe
gb|ABD69864.1|	*Rhodoferax ferrireducens* T118	Rhodoferaxfer
gb|ABX37371.1|	*Delftia acidovorans* SPH-1	Delftia
gb|ABE44227.1|	*Polaromonas* sp. JS666	Polaromonasbe
gb|ABD70831.1|	*Rhodoferax ferrireducens* T118	Rhodoferaxfer

Table 6.2 Bacterial species names and the corresponding truncated names in the order of appearance in the list of phasin sequences in the alignment with Y88^T (Fig. 6.2).
No accession number (NA) is yet available for Y88^T.

Accession number	Species	Truncated species name as it appears in phasin sequence alignment
ref|YP_935317.1|	*Azoarcus* sp. BH72	Azoarcusbeta
gb|AAZ48152.1|	*Dechloromonas aromatica* RCB	Dechloromonas
gb|ACO78555.1|	*Azotobacter vinelandii* DJ	Azotobactervi
gb|AAP85954.1|	*Ralstonia eutropha* H16	Ralstoniaeutb
ref|ZP_02382729.1|	*Burkholderia ubonensis* Bu	Burkholderia
ref|YP_002005766.1|	*Cupriavidus taiwanensis*	Cupriavidusta
gb|ABE35627.1|	*Burkholderia xenovorans* LB400	Burkholderiaxe
gb|EAP73279.1|	*Ralstonia solanacearum* UW551	Ralstoniasola
gb|AAQ59041.1|	*Chromobacterium violaceum* ATCC 12472	Chromobacteri
gb|ABE35234.1|	*Burkholderia xenovorans* LB400	Burkholderiax
gb|ABF11905.1|	*Ralstonia metallidurans* CH34	Ralstoniametb
ref|YP_002254322.1|	*Ralstonia solanacearum* MolK2	Ralstoniasola
gb|ACI53681.1|	*Rhizobium leguminosarum* bv. trifolii WSM2304	Rhizobiumlegu
gb|ABE31092.1|	*Burkholderia xenovorans* LB400	Burkholderiax
gb|ABC24078.1|	*Rhodospirillum rubrum* ATCC 11170	Rhodospirilal
gb|AAD02215.1|	*Pseudomonas putida*	Pseudomonaspu

Table 6.2 (cont.) Bacterial species names and the corresponding truncated names in the order of appearance in the list of the phasin sequences in the alignment with Y88^T (Fig. 6.2).

confirmed that no other phasins, with the exception of the sphingomonad phasins, had this low-complexity, alanine/proline-rich segment. A 1000 iteration neighbour-joining bootstrap phylogenetic analysis (CLC Main Workbench 4.1.2, CLC Bio) confirmed the grouping of the Y88^T phasin exclusively with the four other known sphingomonad phasins with 100% bootstrap values (Fig. 6.4). The alanine/proline-rich N-terminal segment first determined for Y88^T in this study therefore appears to be unique among the sphingomonads (Fig. 6.4).

Bioinformatic secondary structure analysis of the Y88^T phasins revealed that the two 13 kDa phasins and the C-terminal segment of the 34 kDa phasin from Met203 to the end of the sequence are predicted to be composed almost entirely of α-helices (Fig. 6.5A). Secondary structure predictions for the 13 kDa *R. eutropha* phasin (PhaP1) by Neumann et al., (2008) also predicted mainly α-helical domains across the entire phasin sequence that these authors suggest may be implicated in the binding of the entire *R. eutropha* PhaP1 to the PHB granule. Although there is little absolute sequence conservation among the phasins, the proteins are structurally very similar in that they all comprise amphipathic α-helices (Pötter et al., 2005). This structure is consistent with the role of these proteins in binding the large hydrophobic PHA granules while providing a hydrophilic solvent face, presumably to maintain the overall solubility of the PHA granule. Helical wheel analyses of phasins from Y88^T in this study (Fig. 6.6) and *R. eutropha* (Fig.6.7), predict that the common portions of both the *R. eutropha* and Y88^T phasins (Figs. 6.5A, 6.5B) are amphipathic (Figs 6.6. 6.7). The predicted amphipathicity may explain the lack of an identified conserved, specific PHA-binding domain in the *R. eutropha* phasin. It seems likely from its predicted amphipathic helical structure that the phasin binds PHA relatively non-specifically through simple hydrophobic interactions, while providing a soluble solvent interface on the other sides of these predicted α-helices.

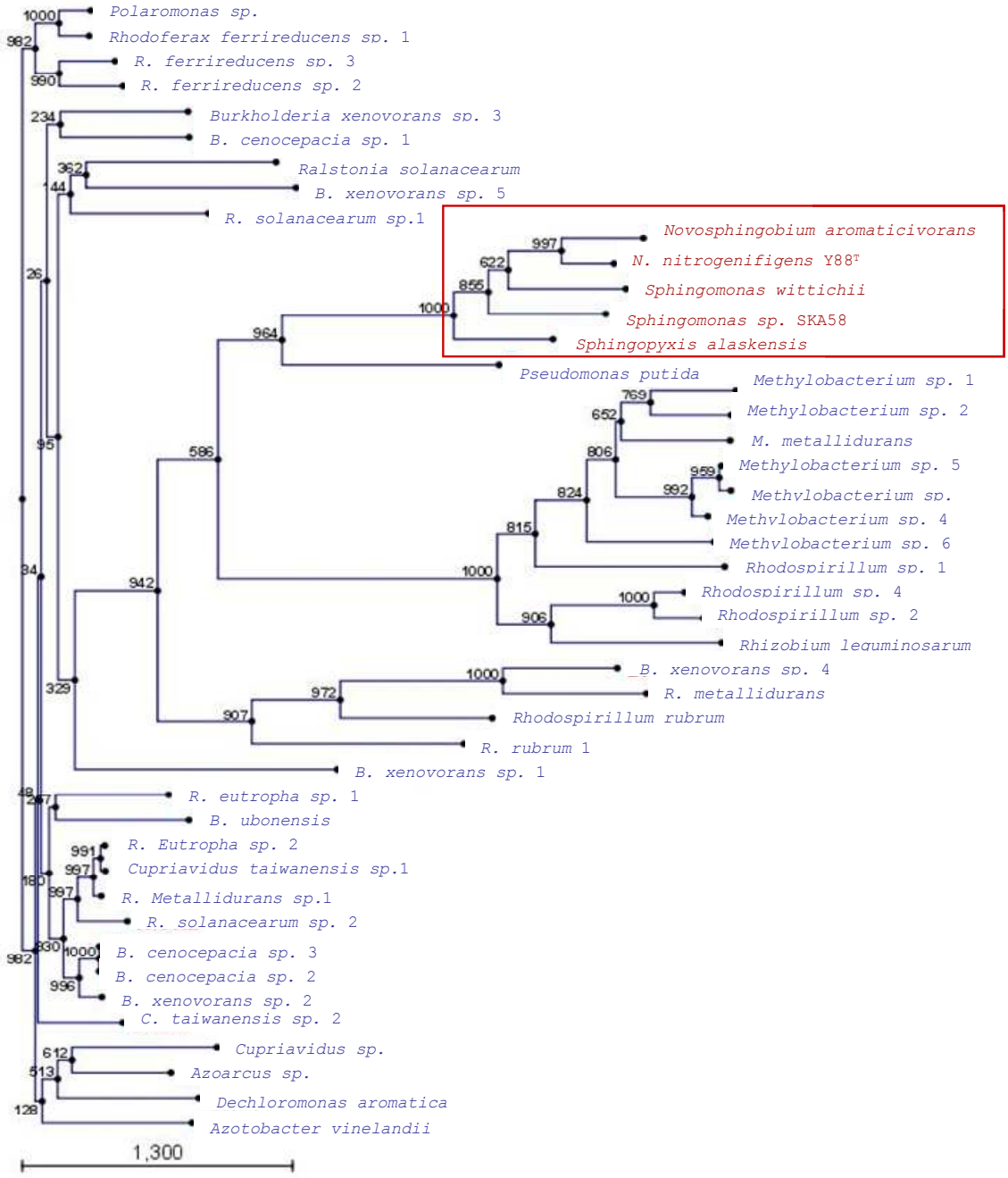


Figure 6.4 Neighbour-joining bootstrap analysis of the Y88^T phasin with 44 phasins from Genbank.

The phylogenetic grouping of the sphingomonad phasins are highlighted (red box). Bootstrap values from 1000 iterations are given at the nodes of each branch. Phylogenetic tree was generated using CLC Main Workbench 4.1.2 (CLC Bio). Branch lengths are given in terms of expected number of substitutions per amino acid site. 1,300 therefore translates to 1.300 (default setting is a comma and needs to be converted to a point) which means that in this tree, branch lengths are based on 1.3 number of substitutions per amino acid site.

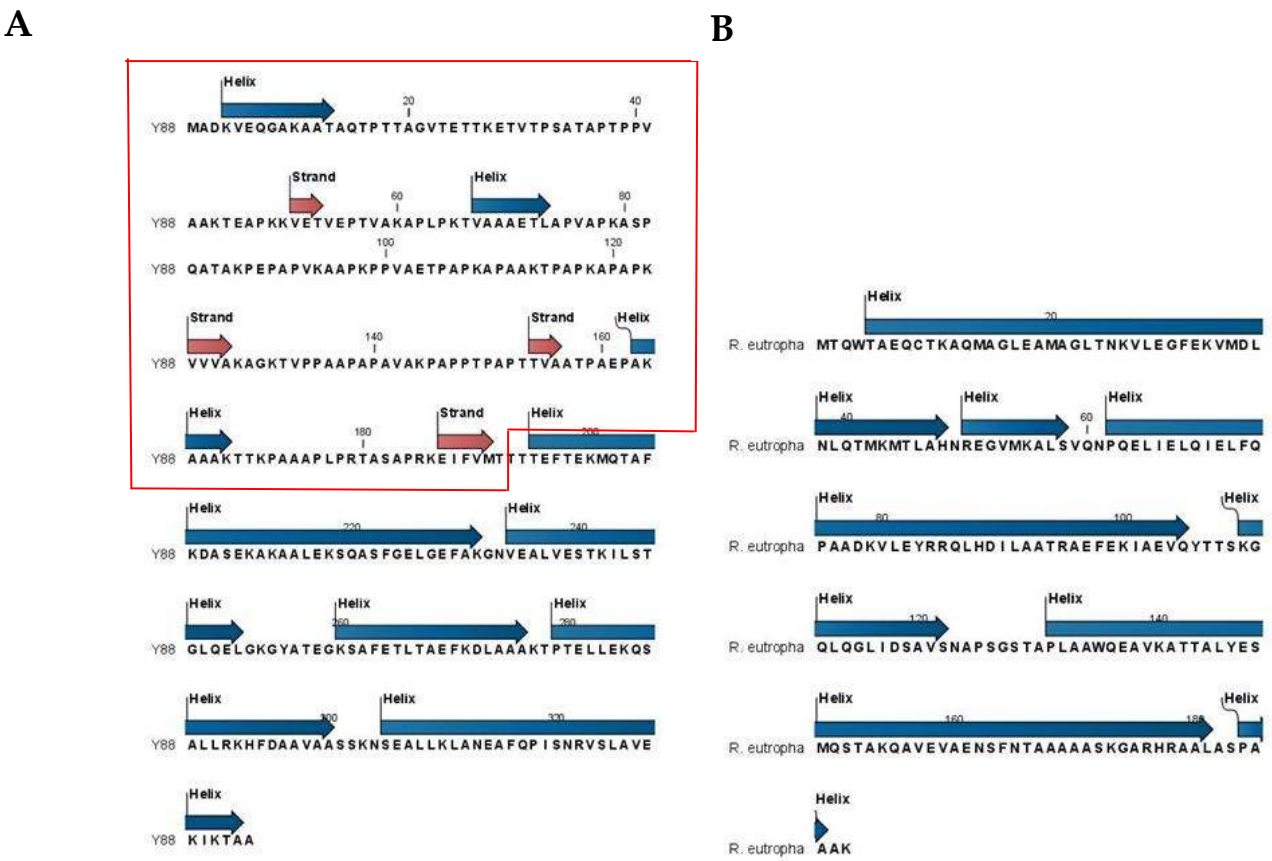


Figure 6.5 Secondary structure predictions for the Y88^{T} and *R. eutropha* phasins.

The Y88^{T} (A) and *R. eutropha* (B) phasins showing the additional N-terminal, low complexity, alanine/proline-rich portion of the Y88^{T} phasin (A; red box) that is common to the sphingomonads. Secondary structure predictions were generated using CLC Main Workbench 4.1.2 (CLC Bio).

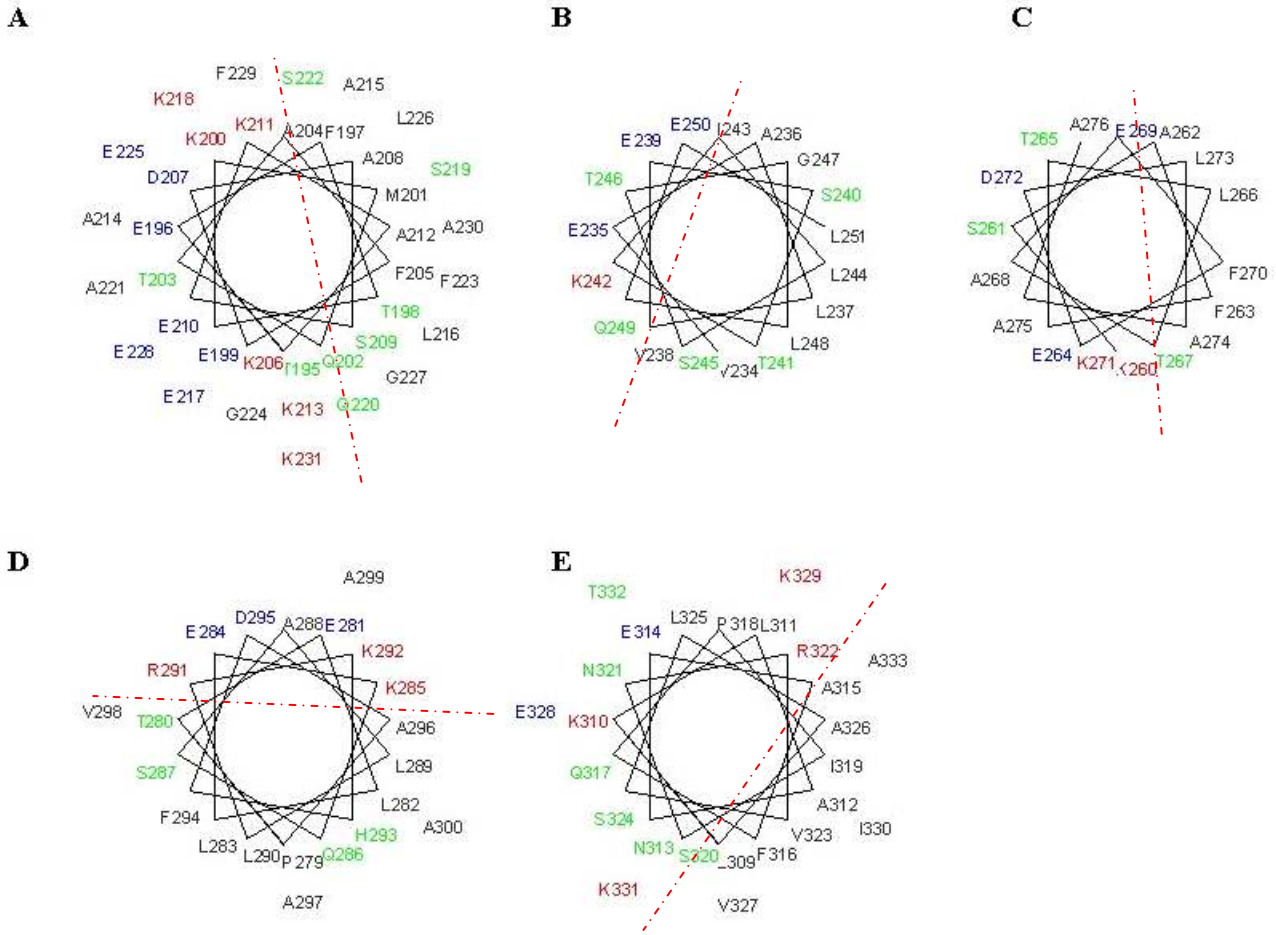


Figure 6.6 Helical wheel predictions of the Y88^{T} phasin

(A) amino acids 195 to 231, (B) amino acids 234 to 251, (C) amino acids 260 to 276, (D) amino acids 279-300, (E) amino acids 309 to 333 of the predicted α-helical secondary structures of the Y88^{T} phasin showing amphipathicity: the charged (acidic amino acids – blue, basic amino acids – brown), hydrophilic uncharged amino acids (green) and the hydrophobic amino acids (black). The characteristic hydrophobic "face" of each structure is apparent to the right of the red line (A, B, C, E) and below the red line (D). Helical wheel diagrams were generated using WinPep 3.01 (Hennig, 1999).

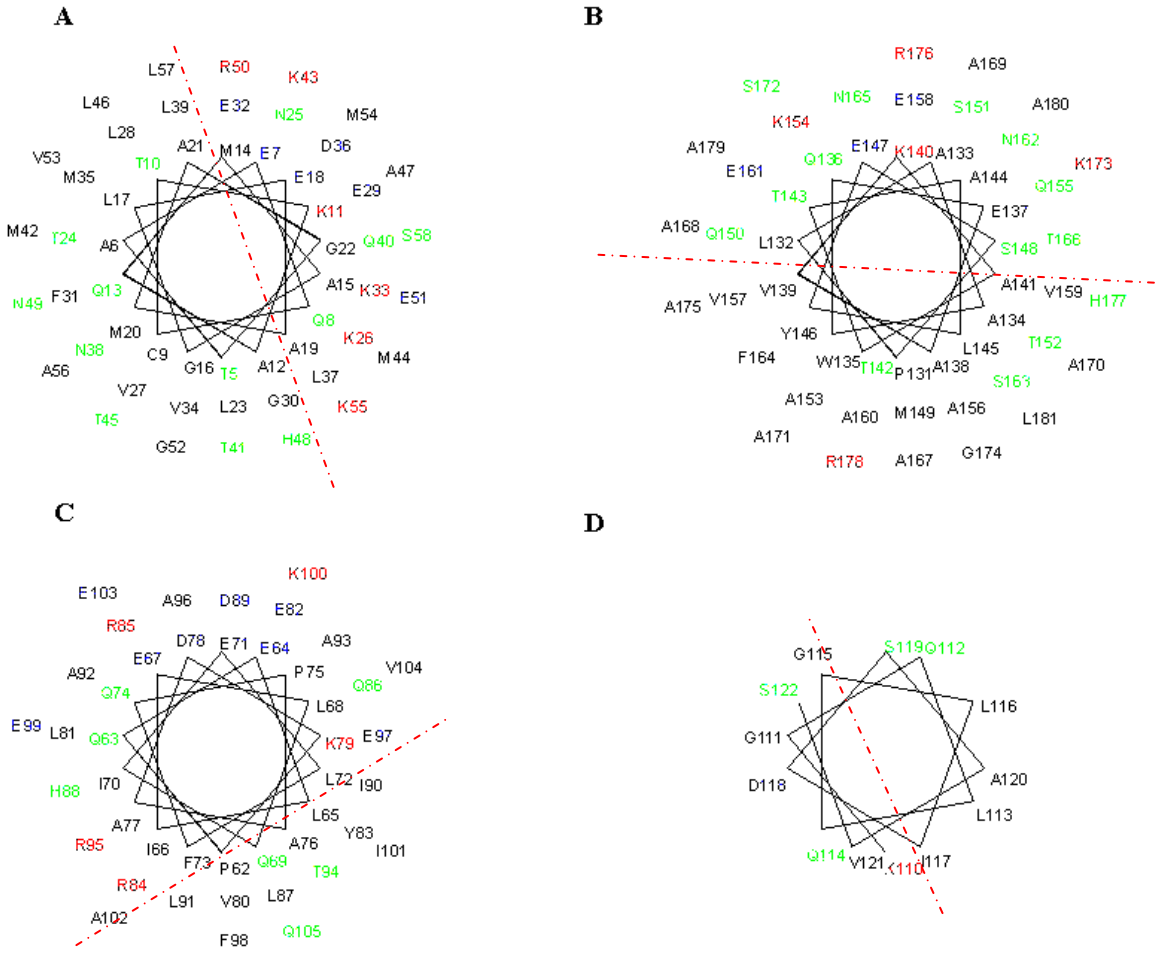


Figure 6.7 Helical wheel predictions of the *R. eutropha* phasin

(A) amino acids 5 to 58, (B), amino acids 131 to 181, (C) amino acids 62-105, (D) amino acids 110 to 122, of the predicted α-helical secondary structures of the *R. eutropha* phasin showing amphipathicity: the charged (acidic amino acids – blue, basic amino acids – brown), hydrophilic uncharged amino acids (green) and the hydrophobic amino acids (black). The characteristic hydrophobic "face" of each structure is apparent to the left of the red line (A), to the right of the red line (D) and below the red line (B, C). Helical wheel diagrams were generated using WinPep 3.01 (Hennig, 1999).

Not surprisingly, given the high proline (an amino acid residue well known to disrupt secondary structure) content of the N-terminal segment of the 34 kDa phasin in Y88^T, this portion of the Y88^T phasin shows very little predicted α-helical secondary structure in stark contrast to the C-terminal segment. Secondary structure predictions for the other four sphingomonads examined in this study also predicted almost no α-helical secondary structure for the N-terminal portions of their phasins also consistent with the proline-rich primary sequences. Once again, secondary structure for the C-terminal portions of the phasins from these sphingomonads showed almost entirely amphipathic α-helical structure similar to that of Y88^T and *R. eutropha*. Thus the sphingomonad N-terminal segment may serve a separate functional role to the "core" phasin. This idea is supported by the facts that i) the low M_r forms appear to be perfectly functional phasins by all known relevant criteria (M_r, sequence homology, secondary structure, amphipathicity), yet ii) the high M_r form of the Y88^T phasin is as abundant as the lower M_r forms, regardless of acidic or neutral growth conditions and iii) the N-terminal sequence is conserved among the known sphingomonads. It is possible that the N-terminal alanine/proline-rich segment may also be involved in binding of the phasin to the PHB granule. It is possible that this high M_r form of phasin may play some role in the copious quantities of PHB produced by Y88^T regardless of the growth conditions in this study, the high molecular weight of the PHB polymers, or both, perhaps by altering the polymer conformation within the granule. Alternatively, since Y88^T has only one phasin gene compared to the four identified in *R. eutropha*, the "additional" alanine/proline-rich N-terminus in the high M_r form may have a functional role analogous to, but not yet defined for the additional phasins in *R. eutropha*.

Chapter 7: Summary and Discussion

7.1 Y88^T: Microbiology, Nitrogen Fixation and Carbon Utilisation

The capacity to fix nitrogen is characteristic of sphingomonads such as *Sphingomonas azotifigens* and *S. trueperi*, but in the main, studies on nitrogen fixation in the sphingomonads, and particularly the *Novosphingobia*, are lacking. Additionally, the intracellular accumulation of PHA is not well understood in the genus *Sphingomonas* (Godoy et al., 2003). The intracellular PHA content of only four sphingomonads, *Sphingopyxis chilensis*, *Sphingopyxis alaskensis*, *Sphingomonas sp*. and *N. aromaticivorans*, had been described in any detail prior to a recent study of Y88^T (Addison et al., 2007). More importantly for this work, studies on the combined ability of bacteria to both accumulate PHA and fix nitrogen are generally lacking and those that have been published have generally focused on the *Rhizobia*. Of the 13 *Novosphingobium* species described, the combined capabilities of PHB production and nitrogen fixation appears, at the time of writing, to be unique to Y88^T.

Y88^T growth was investigated under optimal conditions for nitrogen fixation and PHB accumulation in a minimal nitrogen medium in batch mode with glucose as the preferred carbon source. A reproducible growth profile for Y88^T was achieved for growth conditions at 30°C. Under these conditions, Y88^T grew in a glucose-enriched environment, fixed nitrogen and produced PHB, the size of which exceeded 1,000,000 Da (Chapter 3). This finding is of particular significance because high molecular weight biopolymers such as PHB are desirable from a commercial standpoint to produce higher quality PHAs, as well as from an environmental standpoint as biodegradable replacements for conventional plastics. Y88^T naturally produces a PHB that exceeds the upper range of 50,000 to 1,000,000 Da generally produced by bacteria (Madison & Huisman, 1999). The type (precursor monomer) and molecular weight produced by bacteria varies and is dependent on the species and the carbon source utilised. Y88^T produces high concentrations of PHB with glucose as its

carbon source, and can do so irrespective of DO^{low} or DO^{high} concentrations and within a pH window at least between pH 4.5 and 7.2. PHB production in batch culture by $Y88^T$ occurred irrespective of whether or not the cells fixed nitrogen.

$Y88^T$ also synthesised an extracellular polysaccharide during PHB production. Such polysaccharides are generally considered disadvantageous from a biotechnological standpoint, since they shunt carbon away from PHA and thus decrease PHA yields (Madison & Huisman, 1999). Despite $Y88^T$ producing an exopolysaccharide (confirmed by ^{13}C-NMR analysis), it still produced up to 80% of its dry weight as PHB.

$Y88^T$ proteomic profiles, as expected, included a majority of proteins that showed no significant difference in abundance between growth conditions. In the following sections patterns of protein change are interpreted as indicators of change in physiological state, and potentially also as indicators of change in rate although generally, it is necessary to validate such hypotheses. Proteins that were not detected on the 2-D gels are also expected to participate in $Y88^T$ function although below the level of detection by 2-D DIGE in this study.

7.2 $Y88^T$ growth profiles

$Y88^T$ showed reproducible growth, nitrogen fixation and PHB production in a glucose-enriched, minimal nitrogen, medium. Comparisons of $Y88^T$ with the nitrogen-fixing, PHB-producing, *A. vinelandii* and the PHB-producing WP01, demonstrated superior growth and PHB by $Y88^T$ in a glucose-enriched minimal nitrogen medium in batch culture. Initial supplementation of the medium with ammonium promoted the early growth of $Y88^T$. Nitrogenase activity was not detected in the presence of added ammonium as determined by acetylene reduction assays (Chapter 3). Proteomic signatures for $Y88^T$ during ammonium-supplemented growth included the absence of the nitrogenase proteins on the 2-D gels (Chapter 4). Once the cells had exhausted the ammonium in the medium, $Y88^T$ began fixing nitrogen, as determined by i) acetylene reduction assays, ii) the accumulation of nitrogen above that provided by initial culture conditions (Chapter 5) and iii) confirmed by proteomic analysis (Chapters 4, 5)

with the nitrogenase enzymes detected on the 2-D gels. Without the initial nitrogen supplemented in the form of ammonium, Y88^{T} growth was lower than that in medium supplemented with small amounts of ammonium, despite its ability to fix molecular nitrogen. Y88^{T} therefore requires small amounts of supplemented nitrogen that can be in the form of ammonium in batch culture upon inoculation to initiate and achieve high growth. One possible explanation for such a phenomenon might be that during early growth following inoculation, the initial DO concentration in the medium may be too high to allow the bacteria to immediately fix nitrogen and they must initially deplete the DO in the batch cultures through respiration. A certain initial amount of growth that is dependent on supplemented ammonium may be required to achieve this. Alternatively, a quorum sensing mechanism may also play a role in nitrogen fixation (Brelles-Mariño & Bedmar, 2001). Given the oxygen lability of nitrogenase, there may be strength in numbers for the bacteria in nitrogen fixation because of the protective effects of extracellular polysaccharides used by aerobic nitrogen fixers (Dong et al., 2002) and quorum sensing provides a mechanism by which the cells could recognise when favourable conditions were present.

Y88^{T} growth in batch culture allowed an examination of glucose utilisation, ammonium assimilation and nitrogen fixation as well as the response of the Y88^{T} proteins to the growth conditions in culture. Moreover, growth in bioreactors allowed examination of the differential responses of Y88^{T} to controlled DO^{low} and DO^{high} at different pH during changes in the nitrogen source in glucose-enriched medium. Across the various growth conditions, Y88^{T} assimilated similar amounts of nitrogen and carbon except under DO^{high} conditions, where the cells consumed more glucose. This increased glucose consumption did not contribute to increased intracellular carbon levels, since C:N ratios and biomass were similar across all growth conditions examined. Increased respiration at DO^{high} growth is a possible explanation for this apparent contradiction. Alternatively, the "missing" carbon may have been utilised for exopolysaccharide production, a common oxygen protective

mechanism among aerobic nitrogen-fixing bacteria. Consistent with this hypothesis, ^{13}C NMR analysis of extracellular material from cell pellets of bacteria grown under high oxygen conditions revealed a spectrum consistent with a carbohydrate, likely consisting of glucose subunits (as mentioned in Chapter 3).

The similarity of the C:N ratio in cells grown in both DO^{low} and DO^{high} conditions, despite the depletion of the ammonium in the medium, suggested that nitrogen fixation was occurring even under DO^{high} conditions. A greatly decreased, but still detectable acetylene peak observed for growth under DO^{high} meant that this hypothesis could not be ruled out. Thus, $Y88^{T}$ may have a mechanism of oxygen protection that still enables it to fix enough nitrogen to maintain cellular processes, despite downregulation of its nitrogenase under DO^{high} conditions. Additional work is required to better understand $Y88^{T}$ growth and metabolism under DO^{high} conditions.

Comparative analysis of the $Y88^{T}$ proteomic signatures revealed notable differences between the nitrogen-fixing physiology of the cell (ammonium-depleted conditions, "Fix^{on}") and ammonium-assimilatory physiology (ammonium-supplemented conditions, "Fix^{off}"). Proteins directly involved in and associated with nitrogen fixation (see section 7.3.2) constituted most of the proteins that showed increased abundance for the Fix^{on} state. The abundance of these proteins was similar for the Fix^{on} state in batch flasks and in bioreactors at DO^{low}. In contrast, these proteins were present at very low abundance for the Fix^{off} state in batch flask culture as well as batch bioreactor culture under DO^{high} concentrations. The number of proteins apparently specific to the Fix^{on} state under controlled DO conditions decreased from 38 in the proteomic signatures for growth in acidic media to 10 in the proteomic signatures for growth in neutral media (Chapter 5, Section 5.2.1.5, Figs 5.6, 5.10, Chapter 7, see following section), suggesting that the observed changes in abundance of these proteins was due to the pH of the growth medium rather than a response to nitrogen depletion in the media. Consistent with this notion, this subset of proteins did

not include those directly involved in nitrogen fixation (Chapter 5, Fig. 5.6, 5.10).

7.3 Nitrogen metabolism

7.3.1 Assimilation of ammonium

In free-living nitrogen-fixing prokaryotes, an abundance of fixed nitrogen in the form of ammonium or other compounds such as nitrate or amino acids is reported to affect nitrogen fixation by inhibiting either the synthesis or the activity of nitrogenase (Klugkist & Haaker, 1984; Klassen et al., 2005). Nitrogen fixation and ammonium metabolism are regulated in a coordinated response to the type and amount of nitrogenous compound available. This response involves a network of proteins that are regulated by post-translational modifications in response to nitrogen availability.

It is well known that exogenous ammonium rapidly and reversibly inhibits nitrogenase activity in a variety of nitrogen-fixing bacteria. In contrast, the ammonium ions produced from the reduction of molecular nitrogen by nitrogenase are assimilated directly into anabolic cellular processes beginning with amino acid synthesis via the coordinated activities of GS and GOGAT. The activity of GS is controlled by the nitrogen source of the growth medium and is high in media in which the nitrogen source is growth-rate limiting and low in media containing an excess of ammonium. Extracellular ammonium is incorporated into glutamine, which in turn regulates GS activity indirectly via product inhibition through the nitrogen regulatory proteins NtrC and P_{II} (Ninfa & Atkinson, 2000). The GS-GOGAT system links nitrogen metabolism with carbon metabolism, as GS catalyses the incorporation of newly fixed free ammonium into glutamine, and GOGAT transfers the glutamine amide nitrogen to the AKG carbon skeleton resulting in the formation of two molecules of glutamate (Hodges, 2002). In this system, the intracellular glutamine concentration is the critical regulatory point for the cell in sensing its carbon to nitrogen ratio (Hodges, 2002).

In bacteria, GOGAT occurs as two distinct isoforms: ferredoxin (Fd)-GOGAT and NADPH-GOGAT, the latter of which is an α/β heterodimer. In $Y88^T$, GS and the β-subunit of NAPDH-GOGAT were present on all 2-D gels examined irrespective of the growth conditions. A two-fold increase in NADPH-GOGAT abundance at DO^{low} relative to DO^{high} concentrations may suggest an increased assimilation of ammonium as a consequence of higher levels of nitrogen fixation at DO^{low}, consistent with the larger ethylene peak observed in the acetylene reduction assay under these growth conditions (Chapter 5).

Two types of GS are known: GS type I (GSI) and GS type II (GSII) (Darrow & Knotts, 1977). Analysis of the $Y88^T$ genomic sequence confirms the presence of genes for both these proteins. Examination of $Y88^T$ 2-D gels revealed four abundant adjacent GSI spots (under all growth conditions examined) that migrated at similar M_r but at different pIs. GSII was never identified from any of the gels. This array of GSI proteins is indicative of some form of post-translational modification. Such modifications of a protein are often a means to regulate enzyme activity. In other species, GSI is usually modified by adenylylation under microaerobic, ammonium-excess growth conditions, which reversibly inactivates GS and results in a downregulation of GSII (Rao et al., 1978). In $Y88^T$, the abundance of GSI did not appear to be significantly altered by the switch from Fix^{off} to Fix^{on} nor by the cells growing in DO^{high}. These observations may be consistent with an adenylylation regulatory mechanism, which alters the protein's activity, but not its abundance. Furthermore, the 329 Da difference in size of such a modification might not be detectable, given the abundance and relatively high M_r of GSI. However, adenylylation cannot explain the pI differences of the four GSI isoforms, since monoadenylylation at a specific tyrosine residue is the basis of the regulation (Selao et al., 2008) and thus the modification does not alter the protein charge. The differences in pI of the proteins suggest that the modifications are from low M_r charged adducts. The nature of these adducts and their functional significance has yet to be elucidated.

7.3.2 Nitrogen fixation in Y88^T

Control of nitrogen fixation involves regulatory networks that are still not fully understood (Selao et al., 2008). Although synthesis of the nitrogenase enzyme is known to be under tight transcriptional regulation in all nitrogen-fixing bacteria studied (Martinez-Argudo et al., 2004), in certain bacteria such as *R. rubrum*, it is also regulated at the metabolic level through the GS/GOGAT system. The proteomic signatures of Y88^T in the Fix^{on} state suggest that the nitrogen regulatory system in this organism enables the induction of nitrogenase with no concomitant change in abundance of proteins such as GS and P_{II} under both Fix^{on} and Fix^{off} states.

In the Fix^{on} state at high OD_{600}, the increased abundance of the Y88^T nitrogenase apoproteins (NifDHK) together with proteins known to be involved in the formation and maturation of the nitrogenase apoproteins (NifU) and electron transfer to the nitrogenase complex (electron transfer flavoprotein, ETF-FixAB) suggests that these proteins may form part of a cluster of nitrogen-fixing and accessory proteins that are coordinately regulated to sustain high levels of nitrogenase activity. Their increased abundance is consistent with the large ethylene peaks observed in the nitrogenase assays (Chapters 3, 5). The coordinated increased abundance of ETF-FixAB for the Fix^{on} state suggests a possible role in supplying reductants to the energy-demanding nitrogenase reaction (Selao et al., 2008). The coordinated decreased abundance of ETF-FixAB for the Fix^{off} state together with NifU is significant in that a decreased supply of reductants from ETF-FixAB may result in the requirements for nitrogenase activity not being met. This would explain the observed coordinated decrease in abundance of the cluster of nitrogen-fixing proteins since less electron transfer may result in reduced nitrogenase turnover requiring less NifU to support a decreased rate of nitrogenase synthesis.

The two forms of Y88^T NifH identified at different M_r and pIs are consistent with two NifH forms observed in a *Gloeothece* cyanobacterium (Gallon et al., 2000), *R. rubrum* (Selao et al., 2008) and an *Azolla* cyanobiont (Ekman et al.,

2008). In *Anabaena* and *R. rubrum*, one of these forms is known to result from ADP-ribosylation. In contrast, ADP-ribosylation cannot account for the second Y88^{T} isoform observed on 2-D gels. Although the difference in M_r between the two Y88^{T} NifH forms matched the difference that would be expected from the addition of an ADP-ribosyl adduct, examination of the mass spectrum of the peptide containing the site of ADP-ribosylation, the Arg101 residue, did not reveal any evidence for this modification. Palmitoylation of the same conserved peptide is suggested to be the modification found in *Gloeothece* and *Azolla* (Gallon et al., 2000; Ekman et al., 2008), but not only was no evidence for this modification found in the Y88^{T} MS data, such a modification is also not consistent with the difference in M_r and pI observed for the two NifH forms in Y88^{T}. Therefore the Y88^{T} NifH may be modified differently to NifH proteins from other species.

Notably, the more basic, higher M_r form of NifH was barely discernable on 2-D gels for the Y88^{T} Fix^{on} state at DO^{low} in neutral growth medium. Despite this apparent decreased abundance, nitrogenase activity was not affected, as similar ethylene peaks were obtained in the acetylene reduction assays for cells grown under both acidic and neutral growth conditions at DO^{low}. The more basic form of NifH observed may therefore be an inactive, possibly modified form that is targeted for degradation as suggested to be the case for the modified form in *Gloeothece* and *Azolla* (Gallon et al., 2000; Ekman et al., 2008). The nature of any potential modification is unclear and requires more detailed biochemical analysis.

The reduced abundance of the Y88^{T} NifHDK at DO^{high} concentrations irrespective of pH suggests that, like nitrogenase from most bacteria, the Y88^{T} nitrogenase appears to be sensitive to DO^{high} levels. However, despite decreased abundance of the Y88^{T} nitrogenase under such conditions, it appears that there may still be sufficient nitrogenase to support growth since OD_{600}, biomass and C:N ratios were similar regardless of the DO concentrations. Alternative nitrogenase enzymes that are insensitive to DO^{high} are utilised by

some nitrogen-fixing species (Zhao et al., 2006). Such a hypothesis was considered for Y88^{T}, but examination of its genome revealed no candidates for a homologue to such an alternative nitrogenase. However, the Y88^{T} genome sequence did reveal the presence of a gene encoding the nitrogenase-stabilising protein, NifW. NifW is considered to be a part of the oxygen-protection complex that involves the reversible inactivation of nitrogenase via a conformational change, thereby protecting the enzyme from oxygen (Moshiri et al., 1995; Kim & Burgess, 1996; Lee et al., 1998). Such a reversible protection mechanism for the nitrogenase enzyme might explain the apparent paradox between the decreased abundance of the NifHDK at DO^{high} and the continued growth of the cells, as a small proportion of nitrogenase may cycle in and out of the active state, thus supporting growth.

7.4 Central metabolism, energy and cellular redox state

Key proteins involved in the ED pathway, EM pathway and OPP glucose catabolic pathways such as KDPG aldolase (a diagnostic enzyme of the ED pathway), glyceraldehyde-3-phosphate dehydrogenase (G3P), fructose-1,6-bisphosphate aldolase (a key enzyme of the EM pathway), enolase, phosphoglycerate kinase, 6-phosphogluconate dehydrogenase (6PGDH, a diagnostic enzyme for the OPP pathway) and 6-phosphogluconate dehydratase, amongst others, detected on all the 2-D gels irrespective of the growth conditions suggest that Y88^{T} utilises three interrelated catabolic pathways to metabolise glucose. Of these enzymes, two EM pathway enzymes, glyceraldehyde-3-phosphate (G3P) dehydrogenase and fructose-1,6-bisphosphate aldolase were the only proteins to show altered abundance: G3P dehydrogenase increased at low growth and fructose-1,6-bisphosphate aldolase increased at high growth combined with DO^{low}. Since the proteins from these three pathways generally were present on all 2-D gels regardless of physiological state, they presumably are required to meet metabolic demands. The apparent differential abundance of fructose-1,6-bisphosphate aldolase and G3P dehydrogenase may suggest that there is differential control of the EM pathway.

Based on proteomic profiling of Y88^T growth on glucose-enriched, ammonium-supplemented and ammonium-depleted media, Y88^T appears to have the ability to shunt glucose through the OPP and ED pathways. In Y88^T, 6PGDH and KDPG aldolase were always abundantly expressed regardless of low or high growth and DO^{low} or DO^{high} concentrations, together with another enzyme of the ED pathway, 6-phosphogluconate dehydratase.

The metabolic fates of glucose in the three pathways are different, but yield complementary byproducts of importance to the cell. Ignoring the fate of the carbon for the moment, the EM pathway yields two ATP and two NADH for every glucose molecule processed. In contrast, the OPP and ED pathways yield reducing equivalents instead of ATP, with either two NADPH, or one NADH and one NADPH for every glucose molecule processed in the OPP and ED pathways, respectively (Fig. 7.1). The OPP and ED pathways are therefore costly in terms of energy production. The question thus arises as to what benefit is derived from the utilisation of alternative pathways that result in decreased energy production? Certainly, the OPP pathway is required as the principal source of NADPH for anabolic biochemistry. Coupled with NADP$^+$-dependent isocitrate dehydrogenase, another major route to NADPH, detected under all Y88^T growth conditions, it would appear that the cells are well-geared toward the production of NADPH. However, it is not as simple to explain the presence of the ED pathway enzymes, since this pathway is only 50% as efficient as the OPP pathway in producing NADPH as well as 50% less efficient than the EM pathway in producing ATP (Tauro et al., 1986). It is possible that one of the intermediates in the ED pathway is required as a substrate for a different aspect of metabolism in Y88^T, but as it appears that the ED pathway generally serves simply as an alternate oxidative route to pyruvate (Fig. 7.1) (Peekhaus & Conway, 1998), it might be important for the production of some amino acids after conversion to malate or oxaloacetate via carboxylation reactions. It seems unlikely that the ED pathway would be required by Y88^T for NADH production given a highly active aerobic metabolism. However, the ED pathway also generates NADPH, albeit half less than that of the OPP pathway, and it is

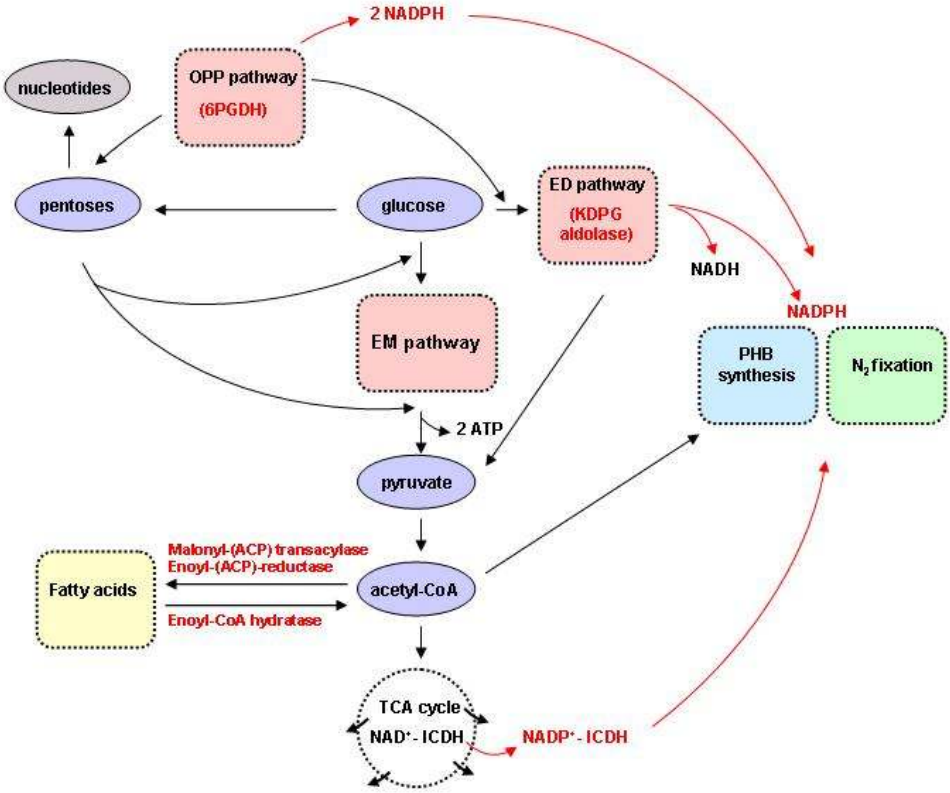


Figure 7.1 Schematic of proposed carbon and reducing equivalent pathways relating to central metabolism, nitrogen fixation and PHB accumulation in Y88^T.

Proposed pathways for PHB and reducing equivalents PHB pathway in Y88^T including possible alternative routes to PHB production, based on proteomic identification of enzymes (red) involved in the relevant pathways. Carbon pathways are shown as *black solid lines* and reducing equivalent pathways are shown as *red solid lines*.

possible that it is in this aspect that the ED pathway is required by Y88^T, due to requirement for NADPH possibly unmet by other metabolic means. One possible circumstance that maximal quantities of NADPH might be required for is in cellular responses to oxidative stress. As an aerobic organism, Y88^T generates NADH from the TCA cycle, which is the initial electron donor to the oxidative phosphorylation electron transport chain. Passage of electrons through the oxidative phosphorylation electron transport chain results in the formation of toxic byproducts in the form of reactive oxygen species (ROS) such as superoxide ($O_2^{\bullet -}$) and H_2O_2. Singh *et al*., (Singh et al., 2007) suggested that increased synthesis of NADPH and decreased synthesis of NADH is a protective response against oxidative stress in *Pseudomonas fluorescens* through mechanisms associated with regeneration of reduced glutathione and recharging protective enzymes including catalase, glutathione peroxidase and SOD. Under highly oxidative conditions, an imbalance in NADH/NADPH ratio can occur, resulting in an "oxidative milieu", a condition conducive to cellular dysfunction and the expression of enzymes that promote NADPH production, primarily glucose-6-phosphate dehydrogenase (G6PDH), $NADP^+$-dependent isocitrate dehydrogenase and malic enzyme induced to counteract this condition (Singh et al., 2007). Therefore, in growth under highly oxidising conditions, one might expect increased abundance of enzymes involved in NADPH-producing pathways in Y88^T.

In this study, none of the key enzymes involved in NADPH production such as the $NADP^+$-dependent isocitrate dehydrogenase or G6PDH that produces most of the NADPH (Marino et al., 2007; Singh et al., 2007) showed altered abundance under any conditions. Consistent with the apparent lack of an induced oxidative response at DO^{high} similar amounts of SOD were observed on 2-D gels at both DO^{low} and DO^{high} concentrations. However, the possibility that DO^{high} may be stressful to the cells at least in early growth is suggested by the two-fold differential abundance of alkyl hydroperoxide reductase, known to be produced under oxidative stress, between DO^{low} and DO^{high} concentrations at low growth. In Y88^T, a highly reductive environment promoted in a glucose-

enriched environment by high levels of NADPH produced by expression of the different glucose catabolic pathways, may be the reason why an apparently minimal induced oxidative response to DO^{high} is evident.

If oxidative stress cannot explain the need for multiple NADPH-producing pathways, what other possible physiological basis might this observation have? Under the growth conditions used in these experiments, $Y88^T$ is essentially constitutively producing PHA and nitrogen fixation is required after the initial ammonium supply is depleted. Therefore, demand for NADPH would be essentially constant and $Y88^T$ may be utilising these pathways, including the ED pathway, to meet this demand (Fig. 7.1). This model assumes that substrate flux through the OPP pathway and $NADP^+$-dependent isocitrate dehydrogenase becomes limiting for NADPH production. Activation of the ED pathway would provide another route to NADPH, albeit a less efficient one. If such a hypothesis proved correct, improvement of reductive flux through the OPP pathway and/or $NADP^+$-dependent isocitrate dehydrogenase could be biotechnological targets for improvement of the organism as a bioplastic producer. Regardless, it appears that under the growth conditions used in these experiments, $Y88^T$ must distribute its glucose among these pathways and maintain a high metabolic flux to meet its energy and redox demands, likely explaining the essentially omnipresence of $NADP^+$-dependent isocitrate dehydrogenase, the enzymes from the EM and OPP pathways, and possibly the ED pathway.

7.5 PHB production

Several interesting phenomena were observed in the $Y88^T$ proteomic signatures that may explain the abundance of PHB-related proteins observed on all 2-D gels irrespective of growth conditions. Enzymes involved in the PHB synthesis pathway via acetyl-CoA (β-Ketothiolase and NADPH-dependent acetoacetyl-CoA reductase), were observed at constant high abundance on all 2-D gels regardless of growth conditions. Acetoacetyl-CoA reductase, which catalyses the second reaction in the PHB biosynthesis pathway, was observed as two forms at the same M_r but different pIs on the $Y88^T$ 2-D gels. Examination of the

Y88^T genome confirmed only one gene encoding this protein and therefore the second form is likely due to post-translational processing. Bacteria that naturally produce PHB also synthesise phasin, a protein which is known to cover the surface of PHB granules (Wieczorek et al., 1995; Pötter et al., 2004; Neumann et al., 2008), stabilise the granules and prevent the coalescence of separated granules in the cytoplasm and nonspecific binding of other proteins to the hydrophobic granule surfaces (Neumann et al., 2008). Despite Y88^T possessing only one gene encoding a phasin protein, three forms of the Y88^T phasin were present at different M_r and pIs suggesting post-translational processing (Chapter 6). The Y88^T more basic low M_r form was the most abundant protein on all 2-D gels under all growth conditions whereas the more acidic low M_r form was a minor component under neutral growth conditions but highly abundant during acidic growth. The Y88^T phasin appears to differ from the *R. eutropha* phasins in possessing a low-complexity, alanine- and proline-rich N-terminal segment which is absent from the Y88^T lower M_r forms as well as in *R. eutropha* and all known phasins from other species, excluding the sphingomonads. This N-terminal segment was found to be exclusive to all the sphingomonads examined (Chapter 6). The higher M_r form observed only on the basic 2-D gels, was also highly abundant under all growth conditions. The presence of the three forms of the Y88^T phasin on all 2-D gels at high growth is consistent with the high amount of PHB produced and the presence of high levels of phasin in Y88^T may therefore affect the amount of PHB produced. Thus all three forms may be implicated in the ability of Y88^T to produce high concentrations of PHB irrespective of the growth conditions since phasin is known to affect the size and number of PHB granules in the cell in *R. eutropha* H16 (Kuchta et al., 2007). The role of the *R. eutropha* H16 phasin (PhaP1) has also been linked to PHB degradation and is therefore important in PHB accumulation. Interestingly, 3-hydroxybutyryl-CoA dehydrogenase, involved in the degradation of PHB, was present on all 2-D gels at unaltered abundance, irrespective of growth conditions, like the PHB synthesis enzymes.

Although Y88^T produces more PHB at later stages of growth and during the stationary phase, high concentrations of PHB were evident throughout the growth cycle. No significant change in the abundance of phasin was observed between DO^{high} and DO^{low} for both the acidic or neutral conditions. However, a significant increase in abundance for all three forms of the Y88^T phasin was detected between high and low OD_{600} during acidic growth. Also, a significant increase in abundance between low and high growth was determined for the more acidic, lower M_r form at DO^{low} for the Fix^{on} state, although this was not the case for low and high growth at DO^{high} during growth in neutral medium.

Interestingly, enoyl-CoA hydratase, an enzyme involved in fatty acid β-oxidation, showed significantly increased abundance at high growth relative to low growth at DO^{low} concentrations in the neutral growth medium, which was the case regardless of DO concentration. The higher abundance of enoyl-CoA hydratase in the later stages of growth may indicate a redirection of carbon metabolism from growth-related activities (i.e., cell membrane synthesis) to carbon storage (i.e., PHB) (Fig. 7.2) as the cultures approach stationary phase (Tsuge et al., 2000). Other fatty acid metabolic enzymes, such as enoyl-(acyl-carrier-protein) reductase, which catalyses a key regulatory step in the fatty acid biosynthesis pathway, were present on the Y88^T 2-D gels, as was malonyl CoA-acyl carrier protein (ACP) transacylase, but neither of these proteins showed altered abundance. The latter of these two enzymes catalyses the transfer of the malonyl moiety from ACP to CoASH, in so doing converting (R)-3-hydroxyacyl-ACP to the corresponding CoA derivative. This enzyme links fatty acid *de novo* biosynthesis to PHB production (Rehm et al., 1998) but whether or not this relationship holds in Y88^T is unknown (Fig. 7.2).

The detected proteins are therefore consistent with functions of biosynthetic pathways geared to apparently maximise PHB production. This raises the question: why are such high levels of PHB required if PHB is essentially an energy reservoir and glucose is present in excess in batch culture? PHAs are often synthesised when one or more other nutrients are limiting. However,

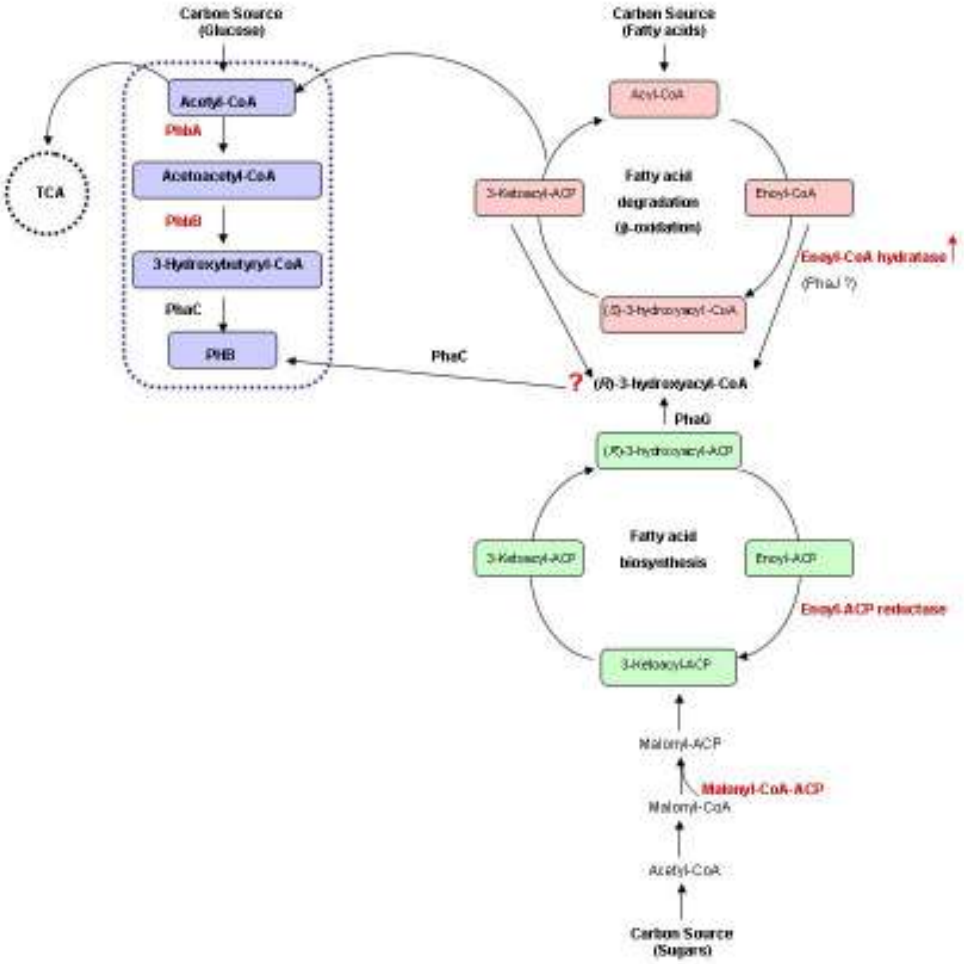


Figure 7.2 Schematic of proposed PHB pathway in Y88^T including possible alternative routes to PHB production, based on proteomic identification of enzymes (red) involved in the relevant pathways. Adapted from Tsuge *et al.*, (2002).

Y88^T synthesises PHB despite growth in medium which supports culture growth, so a nutrient limitation response is unlikely. The *Novosphingobia*, including Y88^T, are generally found in soil and aqueous environments where substrates like glucose (in free form) are presumably not generally found in abundance. Indeed, species from this genus, such as *N. aromaticivorans*, harbour genes that enable them to metabolise unusual carbon substrates. As mentioned previously, Y88^T was isolated from a pulp mill waste stream and glucose, although readily utilisable by Y88^T, may not be the predominant carbon substrate in this environment. Thus, the explanation for Y88^T converting glucose to PHB may simply be an ecological one, which confers a competitive advantage to the species. In other words, Y88^T, upon encountering a rich supply of glucose, may simply be metabolically geared to convert it to PHB in times of "feast" to quickly sequester the glucose away from competitors for later use in times of "famine". Such growth substrate conversion to confer competitive advantage is not unprecedented (Serafim et al., 2004), but further work is required to elucidate the basis for this biotechnologically useful phenomenon.

7.6 Amino acid and nucleic acid metabolism proteins

Enzymes catalysing the biosynthesis of various amino acids and nucleic acids as well as those known to be involved in protein repair and degradation were detected on the Y88^T 2-D gels (Table 7.1). By and large, these enzymes did not vary in their abundance. Their unaltered states possibly reflect their "housekeeping" roles in cellular function.

Another protein observed, saccharopine dehydrogenase, is an intriguing hit to an *Arabidopsis* protein and appears to be multifunctional since it is also involved in lysine catabolism (Zhu et al., 2001).

7.7 Outer membrane proteins

Different outer membrane proteins showed altered abundance for the acidic growth conditions (OmpA/MotB) compared to the neutral growth conditions (OmpW) at high growth and DO^{low}. Additionally, MotA/TolQ/ExbB proton channel protein was present at constant abundance for the DO^{low} and DO^{high},

Enzyme	Cellular Function
L-Isoaspartate-(D-aspartate) O-methyltransferase	repair non-enzymatically damaged protein
polyribonucleotide nucleotidyltransferase	RNA processing and mRNA degradation
DNA polymerase III beta subunit	DNA Replication
dihydroorotate oxidase	pyrimidine biosynthesis
phosphoserine aminotransferase	serine and pyridoxine biosynthesis
δ-Glutamyl phosphate reductase	proline biosynthesis
homoserine dehydrogenase	amino acid biosynthesis
phenylalanyl-tRNA synthetase, α-subunit	protein synthesis
saccharopine dehydrogenase	lysine catabolism, carbon/energy supply

Table 7.1 $Y88^T$ enzymes involved in protein and nucleic acid synthesis that showed unaltered states irrespective of the growth conditions.

acidic and neutral conditions. These proteins may play a role in exopolysaccharide export or the transport of peptides across the membrane and their altered appearance may be related to the osmolality of the culture medium as suggested for *E. coli* (Han & Lee, 2003). They may therefore alter certain aspects of the outer membrane permeability of Y88^{T} with increase in cell density and at DO^{low} conditions under acidic or neutral growth conditions as suggested by the altered abundance of porin proteins.

7.8 Other

Y88^{T} proteomic profiles included a majority of proteins that showed no significant difference in abundance between growth conditions. The fact that many proteins show visually detectable alterations in abundance but are not determined by BVA analysis to have significantly altered abundance does not imply that these changes lack biological significance. Each 2-D DIGE gel is essentially a "snapshot" of the protein state at the moment of culture sampling and the differences observed between two sampling points reflect only the physiological states of the cells at those moments. Proteins that show more subtle alterations in abundance may show differences above the BVA threshold at other sampling times. Essentially, these changes are likely indicative of constantly changing physiological conditions that presumably mask small transient changes maintaining cellular functions related to general and more specialised pathways, such as PHB production.

7.9 Future prospects

The work described here has highlighted important proteins involved in and associated with the major core metabolic, nitrogen fixation and PHB synthesis pathways as well as possible alternative pathways to increase PHB production. This work provides important direction for future studies of PHB production in nitrogen-fixing bacteria. Particular proteins detected here could be the focus of investigations for biochemical characterisation to verify function, to examine regulatory details of expression and to measure specific enzyme activities in full-scale PHB production systems (Wilmes et al., 2008). Also, this project provides structural and functional insight into Y88^{T} nitrogen fixation and PHB

production in batch culture. Further work directed toward optimising Y88^T growth in continuous culture would be advantageous from a production scale perspective to maintain Y88^T in a steady state of continuous PHB production.

The data from this work suggest the possibility that Y88^T may prove to be a useful production strain in pure culture. Experiments on PHB synthesis and extraction from large-scale cultures for the purpose of testing the PHB for its usefulness in manufacturing processes are underway (A-M Smit, G Lloyd-Jones & A. Fernyhough, unpublished results). Biotechnological improvement of Y88^T as a potential production strain is another possibility and several targets for improvement of the strain are noted here. Regardless of the outcome of such experiments, the concept of the use of a nitrogen-fixing PHA-producing bacterium either as a low cost means of converting waste streams to a useful product or as a production strain in pure culture has been shown to be plausible here.

The ubiquity of key proteins such as the Y88^T phasin, regardless of growth conditions, suggests that the PHB synthesis capability of Y88^T is highly robust and that, unlike many other PHB-producing bacteria, Y88^T is capable of producing PHB under a wide range of conditions. Phasin abundance appeared to be positively correlated to the quantity of intracellular PHB, which accumulated most rapidly in the late stages of growth. Based on these observations, maintenance of Y88^T cells in a state of growth akin to late log phase could allow for continuous optimal production of PHB. In a production environment, phasin has the potential to be utilised as a biomarker to maintain the cells in such a state and maximisation of phasin expression would presumably correlate with the maximisation of PHB accumulation.

Further studies using recombinant *E. coli* and the Y88^T phasin would be beneficial in determining the functional role of phasins and their role in PHB production. In this regard, *in situ* experiments involving antibody probing and immunogold detection of the Y88^T phasin (with and without the

sphingomonad-specific N-terminus) to determine where in the cell the intact phasin and the truncated phasin are located, may reveal the functional significance, if any, of the N-terminal segment of the $Y88^T$ phasin in PHB production, and for sphingomonads in general.

Further studies on the $Y88^T$ nitrogenase, in particular the NifH apoprotein, is required to determine whether the $Y88^T$ NifH undergoes post-translational modification as established for several other nitrogen fixers and what the functional significance, if any, of this would be. Furthermore, ^{15}N isotope labeling experiments to trace the incorporation of nitrogen in the cell during DO^{high} growth conditions, thereby increasing the nitrogen content in the cell and confirming that nitrogen fixation occurs irrespective of DO^{high}, will irrefutably confirm whether or not nitrogenase is active or inactive under DO^{high}. An alternative option would be to monitor the amount of nitrogenase present throughout the growth process by employing immunoblotting at different times corresponding to that period of growth when $Y88^T$ is known to fix nitrogen and not to fix nitrogen.

Currently, international focus in biotechnology has shifted to producing next-generation biofuels using engineered microorganisms to produce fossil fuel replacements (Keasling & Chou, 2008). A more far-reaching implication of understanding the role of $Y88^T$ proteins in the PHB synthetic pathway is the prospect of the development of bioproducts other than PHB. Since $Y88^T$ has the ability to naturally produce PHB to high levels in its cells, excess amounts of the 4-carbon intermediate, 3-hydroxybutyryl-CoA, generated in the initial steps of the PHB synthetic pathway, could be redirected through metabolic engineering to produce a "next-generation" biofuel such as butanol (Keasling & Chou, 2008). Additionally, the use of early intermediates in the synthesis of alternative PHAs with varying alkyl groups or chain lengths, offer the potential of even more carbon-rich biofuels with improved properties over biobutanol. The natural propensity of $Y88^T$ to metabolise glucose to preferentially produce PHB offers a substantial biotechnological advantage in that the metabolic

engineering of this microorganism can focus on what to do with the high quantities of its early synthetic precursors, be it to make bioplastic or biofuel. Although this study initially set out to understand the interplay of PHA accumulation and nitrogen fixation of Y88^{T} to enhance bioremediation processes, the real value of this organism may thus lie in its potential in pure culture as a production strain for reduced-cost biomaterials.

Chapter 8 Appendices

8.1 Appendix A Additional proteins (following six pages) not all discussed in chapters 4-7 identified by Mascot-based searches of LC-MS/MS data.

Supporting peptide data confirming positive identifications are shown. Protein identifications represent proteins at DO^{low} / DO^{high}, constant pH, linear range pH 4-7. See 2-D preparative gel (Fig. A.1 below) for spots matching identified proteins and corresponding spots in this table.

Spot number	Protein ID	Two highest scoring unique peptides (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage
A1	Nitrogenase molybdenum-iron beta chain	R.LAGVEIPDDLAK.E(65)	58444	127	15	5.37	9
		R.FVMEVAR.L(46)					
A2	glutamyl-tRNA(Gln) amidotransferase, B subunit	K.MLETGDAPGSIVER.E (76)	54554	178	5	5.25	8
		R.ADVNVSVR.R (70)					
	pyridine nucleotide-disulphide oxidoreductase, class-II, active site	M.PILDAATTAQLK.A(78)	55268	347	21	5.31	12
		R.MDLAQIVAK.L(53)					
	GMP synthase	K.TFIDVFEEEAR.Q (30)	57278			5.43	2
A3	D-3-phosphoglycerate dehydrogenase	R.VTVGTDEGEK.A (45)	56092	93	6	5.44	3
A4	hypothetical protein BRADO5426	R.ISDTAFGTLEAAR.R (85)	22861	187	5	4.93	18
		M.SATTSPSLETVR.I (72),					
	Inorganic diphosphatase	R.SPFVPGAVVR.A (43)	20089	90	5	4.9	10
		R.IVVEAIER.E (39)					
A5	nitrogenase iron protein	K.STTSQNTLAALADLGQR.I(86)	31859	205	16	4.94	20
		R.LGGLICNER.Q(63)					
	GrpE protein	R.DILSVADNLAR.A (66)	20712	63	1	4.86	5
	Inorganic diphosphatase	R.SPFVPGAVVR.A (35)	20089	337	47	4.9	5
A6	glutamine synthetase	R.LVPGFEAPVLLAYSAR.N (81)					
		K.ALNAFTNPTTNSYK.R (70)					
	Glucose-6-phosphate isomerase	R.QLFEGETINNTEGR.A (76),	42356	198	3	178	15
		R.AAFADLVYTQALGCQTR.A (73)					
	chaperonin GroEL	R.AAVEEGIVTGGGTALLYATK.A (103)	57484	198	3	5.09	7
		K.LENVTLGMLGQAK.K (89)					
A7	glutamine synthetase, type I	R.INDEEIEWVDLR.F (72)	52156	331	10	4.94	23
		K.GGYFPVAPVDSLVDIR.G (70)					
	anthranilate synthase component I	R.LAEAIELAR.Q (71)	54709	68	1	4.93	4
		R.LIADTETPVGAALK.L (68)					
	glutamyl-tRNA(Gln) amidotransferase, A subunit	M.TNLTEFGIAAIR.D (69)	52826			4.92	2
A8	glutamine synthetase, type I	R.INDEEIEWVDLR.F (66), ,	52156	262	10	4.94	25
		K.ALNAFTNPTTNSYK.R (64)					
A9	glutamine synthetase, type I	K.AINAFTNPTTNSYK.R (75)	54554	178	5	5.25	8
		R.LVPGFEAPVLLAYSAR.N (73)					
	glutamine synthetase, type I	K.IDDIELPTNSNR.V (69)	52267	125	12	4.94	15
		K.GGYFPVAPVDSLVDIR.G (61)					
A10	Argininosuccinate synthase	K.GQALVVGR.K (59)	45192	79	3	5.37	5
		R.GIESITLDR.G (46)					
A11	translation elongation factor Tu	K.GSALAALEGRDDAIGKDSIK.E (52)	43045	129	6	5.16	11
		R.KLLDQGEAGDNVGALIR.G (45)					
	3-isopropylmalate dehydrogenase	K.GLANPMATILSAAMLLR.H (84)	37267	112	3	5.19	7
		R.TLADGILGR.D (42)					
A12	oxidoreductase, 2OG-Fe(II) oxygenase family	R.MAVAADLGK.A(50)	39009	138	6	5.07	2
	Ribonucleoside-diphosphate reductase	R.LFQAFVQER.D (56)	41366	71	2	4.99	4
		R.NLLTQIFR.F (44)					

Spot number	Protein ID	Two highest scoring unique peptides (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage
B1	NifU	R.TNGLTSIDEVTNYTK.A(110)	33542	1139	29	4.72	22
		R.VNAEMVAEGALAPER.A(82)					
	OmpA/MotB	K.YNLGLSAR.R (39)	11983	48	2	5.76	7
B2	electron transfer flavoprotein, alpha subunit	K.LGAAIGASR.A(60)	31153	141	3	4.8	14
		R.AAVDAGYVPNDYQVGQTGK.I(50)					
	OmpA/MotB	K.YNLGLSAR.R (45)	11983	96	6	5.76	37
		R.NASVQSYLTAHGIPAAEIATK.A (38)					
B3	translation elongation factor Tu	K.LIAPIAMDEGLR.F (58)	42934	77	3	5.16	10
		R.TVGSGVVSK.I (44)					
	OmpA/MotB	K.YNLGLSAR.R (42)	11983	96	6	5.76	18
B4	adenosylhomocysteinase	K.VACVAGFGDVGK.G(80)	50877	788	38	5.33	20
		R.DEFGPTQPLK.G(47)					
B5	putative glutaredoxin family protein	R.GEFVGGCDIVR.E (47)	12194	48	1	4.55	9
B6	OmpA/MotB	K.YNLGLSAR.R (43)	11983	77	5	5.76	7
B7	UspA	K.DAGLVVMGLFGK.S (71)	27836	266	7	5.03	16
		R.KGTIAATIEDFAR.E(68)					
	4-diphosphocytidyl-2C-methyl-D-erythritol kinase	R.LPVAAGLGGGSADAGAIFR.L (141)	28778	141	1	5.08	7
B8	hypothetical protein Saro_2615	K.AAYLQPNFSSTK.D(58)	29433	267	14	5.92	7
		K.TDFGEPFIK.V(33)					
B9	3-oxoacid CoA-transferase, subunit A	R.AGGAGIPGFYTK.T (60)	25869	143	4	5.43	9
		R.MIVGSPYDKK.I (41)					
B10	Phasin	K.SQASFGELGEFAK.G (91)	33963	438	16	9.56	24
		K.SAFETLTAEFK.D (74)					
B11	ribosomal protein L7/L12	K.LTVLEAADLAK.A (83)	12683	323	9	4.65	24
		K.IEEAGGTVEIK.X (78)					
	Phasin	K.SQASFGELGEFAK.G (67)	33963	233	5	9.56	3
B12	ribosomal protein L7/L12	K.LTVLEAADLAK.A (74)	12683	422	12	4.65	28
		R.AITSLGLTEAK.A (66)					
	pyruvate dehydrogenase	K.TLADLAQVTEMK.T (75)	20244			5.12	6
	protein of unknown function DUF179	R.AEGIDPSHLVAATGR.A (34)	19809			5.07	8
	protein-L-isoaspartate(D-aspartate) O-methyltransferase	R.VLVIGKPGAYLAAVTAK.V(47)	20640	64	2	5.03	12
		R.AMIESQLR.V(44)					
	ribosomal protein L7/L12	K.LTVLEAADLAK.A (80)	12683			4.65	8
	translation elongation factor Tu	K.TTLTAAITK.V (47)	43045			5.16	2
C2	protein-L-isoaspartate(D-aspartate) O-methyltransferase	R.VLVIGKPGAYLAAVTAK.V (59)	20640	64	2	5.03	12
		R.AMIESQLR.V (44)					
	Inorganic diphosphatase	R.SPFVPGAVVR.A (46)	20089			4.9	9
	2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase	R.ILHTPMR.Y (26) K.GEPDVVAIEAAAR.A (61)	20779			5.15	6
C3	nitrogen regulatory protein P-II	K.IFVLDLASATR.I (89)					
		K.GQTEIYR.G (46)					
	nitrogen regulatory protein P-II	K.IFVLDLASATR.I (87),	12090	236	5	5.13	30
		R.EALSGLGVAGMTVSEVK.G (76)					
	Inorganic diphosphatase	R.SPFVPGAVVR.A (41)	20089			4.9	5
C4	isocitrate dehydrogenase	K.VQNPIVEIDGDEMTR.I (66)					
		R.LIDDMVASALK.W (61)					
	isocitrate dehydrogenase, NADP-dependent	K.YYDLSVQNR.D (62)	46087	114	3	5.28	8
		R.LIDDMVASALK.W (60)					
		K.VKNPIVELDGDEMTR.I (56)					
	Methionine adenosyltransferase	R.HSGAAPFLEPDAK.S (37)	43503			5.27	3

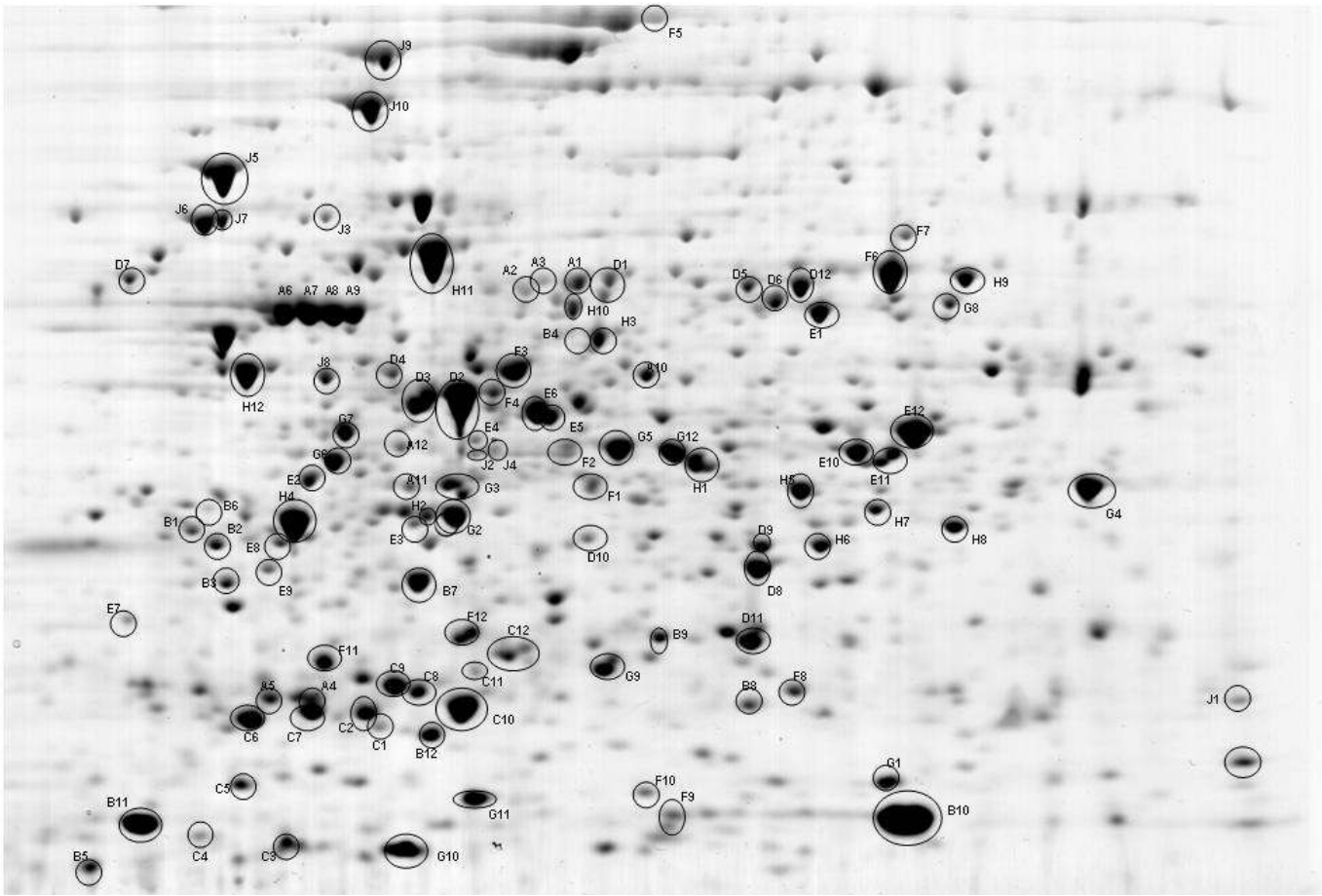
Spot number	Protein ID	Two highest scoring unique peptides (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage
C5	alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen	K.GIDEVACTAVNDPFVMK.A (67)					
	alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen [	K.ALGLVLDGSGFGLGQR.G (84)	16726	141	5	4.84	36
		K.GVDEIACTAVNDPFVMK.A (75)					
C6	chromosome segregation protein SMC	R.LSDLTSEIR.R (75)					
	alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen	K.FVDLTDADIAGK.W (87)	20864	278	8	4.79	17
		K.LKPFANTAFYQGK.F (44)					
C7	inorganic pyrophosphatase	K.KSGALFVDR.L (67)					
	Inorganic diphosphatase	K.ESGALFVDR.I (67)	20089	163	8	4.9	30
		R.VGTWGGAEDAR.R (59)					
	Electron transfer flavoprotein beta-subunit	R.TALAMGADR.A (32)	26104	32	1	5.5	3
C8	pantoate--beta-alanine ligase	MSPATLLR.S (58)					
	isochorismatase hydrolase	K.SIDATLLR.N (63)	23573	157	7	5.06	11
		R.TETYEHTTGIAR.K (57)					
C9	ATP-dependent Clp protease proteolytic subunit 3	R.GMASDIEIQAR.E (75)					
	isochorismatase hydrolase	K.SFSGPMFEEVTAPFPGK.A (66)	23573	86	3	5.06	10
	ATP-dependent Clp protease, proteolytic subunit ClpP	R.GMASDIEIQAR.E (66)	25292	67	1	4.92	4
C10	ATP-dependent Clp protease, proteolytic subunit ClpP	R.GMASDIEIQAR.E (76)					
		R.SFDIYSR.L (29)					
	transaldolase, putative	K.IGADVITAPPAVIK.G (58)	23406	74	2	5.19	11
		K.AGATFVSPFVGR.H (47)					
	ATP-dependent Clp protease, proteolytic subunit ClpP	R.GMASDIEIQAR.E (71)	25292	72	1	4.92	4
	2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase	K.GEPDVVAIEAAAR.A (67)	20779	129	3	5.15	6
C11	possible glycosyltransferase	R.SLVDGVTAQLFR.D (49)					
C12	translation elongation factor Tu	K.LLDQGEAGDNVGALIR.G (85)	43045	401	12	5.16	11
		K.LIAPIAMDEGLR.F (67)					
	enoyl-(acyl-carrier-protein) reductase	R.VLASELGGQGIR.V (113)	29160	139	7	5.45	7
		K.VMADLAVR.W (48)					
D2	GMP synthase	R.IVAVSDGAPFAVIADDTR.K (101)	57278	243	15	5.43	12
		R.LGEAEQVVGLFR.N (71)					
	prolyl-tRNA synthetase	K.GVPLVLEIGGR.D(64)	57944	181	4	5.39	4
		K.ALVSQSVFDEK.V(61)					
	Glycine dehydrogenase (decarboxylating)	R.IDLAALLER.L (79)	55278	127	4	5.34	6
		K.AALDQFIAAMR.D (68)					
	ribosomal protein L9	K.GVPLVLEIGGR.D (67)	58055			5.39	4
		K.ALVSQSVFDEK.V (62)					
D2	translation elongation factor Tu	K.LLDQGEAGDNVGALIR.G (86)	43045	591	19	5.16	25
		R.QVGVPALVVYMNK.V (77)					
D3	DNA polymerase III, beta subunit	R.KLPDGSQVSLETADNR.M(100)	39096	126	9	4.9	9
	nucleotide sugar dehydrogenase	K.QLLANYSEVPQNLIR.A(85)	43208	1280	33	5.07	22
		R.LGTEQVIFSPEFLR.E(82)					
D4	aspartate kinase	K.VSVLIDSDETELAVR.V (87)	44608	174	4	4.99	5
		K.ISVVGVGMR.S (54)					
	translation elongation factor Tu	K.LIAPIAMDEGLR.F (58)	43045	119	4	5.16	8
		K.TTVTGVEMFR.K (55)					
D5	hydro-lyases, Fe-S type, tartrate/fumarate subfamily, alpha region (fumarase)	K.DAIAQILTNSR.M (78)	55452	119	4	5.76	9
		R.GPQTDIER.L (58)					
	peptidase M16-like	K.ASDGNAQDLR.V (47)	103631			5.75	1
D6	aldehyde dehydrogenase (NAD+)	K.NLLVSYSPK.A (53)					
	chloroacetaldehyde dehydrogenase	R.FENFIGGK.W (50)	56306	135	9	5.56	7
		K.NLLVSYSPK.K (41)					
	hydro-lyases, Fe-S type, tartrate/fumarate subfamily, alpha region	K.DAIAQILTNSR.M (60)	55452			5.76	4
D7	trigger factor	K.VPGEFVANDDFAK.Q (73)	55114	296	8	4.71	18
		R.FIPGFEEQLVGVK.A (71)					
	cell division protein FtsZ	R.AIANPLLDGVSMQGAK.G (73)	51291			4.65	7

Spot number	Protein ID	Two highest scoring unique peptides (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage
D8	ribose-phosphate pyrophosphokinase	K.LVANLITTAGADR.V (92)	33700	538	16	5.55	27
		K.ELVITDSILPTEATR.E (88)					
	succinyl-CoA synthetase, alpha subunit	R.KGSVGVVSR.S (66)	30663	296	29	5.54	21
		K.IGIMPGSIFR.K (61)					
D9	probable 3-hydroxybutyryl-coa dehydrogenase protein	K.IVAADAEAALSR.I (86)	30939	175	10	5.42	22
		K.VLAPQAILASNTSSIPITR.M (51)					
	succinyl-CoA synthetase, alpha subunit	R.KGSVGVVSR.S (64)	30663	202	18	5.54	9
		M.SILVNKDTK.V (38)					
	ribose-phosphate pyrophosphokinase	K.LVANLITTAGADR.V (92)	33700	90	1	5.55	4
	UDP-glucose pyrophosphorylase	K.AVFPVAGLGTR.F(56)	32293	141	5	5.64	3
	Nucleotidyl transferase	K.AVFPVAGLGTR.F (58)	32404			5.62	3
D10	putative glutathione S-transferase YghU	R.AVGAPENQLR.E (49)	33744	94	7	5.35	9
		R.LLHVLDTQLGK.S (34)					
	succinyl-CoA synthetase, alpha subunit	R.KGSVGVVSR.S (47)	30663			5.54	3
D11	Enoyl-CoA hydratase/isomerase	K.AFAAFEADPDQR.C (60)	28551	75	2	5.44	7
		K.FGQPEVK.L (40)					
D12	fumarase	K.DAIAQILTNSR.M (68)	55452	197	11	5.76	10
		R.ASVLTDPAFTR.R (50)					
E1	Hypothetical, possible membrane	R.IGLAKTLR.R (47)					
		R.LGIAARDR.D (47)					
E1	Dihydrolipoamide dehydrogenase	K.VTGAEVKDGK.A (46)	49056	93	4	5.63	6
		K.VGKFPMLANSR.A (37)					
E2	phosphoribosylformylglycinamidine cyclo-ligase	R.LLIDHLIEPTR.I (51)	39054	126	12	4.92	12
		R.VIAGIAEGCR.I (44)					
	translation elongation factor Tu	K.LLDQGEAGDNVGALIR.G (92)	43045	197	4	5.16	7
		K.LIAPIAMDEGLR.F (60)					
	fructose-1,6-bisphosphatase, class II	R.VTEAAAIGASR.L (90)	34829	113	4	4.89	8
		K.AADAAAVEAMR.A (80)					
E3	malonyl CoA-acyl carrier protein transacylase	K.VGMGADLAAASAVAR.E (93)	32504	372	15	5.05	13
		R.AFVFPGQGSQK.V (61)					
	two component LuxR family transcriptional regulator	R.AVQASTVGLDATGR.M (62)	33664	88	2	5.12	2
	fructose-1,6-bisphosphatase, class II	R.VTEAAAIGASR.L (88)	34829	84	3	4.89	3
E4	biotin synthase [Novosphingobium aromaticivorans DSM 12444] gb|ABD2	K.LLTAANAGDDADSAMFAR.L (83)	38806	211	7	5.26	11
		K.IDDIEFVR.T (58)					
E4	NADH:flavin oxidoreductase/NADH oxidase	R.LLDDYALAAENAK.K (76)	39525	156	5	5.19	8
		R.ILMAPLTR.G (34)					
	Myo-inositol-1-phosphate synthase	R.GFIVADAPEATK.E (64)	39788	123	3	5.11	5
		R.VLSSLFGAR.G (57)					
	succinyl-CoA synthetase, beta subunit	K.EVNLSVPLVVR.L (50)	42798			5.36	4
E5	S-adenosylmethionine synthetase	K.GIMDTDGNWAPGVPEEV R.V(74)	43392	210	6	5.27	9
		K.NIVAAGLAHR.C(48)					
	isocitrate dehydrogenase, NADP-dependent	R.DATADQITIDSANAIK.E (99)	46087			5.28	8
		R.LIDDMVASALK.W (73)					
	Methionine adenosyltransferase	K.GIMDTDGNWAPGVPEEVER.V (86)	43503	210	5	5.27	10
		K.NIVAAGLAHR.C (49)					
	NADH dehydrogenase I, D subunit	K.LLNVEVPLR.A (44)	46315	62	2	5.36	2
E6	isocitrate dehydrogenase, NADP-dependent	R.LIDDMVASALK.W (76)	46087	215	10	5.28	8
		K.YYDLSVQNR.D (59)					
	GTP-binding protein TypA	R.FSVNDSPFAGR.E (65)	67658	65	1	5.08	1
	Methionine adenosyltransferase	K.ILQQLAADR.H (62)		95	2		
E7	chaperonin GroEL (Class 1 heat-shock protein chaperonin)	K.ALAGLTGANEDQTR.G(96)	57484	1226	72	5.09	12
		K.AAGVIDPTK.V(45)					
E8	malate dehydrogenase, NAD-dependent	R.GGGGEIVALLK.T (85)	33461	289	16	4.92	14
		K.VVGMAGVLDSAR.F (79)					
	translation elongation factor Tu	K.LIAPIAMDEGLR.F (65)	43045	311	17	5.16	10
		R.TVGSGVVSK.I (45)					

Spot number	Protein ID	Two highest scoring unique peptides (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage
E9	translation elongation factor Tu	R.TVGSGVVSK.I (56)	43045	120	6	5.16	8
		K.GSALAALEGR.D (44)					
	malate dehydrogenase, NAD-dependent	R.GGGGEIVALLK.T (58)	33461	141	5	4.92	6
		K.VIEIALDDTAK.A (45)					
	putative xylose isomerase-like protein	R.LPADPFAR.R (43)	32468			4.92	6
	protein of unknown function DUF28 [	K.ALGEAEGVK.L (36)	26975			4.97	3
	OmpA/MotB	K.AFGEANPR.V (48)	11983	166	8	5.76	38
		R.NASVQSYLTAHGIPAAEIATK.A (40)					
	alcohol dehydrogenase superfamily, zinc-containing	R.FGMTDFLNTR.G (58)	43051	74	2	5.5	4
		R.TALEACHR.G (42)					
	beta-ketothiolase	M.SDVFLLSAAR.T (87)	40702	88	2	5.7	2
E11	OmpA/MotB	K.TLNVDNQTITAVGGR.T (84)	25899	57	3	7.03	11
		R.LEGQVGYLDAGNK.T (35)					
	OmpA/MotB	K.YNLGLSAR.R (44)	11983	60	3	5.76	14
		K.AFGEANPR.V (42)					
	alcohol dehydrogenase superfamily, zinc-containing	R.AAVAFEPK.K (42)	43051	82	2	5.5	5
E12	Acetyl-CoA C-acetyltransferase	R.SGVAAEKFDR.V (79)	41038	530	15	5.8	14
		R.AMGTFAQDTANAYQLTR.E (73)					
	hypothetical protein Saro_1919	K.GGIDADKDR.V (48)	46531	69	3	8.37	2
F1	hypothetical protein Saro_1919	R.ITEQGLSTGALK.A (80)	46531	102	4	8.37	5
		R.RFPGEAVR.I (38)					
	cobalt chelatase, pCobS small subunit	R.VLVAEYYQR.V (52)	36373	62	2	5.34	5
		K.LTLLDQNR.V (37)					
	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	K.SDLLIVVGAR.N (55) - 1 good peptide	35135	54	1	5.34	3
F2	phosphoserine aminotransferase	K.EGAAYDIAGYR.D (67)	41426	201	7	5.33	8
		K.VLGGEGGHGVIILGPR.A (44)					
	succinyl-CoA synthetase, beta subunit	K.LSFDSNALYR.H (63)	42798	178	4	5.36	5
		M.DGQLLVLDAK.L (61)					
F3	2-oxoglutarate dehydrogenase, E2 component, dihydrolipoamide succinyltransferase	K.LGFMSFFAK.A (33)	42939	69	20	5.37	8
		K.AACLALKDIPAVNAR.I (24)					
F4	translation elongation factor Tu	K.LLDQGEAGDNVGALIR.G (91)	43045	334	8	5.16	9
		R.QVGVPALVVYMNK.V (66)					
	2-oxoglutarate dehydrogenase, E2 component, dihydrolipoamide succinyltransferase	R.AVLEYGIDPATVK.G (54)	42939	52	2	5.37	3
F5	TonB-dependent receptor	R.DYSACSLASATK.C (74)	107722	201	8	6.02	4
		K.SVAPVTSINNEDIK.T (62)					
F6	succinate CoA transferase	K.VMDAAAAASLIK.S(72)	55310	246	8	5.79	8
		R.NAYISIFMAPSTAK.G(64)					
	Acetyl-CoA hydrolase	K.VMDAAAAASLIK.S (74)	55421			5.79	8
		R.NAYISIFMAPSTAK.G (66)					
	ATP synthase F1, alpha subunit	K.AIDALVPIGR.G (57)	55106			5.84	8
		R.DSKDLGDATK.A (53)					
F7	Electron-transferring-flavoprotein dehydrogenase	K.GAVMGVATGDMGVAR.D (59)	59647	143	7	5.89	10
		K.YTLFAEGAR.G (56)					
	ATP synthase F1, alpha subunit	R.ADILAEIR.D (57)	55106	148	5	5.84	3
		K.AIDALVPIGR.G (50)					
F8	glutathione S-transferase-like	R.SHPVLAPSAGK.V (45)	24700			5.64	4
F9	Phasin	K.ILSTGLQELGK.G (81)	33963	233	5	9.56	13
		K.SQASFGELGEFAK.G (78)					
f10	transcriptional regulator, ModE family	M.SVATSGGAPR.A (46)	16037			5.27	6
F11	ribosomal protein L9	R.AASNAGHLYGSVSVR.D(102)	21643	5027	247	4.84	76
		R.IEKLGTIGDEVTVK.D(69)					
F12	adenylate kinases	R.TAAQAESLDGILASR.G (91)	23112	192	5	5.08	19
		K.TAPILPIYEAR.G (44)					
	NADH-quinone oxidoreductase, E subunit	R.SAVMPLLDLAQR.Q(69)	24975	172	9	5.16	9
		R.YLDMPIIR.V(30)					
	ribosomal protein L9	K.LGTIGDEVTVK.D (69)	21643	167	9	4.84	19
		MQIILLER.I (55)					

Spot number	Protein ID	Two highest scoring unique peptides (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage
G1	peptidyl-prolyl cis-trans isomerase, cyclophilin type	K.LVLSLDSGDVVIK.L (91)	16756	324	6	5.84	16
	Phasin	K.SQASFGELGEFAK.G (82)	33963			9.56	6
	NADH-quinone oxidoreductase, E subunit [	R.SAVMPLLDLAQR.Q (62)	25086	114	3	5.16	5
G2	translation elongation factor Ts [Novosphingobium aromaticivorans DSM 1	K.TPIAQVVEAAGK.A (91)	32125	551	24	5.09	29
		K.ENALLSQLFVMDNK.T (77)					
G3	translation elongation factor Ts	R.TAAEGLVGVAVAGTK.G (103)	32125	114	4	5.09	12
		K.IALVDYVR.F (41)					
	putative sugar decarboxylase	R.AVSNFLSQALTNGK.L (59)	36947			5.11	10
		R.CTESLLFEMHR.T (37)					
	translation elongation factor Tu	K.GSALAALEGR.D (45)	43045			5.16	5
		R.DDAIGKDSIK.E (39)					
G4	glyceraldehyde-3-phosphate dehydrogenase, type I	R.VPTPNVSVVDLTFVPK.R (81)	36036	231	13	6	24
		R.AAAMSMIPTTTGAAR.A (69)					
	transcriptional regulator, LacI family protein	R.MLVDLLTR.R (52)	33561	58	2	5.79	2
	putative sugar decarboxylase	R.AVSNFLSQALTNGK.L (48)	36947			5.11	4
G5	succinyl-CoA synthetase, beta subunit	K.LNGEFPANFLDVGGGATK.E (69)	42798	361	11	5.36	16
		R.LYVTDGVDIAK.E (62)					
	Aspartate-semialdehyde dehydrogenase, USG-1 related	R.EMLNILAER.E (33)	38621			5.68	2
G6	DNA-directed RNA polymerase, alpha subunit	R.VLLSSLQGAAVTSIR.I (94)	38947	302	11	4.94	18
		K.NDNIIYIGDLVQK.T (73)					
G7	molybdenum cofactor synthesis domain-containing protein [Caulobacter sp	R.IVGLPGNPSSALVTAR.L(68)	41264	114	2	4.97	4
	Phosphoglycerate kinase	R.EDLNVPMADGK.V (50)	41203			4.97	11
		K.TLDDIADVTGK.A (43)					
	Molybdopterin binding domain:MoeA, N-terminal, domain I and II:MoeA, C-terminal, domain IV	R.IVGLPGNPSSALVTAR.L (70)	41375			4.97	4
	putative oxidoreductase	R.MAVAADLGK.A (50)	39120			5.07	2
G8	glutamate synthases, NADH/NADPH, small subunit	R.VTGENDVTGLECVR.V (80)	52955	326	18	5.89	18
		R.FAMEFLTQQNK.R (65)					
G9	glutamate synthases, NADH/NADPH, small subunit	R.AVDEALMGVSELPR.X (58)	52955	75	2	5.89	4
		R.NAGDDEIR.A (44)					
	nitrogen regulatory protein P-II	R.EALSGLGVAGMTVSEVK.G (77)	12090	266	8	5.13	25
G10	chaperonin Cpn10	K.TAGGIIIPDSAK.E (73)	10262	317	14	5.12	59
		K.VSGEDLLIMK.E (65)					
	nitrogen regulatory protein P-II	R.EALSGLGVAGMTVSEVK.G(70)	11979	70	1	5.14	15
G11	Nucleoside-diphosphate kinase	K.MLEEAGLR.V (49)	15516	79	4	5.12	19
		R.TFSIIKPDATRR.N (27)					
G12	Aspartate-semialdehyde dehydrogenase, USG-1 related	K.GAALNAVQIAELLGR.R (68)	38621	173	9	5.68	14
		R.EMLNILAER.E (45)					
	Amine dehydrogenase	R.GFVQAGTGVYDSATGK.Y (77)	42431	113	2	5.75	8
		K.AQMTMLTYTGLVYTAK.L (73)					
H2	Tyrosine transaminase	R.INIAGLTTGNIEK.F (83)	47192	261	8	5.96	12
		R.VGAFYAVTSDADALAR.A (73)					
	hypothetical protein Saro_1919	R.EYVAYLQAIDAR.R (80)	46531	319	17	8.37	11
		K.TLLVSAMLGQNAYSAK.D (72)					
	Aspartate-semialdehyde dehydrogenase, USG-1 related	R.AIFVGDPVEPK.K (38)	38621	38	1	5.68	3
H2	translation elongation factor Ts	K.ENALLSQLFVMDNK.T (102)	32125	385	13	5.09	24
		K.ALDETGGDFEAAVDALR.A (94)					
	Fructose-bisphosphate aldolase	K.EILAQLDALPEGTQVMLK.L (94)	32291	179	5	5.06	13
		K.GFIAALDQSGGSTPK.A (75)					
H3	S-adenosylhomocysteine hydrolase	K.VACVAGFGDVGKGSAASLR.N(89)	50877	158	4	5.33	6
		K.SADYENHVYVLPK.H(74)					
	threonine synthase	K.IMQPFVGDSLTPER.L(75)	50463	910	79	5.35	6
		R.LLFDAAGR.D(42)					
	hypothetical protein Saro_1190	K.VACVAGFGDVGK.G (79)	50877			5.33	9
		R.IFDWDDGSGR.T (69)					
	translation elongation factor Ts	R.TAAEGLVGVAVAGTK.G (95)	32125	118	5	5.09	12
		K.TPIAQVVEAAGK.A (41)					
	acetyl-CoA carboxylase, biotin carboxylase	K.AAFGDATVYLEK.Y (44)	49787	44	1	5.53	2

Spot number	Protein ID	Two highest scoring unique peptides (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage
H4	malate dehydrogenase, NAD-dependent	K.IALIGAGNIGGTLAHLAAQK.E (86)	33461	506	16	4.92	33
		K.VIEIALDDTAK.A (80)					
H5	ketol-acid reductoisomerase	R.YSISNTAEYGDITTGPR.L (113)	36252	274	8	5.57	14
		R.ADVDVIMIAPK.G (71)					
H6	UDP-glucose pyrophosphorylase	K.AVFPVAGLGTR.F (83)	32293	151	4	5.64	8
		R.SDMAADVR.A(70)					
	Nucleotidyl transferase	K.AVFPVAGLGTR.F (71)	32404	720	26	5.62	6
		R.SDMAADVR.A (66)					
	ketol-acid reductoisomerase	R.ADVDVIMIAPK.G (61)	36141	108	3	5.57	3
		R.VLEDIQAGR.F (50)					
H7	putative nucleotide di-P-sugar epimerase or dehydratase	R.GMVGGAIVR.K (42)	35012	97	14	5.76	8
		R.KLMDVAR.L (41)					
H8	cysteine synthases	R.VLGFNYDTGER.Y (60)	32414	217	7	9.52	14
		MKADSVLATIGK.T (59)					
H9	Peptidase S1C, Do	R.IQPLSDDLAASLGLPK.R (93)	51818	355	11	6.06	15
		R.TGEFVQAVEPAQGAAK.A (88)					
H10	aldehyde dehydrogenase	K.AFVSYSQTTLDER.K (70)	50471	123	7	5.24	5
H11	chaperonin GroEL	K.LENVTLGMLGQAK.K (91)	57484	341	10	5.09	16
H12	Phosphopyruvate hydratase	K.AGDGTTTATVLAQAIVR.E (79) R.SGETEDSTIADLAVATNCGQIK.T (114)	45276	400	29	4.75	23
		K.GLATAVGDEGGFAPNLSSTR.D (98)					
	chaperonin GroEL	K.LENVTLGMLGQAK.K (89)	57484	194	7	5.09	6
		K.SVAAGINPMDLK.R (52)					
	phosphoribosylamine--glycine ligase	K.KGGLIAGIDAAQADGAK.V (46)	44602			4.8	3
J1	OmpW	K.LLATGVLADGNIDTVR.S(94)	25096	251	6	9.1	4
		K.TPYSLTFDAK.K(71)					
J2	DNA-directed RNA polymerase, alpha subunit	R.LQLSATGPGEVR.A(70)	38947	219	6	4.94	3
J3	heat shock protein Hsp90	K.SDITSEEYTDFYR.S(95)	68507	1597	52	4.94	19
		R.EMIQESPLLSAIR.K(75)					
J4	phosphoserine aminotransferase	K.EGAAYDIAGYR.D(81)	41426	623	54	5.33	29
J5	Chaperone DnaK (Class 1 heat-shock protein molecular chaperone)	K.MLGQFDLVGIPAAR.R(113)	68013	10913	320	4.79	45
		K.KAIEGGDTDEINAK.A(112)					
J6	trigger factor (prolyl isomerase)	K.LVVPVSDAEVDEAVAR.I(102)	55114	6669	216	4.71	43
		K.AAAEETPAGEAGEATEEAAPK.K(103)					
J7	NusA antitermination factor	R.AELLAIANAVASEK.M(90)	61419	420	11	4.73	3
		R.ALEALER.R(56)					
J8	Saccharopine dehydrogenase	K.TNIGDIATGTK.D(72)	44496	1333	44	4.92	11
		K.AVLPEPASLGATTK.G(72)					
J9	Pyruvate, phosphate dikinase	R.TFGDAANPLLVSVR.S(93)	96595	1290	41	5.04	11
		R.AIFEAALDVAQK.S(85)					
J10	translation elongation factor G	K.YCVQSIIDR.L(67)	76577	2113	133	4.94	17
		R.LSAEDPSFR.V(63)					



8.2 Figure A.1 A representative 2-D preparative gel showing additional proteins identified in the pH 4-7 linear range and 6-100 kDa range using LC-MS/MS and not all discussed in chapters 4-7.

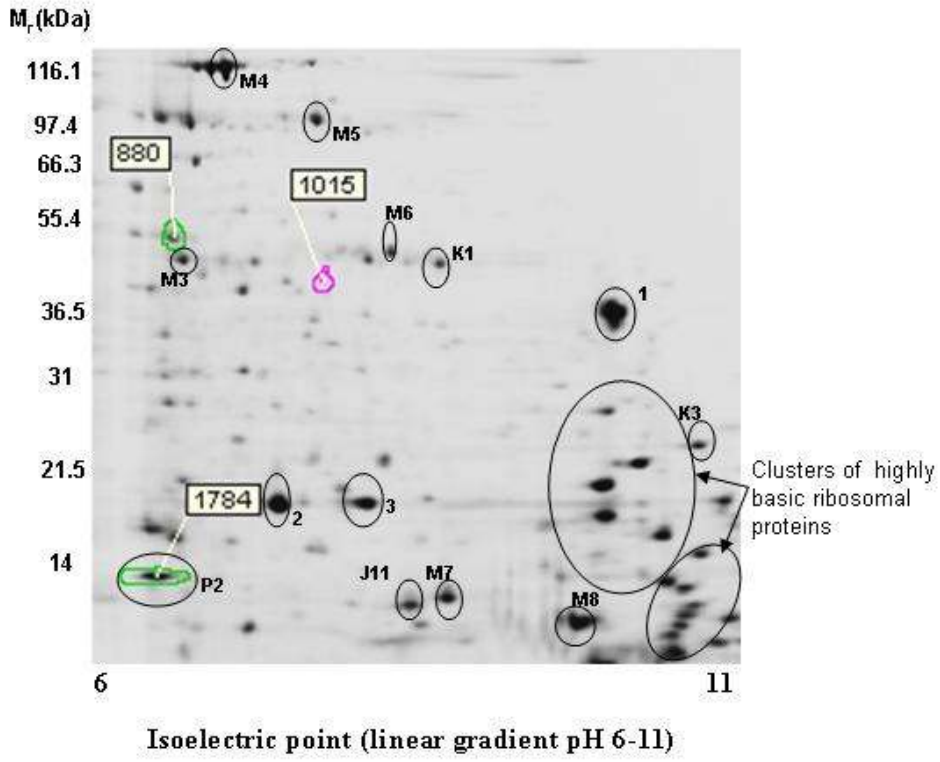
Spot numbers match spots represented in 8.1 (previous pages). Mass spec data and peptide matches are given in 8.1, Appendix A table.

8.3 Appendix B Additional protein identifications not all discussed in Chapters 4-7.

Protein identifications represent proteins at low / high dO_2, constant pH in the pH 6-11 linear range.

Spot number	Protein ID	Two highest scoring unique peptides (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage
J11	Endoribonuclease L-PSP	K.YQSNPGALILDGAEVK.A(94)	18095	2311	67	9.02	43
		K.LDFAGANEGFK.Q(82)					
K1	amine oxidase, flavin-containing	K.SLNAGALDTAVTAEDK.A(127)	58264	220	16	8.65	7
		K.VQILEFQNR.A(44)					
K2	ribosomal protein L10	R.NLGMTVAQSTALR.T(86)	17838	792	19	9.13	20
		K.ALATMPSLDELR.A(59)					
K3	ribosomal protein S5	R.ATFDALQEQTSPK.S(96)	25190	2732	77	10.16	36
		R.AVFESLGVHDVVTK.S(94)					
K4	translation initiation factor IF-2	K.AGDTLEVFEVEER.A(83)	95265	660	20	8.48	8
		K.VGDIFVVGTQSGR.V(85)					
K5	6-phosphogluconate dehydratase	R.AAYLDLIDR.A(70)	64604	561	19	6.26	6
		R.EDAPAPIQAMGVGR.E(69)					
K6	ribosomal protein L6	R.TLVSNLMTGVTEGYTK.T(120)	19366	3216	60	9.77	33
		K.TPDNTTVLISGIDK.Q(96)					
K7	succinate dehydrogenase and fumarate reductase iron-sulfur protein	K.SELDPSLTFR.R(57	30625	99	3	7.59	3
K8	Dihydroorotate oxidase	R.MGFNNDGSEAVAQR.L(90)	36996	540	11	9.64	6
		R.GVSQTPIFLK.L(61)					
K9	ribosomal protein S9	R.LVINQPFQVAGR.D(60)	18796	925	51	10.31	30
		K.GGGLSGQAGAVK.H(54)					
K10	ribosomal protein L11	K.DLNANDIDQATK.I(81)	15303	1668	86	9.68	28
		R.SQLAAIAEAK.M(74)					
K11	ribosomal protein L17	K.KLFEVLAER.Y(71)	15927	2167	162	9.86	39
		R.NMSAALIKHEQITTTLPK.A(53)					
K12	ribosomal protein L24	K.VVVGGVNIATR.H(86)	11329	2523	117	10.28	67
		R.FEAPLHVSK.V(49)					
	Glycine hydroxymethyltransferase	R.SPFVTSGIR.L(56)	47410	2534	86	6.13	16
		R.SGIILTNDEDLAK.K(58)					
	Citrate synthase I	R.SGTLGPDVVDIR.K(75)	48054	1007	46	6.13	10
		R.IFILHADHEQNASTSTVR.L(60)					
M1	ribosomal protein L5	K.GKYDVEIAQAMQAK.F(69)	21946	5455	323	9.45	28
		K.VQTAAAEMEK.I(54)					
M2	ribosomal protein L3	K.AEVPLKAEVAEFR.V(79)	26075	2834	185	9.71	10
		R.TQQNLEIVR.T(66)					
M3	inosine-5'-monophosphate dehydrogenase	K.AAMGYTGSATIEDLR.H(109)	51851	1413	34	6.57	14
		K.ISGIPVVEASGK.L(72)					
M4	2-oxoglutarate dehydrogenase, E1 component	R.LLSDINGASDAETR.K(112)	103902	1330	39	6.86	7
		K.VVYDLFEAR.D(55)					
M5	TonB-dependent siderophore receptor	R.TTDAGGFVSNVGIK.R(107)	83919	2076	39	6.98	15
		K.SGYGSYYADSSR.T(95)					
M6	Pyruvate kinase	R.IAAQVEADNAYNAR.I(97)	53208	1228	25	8.29	11
		K.ILATVGPASSSPEMLEK.L(81)					
M7	ribosomal protein L10	R.NLGMTVAQSTALR.T(86)	17838	2204	71	9.13	21
		K.ALATMPSLDELR.A(59)					
M8	ribosomal protein L18	K.AAGITAVVFDR.G(78)	12559	591	13	11.3	18
		K.SIGSNVDAAVK.V(81)					
M9	ribosomal protein S8	R.NGLGITIVSTPR.G(70)	13477	100	2	10.32	9
		K.YFEGQPAIK.H(48)					
M10	TonB-dependent receptor	R.TPSLQIESSFGR.T(94)	84231	1282	28	6.88	9
		R.VATGYLGPAVQDR.V(90)					
M11	TolB-like	R.LTNTPGINIGGSFSPDGR.K(104)	51546	1185	32	8.29	15
		R.FITNGQATALTPR.Y(90)					
M12	Acetyl-CoA C-acetyltransferase	M.AQFSASDPVVILSYAR.T(113)	41038	5479	119	5.8	43
		R.AMGTFAQDTANAYQLTR.E(105)					

Spot number	Protein ID	Two highest scoring unique peptides (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage
P1	malate dehydrogenase, NAD-dependent	K.VIEIALDDTAK.A(75)	33461	4535	184	4.92	36
		K.VVGMAGVLDSAR.F(87)					
P2	OmpW	K.LLATGVLADGNIDTVR.S(94)	25096	851	16	9.1	10
		K.TPYSLTFDAK.K(71)					
1	Phasin	K.AATAQTPTTAGVTETTK.E(86)	33963	1295	419	9.56	55
		K.SQASFGELGEFAK.G(85)					
2	Acetoacetyl-CoA reductase	R.GIGEAISLALKEQGR.Q(98)	25370	8382	339	6.92	47
		R.QVVANYAGNEEK.A(73)					
3	hypothetical protein Saro_1378	R.TQAAQIQQIQEAGER.E(112)	24314	9327	274	9.1	49
		K.AQIQAQIQPLVAK.Y(93)					
	Acetoacetyl-CoA reductase	K.AFTEATGIPTVR.W(80)	25370	539	12	6.92	22
		R.QVVANYAGNEEK.A(47)					



8.4 Figure B.1. A representative 2-D DIGE gel showing additional proteins identified in the pH 6-11 linear range using LC-MS/MS and not all discussed in chapters 4-7.

Spot numbers correspond to the spot numbers represented in the previous table. The green circles represent proteins with increased abundance for one physiological state and the pink circle represents the protein with increased abundance for the opposing physiological state as determined by BVA analysis.

8.5 Appendix C Additional protein identifications not all discussed in Chapters 4-7.

Protein identifications represent proteins at low / high dO2, pH not controlled, linear range pH 4-7.

Protein ID	Two highest scoring unique peptides (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage
chaperonin GroEL	K.VGGATEVEVK.E(72)	57484	571	29	5.09	6
	K.SVAAGINPMDLK.R(53)					
	K.LAGGVAVIK.V(47)					
chaperonin GroEL	K.ALAGLTGANEDQTR.G(85)	57484	3666	170	5.09	20
	K.SVAAGINPMDLK.R(63)					
	K.VIEDLKGR.S(38)					
	K.LAGGVAVIK.V(64)					
chaperonin GroEL	K.ALAGLTGANEDQTR.G(87)	57484	2302	91	5.09	12
	K.EGVITVEEAK.G(61)					
	K.SVAAGINPMDLK.R(57)					
NifU	K.VDEPSQTISDAR.F(83)	33542	522	13	4.72	3
ATP synthase F1, beta subunit	R.TIAMDSTDGLTR.G(68)	51090	522	13	4.82	2
NifU	K.VDEPSQTISDAR.F(73)	33542	334	10	4.72	3
ATP synthase F1, beta subunit	R.TIAMDSTDGLTR.G(68)	51090	94	2	4.82	2
OmpA/MotB	R.VPTADGVR.E(46)	11983	70	12	5.76	14
phasin	K.GNVEALVESTK(84)	33852	9708	319	9.56	23
	K.ILSTGLQELGK.G(81)					
Fructose-bisphosphate aldolase	K.NNGIIASFSR.A(64)	32291	147	17	5.06	5
	R.ALLEDLR.H(32)					
chaperonin GroEL	K.EGVITVEEAK.G(45)	57484	216	9	5.09	5
	K.SVAAGINPMDLK.R(50)					
nitrogenase molybdenum-iron protein beta chain	R.FVMEVAR.L(41)	44210	110	65	5.37	1
hypothetical protein BRADO5426	M.SATTSPSLETVR.I(61)					
alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen	R.EGVGLADR.G(45)	20864	169	21	4.79	7
	R.NAEELVR.K(47)					
alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen	R.NAEELVR.K(48)	20864	168	19	4.79	7
nitrogenase iron protein	R.LGGLICNER.Q(68)	31859	397	45	4.94	12
	K.YANSGGVR.L(42)					
chaperonin Cpn10	K.TAGGIIIPDSAK.E(83)	10262	4116	440	5.12	64
	K.VSGEDLLIMK.E(80)					
	K.EKPAEGEIVAVGTGTR.A(48)					
Superoxide dismutase	R.TNTAIEGTELADK.S(78)	22197	577		5.22	10
	K.NVRPDYLK.Q(33)					
2-keto-3-deoxy-phosphogluconate aldolase	R.TPAGLEAIR.A(86)	20668	480	21	5.16	7
	R.VLEVTLR.T(39)					
ribosomal protein L7/L12	K.IEEAGGTVEIK.X(81)	12683	2687	164	4.65	43
	R.AITSLGLTEAK.A(77)					
	K.IQVIKEVR.A(61)					
	K.AEAEDIKK.K(50)					
alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen	K.FVDLTDADIAGK.W(83)	20864	1291	168	4.79	14
	R.EGVGLADR.G(72)					
	R.NAEELVR.K(48)					
2-oxoglutarate dehydrogenase, E2 component, dihydrolipoamide succinyltransferase	K.EDVIAAAAQAK.A(90)	42828	1199	89	5.37	14
	K.EAIEEPTR.L(39)					
glyceraldehyde-3-phosphate dehydrogenase, type I	R.AAAMSMIPTTTGAAR.A(99)	36036	2281	112	6	17
	R.AVGEVLPELK.G(68)					
	R.VLISAPAK.N(43)					
	K.VAINGFGR.I(44)					
ATP synthase F1, alpha subunit	R.VVDALGNPIDGK.G(88)	55106	2051	128	5.84	14
	R.DSKDLGDATK.A(72)					
Acetyl-CoA hydrolase	K.VMDAAAAASLIK.S(72)	55421	1396	35	5.79	8
	R.IQNGIGGSGDFAR.N(65)					
nitrite and sulphite reductase 4Fe-4S region	R.VAEFSDQVAR.R				5.83	8
aromatic amino acid aminotransferase	R.TGDGATPVFK.A(66)	47081	416	20	5.97	6
	R.AMANGATIAR.A(52)					

Protein ID	Two highest scoring unique peptides (peptide score)	Mr	Score	Queries matched	pI	Sequence coverage
aspartate-semialdehyde dehydrogenase	R.AIFVGDPVEPK.K(64)	38510	332	24	5.68	12
	K.KILDPAVK.V(38)					
isocitrate dehydrogenase, NADP-dependent	K.TIEAEAAHGTVTR.H(104)	46087	1267	51	5.28	14
	R.LIDDMVASALK.W(72)					
	K.YYDLSVQNR.D(56)					
	K.CATITPDEAR.V(43)					
MotA/TolQ/ExbB proton channel	.MLIQTLAAAATSAAPQNK.F(48)	26220	169	23	9.47	10
2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	R.ADVTPASADVR.E(66)	28970	492	20	5.34	7
	K.TGINELLRD.-(41)					
Nucleoside-diphosphate kinase	R.DIMGATNPANADAGTIR.K(104)	15516	2463	258		
type I phosphodiesterase/nucleotide pyrophosphatase	R.AVTAGTVTVGTGR.F(100)	58770	2369	115	6.18	16
	K.TAIEENEAVDK.V(68)					
	R.LNQQALPQAAR.A(72)					
	R.AQFTGGLAR.L(52)					
TonB-dependent receptor, plug	K.TDASFVVNGR.L(65)	86259			6.44	6
	R.GIVPFDQNQTAR.D(48)					
Polyribonucleotide nucleotidyltransferase	K.VLEIDNR.G(36)	83008	738	46	5.34	3
	R.ANALNAAR.D(75)					
glutamine synthetase, type I	K.ALNAFTNPTTNSYK.R(81)	52267	648	59	4.94	7
aconitate hydratase 1	K.AAAQFGDISR.L(43)	96728	43	1	5.32	1
trigger factor	R.SEPLAAAQLR.A(73)	55114	2309	158	4.71	17
	K.NFADAGEDAVAK.D(68)					
	MQIVETSNEGLK.R					
	MQIVETSNEGLKR.A					
	R.APLYEDK.V(41)					
Chaperone DnaK	K.AIEGGDTDEINAK.A(92)	68013	987	62	4.79	4
	K.VIENSEGAR.T(70)					
ATP synthase F1, beta subunit	R.TIAMDSTDGLTR.G(78)	51090	1189	56	4.82	10
	K.IGLFGGAGVGK.T(61)					
	K.VIDLLAPYAK.G(40)					
Saccharopine dehydrogenase	K.TNIGDIATGTK.D(76)	44496	692	46	4.92	10
	K.AVLPEPASLGATTK.G(53)					
Inorganic diphosphatase	R.IVVEAIER.E(58)	20089	337	47	4.9	14
	R.SPFVPGAVVR.A(54)					
protein-L-isoaspartate(D-aspartate) O-methyltransferase	K.LGAVLADGGR.I(80)	20751	1143	66	5.02	21
	R.AMIESQLR.V(38)					
Inorganic diphosphatase	R.SPFVPGAVVR.A(48)	20089	284	52	4.9	10
	R.IVVEAIER.E(34)					
2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase	R.TPAGLEAIR.A(64)	20779	246	23	5.15	7
	R.VLEVTLR.T(35)					
ribosomal protein L9	K.LGTIGDEVTVK.D(73)	21643	2063	106	4.85	51
	R.IEAENAAR.R(64)					
	R.AASNAGHLYGSVSVR.D(55)					
	R.VALHPEVAVTVK.A(47)					
nitrogenase molybdenum-iron protein alpha chain	M.SVSTPTTIQEVK.D(60)	56185	171	7	6.3	3
	R.SMNYISR.H(38)					
phasin	K.AATAQTPTTAGVTETTK.E(101)	33963	5802	214	9.56	57
	K.TVAAAETLAPVAPK.A(92)					
	K.SQASFGELGEFAK.G(87)					
	K.GNVEALVESTK.I(84)					
	K.TTKPAAAPLPR.T(73)					
TonB-dependent receptor	K.ADAGDAGSAIIVTAK.T(108)	89707	1742	67	6.66	12
	R.SATAVTGAEIQK.I(76)					
	K.NYSVYPIAPR.Q(57)					
TonB-dependent receptor, plug	R.SQTFNSFSPEVVR.S(85)	86259	1166	39	6.44	12
	K.TDASFVVNGR.L(73)					
	R.GIVPFDQNQTAR.D(57)					
Acetoacetyl-CoA reductase	R.QVVANYAGNEEK.A(71)	25370	3241	145	6.92	21
	K.SGIHGFTK.A(65)					
	R.VAIVTGGTR.G(60)					
	M.ARVAIVTGGTR.G(50)					
hypothetical protein Saro_1378	R.TQAAQIQQIQEAGER.E(86)	24314	3264	132	9.1	33
	K.NAVQSQIPTTYK.A(85)					
	K.VTLVLDQR.V(66)					
	R.KVTLVLDQR.V(62)					
	K.SPKPDQAALR.T(39)					
Acetoacetyl-CoA reductase	R.QVVANYAGNEEK.A(70)	25370	1276	55	6.92	17
	R.VAIVTGGTR.G(60)					
ribosomal protein L18	K.SIGSNVDAAVK.V(60)	12559	170	7	11.3	11

References

Addison S. L., Foote S., Reid N., Lloyd-Jones G., *Novosphingobium nitrogenifigens* sp. nov., a polyhydroxyalkanoate-accumulating diazotroph isolated from a New Zealand pulp and paper wastewater. *Int J Syst Evol Microbiol* 2007, *57*, 2467-2471.

Agabian N., Unger B., *Caulobacter crescentus* cell envelope: effect of growth conditions on murein and outer membrane protein composition. *J Bacteriol* 1978, *133*, 987-994.

Anderson A. J., Dawes E. A., Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. *Microbiol Rev* 1990, *54*, 450-472.

Anderson A. J., Haywood G. W., Dawes E. A., Biosynthesis and composition of bacterial poly(hydroxyalkanoates). *Int J Biol Macromol* 1990, *12*, 102-105.

Auman A. J., Speake C. C., Lidstrom M. E., nifH sequences and nitrogen fixation in type I and type II methanotrophs. *Appl Environ Microbiol* 2001, *67*, 4009-4016.

Bae T.-J., Kim M.-S., Kim J.-W., Kim B.-W., *et al.*, Lipid raft proteome reveals ATP synthase complex in the cell surface. *Proteomics* 2004, *4*, 3536-3548.

Baev M., Baev D., Jansco Radek A., Campbell J., Growth of *Escherichia coli* MG1655 on LB medium: monitoring utilization of amino acids, peptides, and nucleotides with transcriptional microarrays. *Appl Microbiol Biotechnol* 2006, *71*, 317-322.

Baillon M., van Vliet A., Ketley J., Constantinidou C., Penn C., An iron-regulated alkyl hydroperoxide reductase (AhpC) confers aerotolerance and oxidative stress resistance to the microaerophilic pathogen *Campylobacter jejuni*. *J Bacteriol* 1999, *181*, 4798-4804.

Barrios A. F. G., Zuo R., Hashimoto Y., Yang L., *et al.*, Autoinducer 2 Controls Biofilm Formation in *Escherichia coli* through a Novel Motility Quorum-Sensing Regulator (MqsR, B3022). *J Bacteriol* 2006, *188*, 305-316.

Bastiaens L., Springael D., Wattiau P., Harms H., *et al.*, Isolation of adherent polycyclic aromatic hydrocarbon (PAH)-degrading bacteria using PAH-sorbing carriers. *Appl Environ Microbiol* 2000, *66*, 1834-1843.

Basu J., Protein palmitoylation and dynamic modulation of protein function. *Curr Sci (Bangalore)* 2004, *87*, 214-217.

Bosch G., Skovran E., Xia Q., Wang T., *et al.*, Comprehensive proteomics of *Methylobacterium extorquens* AM1 metabolism under single carbon and nonmethylotrophic conditions. *Proteomics* 2008, *8*, 3494-3505.

Brelles-Mariño G., Bedmar E. J., Detection, purification and characterisation of quorum-sensing signal molecules in plant-associated bacteria. *J Biotechnol* 2001, *91*, 197-209.

Brigle K. E., Newton W. E., Dean D. R., Complete nucleotide sequence of the *Azotobacter vinelandii* nitrogenase structural gene cluster. *Gene* 1985, *37*, 37-44.

Brotz-Oesterhelt H., Bandow J. E., Labischinski H., Bacterial proteomics and its role in antibacterial drug discovery. *Mass Spectrom Rev* 2005, *24*, 549-565.

Bukau B., Horwich A. L., The Hsp70 and Hsp60 Chaperone Machines. *Cell* 1998, *92*, 351-366.

Burris R. H., Nitrogenases. *J Biol Chem* 1991, *266*, 9339-9342.

Burris R. H., Roberts G. P., Biological nitrogen fixation. *Annu Rev Nutr* 1993, *13*, 317-335.

Busse H., Hauser E., Kämpfer P., Description of two novel species, *Sphingomonas abaci* sp. nov. and *Sphingomonas panni* sp. nov. *Int J Syst Evol Microbiol* 2005, *55*, 2565-2569.

Büttner K., Bernhardt J., Scharf C., Schmid R., *et al.*, A comprehensive two-dimensional map of cytosolic proteins of *Bacillus subtilis*. *Electrophoresis* 2001, *22*, 2908-2935.

Byrom D., Production of Poly-β-hydroxybutyrate: Poly-β-hydroxyvalerate Copolymers. *FEMS Microbiol Rev* 1992, *103*, 247-250.

Chen H., Talaty N. N., Takats Z., Cooks R. G., Desorption electrospray ionization mass spectrometry for high-throughput analysis of pharmaceutical samples in the ambient environment. *Anal Chem* 2005, *77*, 6915-6927.

Chen J. Y., Liu T., Zheng Z., Chen J. C., Chen G. Q., Polyhydroxyalkanoate synthases PhaC1 and PhaC2 from *Pseudomonas stutzeri* 1317 had different substrate specificities. *FEMS Microbiol Lett* 2004, *234*, 231-237.

Chen N., Sun W., Deng X., Hao Y., *et al.*, Quantitative proteome analysis of HCC cell lines with different metastatic potentials by SILAC. *Proteomics* 2008, *8*, 5108-5118.

Chiang P., Gordon R., Tal J., Zeng G., *et al.*, S-Adenosylmethionine and methylation. *FASEB J* 1996, *10*, 471-480.

Choi J., Lee S., High-level production of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) by fed-batch culture of recombinant *Escherichia coli*. *Appl Environ Microbiol* 1999, *65*, 4363-4368.

Chu K. H., Alvarez-Cohen L., Effect of nitrogen source on growth and trichloroethylene degradation by methane-oxidizing bacteria. *Appl Environ Microbiol* 1998, *64*, 3451-3457.

Clark B. N., Gutstein H. B., The myth of automated, high-throughput two-dimensional gel analysis. *Proteomics* 2008, *8*, 1197-1203.

Cochrane J. C., Strobel S. A., Riboswitch effectors as protein enzyme cofactors. *RNA* 2008, *14*, 993-1002.

Cote R., Gherna R., in: Gerhardt, P., Murray, R., Wood, W., Kreig, N., (Eds.), *Methods for general and molecular bacteriology*, ASM Press, Washington, DC 1994, pp. 155-178.

Darrow R. A., Knotts R. R., Two forms of glutamine synthetase in free-living root-nodule bacteria. *Biochem Biophys Res Commun* 1977, *78*, 554-559.

Davis D., Doudoroff M., Stanier R., Mandel M., Proposal to reject the genus *Hydrogenomonas*: taxonomic implications. *Int J Syst Bacteriol* 1969, *19*, 375–390.

Davis J. B., Coty V. F., Stanley J. P., Atmospheric Nitrogen Fixation by Methane-Oxidizing Bacteria. *J Bacteriol* 1964, *88*, 468-472.

Dawes E. A., Senior P. J., The role and regulation of energy reserve polymers in micro-organisms. *Adv Microb Physiol* 1973, *10*, 135-266.

De Mot R., Vanderleyden J., The C-terminal sequence conservation between OmpA-related outer membrane proteins and MotB suggests a common function in both gram-positive and gram-negative bacteria, possibly in the

interaction of these domains with peptidoglycan. *Mol Microbiol* 1994, *12*, 333-334.

De Oliveira V. C., Maeda I., Delessert S., Poirier Y., Increasing the carbon flux toward synthesis of short-chain-length--medium-chain-length polyhydroxyalkanoate in the peroxisome of *Saccharomyces cerevisiae* through modification of the beta-oxidation cycle. *Appl Environ Microbiol* 2004, *70*, 5685-5687.

Dean D. R., Setterquist R. A., Brigle K. E., Scott D. J., *et al.*, Evidence that conserved residues Cys-62 and Cys-154 within the *Azotobacter vinelandii* nitrogenase MoFe protein α-subunit are essential for nitrogenase activity but conserved residues His-83 and Cys-88 are not. *Mol Microbiol* 1990, *4*, 1505-1512.

Dilworth M. J., Eady R. R., Eldridge M. E., The vanadium nitrogenase of *Azotobacter chroococcum*. Reduction of acetylene and ethylene to ethane. *Biochem J* 1988, *249*, 745-751.

Dilworth M. J., Eldridge M. E., Eady R. R., The molybdenum and vanadium nitrogenases of *Azotobacter chroococcum*: effect of elevated temperature on N_2 reduction. *Biochem J* 1993, *289 (Pt 2)*, 395-400.

Dilworth M. J., Fisher K., Kim C.-H., Newton W. E., Effects on Substrate Reduction of Substitution of Histidine-195 by Glutamine in the α-Subunit of the MoFe Protein of *Azotobacter vinelandii* Nitrogenase. *Biochemistry* 1998, *37*, 17495-17505.

Dincturk H., Knaff D., The evolution of glutamate synthase. *Mol Biol Rep* 2000, *27*, 141-148.

Dong Z., Zelmer C. D., Canny M. J., McCully M. E., *et al.*, Evidence for protection of nitrogenase from O_2 by colony structure in the aerobic diazotroph *Gluconacetobacter diazotrophicus*. *Microbiology* 2002, *148*, 2293-2298.

Dos Santos P. C., Smith A. D., Frazzon J., Cash V. L., *et al.*, Iron-sulfur cluster assembly: NifU-directed activation of the nitrogenase Fe protein. *J Biol Chem* 2004, *279*, 19705-19711.

Drozd J., Tubb R., Postgate J., A chemostat study of the effect of fixed nitrogen sources on nitrogen fixation, membranes and free amino acids in *Azotobacter chroococcum*. *J Gen Microbiol* 1972, *73*, 221-232.

Durner J., Böger P., Ubiquitin in the Prokaryote *Anabaena variabilis*. *J Biol Chem* 1995, *270*, 3720-3725.

Durner J., Böhm I., Hilz H., Böger P., Posttranslational modification of nitrogenase. *Eur J Biochem* 1994, *220*, 125-130.

Eady R., in: Gibson, A. H., Newton, W. E., (Eds.), *Current perspectives in nitrogen fixation*, Australian Academy of Science, Canberra 1981, pp. 172-181.

Edgren T., Nordlund S., The fix ABCX genes in *Rhodospirillum rubrum* encode a putative membrane complex participating in electron transfer to nitrgenase. *J Bacteriol* 2004, *186*, 2052-2060.

Ekman M., Tollbäck P., Bergman B., Proteomic analysis of the cyanobacterium of the Azolla symbiosis: identity, adaptation, and NifH modification. *J Exp Bot* 2008, *59*, 1023-1034.

Ekman M., Tollback P., Klint J., Bergman B., Protein expression profiles in an endosymbiotic cyanobacterium revealed by a proteomic approach. *Mol Plant-Microbe Interact* 2006, *19*, 1251-1261.

El-Mansi M., Free CoA-mediated regulation of intermediary and central metabolism: an hypothesis which accounts for the excretion of alpha-ketoglutarate during aerobic growth of Escherichia coli on acetate. *Res Microbiol* 2005, *156*, 874-879.

El-Mansi M., Cozzone A. J., Shiloach J., Eikmanns B. J., Control of carbon flux through enzymes of central and intermediary metabolism during growth of Escherichia coli on acetate. *Curr Opin Microbiol* 2006, *9*, 173-179.

Elias J. E., Haas W., Faherty B. K. G., S. P., Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nat Methods* 2005, *2*, 667-675.

Ellmark P., Ingvarsson J., Carlsson A., Lundin S. B., *et al.*, Identification of protein expression signatures associated with *H. pylori* infection and gastric adenocarcinoma using recombinant antibody microarrays. *Mol Cell Proteomics* 2006, *5*, 1638-1646.

Euzéby J., List of new names and new combinations previously effectively, but not validly, published. *Int J Syst Evol Microbiol* 2006, *56*, 1-6.

Fei H., Chaillou S., Hirel B., Polowick P., *et al.*, Effects of the overexpression of a soybean cytosolic glutamine synthetase gene (GS15) linked to organ-specific promoters on growth and nitrogen accumulation of pea plants supplied with ammonium. *Plant Physiol Biochem* 2006, *44*, 543-550.

Fink A., Chaperone-mediated protein folding. *Physiol Rev* 1999, *79*, 425-449.

Finocchiaro G., Ito M., Ikeda Y., Tanaka K., Molecular cloning and nucleotide sequence of cDNAs encoding the alpha- subunit of human electron transfer flavoprotein. *J Biol Chem* 1988, *263*, 15773-15780.

Fisher K., Dilworth M. J., Kim C. H., Newton W. E., *Azotobacter vinelandii* nitrogenases with substitutions in the FeMo-cofactor environment of the MoFe protein: effects of acetylene or ethylene on interactions with H+, HCN, and CN. *Biochemistry* 2000, *39*, 10855-10865.

Fisher K., Dilworth M. J., Newton W. E., *Azotobacter vinelandii* vanadium nitrogenase: formaldehyde is a product of catalyzed HCN reduction, and excess ammonia arises directly from catalyzed azide reduction. *Biochemistry* 2006, *45*, 4190-4198.

Fitzgerald-Hughes D., Coleman D., O'Connell B., Differentially expressed proteins in derivatives of *Candida albicans* displaying a stable histatin 3-resistant phenotype. *Antimicrob Agents Chemother* 2007, *51*, 2793-2800.

Fleischmann R. D., Adams M. D., White O., Clayton R. A., *et al.*, Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 1995, *269*, 496-512.

Foor F., Janssen K., Magasanik B., Regulation of synthesis of glutamine synthetase by adenylylated glutamine synthetase. *Proc Nat Acad Sci USA* 1975, *72*, 4844-4848.

Fredrickson J. K., Balkwill D. L., Drake G. R., Romine M. F., *et al.*, Aromatic-degrading *Sphingomonas* isolates from the deep subsurface. *Appl Environ Microbiol* 1995, *61*, 1917-1922.

Fu H., Burris R. H., Ammonium inhibition of nitrogenase activity in *Herbaspirillum seropedicae*. *J Bacteriol* 1989, *171*, 3168-3175.

Fu H. A., Hartmann A., Lowery R. G., Fitzmaurice W. P., *et al.*, Posttranslational regulatory system for nitrogenase activity in *Azospirillum* spp. *J Bacteriol* 1989, *171*, 4679-4685.

Fujii K., Satomi M., Morita N., Motomura T., *et al.*, *Novosphingobium tardaugens* sp. nov., an oestradiol-degrading bacterium isolated from activated sludge of a sewage treatment plant in Tokyo. *Int J Syst Evol Microbiol* 2003, *53*, 47-52.

Fuxius S., Eravci M., Broedel O., Weist S., *et al.*, Technical strategies to reduce the amount of "false significant" results in quantitative proteomics. *Proteomics* 2008, *8*, 1780-1784.

Gallon J. R., Cheng J., Dougherty L. J., Gallon V. A., *et al.*, A novel covalent modification of nitrogenase in a cyanobacterium. *FEBS Lett* 2000, *468*, 231-233.

Garret R. H., Grisham C. M., in: *Biochemistry* 2 Edn., Harcourt College Publishers, Fort Worth 1999, pp. 927-948.

Garrity G. M., Bell J. A., Lilburn T., in: Brenner, D. J., Krieg, N. R., Staley, J. T., Garrity, G. M., (Eds.), *Bergey's Manual of Systematic Bacteriology*, Vol. 2 (The Proteobacteria), part C (The Alpha-, Beta-, Delta-, and Epsilonproteobacteria) 2 Edn., Springer, New York 2005, p 1.

Georgiadis M., Komiya H., Chakrabarti P., Woo D., *et al.*, Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*. *Science* 1992, *257*, 1653-1659.

Glaeser S., Kämpfer P., Busse H., Langer S., Glaeser J., *Novosphingobium acidiphilum* sp. nov., an acidophilic salt-sensitive bacterium isolated from the humic acid-rich Lake Grosse Fuchskuhle. *Int J Syst Evol Microbiol* 2009, *59*, 323-330.

Godoy F., Vancanneyt M., Martinez M., Steinbuchel A., *et al.*, *Sphingopyxis chilensis* sp. nov., a chlorophenol-degrading bacterium accumulates polyhydroxyalkanoate, and transfer of *Sphingomonas alaskensis* to *Sphingopyxis alaskensis* comb. nov. *Int J Syst Evol Microbiol* 2003, *53*, 473-477.

Gordon J., Shah V., Brill W., Feedback inhibition of nitrogenase. *J Bacteriol* 1981, *148*, 884-888.

Gottesman S., Proteases and their targets in Escherichia coli. *Annu Rev Genet* 1996, *30*, 465-506.

Govezensky D., Greener T., Segal G., Zamir A., Involvement of GroEL in nif gene regulation and nitrogenase assembly. *J Bacteriol* 1991, *173*, 6339-6346.

Gupta R. S., Mok A., Phylogenomics and signature proteins for the alpha Proteobacteria and its main groups. *BMC Microbiol* 2007, *7*, 106.

Hahne H., Wolff S., Hecker M., Becher D., From complementarity to comprehensiveness - targeting the membrane proteome of growing *Bacillus subtilis* by divergent approaches. *Proteomics* 2008, *8*, 4123-4136.

Hamana K., Miyagawa K., Matsuzaki S., Occurrence of sym-homospermidine as the major polyamine in nitrogen-fixing cyanobacteria. *Biochem Biophys Res Commun* 1983, *112*, 606-613.

Hamano Y., Matsuura N., Kitamura M., Takagi H., A novel enzyme conferring streptothricin resistance alters the toxicity of streptothricin D from broad-spectrum to bacteria-specific. *J Biol Chem* 2006, *281*, 16842-16848.

Han M.-J., Lee S. Y., Proteome profiling and its use in metabolic and cellular engineering. *Proteomics* 2003, *3*, 2317-2324.

Han M., Yoon S., Lee S., Proteome analysis of metabolically engineered *Escherichia coli* producing Poly(3-hydroxybutyrate). *J Bacteriol* 2001, *183*, 301-308.

Han M. J., Lee S. Y., The Escherichia coli proteome: past, present, and future prospects. *Microbiol Mol Biol Rev* 2006, *70*, 362-439.

Haqqani A., Kelly J., Stanimirovic D., Quantitative protein profiling by mass spectrometry using label-free proteomics. *Methods Mol Biol* 2008, *439*, 241-256.

Hardy R. W. F., Holsten R. D., Jackson E. K., Burns R. C., The Acetylene - Ethylene Assay for N2 Fixation: Laboratory and Field Evaluation. *Plant Physiol* 1968, *43*, 1185-1207.

Hartmann A., Burris R. H., Regulation of nitrogenase activity by oxygen in *Azospirillum brasilense* and *Azospirillum lipoferum*. *J Bacteriol* 1987, *169*, 944-948.

Hartmann A., Fu H., Burris R., Regulation of nitrogenase activity by ammonium chloride in *Azospirillum* spp. *J Bacteriol* 1986, *165*, 864-870.

Hecker M., A proteomic view of cell physiology of *Bacillus subtilis*-bringing the genome sequence to life. *Adv Biochem Eng Biotechnol* 2003, *83*, 57-92.

Hecker M., Antelmann H., Büttner K., Bernhardt J., Gel-based proteomics of Gram-positive bacteria: A powerful tool to address physiological questions. *Proteomics* 2008, *8*, 4958-4975.

Hecker M., Volker U., Towards a comprehensive understanding of *Bacillus subtilis* cell physiology by physiological proteomics. *Proteomics* 2004, *4*, 3727-3750.

Hennig L., WinGene/WinPep: User-Friendly software for the analysis of amino acid sequences. *BioTechniques* 1999, *26*, 1170-1172.

Herendeen S. L., VanBogelen R. A., Neidhardt F. C., Levels of major proteins of *Escherichia coli* during growth at different temperatures. *J Bacteriol* 1979, *139*, 185-194.

Hodges M., Enzyme redundancy and the importance of 2-oxoglutarate in plant ammonium assimilation. *J Exp Bot* 2002, *53*, 905-916.

Hong H., Patel D. R., Tamm L. K., van den Berg B., The Outer Membrane Protein OmpW Forms an Eight-stranded beta-Barrel with a Hydrophobic Channel. *J Biol Chem* 2006, *281*, 7568-7577.

Hu Y., Uttamchandani M., Yao S. Q., Microarray: a versatile platform for high-throughput functional proteomics. *Comb Chem High Throughput Screen* 2006, *9*, 203-212.

Huergo L., Souza E., Araujo M., Pedrosa F., *et al.*, ADP-ribosylation of dinitrogenase reductase in *Azospirillum brasilense* is regulated by AmtB-dependent membrane sequestration of DraG. *Mol Microbiol* 2006, *159*, 326-337.

James S., Melnyk S., Pogribna M., Pogribny I., Caudill M., Elevation in S-adenosylhomocysteine and DNA hypomethylation: potential epigenetic mechanism for homocysteine-related pathology. *J Nutr* 2002, *132*, 2361S-2366S.

Joerger R. D., Wolfinger E. D., Bishop P. E., The gene encoding dinitrogenase reductase 2 is required for expression of the second alternative nitrogenase from Azotobacter vinelandii. *J Bacteriol* 1991, *173*, 4440-4446.

Johnson D., Dos Santos P., Dean D., NifU and NifS are required for the maturation of nitrogenase and cannot replace the function of isc-gene products in *Azotobacter vinelandii*. *Biochem Soc Trans* 2005, *33*, 90-93.

Jossek R., Reichelt R., Steinbüchel A., In vitro biosynthesis of poly(3-hydroxybutyric acid) by using purified poly(hydroxyalkanoic acid) synthase of *Chromatium vinosum*. *Appl Microbiol Biotechnol* 1998, *49*, 3258-3266.

Jurat-Fuentes J. L., Adang M. J., A proteomic approach to study Cry1Ac binding proteins and their alterations in resistant *Heliothis virescens* larvae. *J Invertebrate Pathol* 2007, *95*, 187-191.

Kampfer P., Witzenberger R., Denner E. B., Busse H. J., Neef A., *Novosphingobium hassiacum* sp. nov., a new species isolated from an aerated sewage pond. *Syst Appl Microbiol* 2002, *25*, 37-45.

Kaplan M., *Novosphingobium aromaticivorans*: A potential initiator of primary biliary cirrhosis. *Amer J Gastroenterol* 2004, *99*, 2147-2149.

Kawahara K., Kuraishi H., Zähringer U., Chemical structure and function of glycosphingolipids of *Sphingomonas* spp and their distribution among members of the α-4 subclass of Proteobacteria. *J Ind Microbiol Biotechnol* 1999, *23*, 408-413.

Kawai F., Sphingomonads involved in the biodegradation of xenobiotic polymers. *J Ind Microbiol Biotechnol* 1999, *23*, 400-407.

Keasling J. D., Chou H., Metabolic engineering delivers next-generation biofuels. *Nat Biotechnol* 2008, *26*, 298-299.

Kelley S. T., Theisen U., Angenent L. T., St Amand A., Pace N. R., Molecular analysis of shower curtain biofilm microbes. *Appl Environ Microbiol* 2004, *70*, 4187-4192.

Kent H. M., Baines M., Gormal C., Smith B. E., Buck M., Analysis of site-directed mutations in the α-and β-subunits of *Klebsiella pneumoniae* nitrogenase. *Mol Microbiol* 1990, *4*, 1497-1504.

Kichise T., Fukui T., Yoshida Y., Doi Y., Biosynthesis of polyhydroxyalkanoates (PHA) by recombinant *Ralstonia eutropha* and effects of PHA synthase activity on in vivo PHA biosynthesis. *Int J Biol Macromol* 1999, *25*, 69-77.

Kim D. Y., Kim H. C., Kim S. Y., Rhee Y. H., Molecular characterization of extracellular medium-chain-length poly(3-hydroxyalkanoate) depolymerase genes from *Pseudomonas alcaligenes* strains. *J Microbiol* 2005, *43*, 285-294.

Kim S., Burgess B., Evidence for the direct interaction of the nifW gene product with the MoFe protein. *J Biol Chem* 1996, *271*, 9764-9770.

Klassen G., Souza E. M., Yates M. G., Rigo L. U., *et al.*, Nitrogenase switch-off by ammonium ions in *Azospirillum brasilense* requires the GlnB nitrogen signal-transducing protein. *Appl Environ Microbiol* 2005, *71*, 5637-5641.

Kleiner D., Kleinschmidt J., Selective inactivation of nitrogenase in *Azotobacter vinelandii* batch cultures. *J Bacteriol* 1976, *128*, 117-122.

Klose J., Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* 1975, *26*, 231-243.

Klugkist J., Haaker H., Inhibition of nitrogenase activity by ammonium chloride in *Azotobacter vinelandii*. *J Bacteriol* 1984, *157*, 148-151.

Kohler C., Wolff S., Albrecht D., Fuchs S., *et al.*, Proteome analyses of *Staphylococcus aureus* in growing and non-growing cells: a physiological approach. *Int J Med Microbiol* 2005, *295*, 547-565.

Kolibachuk D., Miller A., Dennis D., Cloning, molecular analysis, and expression of the polyhydroxyalkanoic acid synthase (phaC) gene from *Chromobacterium violaceum*. *Appl Environ Microbiol* 1999, *65*, 3561-3565.

Kolkman A., Olsthoorn M. M., Heeremans C. E., Heck A. J., Slijper M., Comparative proteome analysis of *Saccharomyces cerevisiae* grown in chemostat cultures limited for glucose or ethanol. *Mol Cell Proteomics* 2005, *4*, 1-11.

Kosako Y., Yabuuchi E., Naka T., Fujiwara N., Kobayashi K., Proposal of *Sphingomonadaceae* Fam. Nov., consisting of *Sphingomonas* Yabuuchi et al., 1990, *Erythrobacter* Shiba and Shimidu 1982, *Erythromicrobium* Yurkov et al., 1994, *Porphyrobacter* Fuerst et al., 1993, *Zymomonas* Kluyver and van Niel 1936, and

Sandaracinobacter Yurkov et al., 1997, with the type genus *Sphingomonas* Yabuuchi et al., 1990. *Microbiol Immunol* 2000, *44*, 563-575.

Koskinen R., Ali-Vehmas T., Kämpfer P., Laurikkala M., *et al.*, Characterization of *Sphingomonas* isolates from Finnish and Swedish drinking water distribution systems. *J Appl Microbiol* 2000, *89*, 687-696.

Kostyál E., Nurmiaho-Lassila E.-L., Puhakka J. A., Salkinoja-Salonen M., Nitrification, denitrification, and dechlorination in bleached kraft pulp mill wastewater. *Appl Microbiol Biotechnol* 1997, *47*, 734-741.

Kuchta K., Chi L., Fuchs H., Potter M., Steinbüchel A., Studies on the Influence of Phasins on Accumulation and Degradation of PHB and Nanostructure of PHB Granules in *Ralstonia eutropha* H16. *Biomacromolecules* 2007, *8*, 657-662.

Kuhla J., Oelze J., Dependence of nitrogenase switch-off upon oxygen stress on the nitrogenase activity in *Azotobacter vinelandii*. *J Bacteriol* 1988, *170*, 5325-5329.

Kwon H.-R., Kim Y.-C., Nucleotide Sequence and Secondary Structure of 5S rRNA from *Sphingobium chungbukense* JJ77. *J Microbiol* 2007, *45*, 79-82.

Laane C., Krone W., Konings W., Haaker H., Veeger C., Short-Term Effect of Ammonium Chloride on Nitrogen Fixation by *Azotobacter vinelandii* and by Bacteroids of *Rhizobium leguminosarum*. *Eur J Biochem* 1980, *103*, 39-46.

Larsen M. R., Roepstorff P., Mass spectrometric identification of proteins and characterization of their post-translational modifications in proteome analysis. *Fresenius' Journal of Analytical Chemistry* 2000, *366*, 677-690.

Lee I. Y., Kim M. K., Chang H. N., Park Y. H., Regulation of poly-β-hydroxybutyrate biosynthesis in nicotinamide nucleotide in *Alcaligenes eutrophus*. *FEMS Microbiol Lett* 1995, *131*, 35.

Lee S.-H., Pulakat L., Parker K. C., Gavini N., Genetic Analysis on the NifW by Utilizing the Yeast Two-Hybrid System Revealed that the NifW of *Azotobacter vinelandii* Interacts with the NifZ to Form Higher-Order Complexes. *Biochem Biophys Res Commun* 1998, *244*, 498-504.

Lee S. Y., Bacterial Polyhydroxyalkanoates. *Biotechnol Bioeng* 1996, *49*, 1-14.

Lemoigne M., Products of dehydration and of polymerization of β-hydroxybutyric acid. *Bull Soc Chem Biol* 1926, *8*, 770-782.

Lemos J., Luzardo Y., Burne R., Physiologic effects of forced down-regulation of dnaK and groEL expression in *Streptococcus mutans*. *J Bacteriol* 2007, *189*, 1582-1588.

Leys N. M., Ryngaert A., Bastiaens L., Verstraete W., *et al.*, Occurrence and phylogenetic diversity of *Sphingomonas* strains in soils contaminated with polycyclic aromatic hydrocarbons. *Appl Environ Microbiol* 2004, *70*, 1944-1955.

Li H. *Azotobacter vinelandii nitrogenase: multiple substrate reduction sites and effects of pH on substrate reduction and CO inhibition*. PhD Thesis, Virginia Polytechnic Institute and State University, Blacksburg 2002.

Liebergesell M., Sonomoto K., Madkour M., Mayer F., Steinbuchel A., Purification and characterization of the poly(hydroxyalkanoic acid) synthase from *Chromatium vinosum* and localization of the enzyme at the surface of poly(hydroxyalkanoic acid) granules. *Eur J Biochem* 1994, *226*, 71-80.

Lillemoen J., Cameron C., Hoffman D., The stability and dynamics of ribosomal protein L9: investigations of a molecular strut by amide proton exchange and circular dichroism. *J Mol Biol* 1997, *268*, 482-493.

Lin J., Huang S., Zhang Q., Outer membrane proteins: key players for bacterial adaptation in host niches. *Microbes Infect* 2002, *4*, 325-331.

Linkerhagner K., Oelze J., Cellular ATP levels and nitrogenase switchoff upon oxygen stress in chemostat cultures of *Azotobacter vinelandii*. *J Bacteriol* 1995, *177*, 5289-5293.

Linn T., Losick R., The program of protein synthesis during sporulation in *Bacillus subtilis*. *Cell* 1976, *8*, 103-114.

Linsen L., Locherbach J., Berth M., Becher D., Bernhardt J., Visual Analysis of Gel-Free Proteome Data. *IEEE Transactions on Visualization and Computer Graphics* 2006, *12*, 497-508.

Liu Z. P., Wang B. J., Liu Y. H., Liu S. J., *Novosphingobium taihuense* sp. nov., a novel aromatic-compound-degrading bacterium isolated from Taihu Lake, China. *Int J Syst Evol Microbiol* 2005, *55*, 1229-1232.

Ludden P., Lehman L., Roberts G., Reversible ADP-ribosylation of dinitrogenase reductase in a nifD- mutant of *Rhodospirillum rubrum*. *J Bacteriol* 1989, *171*, 5210-5211.

Ludden P., Roberts G., Regulation of nitrogenase activity by reversible ADP ribosylation. *Curr Top Cell Regul* 1989, *30*, 23-56.

Ludden P. W. (2001) Nitrogenase Complex. *In* Zheng, Y., Tickle, C., Jansson, R., Kehrer-Sawatzki, H., Cooper, D. N., Melino, G., Delves, P., Battista, J., Levitan, I., Roberts, K., Bynum, W. F., Harper, D., eds, Encyclopedia of Life Sciences. John Wiley & Sons Ltd.

Luli G., Strohl W., Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. *Appl Environ Microbiol* 1990, *56*, 1004-1011.

Madison L., Huisman G., Metabolic engineering of poly(3-hydroxyalkanoates): from DNA to plastic. *Microbiol Mol Biol Rev* 1999, *63*, 21-53.

Mancinelli R. L., The nature of nitrogen: an overview. *Life Support Biosph Sci* 1996, *3*, 17-24.

Mandon K., Michel-Reydellet N., Encarnacion S., Kaminski P. A., *et al.*, Poly-beta-hydroxybutyrate turnover in *Azorhizobium caulinodans* is required for growth and affects nifA expression. *J Bacteriol* 1998, *180*, 5070-5076.

Marino D., Frendo P., Ladrera R., Zabalza A., *et al.*, Nitrogen fixation control under drought stress. Localized or systemic? *Plant Physiol* 2007, *143*, 1968-1974.

Marouga R., David S., Hawkins E., The development of the DIGE system: 2D fluorescence difference gel analysis technology. *Anal Bioanal Chem* 2005, *382*, 669-678.

Martin D., Reinhold-Hurek B., Distinct Roles of P_{II}-Like Signal Transmitter Proteins and *amt*B in Regulation of *nif* Gene Expression, Nitrogenase Activity, and Posttranslational Modification of NifH in *Azoarcus* sp. Strain BH72. *J Bacteriol* 2002, *184*, 2251-2259.

Martinez-Argudo I., Little R., Shearer N., Johnson P., Dixon R., The NifL-NifA System: a Multidomain Transcriptional Regulatory Complex That Integrates Environmental Signals. *J Bacteriol* 2004, *186*, 601-610.

Merali S., Vargas D., Franklin M., Clarkson Jr A., S-adenosylmethionine and *Pneumocystis carinii*. *J Biol Chem* 2000, *275*, 14958-14963.

Mittendorf V., Bongcam V., Allenbach L., Coullerez G., *et al.*, Polyhydroxyalkanoate synthesis in transgenic plants as a new tool to study carbon flow through beta-oxidation. *Plant J* 1999, *20*, 45-55.

Moeck G. S., Coulton J. W., TonB-dependent iron acquisition: mechanisms of siderophore-mediated active transport. *Mol Microbiol* 1998, *28*, 675-681.

Moshiri F., Crouse B., Johnson M., Maier R., The "nitrogenase-protective" FeSII protein of *Azotobacter vinelandii*: overexpression, characterization, and crystallization. *Biochemistry* 1995, *34*, 12973-12982.

Moskowitz G. J., Merrick J. M., Metabolism of poly-beta-hydroxybutyrate. II. Enzymatic synthesis of D-(-)-beta hydroxybutyryl coenzyme A by an enoyl hydrase from *Rhodospirillum rubrum*. *Biochemistry* 1969, *8*, 2748-2755.

Muller S., Bley T., Babel W., Adaptive responses of *Ralstonia eutropha* to feast and famine conditions analysed by flow cytometry. *J Biotechnol* 1999, *75*, 81-97.

Nakagawa H., Kirimura K., Nitta T., Kino K., *et al.*, Recycle Use of *Sphingomonas* sp. CDH-7 Cells for Continuous Degradation of Carbazole in the Presence of $MgCl_2$. *Curr Microbiol* 2002, *44*, 251-256.

Nakamura J., Straub K., Wu J., Lou L., The glutamine hydrolysis function of human GMP synthetase. Identification of an essential active site cysteine. *J Biol Chem* 1995, *270*, 23450-23455.

Nandi B., Nandy R. K., Sarkar A., Ghose A. C., Structural features, properties and regulation of the outer-membrane protein W (OmpW) of *Vibrio cholerae*. *Microbiology* 2005, *151*, 2975-2986.

Neef A., Witzenberger R., Kämpfer P., Detection of sphingomonads and in situ identification in activated sludge using 16S rRNA-targeted oligonucleotide probes. *J Ind Microbiol Biotechnol* 1999, *23*, 261-267.

Neumann L., Spinozzi F., Sinibaldi R., Rustichelli F., *et al.*, Binding of the major phasin, PhaP1, from *Ralstonia eutropha* H16 to poly(3-hydroxybutyrate) granules. *J Bacteriol* 2008, *190*, 2911-2919.

Niederhoffer E., Naranjo C., Bradley K., Fee J., Control of *Escherichia coli* superoxide dismutase (sodA and sodB) genes by the ferric uptake regulation (fur) locus. *J Bacteriol* 1990, *172*, 1930-1938.

Nikaido H., Microdermatology: cell surface in the interaction of microbes with the external world. *J Bacteriol* 1999, *181*, 4-8.

Ninfa A. J., Atkinson M. R., PII signal transduction proteins. *Trends Microbiol* 2000, *8*, 172-179.

Noda I., Green P. R., Satkowski M. M., Schechtman L. A., Preparation and properties of a novel class of polyhydroxyalkanoate copolymers. *Biomacromolecules* 2005, *6*, 580-586.

Nordlund S., Ludden P., in: Klipp, W., Masepohl, B., Gallon, J., Newton, W., (Eds.), *Genetics and Regulation of Nitrogen Fixation in Free-Living Bacteria*, Kluwer Academic Publishers 2004, pp. 175-196.

O'Farrell P. H., High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 1975, *250*, 4007-4021.

O'Neill H., Mayhew S. G., Butler G., Cloning and Analysis of the Genes for a Novel Electron-transferring Flavoprotein from *Megasphaera elsdenii*. EXPRESSION AND CHARACTERIZATION OF THE RECOMBINANT PROTEIN. *J Biol Chem* 1998, *273*, 21015-21024.

Oeding V., Schlegel H. G., Beta-ketothiolase from *Hydrogenomonas eutropha* H16 and its significance in the regulation of poly-beta-hydroxybutyrate metabolism. *Biochem J* 1973, *134*, 239-248.

Okamoto S., Lezhava A., Hosaka T., Okamoto-Hosoya Y., Ochi K., Enhanced expression of S-adenosylmethionine synthetase causes overproduction of actinorhodin in *Streptomyces coelicolor* A3(2). *J Bacteriol* 2003, *185*, 601-609.

Oke V., Long S. R., Bacteroid formation in the *Rhizobium*-legume symbiosis. *Curr Opin Microbiol* 1999, *2*, 641-646.

Oshiman K.-i., Tsutsumi Y., Nishida T., Matsumura Y., Isolation and characterization of a novel bacterium, *Sphingomonas bisphenolicum* strain AO1, that degrades bisphenol A. *Biodegradation* 2007, *18*, 247-255.

Page W. J., Manchak J., Rudy B., Formation of poly(hydroxybutyrate-co-hydroxyvalerate) by Azotobacter vinelandii UWD. *Appl Environ Microbiol* 1992, *58*, 2866-2873.

Pal R., Bhasin V. K., Lal R., Proposal to reclassify [*Sphingomonas*] *xenophaga* Stolz et al. 2000 and [*Sphingomonas*] *taejonensis* Lee et al. 2001 as *Sphingobium xenophagum* comb. nov. and *Sphingopyxis taejonensis* comb. nov., respectively. *Int J Syst Evol Microbiol* 2006, *56*, 667-670.

Parsons J. F., Calabrese K., Eisenstein E., Ladner J. E., Structure and Mechanism of *Pseudomonas aeruginosa* PhzD, an Isochorismatase from the Phenazine Biosynthetic Pathway. *Biochemistry* 2003, *42*, 5684-5693.

Peekhaus N., Conway T., What's for Dinner?: Entner-Doudoroff Metabolism in *Escherichia coli*. *J Bacteriol* 1998, *180*, 3495-3502.

Peng X., Egashira T., Hanashiro K., Masai E., *et al.*, Cloning of a *Sphingomonas paucimobilis* SYK-6 gene encoding a novel oxygenase that cleaves lignin-related biphenyl and characterization of the enzyme. *Appl Environ Microbiol* 1998, *64*, 2520-2527.

Peters J. W., Fisher K., Dean D. R., Nitrogenase Structure and Function: A Biochemical-Genetic Perspective. *Annu Rev Microbiol* 1995, *49*, 335-366.

Peters J. W., Stowell M. H. B., Soltis S. M., Finnegan M. G., *et al.*, Redox-dependent structural changes in the nitrogenase P-cluster. *Biochemistry* 1997, *36*, 1181-1187.

Pettinari J. M., Chaneton L., Vazquez G., Steinbuchel A., Mendez B. S., Insertion sequence-like elements associated with putative polyhydroxybutyrate regulatory genes in *Azotobacter* sp. FA8. *Plasmid* 2003, *50*, 36-44.

Pettinari M., Vázquez G., Silberschmidt D., Rehm B., *et al.*, Poly(3-hydroxybutyrate) synthesis genes in *Azotobacter* sp. strain FA8. *Appl Environ Microbiol* 2001, *67*, 5331-5334.

Pham D., Burgess B., Nitrogenase reactivity: effects of pH on substrate reduction and CO inhibition. *Biochemistry* 1993, *14*, 13725-13731.

Philmus B., Christiansen G., Yoshida W. Y., Hemscheidt T. K., Post-translational Modification in Microviridin Biosynthesis. *ChemBioChem* 2008, *9*, 3066-3073.

Poirier Y., Erard N., MacDonald-Comber Petetot J., Synthesis of polyhydroxyalkanoate in the peroxisome of *Pichia pastoris*. *FEMS Microbiol Lett* 2002, *207*, 97-102.

Poirier Y., Somerville C., Schechtman L. A., Satkowski M. M., Noda I., Synthesis of high-molecular-weight poly([R]-(-)-3-hydroxybutyrate) in transgenic *Arabidopsis thaliana* plant cells. *Int J Biol Macromol* 1995, *17*, 7-12.

Poirier Y., Ventre G., Caldelari D., Increased flow of fatty acids toward beta-oxidation in developing seeds of *Arabidopsis* deficient in diacylglycerol acyltransferase activity or synthesizing medium-chain-length fatty acids. *Plant Physiol* 1999, *121*, 1359-1366.

Pope M. R., Murrell S. A., Ludden P. W., Covalent modification of the iron protein of nitrogenase from *Rhodospirillum rubrum* by adenosine

diphosphoribosylation of a specific arginine residue. *Proc Natl Acad Sci USA* 1985, *82*, 3173-3177.

Post E., Kleiner D., Oelze J., Whole cell respiration and nitrogenase activities in *Azotobacter vinelandii* growing in oxygen controlled continuous culture. *Arch Microbiol* 1983, *134*, 68-72.

Pötter M., Müller H., Reinecke F., Wieczorek R., *et al.*, The complex structure of polyhydroxybutyrate (PHB) granules: four orthologous and paralogous phasins occur in *Ralstonia eutropha*. *Microbiology* 2004, *150*, 2301-2311.

Pötter M., Muller H., Steinbuchel A., Influence of homologous phasins (PhaP) on PHA accumulation and regulation of their expression by the transcriptional repressor PhaR in *Ralstonia eutropha* H16. *Microbiology* 2005, *151*, 825-833.

Pötter M., Steinbüchel A., Poly(3-hydroxybutyrate) granule-associated proteins: impacts on poly(3-hydroxybutyrate) synthesis and degradation. *Biomacromolecules* 2005, *6*, 552-560.

Povolo S., Tombolini R., Morea A., Anderson A., *et al.*, Isolation and characterization of mutants of *Rhizobium meliloti* unable to synthesize poly-β-hydroxybutyrate. *Can J Microbiol* 1994, *40*, 823-829.

Quillaguaman J., Hashim S., Bento F., Mattiasson B., Hatti-Kaul R., Poly(beta-hydroxybutyrate) production by a moderate halophile, *Halomonas boliviensis* LC1 using starch hydrolysate as substrate. *J Appl Microbiol* 2005, *99*, 151-157.

Ramsay B. A., Ramsay J. A., Cooper D. G., Production of Poly-beta-Hydroxyalkanoic Acid by *Pseudomonas cepacia*. *Appl Environ Microbiol* 1989, *55*, 584-589.

Rao V., Darrow R., Keister D., Effect of oxygen tension on nitrogenase and on glutamine synthetases I and II in *Rhizobium japonicum* 61A76. *Biochem Biophys Res Commun* 1978, *81*, 224-231.

Reeh S., Pedersen S., Neidhardt F. C., Transient rates of synthesis of five aminoacyl-transfer ribonucleic acid synthetases during a shift-up of *Escherichia coli*. *J Bacteriol* 1977, *129*, 702-706.

Rees D. C., Akif Tezcan F., Haynes C. A., Walton M. Y., *et al.*, Structural basis of biological nitrogen fixation. *Philos Transact A Math Phys Eng Sci* 2005, *363*, 971-984; discussion 1035-1040.

Rehm B., Krüger N., Steinbüchel A., A new metabolic link between fatty acid *de novo* synthesis and polyhydroxyalkanoic acid synthesis. The PHAG gene from *Pseudomonas putida* KT2440 encodes a 3-hydroxyacyl-acyl carrier protein-coenzyme a transferase. *J Biol Chem* 1998, *273*, 24044-24051.

Rehm B. H. A., Polyester synthases: natural catalysts for plastics. *Biochem J* 2003, *376*, 15-33.

Reinders J., Lewandrowski U., Moebius J., Wagner Y., Sickmann A., Challenges in mass spectrometry-based proteomics. *Proteomics* 2004, *4*, 3686-3703.

Reinecke F., Steinbüchel A., *Ralstonia eutropha* Strain H16 as Model Organism for PHA Metabolism and for Biotechnological Production of Technically Interesting Biopolymers. *J Mol Microbiol Biotechnol* 2009, *16*, 91-108.

Reitzer L., Nitrogen assimilation and global regulation in *Escherichia coli*. *Ann Rev Microbiol* 2003, *57*, 155-176.

Reusch R., Sparrow A., Gardiner J., Transport of poly-beta-hydroxybutyrate in human plasma. *Biochim Biophys Acta* 1992, *1123*, 33-40.

Ribbe M., Gadkari D., Meyer O., N_2 fixation by *Streptomyces thermoautotrophicus* involves a molybdenum-dinitrogenase and a manganese-superoxide oxidoreductase that couple N_2 reduction to the oxidation of superoxide produced from O_2 by a molybdenum-CO dehydrogenase. *J Biol Chem* 1997, *272*, 26627–26633.

Rich P. R., The molecular machinery of Keilin's respiratory chain. *Biochem Soc Trans* 2003, *31*, 1095-1105.

Riis V., Mai W., Gas chromatographic determination poly-β-hydroxybutyric acid in microbial biomass after hydrochloric acid propanolysis. *J Chromatogr* 1988, *445*, 285-289.

Ritchie G. A., Senior P. J., Dawes E. A., The purification and characterization of acetoacetyl-coenzyme A reductase from *Azotobacter beijerinckii*. *Biochem J* 1971, *121*, 309-316.

Roe M. R., Griffin T. J., Gel-free mass spectrometry-based high throughput proteomics: Tools for studying biological response of proteins and proteomes. *Proteomics* 2006, *6*, 4678-4687.

Rosen R., Büttner K., Schmid R., Hecker M., Ron E. Z., Stress-induced proteins of *Agrobacterium tumefaciens*. *FEMS Microbiol Ecol* 2001, *35*, 277-285.

Rosen R., Sacher A., Shechter N., Becher D., *et al.*, Two-dimensional reference map of *Agrobacterium tumefaciens* proteins. *Proteomics* 2004, *4*, 1061-1073.

Rudolf M., Kroneck P. M., The nitrogen cycle: its biology. *Met Ions Biol Syst* 2005, *43*, 75-103.

Sabra W., Zeng A. P., Lunsdorf H., Deckwer W. D., Effect of oxygen on formation and structure of *Azotobacter vinelandii* alginate and its role in protecting nitrogenase. *Appl Environ Microbiol* 2000, *66*, 4037-4044.

Sandhiya G. S., Sugitha T. C., Balachandar D., Kumar K., Endophytic colonization and in planta nitrogen fixation by a diazotrophic *Serratia* sp. in rice. *Indian J Exp Biol* 2005, *43*, 802-807.

Sarma A., Emerich D., A comparative proteomic evaluation of culture grown vs nodule isolated *Bradyrhizobium japonicum*. *Proteomics* 2006, *6*, 3008-3028.

Schauder S., Shokat K., Surette M. G., Bassler B. L., The LuxS family of bacterial autoinducers: biosynthesis of a novel quorum-sensing signal molecule. *Mol Microbiol* 2001, *41*, 463-476.

Schelin J., Lindmark F., Clarke A., The clpP multigene family for the ATP-dependent Clp protease in the cyanobacterium *Synechococcus*. *Microbiology* 2002, *148*, 2255-2265.

Schilke B., Voisine C., Beinert H., Craig E., Evidence for a conserved system for iron metabolism in the mitochondria of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 1999, *96*, 10206-10211.

Scott J. D., Ludwig R. A., *Azorhizobium caulinodans* electron-transferring flavoprotein N electrochemically couiples pyruvate dehydrogenase complex activity to N_2 fixation. *Microbiology* 2004, *150*, 117-126.

Selao T. T., Nordlund S., Norén A., Comparative Proteomic Studies in *Rhodospirillum rubrum* Grown under Different Nitrogen Conditions. *J Proteome Res* 2008, *7*, 3267-3275.

Senior P., Dawes E., The regulation of poly-beta-hydroxybutyrate metabolism in *Azotobacter beijerinckii*. *Biochem J* 1973, *134*, 225-238.

Senior P. J., Beech G. A., Ritchie G. A., Dawes E. A., The role of oxygen limitation in the formation of poly-β-hydroxybutyrate during batch and continuous culture of *Azotobacter beijerinckii*. *Biochem J* 1972, *128*, 1193-1201.

Senior P. J., Dawes E. A., Poly-β-hydroxybutyrate biosynthesis and the regulation of glucose metabolism in *Azotobacter beijerinckii*. *Biochem J* 1971, *125*, 55-66.

Serafim L. S., Lemos P. C., Oliveira R., Reis M. A., Optimization of polyhydroxybutyrate production by mixed cultures submitted to aerobic dynamic feeding conditions. *Biotechnol Bioeng* 2004, *87*, 145-160.

Sharma L., Kumar Singh A., Panda B., Mallick N., Process optimization for poly-[beta]-hydroxybutyrate production in a nitrogen fixing cyanobacterium, Nostoc muscorum using response surface methodology. *Bioresource Technology* 2007, *98*, 987-993.

Sherman M., Goldberg A., Heat shock-induced phosphorylation of GroEL alters its binding and dissociation from unfolded proteins. *J Biol Chem* 1994, *269*, 31479-31483.

Sherman M. Y., Goldberg A. L., Heat shock in *Escherichia coli* alters the protein-binding properties of the chaperonin groEL by inducing its phosphorylation. *Nature* 1992, *357*, 167-169.

Shi H., Nikawa J., Shimizu K., Effect of modifying metabolic network on poly-3-hydroxybutyrate biosynthesis in recombinant *Escherichia coli*. *J Biosci Bioeng* 1999, *87*, 666-677.

Singh R., Mailloux R., Puiseux-Dao S., Appanna V., Oxidative stress evokes a metabolic adaptation that favors increased NADPH synthesis and decreased NADH production in *Pseudomonas fluorescens*. *J Bacteriol* 2007, *189*, 6665-6675.

Slade A. H., Gapes D. J., Stuthridge T. R., Anderson S. M., *et al.*, N-ViroTech-a novel process for the treatment of nutrient limited wastewaters. *Water Sci Technol* 2004, *50*, 131-139.

Sprent J., Sprent P. *Nitrogen Fixing Organisms: Pure and Applied Aspects*, 2nd Edn. Chapman Hall, London & New York 1990

Sriram V., Du W., Gervay-Hague J., Brutkiewicz R. R., Cell wall glycosphingolipids of *Sphingomonas paucimobilis* are CD1d-specific ligands for NKT cells. *Eur J Immunol* 2005, *35*, 1692-1701.

Staal M., Rabouille S., Stal L., On the role of oxygen for nitrogen fixation in the marine cyanobacterium *Trichodesmium* sp. *Environ Microbiol* 2007, *9*, 727-736.

Steen H., Mann M., The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol* 2004, *5*, 699-711.

Steenhoudt O., Vanderleyden J., *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. *FEMS Microbiol Rev* 2000, *24*, 487-506.

Steinbüchel A., Fuchtenbusch B., Bacterial and other biological systems for polyester production. *Trends Biotechnol* 1998, *16*, 419-427.

Steinbüchel A., Hustede E., Liebergesell M., Pieper U., *et al.*, Molecular basis for biosynthesis and accumulation of polyhydroxyalkanoic acids in bacteria. *FEMS Microbiol Rev* 1992, *9*, 217-230.

Steinbüchel A., Schlegel H., Physiology and molecular genetics of poly(beta-hydroxy-alkanoic acid) synthesis in *Alcaligenes eutrophus*. *Mol Microbiol* 1991, *5*, 535-542.

Storz G., Tartaglia L. A., Farr S. B., Ames B. N., Bacterial defenses against oxidative stress. *Trends Genet* 1990, *6*, 363-368.

Sudesh K., Taguchi K., Doi Y., Effect of increased PHA synthase activity on polyhydroxyalkanoates biosynthesis in *Synechocystis* sp. PCC6803. *Int J Biol Macromol* 2002, *30*, 97-104.

Suzuki S., Hiraishi A., *Novosphingobium naphthalenivorans* sp. nov., a naphthalene-degrading bacterium isolated from polychlorinated-dioxin-contaminated environments. *J Gen Appl Microbiol* 2007, *53*, 221-228.

Taguchi S., Doi Y., Evolution of Polyhydroxyalkanoate (PHA) Production System by "Enzyme Evolution": Successful Case Studies of Directed Evolution. *Macromolecular Bioscience* 2004, *4*, 145-156.

Taguchi S., Nakamura H., Hiraishi T., Yamato I., Doi Y., *In vitro* evolution of a polyhydroxybutyrate synthase by intragenic suppression-type mutagenesis. *J Biochem* 2002, *131*, 801-806.

Takahashi N., Kaji H., Yanagida M., Hayano T., Isobe T., Proteomics: advanced technology for the analysis of cellular function. *J Nutr* 2003, *133*, 2090S-2096S.

Takeuchi M., Hamana K., Hiraishi A., Proposal of the genus *Sphingomonas* sensu stricto and three new genera, *Sphingobium*, *Novosphingobium* and *Sphingopyxis*, on the basis of phylogenetic and chemotaxonomic analyses. *Int J Syst Evol Microbiol* 2001, *51*, 1405-1417.

Takeuchi M., Sakane T., Yanagi M., Yamasato K., *et al.*, Taxonomic study of bacteria isolated from plants: proposal of *Sphingomonas rosa* sp. nov., *Sphingomonas pruni* sp. nov., *Sphingomonas saccharolytica* sp. nov., and *Sphingomonas mali* sp. nov. *Int J Syst Bacteriol* 1995, *45*, 334-341.

Takeuchi M., Sawada H., Oyaizu H., Yokota A., Phylogenetic Evidence for *Sphingomonas* and *Rhizomonas* as Nonphotosynthetic Members of the Alpha-4 Subclass of the Proteobacteria. *Int J Syst Bacteriol* 1994, *44*, 308-314.

Takusagawa F., Kamitori S., Markham G., Structure and Function of S-Adenosylmethionine Synthetase: Crystal Structures of S-Adenosylmethionine Synthetase with ADP, BrADP, and PPi at 2.8 Å Resolution. *Biochemistry* 1996, *35*, 2586-2596.

Tauro P., Kapoor K., Yadav K., in: *An introduction to Microbiology*, New Age International (P) Limited, New Delhi 1986, pp. 160-174.

Tessmer N., König S., Malkus U., Reichelt R., *et al.*, Heat-shock protein HspA mimics the function of phasins *sensu stricto* in recombinant strains of *Escherichia coli* accumulating polythioesters or polyhydroxyalkanoates. *Microbiology* 2007, *153*, 366-374.

Thiel M., Kaschabek S. R., Groning J., Mau M., Schlomann M., Two unusual chlorocatechol catabolic gene clusters in *Sphingomonas* sp. TFD44. *Arch Microbiol* 2005, *183*, 80-94.

Tholl D., Ober D., Martin W., Kellermann J., Hartmann T., Purification, molecular cloning and expression in *Escherichia coli* of homospermidine synthase from *Rhodopseudomonas viridis*. *Eur J Biochem* 1996, *240*, 373-379.

Tiirola M. A., Busse H. J., Kampfer P., Mannisto M. K., *Novosphingobium lentum* sp. nov., a psychrotolerant bacterium from a polychlorophenol bioremediation process. *Int J Syst Evol Microbiol* 2005, *55*, 583-588.

Tsagou V., Kefalogianni I., Sini K., Aggelis G., Metabolic activities in *Azospirillum lipoferum* grown in the presence of NH_4^+. *Appl Microbiol Biotechnol* 2003, *62*, 574-578.

Tsai M., Saier M. J., Phylogenetic characterization of the ubiquitous electron transfer flavoprotein families ETF-a and ETF-b. *Res Microbiol* 1995, *146*, 397-404.

Tsuge T., Metabolic improvements and use of inexpensive carbon sources in microbial production of polyhydroxyalkanoates. *Journal of Bioscience and Bioengineering* 2002, *94*, 579-584.

Tsuge T., Fukui T., Matsusaki H., Taguchi S., *et al.*, Molecular cloning of two (*R*)-specific enoyl-CoA hydratase genes from *Pseudomonas aeruginosa* and their use for polyhydroxyalkanoate synthesis. *FEMS Microbiol Lett* 2000, *184*, 193-198.

Turner M., Yang X., Yin D., Kuczera K., *et al.*, Structure and function of S-adenosylhomocysteine hydrolase. *Cell Biochem Biophys* 2000, *33*, 101-125.

Uchino K., Saito T., Thiolysis of poly(3-hydroxybutyrate) with polyhydroxyalkanoate synthase from *Ralstonia eutropha*. *J Biochem (Tokyo)* 2006, *139*, 615-621.

Ureta A., Nordlund S., Evidence for Conformational Protection of Nitrogenase against Oxygen in *Gluconacetobacter diazotrophicus* by a Putative FeSII Protein. *J Bacteriol* 2002, *184*, 5805-5809.

Uttamchandani M., Wang J., Yao S. Q., Protein and small molecule microarrays: powerful tools for high-throughput proteomics. *Mol Biosyst* 2006, *2*, 58-68.

VanBogelen R., Schiller E., Thomas J., Neidhardt F., Diagnosis of cellular states of microbial organisms using proteomics. *Electrophoresis* 1999a, *20*, 2149-2159.

VanBogelen R. A., Probing the molecular physiology of the microbial organism, *Escherichia coli* using proteomics. *Adv Biochem Eng Biotechnol* 2003, *83*, 27-55.

VanBogelen R. A., Greis K. D., Blumenthal R. M., Tanic T. H., Matthews R. G., Mapping regulatory networks in microbial cells. *Trends Microbiol* 1999b, *7*, 320-328.

Vandamme P., Coenye T., Taxonomy of the genus *Cupriavidus*: a tale of lost and found. *Int J Syst Evol Microbiol* 2004, *54*, 2285–2289.

Vaneechouette M., Kampfer P., De Baere T., Falsen E., Verschraegen G., *Wautersia* gen. nov., a new genus accommodating the phylogenetic lineage including *Ralstonia eutropha* and related species, and proposal of *Ralstonia* [*Pseudomonas*] *syzygii* (Roberts et al. 1990) comb. nov. *Int J Syst Evol Microbiol* 2004, *54*, 317-327.

Vazquez G., Pettinari M., Méndez B., Evidence of an association between poly(3-hydroxybutyrate) accumulation and phosphotransbutyrylase expression in *Bacillus megaterium*. *Int Microbiol* 2003, *6*, 127-129.

Vlamis-Gardikas A., The multiple functions of the thiol-based electron flow pathways of *Escherichia coli*: Eternal concepts revisited. *Biochim Biophys Acta - Gen Subj* 2008, *1780*, 1170-1200.

Voigt B., Schweder T., Becher D., Ehrenreich A., *et al.*, A proteomic view of cell physiology of *Bacillus licheniformis*. *Proteomics* 2004, *4*, 1465-1490.

Voigt B., Schweder T., Sibbald M. J. J. B., Albrecht D., *et al.*, The extracellular proteome of *Bacillus licheniformis* grown in different media and under different nutrient starvation conditions. *Proteomics* 2006, *6*, 268-281.

Wang F., Lee S., Poly(3-Hydroxybutyrate) Production with High Productivity and High Polymer Content by a Fed-Batch Culture of *Alcaligenes latus* under Nitrogen Limitation. *Appl Environ Microbiol* 1997, *63*, 3703-3706.

Wang M., Wu J. Q., Dennison J. B., Bridges A. S., *et al.*, A gel-free MS-based quantitative proteomic approach accurately measures cytochrome P450 protein concentrations in human liver microsomes. *Proteomics* 2008, *8*, 4186-4196.

Webb C., Koutinas A. A., Wang R., in: Zhong, J.-J., (Ed.) *Advances in Biochemical Engineering/Biotechnology*, Vol. 87Springer-Verlag Berlin Heidelberg New York 2004, pp. 195-268.

Wendisch V. F., de Graaf A. A., Sahm H., Eikmanns B. J., Quantitative determination of metabolic fluxes during coutilization of two carbon sources: comparative analyses with *Corynebacterium glutamicum* during growth on acetate and/or glucose. *J Bacteriol* 2000, *182*, 3088-3096.

White D. C., Sutton S. D., Ringelberg D. B., The genus *Sphingomonas*: physiology and ecology. *Curr Opin Biotechnol* 1996, *7*, 301-306.

Wieczorek R., Pries A., Steinbuchel A., Mayer F., Analysis of a 24-kilodalton protein associated with the polyhydroxyalkanoic acid granules in *Alcaligenes eutrophus*. *J Bacteriol* 1995, *177*, 2425-2435.

Wilmes P., Wexler M., Bond P. L., Metaproteomics Provides Functional Insight into Activated Sludge Wastewater Treatment. *PLoS One* 2008, *3*, e1778.

Wolfe A. J., The acetate switch. *Microbiol Mol Biol Rev* 2005, *69*, 12-50.

Xie C. H., Yokota A., *Sphingomonas azotifigens* sp. nov., a nitrogen-fixing bacterium isolated from the roots of *Oryza sativa*. *Int J Syst Evol Microbiol* 2006, *56*, 889-893.

Yabuuchi E., Kosako Y., Fujiwara N., Naka T., *et al.*, Emendation of the genus *Sphingomonas* Yabuuchi et al. 1990 and junior objective synonymy of the species of three genera, *Sphingobium*, *Novosphingobium* and *Sphingopyxis*, in conjunction with *Blastomonas ursincola*. *Int J Syst Evol Microbiol* 2002, *52*, 1485-1496.

Yabuuchi E., Kosako Y., Yano I., Hotta H., Nishiuchi Y., Transfer of two *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. nov: proposal of *Ralstonia pickettii* (Ralston, Palleroni and Doudoroff 1973) comb. nov., *Ralstonia*

solanacearum (Smith 1896) comb. nov. and *Ralstonia eutropha* (Davis 1969) comb. nov. *Microbiol Immunol* 1995, *39*, 897–904.

Yabuuchi E., Yano I., Oyaizu H., Hashimoto Y., *et al.*, Proposals of *Sphingomonas paucimobilis* gen. nov. and comb. nov., *Sphingomonas parapaucimobilis* sp. nov., *Sphingomonas yanoikuyae* sp. nov., *Sphingomonas adhaesiva* sp. nov., *Sphingomonas capsulata* comb. nov., and two genospecies of the genus *Sphingomonas*. *Microbiol Immunol* 1990, *34*, 99-119.

Yang D. C., Im W. T., Kim M. K., Ohta H., Lee S. T., *Sphingomonas soli* sp. nov., a beta-glucosidase-producing bacterium in the family Sphingomonadaceae in the alpha-4 subgroup of the Proteobacteria. *Int J Syst Evol Microbiol* 2006, *56*, 703-707.

Yu Z., Wu X., Li D., Yang S., *et al.*, Enhancement of the production of SAM by overexpression of SAM synthetase in *Pichia pastoris*. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* 2003, *35*, 127-132.

Zakhia F., Jeder H., Willems A., Gillis M., *et al.*, Diverse bacteria associated with root nodules of spontaneous legumes in Tunisia and first report for nifH-like gene within the genera *Microbacterium* and *Starkeya*. *Microb Ecol* 2006, *51*, 375-393.

Zhang J., The role of heterologous nifAc product in the regulation of nif expression in *Agrobacterium tumefaciens*. *Chin J Biotechnol* 1997, *13*, 31-36.

Zhang J., Hao N., Chen G. Q., Effect of expressing polyhydroxybutyrate synthesis genes (phbCAB) in *Streptococcus zooepidemicus* on production of lactic acid and hyaluronic acid. *Appl Microbiol Biotechnol* 2006, *71*, 222-227.

Zhang Y., Burris R., Ludden P., Roberts G., Presence of a second mechanism for the posttranslational regulation of nitrogenase activity in *Azospirillum brasilense* in response to ammonium. *J Bacteriol* 1996, *178*, 2948-2953.

Zhang Y., Burris R. H., Ludden P. W., Roberts G. P., Comparison studies of dinitrogenase reductase ADP-ribosyl transferase/dinitrogenase reductase activating glycohydrolase regulatory systems in *Rhodospirillum rubrum* and *Azospirillum brasilense*. *J Bacteriol* 1995, *177*, 2354-2359.

Zhao J., Shimizu K., Metabolic flux analysis of *Escherichia coli* K12 grown on 13C-labeled acetate and glucose using GC-MS and powerful flux calculation method. *J Biotechnol* 2003, *101*, 101-117.

Zhao Y., Bian S.-M., Zhou H.-N., Huang J.-F., Diversity of Nitrogenase Systems in Diazotrophs. *J Integr Plant Biol* 2006, *48*, 745-755.

Zhou H. W., Guo C. L., Wong Y. S., Tam N. F., Genetic diversity of dioxygenase genes in polycyclic aromatic hydrocarbon-degrading bacteria isolated from mangrove sediments. *FEMS Microbiol Lett* 2006, *262*, 148-157.

Zhu K., Zhao J., Lubman D. M., Miller F. R., Barder T. J., Protein pI Shifts due to Posttranslational Modifications in the Separation and Characterization of Proteins. *Analytical Chemistry* 2005, *77*, 2745-2755.

Zhu X., Tang G., Granier F., Bouchez D., Galili G., A T-DNA insertion knockout of the bifunctional lysine-ketoglutarate reductase saccharopine dehydrogenase gene elevates lysine levels in *Arabidopsis* seeds. *Plant Physiol* 2001, *126*, 1539-1545.

Zwietering M. H., Jongenburger I., Rombouts F. M., Riet K. v. t., Modeling of the Bacterial Growth Curve. *Appl Environ Microbiol* 1990, *56*, 1875-1881.