

If you have discovered material in AURA which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our [Takedown Policy](#) and [contact the service](#) immediately

IMPROVED PRODUCTION OF INDUSTRIALLY
RELEVANT RECOMBINANT PROTEINS

Application of multi-factorial design of experiments
and scalable process modelling

WILLIAM JOSEPH HOLMES

Doctor of Philosophy

ASTON UNIVERSITY

February 2009

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgment.

ASTON UNIVERSITY

IMPROVED PRODUCTION OF INDUSTRIALLY
RELEVANT RECOMBINANT PROTEINS

Application of multi-factorial design of experiments
and scalable process modelling

Name: William Joseph Holmes

Degree: PhD

Year of submission: 2009

Understanding the structures and mechanisms of proteins gives direct insight into cellular function as well as providing targets for the investigation of disease. Since the vast majority of proteins are not sufficiently abundant from natural sources, recombinant overproduction is a universally-recognised solution to obtaining the milligram quantities required for applications as diverse as structural genomics and biopharmaceutical manufacture. The routine achievement of high production yields in any of the currently-available host cell systems continues to be a substantial bottleneck to further progress, and still relies on trial-and-error. As such sub-optimal product yields are often obtained and further experimentation is required for process scale-up. The work described in this thesis focuses on the use of a design-of-experiments approach in a multi-well mini-bioreactor to enable the rapid establishment of high yielding production phase conditions in yeast, which is an increasingly popular host system in both academic and industrial laboratories. Using green fluorescent protein secreted from the yeast, *Pichia pastoris*, a scalable, predictive model of protein yield per cell was derived from 13 sets of conditions each with three factors (temperature, pH and dissolved oxygen) at 3 levels and was directly transferrable to a 7 L bioreactor. This was in clear contrast to the situation in shake flasks, where the process parameters cannot be tightly controlled. By further optimising both the accumulation of cell density in batch and improving the fed-batch induction regime, additional yield improvement was found to be additive to the per cell yield of the model. A separate study also demonstrated that improving biomass improved product yield in a second yeast species, *Saccharomyces cerevisiae*. Investigations of cell wall hydrophobicity in high cell density *P. pastoris* cultures indicated that cell wall hydrophobin (protein) composition changes with growth phase becoming more hydrophobic in log growth than in lag or stationary phases. This is possibly due to an increased occurrence of proteins associated with cell division. Finally, the modelling approach was validated in mammalian cells, showing its flexibility and robustness. In summary, the strategy presented in this thesis has the benefit of reducing process development time in recombinant protein production, directly from bench to bioreactor.

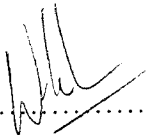
Key words: recombinant protein; predictive model; process optimisation; *Pichia pastoris*; hybridoma

ACKNOWLEDGMENTS

This thesis would not have been possible without the support of my family and friends (especially Helen and Cliff). I would like to thank Roslyn and Rod for their guidance, insight and humour during difficult times. Thanks also go to all those at Alpha Biologics UK and the laboratory at Aston for their comradeship and assistance. Finally, for making the write-up process a little more enjoyable, I thank ROCK.

Aspects of this thesis were carried out in collaboration with others:

- The expression vectors used for creating *Pichia pastoris* and *Saccharomyces cerevisiae* cell lines expressing a novel monoclonal antibody-based influenza vaccine (Chapter 4) were supplied by Alpha Biologics UK
- The hybridoma HD1 cell-line, expressing an anti-tumour factor antibody (Chapter 4) was supplied by Alpha Biologics UK
- The planning and analysis of the statistical design of experiments (Chapters 3 and 4) were carried out in consultation with Martin Wilks (Cameron International Ltd)
- The planning of the medium design experiments for improving biomass in *S. cerevisiae* cultures was carried out in consultation with Nicklas Bonander (Aston University)
- The planning of the *P. pastoris* cell surface hydrophobicity investigation (Chapter 5) was carried out in consultation with Steve Smith (Aston University)

SIGNED:

FULL NAME: William Joseph Holmes

DATE: 26 MAY 2009

CONTENTS

TITLE PAGE	1
THESIS SUMMARY	2
ACKNOWLEDGMENTS	3
CONTENTS	4
LIST OF TABLES, FIGURES AND EQUATIONS	9
1 INTRODUCTION	14
1.1 Recombinant proteins as pharmaceuticals.....	15
1.2 Host systems for recombinant protein production.....	16
1.2.1 <i>Escherichia coli</i>	16
1.2.2 Yeasts e.g. <i>Pichia pastoris</i> , <i>Saccharomyces cerevisiae</i> , <i>Schizosacharomyces pombe</i>	17
1.2.3 Filamentous fungi e.g. <i>Aspergillus nidulans</i> , <i>Aspergillus niger</i> , <i>Aspergillus sydowii</i> , <i>Aspergillus awamori</i> , and various <i>Fusarium sp.</i> and <i>Trichoderma sp.</i>	17
1.2.4 Insect cell-lines e.g. <i>Spodoptera frugiperda</i> cell lines (sf9, sf21).....	17
1.2.5 Mammalian cell-lines e.g. Chinese Hamster Ovary (CHO), HeLa (cervical cancer cells taken in 1951 from Henrietta Lacks) and hybridoma (immortal cell line created by fusing B-cells and myeloma cells)	17
1.3 Yeast as a host system	18
1.3.1 <i>Pichia pastoris</i> , an industry standard for protein expression	19
1.3.2 <i>S. cerevisiae</i> Alcofree™; a novel respiratory yeast.....	19
1.4 Target proteins.....	20
1.4.1 Green fluorescent protein (GFP)	20
1.4.2 The role of GFP	21
1.4.2.1 Protein engineering of GFP	23
1.4.2.2 GFP as a tool	23
1.4.3 Monoclonal antibody raised against a tumour inducing factor	25
1.4.4 Insulin	25
1.4.4.1 The role of insulin in healthy humans	25
1.4.4.2 Insulin and disease.....	26
1.4.4.3 Insulin as a biotech pharmaceutical.....	29
1.4.4.4 Recombinant production of insulin	30
1.4.5 Human growth hormone.....	30
1.4.5.1 Role of human growth hormone in healthy humans	31
1.4.5.1.1 Role of hPGH	31
1.4.5.2 Synthesis/structure of native hPGH.....	31
1.4.5.3 Role of hGH	32
1.4.5.4 Human growth hormone and disease.....	32
1.4.5.5 hGH as a biotherapeutic	33
1.4.5.6 Recombinant production of hPGH (rhPGH)	34
1.4.6 Novel monoclonal antibody-based influenza vaccine	34
1.5 Optimisation of culture conditions for the expression of recombinant proteins.....	34
1.5.1 Temperature.....	35
1.5.2 pH	36
1.5.3 Aeration	36
1.5.4 Medium design	37
1.5.5 Anti-foam agents	37
1.5.6 Induction strategy/protocol.....	38
1.5.7 Scale-down modelling and Design of Experiments (DoE)	39
1.5.7.1 Scale down modelling of the fermentation process.....	39

1.5.7.2	Use of statistical Design of Experiments in process optimisation	40
1.6	Project aims and objectives	43
2	MATERIALS AND METHODS	46
2.1	Microbial strains used.....	46
2.1.1	<i>Escherichia coli</i>	46
2.1.1.1	XL1-Blue super-competent cells	46
2.1.1.2	DH5 α ^{97 98 99 100}	46
2.1.1.3	BL21-DE3 ^{101 102}	47
2.1.1.4	TOPO10 ^{99 103}	47
2.1.2	<i>Pichia pastoris</i>	47
2.1.2.1	GS115	47
2.1.2.2	X33	47
2.1.3	<i>Saccharomyces cerevisiae</i>	48
2.1.3.1	W303-1A ^{105 106}	48
2.1.3.2	Alcofree (TM6* prototrophic).....	48
2.1.3.3	Alcofree (TM6* <i>URA</i>)	48
2.2	Mammalian cell lines used	48
2.2.1	Hybridoma cell line HD1	48
2.3	Molecular biology	48
2.3.1	DNA manipulation	49
2.3.1.1	Restriction Digests.....	49
2.3.1.2	Ligations	49
2.3.1.3	Polymerase chain reaction (PCR).....	50
2.3.2	Agarose gel electrophoresis.....	50
2.3.2.1	Pouring agarose gels.....	51
2.3.2.2	Sample preparation	51
2.3.2.3	Running conditions.....	51
2.3.2.4	Visualisation	51
2.3.3	Plasmid preparation (mini/maxi preps)	52
2.3.4	Sequence analysis	52
2.4	Transformation of cells.....	52
2.4.1	Preparation of competent cells	53
2.4.1.1	<i>E. coli</i>	53
2.4.1.2	<i>P. pastoris</i>	53
2.4.2	Heat shock transformation of <i>E. coli</i> DH5 α cells.....	54
2.4.3	Electroporation	54
2.4.3.1	Electroporation of <i>E. coli</i> cells	54
2.4.3.2	Electroporation of <i>P. pastoris</i> cells	55
2.4.4	Lithium acetate transformation of <i>S. cerevisiae</i> cells.....	55
2.4.4.1	Carrier DNA	56
2.4.5	Modified Heat shock lithium acetate transformation of <i>S. cerevisiae</i> cells.....	56
2.5	Clone selection	56
2.5.1	Confirmation of DNA uptake / integration (colonies on selective agar).	57
2.5.2	Expression confirmation (screening small cultures)	57
2.6	Storage of strains (wild-type and transformed) by glycerol stock.....	57
2.7	Analytical methods	57
2.7.1	Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).....	57
2.7.2	SDS-PAGE protein band visualisation by InstantBlue stain.....	58
2.7.3	SDS-PAGE protein band visualisation by silver stain	58
2.7.4	Western blotting	58
2.7.5	Mass spectrometry	58
2.7.6	GFP _{uv} fluorescence analysis	59

2.7.7	Antibody quantification by Protein-A purification	59
2.7.8	Optical density at 595 and 600 nm wavelength (OD _{595/600})	59
2.7.9	Cell surface hydrophobicity (CSH) assays	60
2.7.9.1	Microbial Attachment to Hydrocarbons (MATH) assay	60
2.7.9.2	Hydrophobic interaction chromatography (HIC) assay	61
2.7.10	Ethanol concentration assay (UV-method)	61
2.8	Shake flask (Erlenmeyer flask) cultures	61
2.9	Media	62
2.9.1	LB	62
2.9.2	YPD	62
2.9.3	BMGY (Buffered Glycerol-complex Medium).....	62
2.9.4	BMMY (Buffered Methanol-complex Medium).....	64
2.9.5	2xCBS.....	65
2.9.6	CSM (Complete synthetic medium).....	66
2.9.7	BSM (Basal salts medium).....	66
2.9.8	SD ^{-URA}	67
2.9.9	SOC	67
2.9.10	ExCell 302	68
2.10	Antibiotics	68
2.11	Applikon Micro 24 bioreactor	68
2.11.1	Installation	70
2.11.2	Programming/ run set-up	71
2.12	7 L vessel Fermentor/ ADI1030 control unit (Scale-up experiments)	71
2.12.1	Fermentor set-up/ operation.....	71
2.12.2	Applikon ADI1030 set-up and operation	72
2.12.3	Applikon BioXpert Lite software	72
2.13	MiniTab statistical software	73
2.14	Model building and refinement	73
3	OPTIMISATION OF RECOMBINANT PROTEIN YIELD USING PARALLEL MINI-BIOREACTORS AND A DESIGN-OF-EXPERIMENTS MODELLING APPROACH.....	75
3.1	Optimisation of GFP expression in <i>P. pastoris</i>	77
3.1.1	GFP as the target protein	78
3.1.2	Strain creation.....	78
3.1.3	Strain characterisation	80
3.1.3.1	Growth curves	82
3.1.4	Fluorescence analysis	84
3.2	Initial optimisation of GFP yield using a design-of-experiments approach	85
3.2.1	The Applikon Micro 24 Bioreactor	85
3.2.2	Why use Design of Experiments?	88
3.2.3	Normalised yield allows additional flexibility in process optimisation ..	90
3.2.4	Data collection for predictive model	90
3.2.5	Building a predictive model for yield.....	93
3.2.6	Validating the model on a small scale	95
3.2.7	Improving the predictive model: generating DoE2	96
3.2.7.1	Data collection for predictive model DoE2	97
3.2.7.2	Building a predictive model for yield (DoE2).....	100
3.2.7.3	Testing the DoE2 model at small scale (check points).....	104
3.3	Is the predictive model scalable?.....	106
3.3.1	Bench-scale (7L vessel) experimental set-up	107
3.3.2	Initial DoE2 scale-up bioreactor experiments	109
3.3.2.1	DoE2 scale-up conclusions from initial runs.....	113

3.4	Further improvements are possible to the DoE modelled process	115
3.4.1	Increasing pre-induction biomass with DoE2 optimised process conditions	115
3.4.2	Increasing pre-induction biomass and using a mixed feed induction with DoE2 optimised process conditions.....	119
3.5	Use of predictive scale-down/up modelling and micro-bioreactors for process development	125
4	PRODUCTION OF INDUSTRIALLY RELEVANT PROTEINS.....	128
4.1	Extension of the DoE approach to a hybridoma cell-line.....	128
4.1.1	Cell-line characterisation.....	129
4.1.1.1	Growth profiling.....	130
4.1.1.2	Product detection	131
4.1.2	The Applikon Micro 24 Bioreactor for mammalian cell culture.....	133
4.1.3	DoE plan for hybridoma HD1 yield optimisation (DoE3)	135
4.1.4	Data collection for predictive model building.....	135
4.1.5	Building the DoE3 predictive model.....	138
4.1.6	Testing the model at small scale.....	140
4.1.7	Model scalability	144
4.1.7.1	Bench-scale (7L vessel) experimental set-up	144
4.1.7.2	DoE3 scale-up experiments.....	146
4.1.7.3	DoE3 scale-up antibody yield results	151
4.1.8	Conclusions on DoE3 model scalability.....	152
4.2	Examination of additional industrially-relevant target proteins to the DoE process in yeast.....	152
4.2.1	Insulin	153
4.2.2	hPGH	159
4.2.3	Monoclonal antibody based influenza vaccine.....	161
4.3	Conclusions from yeast strain creation.....	165
5	OPTIMISING ADDITIONAL FACTORS: INVESTIGATIONS OF THE INFLUENCE OF CELL SURFACE HYDROPHOBICITY AND ANTIFOAMS ON PRODUCTIVITY IN <i>PICHIA PASTORIS</i> CULTURES.....	167
5.1	Investigation of cell wall hydrophobicity in high cell density <i>Pichia pastoris</i> cultures	168
5.1.1	Cell surface hydrophobicity; Important aspects	168
5.1.2	Determination of CSH	169
5.1.3	Preliminary CSH investigation with shake flask cultures	170
5.1.4	High Cell Density 7 L bioreactor culture	173
5.1.5	Implications for high cell density fermentations	181
5.2	Rationalisation for the use of antifoam agents	182
5.2.1	Antifoam evaluation with an Fc fusion protein as an indication of yield	183
5.2.2	Antifoam evaluation with green fluorescent protein (GFP _{uv}) as an indication of yield	185
5.3	Many factors can affect a protein production process.....	189
6	IMPROVING BIOMASS AND VOLUMETRIC PRODUCT YIELD IN WILD TYPE AND A NOVEL ALCOFREE STRAIN OF <i>SACCHAROMYCES CEREVISIAE</i>	191
6.1	Constitutive expression and growth; linking yield to biomass.....	192
6.2	Medium design	193
6.2.1	Non-fermentable carbon sources.....	193
6.2.2	Optimisation of CSM media.....	196
6.2.2.1	Carbon source: Micro 24 experiments	197
6.2.3	Nutrient addition: Micro24 experiments	201
6.3	Conclusions on improving <i>S. cerevisiae</i> biomass	206

7	CONCLUSIONS	208
8	REFERENCES	217
9	APPENDICES	224
9.1	Expression vector construction strategy	224
9.1.1	MPI	224
9.1.2	hPGH	227
9.2	Bioforum Europe article	231
9.3	Poster presentations	233
9.3.1	Recombinant protein production (RPP) conference 2006, Barcelona, Spain.	233
9.3.2	Poster presentations at Aston University, School of life and health sciences post-graduate research days	234
9.3.2.1	2007	234
9.3.2.2	2008 (Additional handout information for poster presentation).....	235
9.3.3	BioProcessUK meeting Cardiff, United Kingdom, 2007. (Travel bursary awarded)	236
9.4	Applikon Biotechnology; Fermentation and cell culture symposium 2008 (Liverpool). Use of the Micro 24 system in predictive scale-down modelling of recombinant protein expression in <i>P. pastoris</i> and a hybridoma cell-line. Slides for presentation given as an invited speaker.....	237
10.	DATA CD	
10.1.	Expression vector sequence data	
10.1.1.	pYX212GFPuv Clone 5.doc	
10.1.2.	pYX212GFPuv Clone 8.doc	
10.1.3.	pYX212hGH2 Clone 7.doc	
10.1.4.	pYX212hGH2 Clone 8.doc	
10.1.5.	pYX212hGH2 Clone 11.doc	
10.1.6.	pYX212hGH2 Clone 12.doc	
10.1.7.	pYX212MPI Clone 1.doc	
10.1.8.	pPICZalphaAGFPuv Clone 2.doc	
10.1.9.	pPICZalphaAMPI Clone 2.doc	
10.1.10.	pPICZalphaAhGH2 Clone 1.doc	
10.2.	Mass spectrometry data	
10.2.1.	GFPuv lower band analysis.doc	
10.2.2.	GFPuv upper band analysis.doc	
10.3.	Manuscript placeholder: Developing a scalable model of recombinant protein yield from <i>Pichia pastoris</i> : the influence of culture conditions, biomass and induction regime.txt	

LIST OF TABLES, FIGURES AND EQUATIONS

Table 1.1. Comparison of some characteristics of the yeasts <i>S. cerevisiae</i> and <i>P. pastoris</i> .	18
Figure 1.1. Process of green light emission from <i>A. victoria</i> .	21
Figure 1.2. GFP chromophore (p-hydroxybenzylideneimidazolinone) formation of active state.	23
Figure 1.3. Enzymatic conversion of proinsulin to mature insulin and C-peptide.	28
Figure 1.4. Experimental design space suggested by the Box-Behnken DoE model.	42
Figure 2.1. Micro 24 micro-bioreactor (Applikon Biotechnology Ltd.) in use at Aston University. Labels highlight the required ancillary equipment for running the Micro 24.	69
Figure 2.2. Micro 24 micro-bioreactor (Applikon Biotechnology Ltd.) in use at Aston University. The incubation cassette, well closures and shaking platform are highlighted.	70
Figure 2.3. 7 L bioreactor with Applikon ADI1030 control unit and ancillary equipment, in use at the Alpha Biologics laboratory.	72
Figure 3.1. Typical biomass (OD ₅₉₅) and dissolved oxygen (DO; %) profiles for a fed-batch fermentation of <i>P. pastoris</i> .	77
Figure 3.2. Plasmid map of pPICZαA-GFP _{uv} generated in Clone Manager v5.02.	79
Figure 3.3. SDS-PAGE and anti-His ₆ Western blot analysis of <i>P. pastoris</i> X33GFP _{uv} culture supernatants.	81
Figure 3.4. Snake-plot of the GFP _{uv} amino-acid sequence showing the mass spectrometry result fragments.	82
Figure 3.5. Growth profiles of <i>P. pastoris</i> X33GFP _{uv} clones 3, 4 and 5 in non-induced (BMGY) and induced (BMMY) shake flask cultures.	83
Figure 3.6. Scatter plot showing the relationship between <i>P. pastoris</i> X33GFP _{uv} culture supernatant fluorescence (RFU) and SDS-PAGE band densitometry.	84
Figure 3.7. Calibration plot for GFP _{uv} fluorescence (RFU) to GFP _{uv} protein concentration (mg/mL).	85
Figure 3.8. Applikon's Micro 24 bioreactor system in use at the Aston laboratory, highlighting the required ancillary equipment and bioreactor unit size.	86
Figure 3.9. Applikon's Micro 24 bioreactor system in use at the Aston laboratory.	87
Figure 3.10. Example of the Micro 24 control software display.	88
Figure 3.11. Representation of the Box-Behnken design space.	89
Table 3.1. Experimental matrix showing coded (-1, 0, 1) and non-coded set-points required for the Box-Behnken DoE1 predictive model.	90
Table 3.2. Specification of the input factors and measurable outputs for the initial model building experiments.	92
Table 3.3. Initial predictive model terms and coefficients based on the model-building data shown in Table 3.2.	93
Equation 3.1. Refined DoE1 predictive model.	93
Table 3.4. ANOVA analysis of DoE1 model.	94
Figure 3.12. Response optimisation for DoE1 model as predicted by Minitab software.	94
Figure 3.13. Response surface contour plot for DoE1 model showing normalised yield (RFU/OD ₅₉₅) changing with each of the factors.	95
Figure 3.14. Response surface wire-frame plot for DoE1 model showing normalised yield (RFU/OD ₅₉₅) changing with temperature and pH.	95
Figure 3.15. DoE1-predicted normalised yield compared with experimental normalised yield in Micro 24 cultures.	96
Table 3.5. Experimental matrix showing coded (-1, 0, 1) and non-coded set-points required for the Box-Behnken DoE2 predictive model.	97
Table 3.6. Specification of the input factors and measurable outputs for the DoE2 model building experiments.	98

Table 3.7. Input factors for the DoE2 model with resulting normalised GFP _{uv} yields converted from fluorescence values (RFU) to amount of GFP _{uv} protein (µg).....	99
Figure 3.16. Relationship between biomass (OD ₅₉₅) and GFP _{uv} yield (µg) for the cultures in Table 3.7.	100
Equation 3.2. Refined DoE2 predictive model.....	101
Table 3.8. ANOVA results for the DoE2 model.	101
Figure 3.17. Response optimisation for DoE2 as predicted by Minitab software.....	102
Figure 3.18. Response surface contour plot for DoE2 model showing normalised yield (RFU) changing with each of the factors.	103
Figure 3.19. Response surface wire-frame plot for DoE2 model showing normalised yield (RFU) changing with temperature and pH.	103
Figure 3.19. Response surface wire-frame plots for DoE1 and DoE2 showing how the two design spaces fit together demonstrating the full range of yield conditions.	104
Table 3.9. Factor settings (T, pH and DO) and yield results (ng/OD ₅₉₅) for the check point conditions used to validate the DoE2 model in the Micro 24.	105
Figure 3.20. Normalised GFP yield (µg/OD ₅₉₅) as predicted by the DoE2 model.....	105
Figure 3.21. Relationship between predicted and experimental normalised yields from Micro 24 cultures from Table 3.9 and Figure 3.20.	106
Figure 3.22. Schematic overview of a typical scale-up bioreactor experiment.....	107
Figure 3.23. Set-up of 7 L (total volume) stirred tank bioreactor used for the <i>P. pastoris</i> X33GFP _{uv} scale-up experiments.	108
Table 3.10. Initial DoE2 scale up experimental run descriptions and yield data.	110
Figure 3.24. Initial DoE2 scale up culture profiles for fed-batch <i>P. pastoris</i> X33GFP _{uv} experiments, using the conditions indicated by Table 3.10.....	111
Figure 3.25. GFP yield data from fed-batch <i>P. pastoris</i> X33GFP _{uv} DoE2 scale-up experiments using conditions described in Table 3.10.	112
.....	113
Figure 3.26. Comparison of normalised yield (ng/mL/OD ₅₉₅) data at 48 h post-induction for DoE2 model predictions, small-scale Micro 24 experimental runs and large-scale (7 L Bioreactor) experimental runs.	113
Figure 3.27. Total GFP _{uv} yields (RFU) profiles for the initial DoE2 scale-up bioreactor runs.	114
.....	114
Figure 3.28 Culture profile for fed-batch <i>P. pastoris</i> X33GFP _{uv} DoE2 scale-up experiment (f007) using “DoE2 predicted optimal” conditions (21.5 °C, pH 7.6, 90 % DO) with 40 g/L glycerol in batch phase,	117
Figure 3.29 GFP _{uv} yield data from (f007) fed-batch <i>P. pastoris</i> X33GFP _{uv} DoE2 scale-up experiment using “DoE2 predicted optimal” conditions (21.5 °C, pH 7.6, 90 % DO) with 40 g/L glycerol in batch phase.....	118
Figure 3.30. Total GFP _{uv} yields (RFU) profiles for the initial DoE2 scale-up and increased pre-induction biomass bioreactor runs.	118
Table 3.11 DoE2 scale up experimental run descriptions and yield data for cultures using improved medium composition (40 g/L glycerol) and mixed feed (60 % w/v sorbitol, 40 % v/v methanol) induction.....	121
Figure 3.31 DoE2 scale up culture profiles for fed-batch <i>P. pastoris</i> X33GFP _{uv} experiments, using improved medium composition (40 g/L glycerol) and mixed feed (60 % sorbitol, 40 % methanol) induction.....	122
Figure 3.32 GFP _{uv} yield data from fed-batch <i>P. pastoris</i> X33GFP _{uv} DoE2 scale-up experiments using improved BMGY medium, mixed feed induction and process conditions described in Table 3.11.....	123
Figure 3.33. Total GFP _{uv} yields (RFU) profiles for the DoE2 scale-up increased pre-induction biomass and mixed feed induction bioreactor runs.	124

Figure 3.34 Comparison of normalised yield ($\mu\text{g}/\text{mL}/\text{OD}_{595}$) data at harvest (48-50 h post-induction) for DoE2 model predictions, small-scale Micro 24 experimental runs and large-scale (f002-f011) experimental runs.....	125
Figure 3.35. Overview of the optimisation process used for maximising GFP _{uv} expression from <i>P. pastoris</i> X33GFP _{uv} bioreactor cultures as well as other target proteins and expression systems.....	126
Table 4.1. Culture descriptions for hybridoma HD1 cell-line characterisation experiments.....	130
Figure 4.1. Example of the growth profiles for T75 flat flask (culture B) batch culture of hybridoma HD1 cells over 12 days.....	130
Figure 4.2. Example of the growth profiles for 250 mL shake flask (culture D) batch culture of hybridoma HD1 cells over 12 days.....	131
Figure 4.3. Coomassie stain visualised SDS-PAGE image of non-reduced supernatant sample from 250 mL shake flask (culture D) batch culture of hybridoma HD1 cells over 12 days (0-288 hours).....	133
Figure 4.4. Coomassie stain visualised SDS-PAGE image of reduced supernatant sample from 250 mL shake flask (culture D) batch culture of hybridoma HD1 cells over 12 days (0-288 hours).....	133
Table 4.2. Experimental matrix showing coded (-1, 0, 1) and non-coded set-points required for the Box-Behnken DoE3 predictive model for optimising antibody yield from hybridoma HD1 hybridoma cell culture.....	135
Table 4.3. Specification of the input factors and measurable outputs for the hybridoma HD1 model building experiments.....	137
Table 4.4. Refined DoE3 predictive model terms, coefficients and associated <i>p</i> -values based on the model building data shown in Table 4.3.....	138
Equation 4.1. Refined DoE3 predictive model for Yield ($\mu\text{g}/\text{mL}$).....	138
Table 4.5. ANOVA analysis of DoE3 model.....	139
Figure 4.5. Response surface contour plot for DoE3 model.....	139
Figure 4.6. Response surface wire-frame plot for DoE3 model.....	140
Figure 4.7. Response optimisation for DoE3 model as predicted by Minitab software.....	140
Figure 4.8. DoE3 model predicted and experimental yield from model confirmation Micro 24 cultures.....	141
Figure 4.9. XY scatter plot of DoE3 model predicted and experimental yield ($\mu\text{g}/\text{mL}$) from model confirmation Micro 24 cultures (as shown in Figure 4.12).....	142
Figure 4.10. DoE3 model predicted and experimental yield ($\mu\text{g}/\text{mL}$) from model confirmation Micro 24 cultures. Check point conditions are outside the model design space.....	143
Figure 4.11 DoE3 model predicted and experimental yield from model confirmation Micro 24 cultures. Showing antibody yield ($\mu\text{g}/\text{mL}$) from optimised conditions and both internal and external design space check point conditions.....	143
Table 4.6. DoE3 scale-up experimental run conditions.....	144
Figure 4.12 Set-up of 7 L (total volume) stirred tank bioreactor used for the mammalian cell culture of the hybridoma HD1 cell-line for DoE3 scale-up experiments.....	146
Table 4.7. DoE3 scale up experimental run descriptions and yield data.....	149
Figure 4.13. DoE3 scale up culture profiles for fed-batch hybridoma (hybridoma HD1) cell culture experiments, using the conditions indicated by Table 4.7.....	150
Figure 4.14. Antibody yields from the bi-phase DoE3 scale-up experiments BR01 and BR03.....	151
Figure 4.15. Schematic of the clone creation strategy.....	153
Figure 4.16. Representations of the three expression vectors used for the production of secreted MPI in <i>P. pastoris</i> , <i>S. cerevisiae</i> Alcofree and <i>S. cerevisiae</i> W303-1A.....	155
Figure 4.17. Alcofree MPI clone selection analysis of 100x concentrated supernatant.....	157

Figure 4.18. Growth curves of <i>S. cerevisiae</i> Alcofree pYX212- α -MPI clone 1 and 6 in baffled 500 mL shake flasks with 50 mL 2xCBS medium over 97 h.....	158
Figure 4.19. Enhance colour image of coomassie stained SDS-PAGE gel. Samples were taken at 24 h intervals during shake flask cultures (See Figure 6.18) of Alcofree MPI clones 1 and 6 to evaluate MPI expression.....	158
Figure 4.20. Silver stained SDS-PAGE gel. Samples were taken at 24 h intervals during time course shake flask culture (See Figure 4.18) of Alcofree MPI clones 1 and 6 to evaluate MPI expression.	159
Figure 4.21. Image of a photographic film developed from ECL detection of a Western Blot. Samples are of 100x concentrated supernatants of <i>S. cerevisiae</i> Alcofree clones.	163
Figure 5.1. Growth curves (OD ₆₀₀) for <i>P. pastoris</i> X33GFP _{uv} shake flask cultures grown in non-inducing (BMGY) and inducing (BMMY) media.	170
Figure 5.2. Non-induced (BMGY) <i>P. pastoris</i> X33GFP _{uv} shake flask culture. Relative CSH and cell density (OD ₆₀₀) at early/mid and late log phases of growth curve are shown	171
Figure 5.3. Induced (BMMY) <i>P. pastoris</i> X33GFP _{uv} shake flask culture. Relative CSH and cell density (OD ₆₀₀) at early/mid and late log phases of growth curve are shown.....	172
Figure 5.4. Trace from BioXpert Lite monitoring software of <i>P. pastoris</i> X33GFP _{uv} 7 L bioreactor culture with model predicted optimal expression conditions as described in Chapter 3.	175
Figure 5.5. Non-induced (BMGY batch) and induced (methanol fed-batch) <i>P. pastoris</i> X33GFP _{uv} 7 L bioreactor culture, as shown in Figure 5.4. Relative CSH (%) and cell density (OD ₅₉₅) of non-induced and induced phases of the culture are shown.....	176
Figure 5.6. Trace from BioXpert Lite monitoring software of (f008) <i>P. pastoris</i> X33GFP _{uv} 7 L bioreactor culture with model-predicted optimal expression conditions and mixed feed induction as described in Chapter 3.....	177
Figure 5.7. Non-induced (BMGY batch) and induced (methanol fed-batch) <i>P. pastoris</i> X33GFP _{uv} 7 L bioreactor culture, as shown in Figure 5.6. Relative CSH (%) and cell density (OD ₅₉₅) of non-induced and induced phases of the culture are shown.....	178
Figure 5.8. Trace from BioXpert Lite monitoring software of (f009) <i>P. pastoris</i> X33GFP _{uv} 7 L bioreactor culture with standard process conditions (Invitrogen handbook) and mixed feed induction.	179
Figure 5.9. Non-induced (BMGY batch) and induced (methanol fed-batch) <i>P. pastoris</i> X33GFP _{uv} 7 L bioreactor culture, as shown in Figure 5.8. Relative CSH (%) and cell density (OD ₅₉₅) of non-induced and induced phases of the culture are shown.....	180
Figure 5.10. Structure of polypropylene glycol (Struktol P2000 antifoam).....	183
Table 5.1. Summary of antifoam influence on yeast (<i>S. cerevisiae</i> Alcofree and <i>P. pastoris</i>) growth and recombinant protein production (Fc fusion protein) in shake flask cultures.	184
Figure 5.11. XY scatter plot of GFP _{uv} protein concentration ($\mu\text{g/mL}$) and associated fluorescence (RFU).....	186
Figure 5.12. Normalised GFP _{uv} yield at 48 h post induction from triplicate shake flask cultures of <i>P. pastoris</i> X33GFP _{uv} containing a range of Fluka P2000 antifoam concentrations.	187
Figure 5.13. Normalised GFP _{uv} yield at 48 h post induction from triplicate shake flask cultures of <i>P. pastoris</i> X33GFP _{uv} containing a range of Struktol SB2121 antifoam concentrations.....	187
Figure 5.14. Fluorescence blanking experiments with Fluka SB2121 antifoam additions (white bars); BMMY medium, culture supernatant with existing GFP _{uv} (spotted bars).	188
Figure 6.1. Growth profiles of <i>S. cerevisiae</i> Alcofree 2xCBSE shake flask cultures.	195
Figure 6.2A. Growth curves (OD ₆₀₀) for <i>S. cerevisiae</i> 250 mL baffled shake flask cultures with 50 mL of 1 / 2 % YPE medium.	196
Figure 6.2B. Ethanol depletion curves for <i>S. cerevisiae</i> shake flask cultures shown in Figure 6.2A.	196

Table 6.1. Micro 24 well information for alternative carbon source experiment.....	198
Figure 6.3. Micro 24 culture growth profiles for <i>S. cerevisiae</i> Alcofree and W303-1A strains grown on CSM and CBS media with variable amino acid and carbon source compositions (as indicated by Table 6.1).....	199
Figure 6.4. Micro 24 culture growth profiles for <i>S. cerevisiae</i> W303-1A grown on CSM and CBS media with variable carbon source compositions (well contents as indicated by Table 6.1).....	200
Figure 6.5. Ranked 2xCBS and CSM media carbon source composition, based on <i>S. cerevisiae</i> Alcofree Micro 24 culture biomass at 48 h post inoculation.....	200
Table 6.2. Initial Micro 24 incubation cassette well layout for the nutrient addition and novel media cultures.....	201
Table 6.3. Micro 24 well information for nutrient addition and novel medium experiment.....	202
Figure 6.6. Micro 24 culture growth profiles for <i>S. cerevisiae</i> W303-1A grown in media as indicated by Table 6.3.....	203
Table 6.4. Micro 24 well information for repeat nutrient addition and novel medium experiment.....	204
Figure 6.7. Repeat Micro 24 culture growth profiles for <i>S. cerevisiae</i> W303-1A grown in media as indicated by Table 6.4.....	205
Figure 6.8. Combined biomass data from <i>S. cerevisiae</i> W303-1A Micro 24 cultures at 48 hours post inoculation.....	206
Table 7.1. Thesis objectives and current status.....	209
Figure 7.1. Overview of the optimisation process used for maximising GFP expression from <i>P. pastoris</i> X33GFP bioreactor cultures as well as other target proteins and expression systems.....	211

CHAPTER 1

1 INTRODUCTION

This thesis is concerned with the production of therapeutically-relevant recombinant proteins and particularly in understanding how to design more effective strategies for optimising production parameters and shortening the development time frame. The concept of scale down and its relevance to large scale coupled with multiple simultaneous experimentation was a driver for this work.

Since the vast majority of therapeutic proteins are not sufficiently abundant from natural sources, recombinant overproduction has become a universally-recognised solution to obtaining the required protein. These requirements are often multi-gram quantities required for applications as diverse as structural genomics and biopharmaceutical manufacture.¹ Routinely achieving high production yields in any of the currently-available host cell systems continues to be a substantial bottleneck to further progress and still relies on a directed trial-and-error approach. As such, sub-optimal product yields are often obtained and further experimentation is required for process scale-up. The work described in this thesis focused on the use of a design-of-experiments approach in a multi-well mini-bioreactor to enable the rapid establishment of high yielding production conditions in yeast and mammalian cell-lines.

The range of possible host cell factories suitable for producing recombinant proteins include microbes such as *Escherichia coli*, *Saccharomyces cerevisiae* and *Pichia pastoris*, as well as mammalian (e.g. CHO, NS0 and BHK) and insect (e.g. Sf9, Sf21, S2 and Tn5B1-4) cell-lines transfected with viral vectors. Each system has its individual benefits and drawbacks in their use. Yeast with its well-established genetic and molecular biological resources, combines the ease and speed of use of bacterial systems with its ability as a eukaryote to secrete post-translationally-modified proteins. As a consequence it is an increasingly popular choice in both academic and industrial laboratories² and was one of the host cells studied in this thesis.

Using green fluorescent protein secreted from *Pichia pastoris*, a scalable, predictive model of protein yield per cell was derived from 13 sets of conditions each with three factors (temperature, pH and dissolved oxygen) at 3 levels which was directly transferrable to a 7L bioreactor (Chapter 3). This was in clear contrast to the situation in shake flasks, where the process parameters could not be tightly controlled. With further optimisation of both the

accumulation of cell density in batch and improvement of the fed-batch induction regime, additional yield improvement was found to be additive to the predictive model.

The design of experiments (DoE) experimentally derived rational process optimisation approach was validated using mammalian cell culture (Chapter 4), highlighting its versatility, flexibility and product robustness. Investigations into the cell wall hydrophobicity of high cell density *P. pastoris* cultures indicated that cell wall, hydrophobin, (protein) composition changed with the growth phase. The cell wall becoming more hydrophobic in log growth than in the lag or stationary phases (Chapter 5). This is possibly due to an increased occurrence of proteins associated with cell division or cell cycle progression. Finally, a separate study demonstrated that improving biomass improved product yield in a second yeast species, *Saccharomyces cerevisiae* (Chapter 6). In summary, the work presented in this thesis has contributed to reducing process development time in recombinant protein production, directly from bench to bioreactor.

1.1 Recombinant proteins as pharmaceuticals

Virtually all biotherapeutic agents in clinical use are 'biotech pharmaceuticals', a biotech pharmaceutical being any medically-useful drug whose manufacture involves micro-organisms or substances that living organisms produce (*e.g.* enzymes, hormones). Most biotech pharmaceuticals are recombinant proteins— that is, produced by genetic engineering.³ One of the earliest biotech pharmaceuticals to be approved for use was insulin in 1982.

Steinberg has classified biotech pharmaceuticals (excluding gene-therapy and membrane proteins) into six categories:³

1. Cytokines *e.g.*: Interferons, Interleukins, Granulocyte-colony stimulating factor (G-CSF) and Granulocyte-macrophage colony-stimulating factor (GM-CSF)
2. Enzymes *e.g.*: Alteplase, Dornase alpha, Imiglucerase
3. Hormones *e.g.*: Insulin, Erythropoietin (EPO), Human Growth Hormone (hGH).

4. Clotting factors *e.g.*: Factor IX, Factor VIII.
5. Vaccines *e.g.*: Hepatitis B.
6. Monoclonal Antibodies *e.g.*: muromonab-CD3 (restrains immune response to transplanted hearts, kidneys and livers).

In 2000 there were eighty-four biotech pharmaceuticals in general medical use, with approximately 60 million patents being held world wide and an estimated 500 examples in clinical trials.⁴ In 2000, thirty-three new medicines received Food and Drug Administration (FDA) approval, 18 % of which were biologics; in 1998, 23 % of the FDA approved medicines were biologics.⁵ Major areas of current research into new biotech pharmaceuticals include vaccines, monoclonal antibody-based products, membrane proteins and gene therapy (*e.g.* anti-sense technology).^{3, 4}

1.2 Host systems for recombinant protein production

The manipulation and exploitation of micro-organisms and eukaryotic cell-lines plays an important role in the production of biotech pharmaceuticals. In particular, the use of recombinant production techniques is gaining an increased level of importance within medicine.⁶

Micro-organisms and eukaryotic cell-lines are used as host systems for protein production each with their own benefits and drawbacks. All host systems must meet stringent criteria for product consistency and cost-effectiveness. With regards to cost, systems that can provide a secreted, active, mature protein cheaply are extremely beneficial. Since most proteins for therapeutic use are of human origin, the corresponding recombinant protein is ideally produced in a host cell as close to its native cell as possible. The host systems currently employed by industry are summarised below.⁶

1.2.1 *Escherichia coli*.

Benefits include a short generation time, ease of handling, the fact that strains of the organism are well characterised both genetically and physiologically. *E.coli* has established well-characterised fermentation techniques and a high capacity for exogenous protein production (>20 % total protein).⁶ Drawbacks include the presence of endotoxins, a low secretion

efficiency (resulting in some proteins forming insoluble inclusion bodies) and limited post-translational modification.

1.2.2 Yeasts e.g. *Pichia pastoris*, *Saccharomyces cerevisiae*, *Schizosacharomyces pombe*.

Benefits include rapid growth on cheap media reaching high-cell-densities (130 g/L), complex post translational processing pathways, and the fact that they are non-pyrogenic and non-pathogenic. The drawbacks are few but the most important are the inappropriate glycosylation and hyperglycosylation of proteins which can be problems and the use of the current methanol inducible system. This brings the potential risk of fire or explosion.

1.2.3 Filamentous fungi e.g. *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus sydowii*, *Aspergillus awamori*, and various *Fusarium sp.* and *Trichoderma sp.*

These have complex post-translational modification apparatus. However, fungal physiology and their genetics are less well-understood than either *E. coli* or yeast. Fungi have a very good defence system in being able to degrade foreign DNA rapidly therefore, expression vector degradation and low transformation rate are also serious drawbacks.

1.2.4 Insect cell-lines e.g. *Spodoptera frugiperda* cell lines (sf9, sf21).

Insect cell-lines can be more stress resistant, productive and easier to handle than mammalian cell-lines these attributes make these cell lines suitable for high-throughput protein expression. However, insect cell lines can have high protease activity and difficulties with the scale-up regarding oxygen supply and carbon dioxide accumulation.⁶

1.2.5 Mammalian cell-lines e.g. Chinese Hamster Ovary (CHO), HeLa (cervical cancer cells taken in 1951 from Henrietta Lacks) and hybridoma (immortal cell line created by fusing B-cells and myeloma cells)

These provide the highest similarity in pattern and capacity of posttranslational modifications of human proteins. However, cultivation is complicated, costly and gives low product titres.

The cell types and expression systems summarised above each have individual benefits and drawbacks. The expression host selected to produce a recombinant protein will depend heavily on the role of the protein, such as bulk enzymes (wood/paper processing, washing

powder additives), drug target scouting, drug for regulatory acceptance *etc.* The requirement for a high quality, authentic recombinant protein may itself be the deciding factor in host selection or it could be a balance between this and cost-of-goods or production time. For the industrial expression of therapeutic proteins such as monoclonal antibodies, mammalian cell-lines (especially CHO) are widely used as they provide authentic molecules and have a history of regulatory (*e.g.* FDA) approval.⁷ The protein needs of academia and industry are not necessarily the same and generated material may be met by using different expression systems. For example, an academic laboratory expressing a protein for structural analysis would need an authentic molecule but not an FDA approved process, whereas industrial production of the same molecule for use as a therapeutic would need FDA approval.

1.3 Yeast as a host system

In choosing a host system, the balance between an inexpensive, easy-to-handle host system and one giving the most authentic environment for the target protein is one of the major challenges in designing a good protein production experiment. The production of recombinant human proteins is an area that is well suited to yeast expression as they provide the benefits of a eukaryotic host, can perform complex post translational modifications but are as easy to handle as the prokaryote, *E. coli*.⁸ Consequently, the yeasts *Pichia pastoris* and *Saccharomyces cerevisiae* were studied as production hosts in the work described in this thesis. Table 1.1 highlights some of the main characteristics of these organisms.

Characteristic	<i>S. cerevisiae</i>	<i>P. pastoris</i>
Industrial application	+	+
Requirement of explosion-proofed equipment	-	+
Food grade (GRAS)	+	-
Secretion efficiency	-	+
Hyperglycosylation	+	-
Episomal vector stability	+	-
Protease activity in secretion vesicles	High	Low

Table 1.1. Comparison of some characteristics of the yeasts *S. cerevisiae* and *P. pastoris*. Adapted from Schmidt, 2004.⁹

S. cerevisiae has been historically used in the production of food stuffs; bread and beer being prime examples. This has led to it being classified as 'Generally Regarded As Safe' (GRAS) by the U.S. FDA.⁸ In 1996 it was the first eukaryote to have its complete genome sequenced; following this, yeast has become a key organism in the field of genomics research.⁹ Its use in food production and the wealth of genetic information now available following the whole genome sequencing has established *S. cerevisiae* as one of the organisms of choice for the production of heterologous proteins.⁶

1.3.1 *Pichia pastoris*, an industry standard for protein expression

In the 1970's the methylotropic yeast *P. pastoris* was developed by the Phillips Petroleum Company as a potential source of Single Cell Protein (SCP) for animal feeds. However, the major oil crisis of that decade resulted in this form of SCP becoming an uneconomical option. Despite this, a wealth of information for achieving continuous high cell density cultures had been amassed.¹ During the 1980's, the Salk Institute Biotechnology/Industry Associates, Inc. (SIBIA, La Jolla, CA) developed *P. pastoris* for the production of recombinant proteins.¹ The highly inducible *AOX1* promoter controls methanol oxidase expression; the first enzyme in the methanol utilization pathway, which accounts for up to 35 % total protein in cells grown on 1% methanol.¹⁰ Using this promoter for induction of recombinant protein expression is a powerful tool.¹ The achievement of high density cultures with high yields (15 g/L murine collagen)⁶ of recombinant proteins has resulted in the *Pichia* system being the most frequently used host for heterologous protein expression in general.^{1, 6, 11} Today *P. pastoris* is available as an "off-the-shelf" heterologous protein expression kit from Invitrogen Ltd. However, *P. pastoris* does not have the classification of GRAS, has a privately-held genome sequence and requires special precautions when handling large amounts of methanol due to risk of fire and its toxicity.^{6, 12}

1.3.2 *S. cerevisiae* Alcofree™; a novel respiratory yeast

In this project, a unique respiratory *S. cerevisiae* Alcofree[™]¹³ has also been examined. Alcofree is a novel *S. cerevisiae* strain produced by Gothia Yeast Solutions (Gothenburg, Sweden). It has a modified glucose uptake system resulting in the cell-line having a fully respiratory metabolism. This was achieved by introducing chimeric hexose transporters limiting the uptake of glucose into the cell.¹³ Therefore, the strain does not produce ethanol under standard fermentation conditions (aerobic cultivation at high glucose concentrations). As the cell line does not exhibit a diauxic shift, more energy is available for protein expression and no alcohol

is present to depress growth rate. This results in a possible increase in recombinant protein expression.

The strain was designed to combine the industrial benefits of both *S. cerevisiae* (well understood genetics and physiology, open to manipulation) and *P. pastoris* (high biomass production). The anticipated industrial applications of the Alcofree principal include: (source www.gothiayeast.com)

- Heterologous protein production; higher yield with fewer by-products
- Production of fine chemicals; novel possibilities for metabolic engineering
- Food industry; yeast-based products with novel properties (flavour, taste)
- Beverage industry; production of alcohol-free or alcohol-reduced wine and beer and development of novel (functional) drinks

1.4 Target proteins

As this project was a CASE studentship, the majority of time was spent in the commercial laboratories of Alpha Biologics Ltd. This facilitated work with target proteins that had a direct industrial relevance either as proven success stories or potential drug product candidates. The initial target proteins were chosen because they had a proven record of production at an industrial scale or there were examples in the scientific literature using expression in yeast. The target proteins selected were green fluorescent protein, an anti-tumour factor antibody, insulin, human growth hormone and a novel monoclonal antibody-based influenza vaccine.

1.4.1 Green fluorescent protein (GFP)

Green Fluorescent Protein (GFP) is a widely-used tool in modern bioscience. It was originally isolated from *Aequorea victoria* commonly known as the crystal jellyfish. These jellyfish are primarily found in the waters off western North America, where Shimomura *et al* (Friday Harbour Laboratories of the University of Washington) collected specimens to harvest the light-emitting tissues. Consequently, GFP was identified as part of the bioluminescent mechanism of *Aequorea victoria*, being an accessory protein to aequorin. The 2008 Nobel

Prize in Chemistry was jointly awarded to Osamu Shimomura, Martin Chalfie and Roger Y. Tse for the discovery and development of GFP.¹⁴

1.4.2 The role of GFP

As GFP was initially isolated from the jellyfish *A. victoria*, its mode of action in this jellyfish is well documented. The generation of green light by the bioluminescent tissues of *A. victoria* is due to the interaction of calcium ions (Ca^{2+}) and the photoprotein, aequorin, which produces blue light. GFP converts this blue light to green and emits it into the jellyfish's environment. *A. victoria* may use bioluminescence as a basic communication method, as a warning for defence against predators or to find or attract prey.

Aequorin (Uniprot entry #P07164) is a Ca^{2+} -dependent bioluminescence photoprotein composed of two units; an apoprotein (apoaequorin) of approximately 22 kDa and coelenterazine (the chromophore), which is a prosthetic group of the luciferin family. In the presence of molecular oxygen these two units form the functional aequorin protein. The EF-hand (alpha-turn-helix domain, found in Ca^{2+} -binding proteins) structures in aequorin act as binding sites for Ca^{2+} . When Ca^{2+} binds, a conformational change occurs resulting in the oxidation of the prosthetic group to its excited state (coelenteramide) and a release of CO_2 . When coelenteramide returns to the ground state a blue light at 470 nm is emitted. GFP shifts aequorin's blue emission peak at 470 nm towards the green part of the spectrum at 510nm. This process is depicted in Figure 1.1 and explained further below.

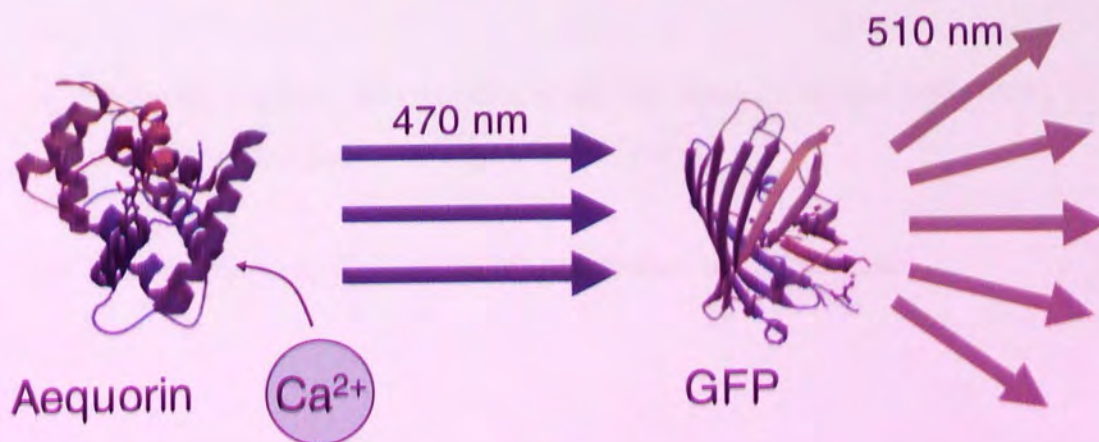


Figure 1.1. Process of green light emission from *A. victoria*. Ca^{2+} binds to coelenteramide, the prosthetic group of aequorin. When this returns to the ground state it emits blue light (470 nm) which is then converted to green light (510 nm) by GFP. Aequorin (DOI 10.2210/pdb1sl8/pdb) and GFP (DOI 10.2210/pdb1ema/pdb) images are from the protein data bank.

GFP has an unusual folding pattern of an 11 stranded anti-parallel β -sheet barrel (also called a β -can), enclosing an α -helix structure which contains the chromophore. Short α -helices are also found at either end of the barrel. GFP contains 238 amino-acids, forming 6 α -helices (24 residues) and 14 β -strands (117 residues; UniProt reference: P42212).

The primary citation for the structure of GFP, solved by X-ray diffraction (1.90 Å resolution) is from Ormo *et al* (1996),¹⁵ PDB accession number 1EMA, using multiple anomalous dispersion of selenomethionine groups to generate phasing information on recombinant GFP. However, GFP was first crystallised in 1974 and had a diffraction pattern reported in 1988.¹⁶ Numerous protein structures have been solved for GFP, both native and with specific mutations, indicating how important GFP has become as a research tool.

The GFP chromophore, p-hydroxybenzylideneimidazolinone, is located in the centre of the β -can structure and is composed of residues 65, 66 and 67; serine, tyrosine and glycine, respectively (Figure 1.2). The cyclised backbone of these residues forms the imadazolidone ring of the chromophore. The fluorophore is not accessible to solvents, but does interact with the side-chains of surrounding residues by hydrogen bonding. GFP chromophore formation occurs as follows, also shown in Figure 1.2. Key features are:

- GFP folds into semi-native conformation
- Imidazolinone is formed by nucleophilic attack of the amide (Gly67) on the carbonyl of residue 65, followed by dehydration.
- Molecular oxygen dehydrogenates the α - β bond of residue 66, putting its aromatic group into conjugation with the imidazolinone.
- The chromophore now has visible absorbance and fluorescence.

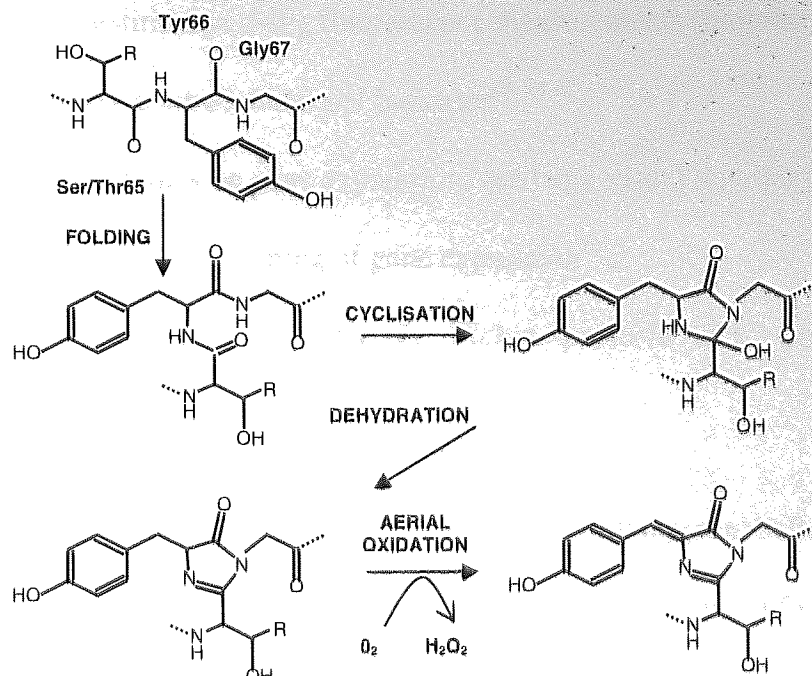


Figure 1.2. GFP chromophore (p-hydroxybenzylideneimidazolinone) formation of active state. Adapted from Tsien (1998)¹⁶

1.4.2.1 Protein engineering of GFP

Numerous mutations have been performed on native GFP to alter the characteristics to better suit specific applications. The PDB, accessible at <http://www.rcsb.org/pdb/home/home.do> provides 183 structures for GFP (accessed on 9th December 2008) many of which show specific mutations that infer non-native traits such as alternative excitation/emission wavelengths (different coloured light; yellow, blue, red fluorescent proteins to allow multiple labelling),^{17, 18} reduced maturation time (quicker detection of gene regulation), codon optimisation (allows improved expression in different species),¹⁹ optimisation for use as a tag in fluorescence activated cell sorting (FACS) analysis,²⁰ improving thermostability (allows the use of GFP at elevated temperatures)²¹ and alternative pH tolerances.²² As new applications are devised, the structure of GFP continues to be altered to better suit the task at hand.

1.4.2.2 GFP as a tool

GFP can be used for imaging intra-cellular processes without the need for invasive techniques or the use of co-factors, as with luciferase systems. There are many examples of the uses of GFP. The majority are directly relevant to biological research, however some are applicable to industrial bio-processing and include:

- Biomass estimation using fluorescence measurements (bioprocess monitoring)²³
- Non-invasive intracellular pH indicator^{22, 24}
- Real-time studies on gene expression, used as a visual reporter (*e.g.* *GAL1*)²⁵
- High throughput screening of gene expression²⁶
- Protein expression reporter in *Pichia pastoris* (has limitations due to pH dependence)²⁷
- Quantitative fluorescence microscopy²⁸
- Real-time monitoring of protein localisation (GFP tagging) cellular and sub-cellular.¹⁹
- Visualisation of physiological processes, including neuronal function, identification of cells for *in vitro* and *in vivo* manipulations *etc.*²⁹
- FACS (fluorescence activated cell sorting) labelling positive clones, cell types, tissue *etc*²⁰
- Monitor protease activity³⁰
- Monitor transcription factor dimerisation³¹
- Ca^{2+} detection³²

Overall the applications of GFP in cell biology can be classified according to its use as a tag or as an indicator. When used as a tag, fluorescence can provide information on gene expression level or sub-cellular localisation due to targeting domains or host proteins which fuse GFP. When used as an indicator, fluorescence is controlled post-translationally by environmental factors, such as pH, or by protein-protein interactions. When GFP is used as an active indicator, fluorescence resonance energy transfer (FRET) between two fluorescent proteins, with different emission spectra, is utilised. FRET will occur when the emission spectra of the two fluorescent proteins overlap so that the emission peak of one (*e.g.* blue fluorescence protein) matches the excitation peak of the second (*e.g.* GFP). The two fluorophores must be within 100 Å of each other for this quantum-mechanical effect to take place. Any biochemical event that changes the spatial relationship of the two fluorophores will also alter the efficiency of FRET. Many of the available GFP mutants were created to give novel FRET characteristics.

1.4.3 Monoclonal antibody raised against a tumour inducing factor

The monoclonal antibody used here was made available as part of an ongoing cGMP development project at Alpha Biologics Ltd, and was the second target protein used in this thesis. The ability to generate antibodies that specifically recognise and bind to tumours makes immunochemotherapy and radioimmunotherapy possible, by binding chemotherapy or radiotherapy agents to an antibody and used to directly target cancerous tissue.³³ This is a very appealing proposition as this will target only those cells which display tumourous characteristics and not 'normal' tissue, as monoclonal antibodies are highly selective in their binding. If these antibodies are labelled, they can also be used for the identification of tumour tissue during traditional surgery.

1.4.4 Insulin

1.4.4.1 The role of insulin in healthy humans

Insulin was chosen as the third target because it is one of the most famous biotech pharmaceuticals and has been produced in yeast on an industrial scale. The peptide hormone insulin plays a number of roles in human metabolism as a whole and on individual cells. The level of insulin in the blood circulation throughout the body, can have exceptionally widespread and varying effects on different cellular processes. Actions on human metabolism include the control of cellular uptake of glucose in muscle and adipose tissue, control of DNA replication and protein synthesis *via* regulation of amino acid uptake and modifying the activity of numerous enzymes *via* allosteric interaction.

Insulin acts upon cells with a range of critical effects:

- Glycogen synthesis in liver and muscle cells. Insulin induces these cells to store glucose in the form of glycogen. Lower levels of insulin results in liver cells converting glycogen to glucose and secreting it into the blood stream. This is the clinical action of insulin.
- Fatty acid synthesis. Insulin causes fat cells to accumulate blood lipids which are then converted to triglycerides. Lowered insulin levels lead to a reverse in this process *i.e.* conversion of triglycerides to blood lipids.

- Esterification of fatty acids. Insulin causes adipose tissue to make triglycerides (fats) from fatty acid esters. Lowered insulin levels lead to a reverse in this process.
- Proteinolysis. Insulin suppresses protein degradation. Lowering insulin levels results in removal of this suppression.
- Lipolysis. Insulin reduces the conversion of fat cell lipid stores to blood fatty acids. Low insulin levels lead to an increased conversion of fat stores to blood fatty acids.
- Gluconeogenesis. Insulin decreases the production of glucose from substrates in the liver. Reduced insulin levels results in increased glucose production.
- Amino acid uptake. Insulin promotes the uptake of amino acids by cells. Low insulin levels lead to an inhibition of their absorption.
- Potassium uptake. Insulin promotes the uptake of serum potassium by cells, low insulin levels lead to an inhibition of this absorption.
- Arterial muscle contraction. Insulin causes arterial wall muscle to relax so increasing blood flow, especially in micro arteries, low insulin levels allows these arteries to contract so reducing blood flow.

1.4.4.2 Insulin and disease

There are several conditions associated with atypical insulin distribution. Diabetes mellitus is the general term referring to states characterised by hyperglycaemia. The major conditions are Type 1, 2 and 3 diabetes (described below), however other conditions do involve insulin *e.g.* insulinoma, Reaven's syndrome (metabolic syndrome) and polycystic ovary syndrome.

- Type 1 diabetes is the autoimmune-mediated destruction of insulin-producing pancreatic beta-cells resulting in an absolute insulin deficiency. There is also the production of antibodies raised against insulin and islet cell proteins. The foremost

treatment for type 1 diabetes is insulin replacement. Without insulin, ketosis and diabetic ketoacidosis can develop and coma or death will result.

- Type 2 diabetes, cells become resistant to insulin and so do not take up the glucose they require. This form of diabetes is highly inherited but environmental factors are also influential. Factors that contribute to susceptibility include obesity, age and physical inactivity.
- Type 3 or gestational diabetes is found in pregnant women. It is a combination of inadequate insulin secretion and responsiveness, resembling type 2 diabetes. It is temporary and fully treatable. However, if untreated there may be issues with the pregnancy, such as macrosomia (significant foetal overgrowth), foetal malformation and congenital heart disease. There is no specific cause but it is believed that hormone levels reduce receptivity of insulin leading to high blood sugar.

Insulin is phylogenetically ancient; its presence is observed in mammals, reptiles, bird and fish with a similar conserved sequence encoding it. Insulin is synthesised from precursor molecules; preproinsulin and proinsulin in the β -cells of the Islets of Langerhans within the pancreas. The proinsulin precursor comprises of B-chain C-peptide A-chain, as shown in Figure 1.3. The C-peptide shows the highest level of divergence between species.³⁴

Proinsulin is processed by enzymatic cleavage to form insulin plus a C-peptide. (Figure 1.3). This is achieved by the action of two endopeptidases; prohormone convertase 2 and 3 (PC2 and PC3) at two dibasic amino acid sites. PC3 cleaves at Arg³¹-Arg³² at the junction of the B/C chains and PC2 cleaves at Lys³⁵-Arg⁶⁵, the A/C chain junction. Both the PC2 and PC3 forms are calcium-dependent and have acidic pH optima; however they differ in their kinetics.³⁴

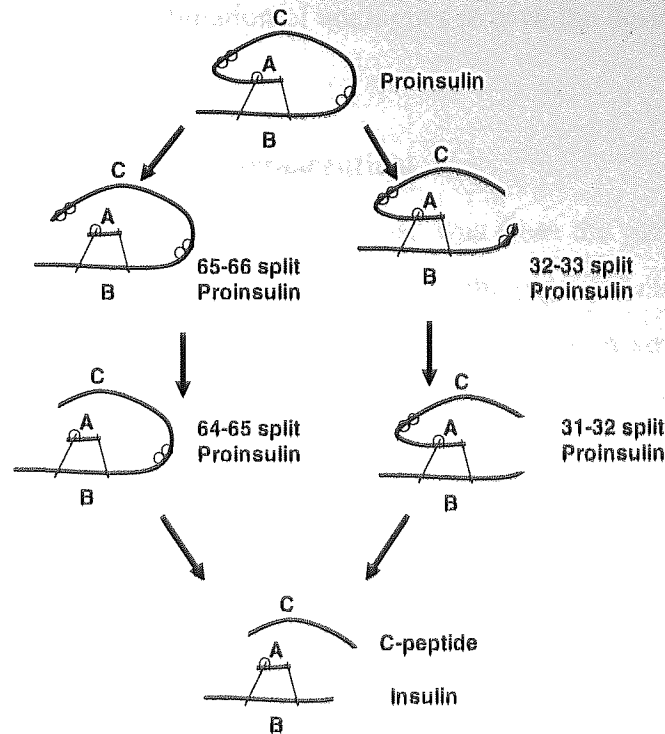


Figure 1.3. Enzymatic conversion of proinsulin to mature insulin and C-peptide. Adpated from Clark (1999)³⁴

The PDB has 279 (accessed on 9th December 2008) model structures relating to the insulin molecule. Many of the structures are of structural mutants; the interest in these mutants lies in their activity and half life when administered as treatment for medical conditions. The structural modifications present in these analogues were engineered into the molecules using recombinant techniques. The analogues possess characteristics that improve their value as a drug over native insulin. Commercially available insulin analogues fall into the following categories:

- Quick-acting; begins working 5 to 15 minutes post injection and remains active for 3 to 4 hours (*e.g.* lispro)
- Short-acting; begins working within 30 minutes post injection and remains active for 5 to 8 hours (*e.g.* native insulin)
- Intermediate-acting; starts working 1 to 3 hours post injection and is active for 16 to 24 hours (*e.g.* NPH or lente insulin)
- Long-acting; starts working 4 to 6 hours post administration and is active for 24 to 28 hours (*e.g.* ultralente)

The dose may contain a combination of analogues to give the desired metabolic response.

1.4.4.3 Insulin as a biotech pharmaceutical

The historical supply of insulin for clinical use was from the pancreas of cows, horses, pigs and fish. Of the animal sources for insulin, that of the pig most closely mimics the regulatory function of human insulin. The high level of homogeneity in structure allows cross-species reactivity. The purity of the extracted insulin was a problem both in relation to allergic reactions to non-insulin substances that were co-purified and viral contamination. Despite these issues with purity, insulin harvested from animal tissue was used to successfully treat diabetes mellitus. The first patient was treated with insulin injections on the 11th January 1922 from material produced by Frederick Banting, Charles Best and James Collip. Unfortunately the material, extracted from foetal calf pancreas, was impure and caused a severe allergic reaction. However, 12 days later following the production of an improved extract, Leonard Thompson a fourteen year old diabetic received a second dose and completely eliminated the glycosuria (elevated levels of glucose in the urine). Banting and the Laboratory Director J.J.R. Macleod received the Nobel Prize in Physiology or Medicine for the work in 1923.

Further improvements to the processes of extracting and purifying insulin from the pancreas of animals led in 1923 to Eli Lilly producing commercial quantities of much more pure bovine insulin for medical use. Despite these improvements, allergic reactions continued to arise although with reduced severity. Bovine and porcine derived insulin were produced by Eli Lilly until 2005.

Recombinant human insulin is now manufactured for widespread clinical applications using genetic engineering techniques. These greatly reduce issues associated with animal derived products i.e. contamination and allergenic response. In 1982 Eli Lilly was the first company to reach market with a product (Humulin; human insulin analogue) manufactured using modern genetic engineering techniques. Genentech and Novo Nordisk also market human insulin analogues, manufactured using genetic engineering. Both *E. coli* and yeast systems are used at production scale for insulin manufacture.

1.4.4.4 Recombinant production of insulin

Since recombinant human insulin was first produced in *Escherichia coli*³⁶ and yeasts,³⁷ it has been possible to use a recombinant insulin analogues rather than an animal derived product in clinical applications.

Expression in *E. coli* is as proinsulin or as separate A- and B- chains. The expressed product formed inclusion bodies (aggregated recombinant protein held intracellularly) which require denaturing using caltropic agents such as high molarity urea or guanidine-HCl followed by a refolding step. This increases the amount of handling required to obtain a functional form of the molecule therefore increasing manufacturing costs and decreasing yields.

Yeast systems have been used to minimise the amount of handling required for a functional molecule. A single chain insulin molecule, with appropriate post-translational modifications, can be secreted in a correctly folded active form at high levels. Yields of up to 1.5 g/L are reported in the literature.³⁸

For efficient expression of insulin homologues in yeast systems (*S. cerevisiae* and *P. pastoris*) the native human proinsulin (A-chain – C-chain – B-chain) cannot be used. The C-peptide must be replaced with a short (tri-peptide,³⁹ pentapeptide⁴⁰) linking sequence forming an A-chain – B-chain fusion protein (Mini Pro-Insulin; MPI).⁴¹ A leader sequence (prepro-peptide), such as the *S. cerevisiae* α -mating factor, is also required to promote secretion of the MPI protein into the culture medium by the secretory pathway.⁴²

By using protein engineering and recombinant technologies, the previously-mentioned insulin analogues can be tailor-made to suit the intended applications, which would not be possible with animal-derived insulin products.

1.4.5 Human growth hormone

Human growth hormone-2 (hGH2) cDNA codes for human placental growth hormone (hPGH) or placental specific growth hormone, which is a member of the somatotrophin/prolactin family of hormones and is the placental variant of pituitary growth hormone. The choice of hPGH and not hGH was made because of cDNA availability. Despite hGH having more of an industrial relevance, this was not seen as a problem as hGH and hPGH are very similar and hPGH is also an important target. Of the 191 residues in the

mature proteins (hGH and hPGH), there are only 13 residue substitutions (6.8% divergence), however, native hPGH shows glycosylation whereas hGH does not.⁴³ Although hGH was not used in the work described here, the physiological role, conditions relating to hGH and its recombinant production is also described on account of its industrial relevance.

1.4.5.1 Role of human growth hormone in healthy humans

1.4.5.1.1 Role of hPGH

In healthy humans hPGH appears in the maternal serum in mid-pregnancy and rises in concentration thereafter to term, being secreted by the placenta. The hPGH is not expressed in the anterior pituitary gland as with growth hormone. The hGH2 gene, along with four other closely-related genes, is located on chromosome 17. They share a remarkable degree of sequence identity, with alternate splicing generating additional isoforms of all five hormones. hPGH stimulates the liver and other tissues to secrete insulin-like growth factor 1 (IGF1). It also stimulates both the differentiation and proliferation of myoblasts. Additionally it stimulates amino acid uptake and protein synthesis in muscle and other tissues. There is evidence that circulatory hPGH levels during pregnancy regulate IGF levels and foetal growth in mothers with Type I diabetes.⁴⁴ However, hPGH has not been shown to possess lactogenic activity.⁴⁵

1.4.5.2 Synthesis/structure of native hPGH

The hGH2 gene product, hPGH, is specifically expressed by the syncytiotrophoblast layer of the placenta in a continuous manner. Pituitary hGH is expressed in a pulsatile 24 hour cycle. Anti-hPGH monoclonal antibody assays have shown hPGH to be found in the maternal circulation from 15-20 weeks to term, but not detectable in the foetal circulation.⁴⁶ Secretion of hPGH has been shown to be inhibited by glucose. Women with gestational diabetes show an increase in glycaemia and a decrease in circulating hPGH following oral glucose load. This suggests that hPGH regulation can protect the foetus against reduced nutrient availability. The hPGH strongly stimulates maternal organs, such as the liver, to switch to gluconeogenesis, lipolysis and anabolism, therefore increasing nutrient availability to the fetoplacental unit.

The structure of hPGH has the following features: 1-26 signal peptide, chain 27-217, single glycosylation site at 166, two disulphide bonds at positions 79 to 191 and 208 to 215.

(Uniprot entry # P01242). As with other members of the somatotropin family hPGH forms a structure with four α -helices.

1.4.5.3 Role of hGH

Human growth hormone is secreted from somatotropes in the anterior pituitary in response to a number of factors including:

Stimulators; growth hormone releasing hormone (GHRH), sleep, exercise, hypoglycaemia, dietary protein levels, elevated androgen levels.

Inhibitors; somatostatin, circulatory hGH and IGF-1 (negative feed-back loop), hyperglycaemia, glucocorticoids and estradiol.

The secreted hGH is found in the circulatory system and performs a number of physiological roles, most notably anabolic growth in children. The mechanism of action is either by direct stimulation of chondrocyte division (found in the cartilage); these are the "growth centres" (epiphyses) for the body's long bones such as arms and legs. The second mechanism is by hGH stimulation of IGF-1 expression in the liver; IGF-1 has a growth-inducing effect in a range of cell types including muscle, cartilage, bone, skin liver, lungs and kidneys.

1.4.5.4 Human growth hormone and disease

Mutations in the hGH2 gene can lead to placental growth hormone/lactogen deficiency (NCBI Entrez GeneID: 2689). It has been noted that hPGH levels are significantly elevated in cases of Down's syndrome and can be used as a screening target, giving a 71.9 % detection rate; a gold standard test provides 65.6%.⁴⁷

Diseases relating to hGH correspond to abnormal, either depressed or elevated, circulatory levels of the hormone. hGH excess is often the result of tumour formation in the pituitary region of the brain, composed of somatotroph cells (somatotroph adenoma). These are generally benign, slow growing tumours which as they grow produce increasing amounts of hGH. High circulatory levels of hGH result in excess bone development (increased thickness) in the jaw, digits and toes, which in turn can lead to carpal tunnel syndrome and muscle weakness. The tumour itself may exert pressure on the surrounding tissue resulting in headaches, impaired vision and reduced secretion of other pituitary hormones.

Deficiencies in hGH are the other major class of medical conditions relating to hGH; the effects of the deficiency are dependant on the age of the sufferer if present from infancy/early childhood or in adults.

Prenatal severe hGH deficiency can lead to hypoglycaemia and jaundice, but does not lead to reduced growth. Severe hGH deficiency in early childhood results in reduced muscle development and altered body composition (ratios of bone:muscle:fat) with a 50 % reduction in rate of growth and physical maturation. If severe hGH deficiency is present and not treated from birth, peak adult height can be as low as 150 cm.

Adult hGH deficiency results in abnormal body composition leading sufferers to have reduced strength and mass in muscles and bones. The effects of adult hGH deficiency are not limited to just reduced physical strength but extend to impairment of cognitive processes (reduced concentration and memory), depression, elevated cholesterol levels and mortality due to cardiovascular disease.

1.4.5.5 hGH as a biotherapeutic

The main therapeutic use of hGH is for the treatment of the various diseases resulting in short stature including hGH deficiencies,^{48, 49} described above. Since recombinant hGH has been available it has also been misused as a doping agent to enhance athletic performance⁵⁰ and has been reported as a "treatment" for old age.⁵¹

During the 1960's purified pituitary hGH began to be used as a treatment for severe hGH deficiency in children. The national agencies, such as the national pituitary agency (NPA), were set up to supervise the collection, purification and distribution of hGH obtained from the pituitary glands removed from cadavers. Treatment using human-derived hGH continued until 1981 when Genentech began trials of a recombinant form of hGH expressed in *E. coli*. This coincided with several reports of people contracting Creutzfeldt-Jakob disease (CJD) after receiving human-derived hGH.⁵² The link between the two was recognised in 1985 and Genentech's rhGH (Protropin) was rapidly granted FDA approval. Other manufacturers quickly released recombinant forms of hGH including Eli Lilly, Pfizer, Novo Nordisk and Serono.

1.4.5.6 Recombinant production of hPGH (rhPGH)

Commercial rhPGH is available from AbD Serotec (Oxford, UK) (Cat# OBT1879G). This rhPGH is expressed in *E. coli* and is primarily for use in enzyme-linked immunosorbant assays (ELISAs) as a standard. Many of the examples in the literature of rhPGH production also use *E. coli* as the expression system,^{43, 53 54} but there are cases where yeasts (*P. pastoris* and *S. cerevisiae*) are used as the production host.^{8, 55-57} In these situations the hGH is used as a target protein to evaluate novel expression technologies (promoter sequences, transcription factors) or novel strains (protease deficiencies).

1.4.6 Novel monoclonal antibody-based influenza vaccine

The final target protein was highly industrially orientated as the work on it formed part of an Alpha Biologics client contract. The client company wanted to demonstrate proof-of-concept that the yeasts *S. cerevisiae* and *P. pastoris* could be used to rapidly express their molecule whilst maintaining efficacy of the end product. If expression were feasible in yeast, this would reduce the time and cost for development and the clinical production when compared to the baculovirus-mediated mammalian expression system currently being used. This could be of particular importance in instances of vaccine production where a rapid turnaround from the point of isolating a newly emergent strain (e.g. during the 2008/2009 season, H1N1 and H3N2-like viruses were predicted to be most prevalent) to supplying medical practitioners with the specific vaccine can provide protection/immunisation to the populous. In this instance the speed would enable the prevention of an epidemic or potential pandemic such as Spanish (H1N1) or Avian (H5N1) flu.

1.5 Optimisation of culture conditions for the expression of recombinant proteins

Optimisation of culture conditions or 'process control parameters' for the expression and production of biotherapeutic agents is a key element to the rapid and cost effective manufacture of these important molecules.⁵⁸ It has previously been shown that optimal production conditions do not necessarily support optimal growth and that point of harvest can be critical for achieving high yields.⁵⁹ Usually, process optimisation is conducted on a case-by-case basis and is thus essentially a trial-and-error-based procedure. Major factors often considered when producing proteins from micro-organisms include pH, temperature, carbon and nitrogen sources and the essential oxygen requirement. Although these are most the frequently investigated factors, there are many other factors which offer further possibilities

for optimising the protein expression process. The factors covered in the work here were temperature, pH, dissolved oxygen (DO), carbon sources, trace element/nutrient additions, antifoam agent and concentration, and induction strategy. There are many other factors which affect recombinant protein expression, such as expression vector design including promoter sequences, leader sequences and codon optimisation which can be manipulated to give increases in yield.^{35, 53, 60} These other factors relate to the molecular biology of the chosen expression strategy, itself a large and important field of research critical for successful production of any target protein.

As there are such a large number of possible avenues for investigation to obtain an optimised process, simply changing one factor at a time is not only time consuming but may also miss critically important interactions between the factors. To circumvent these issues statistical techniques such as statistical design of experiments and response surface methodologies can be employed to develop a process. These techniques provide the maximum amount of data from an experimental period and can estimate factor interactions,⁶¹ resulting in a reduction of required experimentation. A way of further increasing the throughput of the optimisation process includes scalable modelling, where the protein expression experiment is carried out in "micro-bioreactors" (e.g. 6 mL) and directly transferred to large(r) (e.g. >7 L) vessels. The use of micro-bioreactors allows many parallel experiments to be run at once so reducing the overall development time. Using traditional stirred tank bioreactors it may be possible for one operator to run 6 experiments in parallel, however micro-bioreactors could allow >72 parallel experiments. Increases in throughput can make it possible to investigate additional factors in the experimental period which would otherwise be overlooked.⁶²

1.5.1 Temperature

Optimum temperature must be maintained, either by heating or cooling the medium with the use of a probe and a control system coupled to an incubator, heat exchanger or heating jacket. Optimal temperatures for growth are different for each organism, typically 37 °C for *E. coli* and 30 °C for yeasts. Optimising temperature in relation to recombinant protein expression focuses on maximising the yield of protein produced per cell. It is likely that the optimal temperature for cell growth and recombinant protein expression will be different. Low temperature induction conditions in *P. pastoris*^{61, 63, 64} and *E. coli*⁴³ or high temperature (42 °C) induction expression in *E. coli*⁶³ have been shown to provide increases in protein yield. In many examples of recombinant protein expression found in the literature, the same

temperature is used for growth and expression.^{65, 66} In cases where temperature has been optimised for expression, increases in yield can be seen.⁶³

1.5.2 pH

During fermentations, the pH of the medium changes as the organism grows. A defined setpoint can be controlled by the addition of acid, or more usually, base. A pH probe connected to a control unit and a pump maintains a pH set-point. The pH set-point is not only important for the optimal growth of the organism but also for the optimal expression of the recombinant protein. The pH of the medium will also influence the stability of secreted proteins depending on the isoelectric point (pI) of the protein. The pI of a molecule is the pH at which the molecule has no net charge; this is determined by the amino acids displayed on the surface. It has been shown that using a low pH in *P. pastoris* fermentations leads to a decrease in proteolytic degradation of secreted recombinant protein.⁶¹ The medium may have a buffering capacity, which is especially useful in shake flask cultures. For example, standard BMGY/BMMY media contain potassium phosphate buffer adjusted to pH 6.0. The medium component used to provide buffering capacity can be altered depending on the desired pH value; e.g. citrate buffer, phosphate buffer or HEPES. The specific acid and base used for the pH control will be important as different cell types may not have the same tolerances. This gives several variables that can be optimised for the application; pH set-point, buffer range, acid type and concentration, base type and concentration.

1.5.3 Aeration

The demand for oxygen by a micro-organism can be met by aerating the culture medium. The values associated with aeration are typically (%) dissolved oxygen (DO); where 100 % is oxygen saturated (with respect to air, (wrt air)) and 0 % has no oxygen in the medium. Aeration of the medium is normally achieved by sparging sterile air or oxygen-enriched air directly into the medium or into the headspace of a culture. High DO levels may require additional agitation and/or increased gas flow rates to maintain the desired setpoint. Increasing the agitation of the culture (different impellor design, additional impellers, increased agitation rate (rpm)) extends the retention time of the gas bubbles in the culture therefore improving the mass transfer of oxygen from the gas to the medium and hence to the organism.⁶⁷ Sparger (a tube with numerous small holes which gas is passed through to form bubbles) design may also play a role in improving oxygen transfer by having smaller holes and thus smaller bubbles giving a larger surface area for diffusion to occur over.⁶⁸ The DO

set-point for an experiment is likely to be determined by the tolerance of the cells to high DO levels, equipment capabilities/limitations, cost and optimal levels for protein expression or growth.^{67, 69} For instance it may not be economical to run a high cell density fermentation at 100% DO as this may require very high agitation and gas flow rates with pure oxygen. An additional possibility for maintaining high DO is to introduce back-pressure to the gas flow in the system, so aiding diffusion into the liquid phase: this is possible in stainless steel vessels but not possible in a glass fermentation vessel.

1.5.4 Medium design

The type of medium used is dependent on the system's requirements, promoter strategy and cost. There is a wide range of medium recipes available including complex, defined, pH buffered and selective. The medium used for a particular application will reflect a specific organism's preference for carbon and nitrogen sources. The medium should provide all required nutrients without excess which can result in nutrient inhibition. Medium design is the optimisation process used to achieve a medium most suited for the particular application; improving growth, increasing product yield or other goals such as reducing process costs. Spent medium analysis can be used to determine which, if any, medium components are depleted/consumed and which are still in excess after being used for an experiment. It can also identify accumulation of toxic metabolites. From this analysis it is possible to further improve the medium, by either increasing or decreasing the starting concentrations of the relevant nutrients. Due to the potential for having a large number of variables (carbon sources, nitrogen sources, essential and non-essential amino acids, trace elements and other additives such as foetal bovine serum) this subject is well suited to design of experiments methodologies as described in the literature.⁷⁰⁻⁷³

1.5.5 Anti-foam agents

An unfortunate effect of both sparging gas through culture medium at high rates and agitation is the formation of foam. This is a particular problem when surface-active species, such as proteins, are present at high concentrations. Foaming reduces the efficiency of gas exchange at the surface of the culture, as a barrier is formed between the culture and the gases in the headspace of the vessel. Foaming can also be detrimental to the cells in culture. For example, when bubbles burst they exert sheer forces which may damage cells or secreted proteins. They can also cause a loss of cells from the culture to be made by depositing cells above the air/medium interface. Additionally cells and culture medium can be lost to the foam phase

which can lead to a decrease in process productivity. In extreme cases of foaming, 'foam out' can lead to loss of process sterility when rising foam enters the off-gas filters or other ancillaries in the headspace.⁷⁴

In order to minimise the detrimental effects of foaming, an antifoam agent can be included in the culture medium. Antifoams prevent foam forming by reducing the surface tension of the culture so that any bubbles that reach the surface cannot be supported and therefore disperse without bursting.⁵⁸

1.5.6 Induction strategy/protocol

The strategy for inducing recombinant protein expression needs to be defined when designing the expression cassette for the intended cell type. There are many promoter types used including *AOX1* (*P. pastoris* alcohol oxidase 1; methanol inducible promoter), IPTG (system used in *E. coli*, CHO), temperature mediated (*E. coli* temperature drop from 37 to 30 °C) and galactose induction (*S. cerevisiae*). The ideal system would be highly inducible, very tightly regulated (not "leaky"), being acceptable to regulatory bodies (FDA) for use in large scale manufacture and practical for use at large scale.

In the case of the *AOX1* promoter, as used in the work described here, the method of inducing expression can be varied. The standard method of induction uses a feed of 100 % methanol. Alternatives such as mixed feed induction use a mixture containing methanol and sorbitol instead of 100 % methanol and methanol-free induction of *AOX1* promoters has also been reported.^{75, 76} Once the promoter system is determined, the method will require optimisation. In the instance of the *P. pastoris AOX1* system the rate of methanol addition is critical in optimising recombinant protein expression. If the rate of addition is too slow methanol will be limiting so resulting in sub-optimal yield, but if the rate is too fast the cells will not be able to metabolise the available methanol so resulting in a build up of methanol in the medium. Once at a critical concentration this becomes toxic to the cells and they die.^{77, 78} Therefore, the methanol feed rate must be balanced against the rate of uptake and metabolism of the cells and their specific growth rate (μ).⁷⁹ This will be dependant on the culture conditions used during induction i.e. low temperatures may give a slower rate of cell metabolism, so methanol feed rate should be adjusted to meet this.

1.5.7 Scale-down modelling and Design of Experiments (DoE)

Small-scale culture techniques, based on shaken vessels, have been in use since Erlenmeyer flasks (designed for chemistry laboratories) were first adopted for microbial culture in the 1880's and later for the first instance of biotechnology, namely antibiotic production in the 1940's.^{80, 81} More recently, during the 1980-1990's, microtiter plates (MTP) have also found a role as culture vessels, providing increased throughput for applications such as mutant screening and secondary metabolite discovery.⁸⁰ These vessels have recently been the focus of characterisation using an engineering approach to determine the hydrodynamics, mixing and mass transfer characteristics, previously only applied to large stirred tank bioreactors.⁸⁰ The characterisation of these vessels, along with the introduction of more sophisticated micro-bioreactors with elements of monitoring and control,⁸² has begun to allow processes developed at the small scale to be (more) directly transferred to large-scale production (by scalable modelling). Scale-down modelling applies to situations where a process is already in use at the commercial large-scale and requires investigations into process changes. Conducting these experiments in production scale vessels would be undesirable due to economic, strategic or environmental reasons.⁸³ Therefore using small-scale vessels that have equivalent characteristics to the production vessel is ideal.

The combining of methods for scale-down modelling of industrial fermentation processes and statistical design of experiments (DoE) provides the opportunity for developing an optimised up-stream process in a compact time frame. Scale down modelling allows large-scale processes to be performed at a small-scale; perhaps 1000 fold smaller volumes. The attraction of this is the reduced cost and time required for conducting the experiment, which is further enhanced by using multiple parallel experiments. Statistical DoE is an efficient way of planning an experiment to provide data capable of giving objective and valid conclusions, in this case optimising a fermentation in a reduced number of runs compared to investigating each variable separately as with one factor at a time (OFAT) methods. The statistical nature of DoE methods also allows the determination of interactions between independent variables (*e.g.* culture temperature and pH) and how they impact on the response for example protein yield.

1.5.7.1 Scale down modelling of the fermentation process

Scale-down process modelling is used to determine the effects of altering a production scale fermentation process from small-scale experiments. This frees-up production equipment for

product generation and avoids the cost implications (such as the raw materials and utilities.) of performing development runs at a large scale.⁸³ Any product generated, however would not be acceptable as cGMP grade material, due to process generated protocol deviations. Scale-down modelling is applicable to fields such as natural product fermentations (*e.g.* antibiotics, citric acid and enzymes) and recombinant protein expression (*e.g.* hormones, monoclonal antibodies and enzymes) where the commercial scale makes process development activities in those large vessels unfavourable (*e.g.* expensive, time consuming, cumbersome and tie up manufacturing resources).⁸³

Scale-down modelling is essential for process validation and commercial manufacturing including process characterisation and supporting production processes.⁸⁴ The goal of a scale-down experiment is to define a lab-scale system which replicates the large scale process, therefore allowing changes made at the small scale to be reproducibly implemented in the large scale. If the scale-down model produces inaccurate data for the large-scale system, inappropriate conclusions may be drawn resulting in a sub-optimal large-scale process. One of the major factors in constructing a valid scale-down model is the use of appropriate equipment capable of replicating the environment present at the large scale *e.g.* mass transfer coefficients. This is referred to as having a similar reactor or similar reactor geometry *e.g.* aspect ratio. This is simplified by using comparable vessels (*e.g.* stainless steel) with fixed geometry or aspect ratios, such as height to diameter to impeller size and spacing or having fixed flow rates in volume per volume per time.

Once the small scale system is validated against its large scale counterpart, it is possible to begin the procedure of optimisation. Validation might be based upon running a number of fixed conditions in both systems and comparing the outputs (specific growth rate/yield), which would highlight any systematic variation between the systems. As the difference in vessel volume may be 100,000 fold from bench to manufacture. Heterogeneity in the vessels is critical: all locations within the vessel must be subject to the same conditions (*e.g.* temperature, pH and aeration).

1.5.7.2 Use of statistical Design of Experiments in process optimisation

Statistical DoE is a way of planning any experiment with one or more variables to give valid and objective conclusions on the response seen. DoE can be applied to situations where there are controlled inputs with measurable output(s) and the operator wants to efficiently use their

time. Applications include optimisation of feed rates,⁸⁵ fermentation process condition optimisation,⁸⁶⁻⁸⁸ media composition,^{70, 71, 89, 90} inoculum size⁹¹ and age,⁹² harvest point^{92, 93} and yeast biomass.⁹⁴ The DoE here differs from the majority the examples in the literature, as it was implemented using multiple simultaneous small-scale experimentation in a non-traditional micro-bioreactor. The direct transfer of the optimised process conditions from the micro-bioreactor to a traditional larger-scale bioreactor is also uncommon in the literature due to the juvenile status of micro-bioreactor technology.

The statistical basis of DoE is rooted in the concept of process modelling. Producing a process model allows the estimation, prediction, calibration and optimisation of the real-life process. Common process models are seen as 'black boxes' with several input factors (*e.g.* temperature, pH and DO) which can be controlled by the operator and one or more output responses, which can be measured by the operator. Experimental output data, from a number of different input factor combinations, is used to build an empirical model of inputs and outputs.⁹⁵ These models are often linear or quadratic (cubic functions and higher are also possible) and can be used for comparisons, key factor selection, process optimisation, removing process variability, improving process robustness or other outcomes.

The experimental design employed will depend on the number of factors (input variables) and the objectives of the experiment. Available designs include completely randomised designs, randomised block designs, full-factorial designs, fractional factorial designs, Plackett-Burman designs and response surface designs.⁹⁵ Here we utilised a response surface Box-Behnken design requiring each factor to be varied at three levels. This design was selected due to the reduced number of experimental treatment combinations required to generate a predictive model. The drawback of using fewer combinations is an increased number of aliased terms in the model (*e.g.* $[A] = A$ or BEF or CDF and $[AB] = AB$ or EF or ACDE or BCDF). This may be problematic if there is poor understanding of the process and the correct term cannot be identified from the possible aliases. The design space given by a Box-Behnken model with 3 factors at 3 levels is illustrated by Figure 1.4. The design space, which is rotatable, contains regions of poor prediction quality. These "missing corners" may be useful when the experimenter should avoid combined factor extremes. There are also no combinations requiring all 3 factors to be maximal. This property prevents a potential loss of data in those cases.

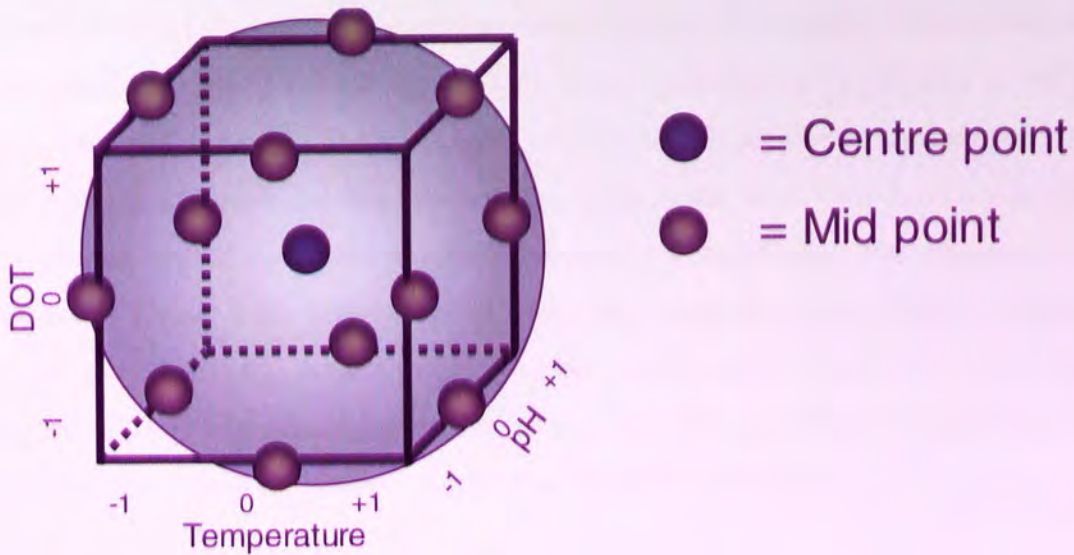


Figure 1.4. Experimental design space suggested by the Box-Behnken DoE model. Each axis of the cube represents an input factor (variable) with the corners and mid-points being the levels for each factor. The design space is shown as a sphere filling the “input cube” which protrudes through each face, with the surface of the sphere being tangential to the mid-point of each edge.

The DoE starts with specifying the objectives and process factors to be investigated, in our case this was what combination of temperature, pH and DO gives the highest yield. A detailed plan of the experiment is then drawn out to provide the most useful information from the experimental period. Sources of variability are identified as much as possible and accounted for to avoid “lurking variables”,⁹⁶ which could lead to a reduced predictive capacity in the model equation.

Response Surface Methodologies (RSMs) allow an experiment to be designed to estimate factor interactions and quadratic effects. This is useful in process optimisation, process troubleshooting and improving the robustness of the process (reducing susceptibility to external uncontrolled influences). There are many types of RSM designs, the choice of which to use depending on the objectives of the experiment, the number of factors to be investigated, available time and equipment and any special considerations such as avoiding a specific set of factor combinations. Some are based on full or fractional factorial layout (Central Composite Designs with circumscribed, inscribed or face centred factor settings) whereas others are completely independent (Box-Behnken).

Equations from the RSM can be displayed in a graphical manner showing the profile of how the response (in this case protein yield) changes with the factors, including interactions. A plot describing the optimised area of the design space will resemble a hill with the optimal

combination of factors (fermentation conditions) at the summit; sub-optimal areas would resemble a portion of the hill-side or valley. The quality of the predictive model is described by analysis of variance (ANOVA) analysis of how well the line of regression for the equation fits with the experimental data. Each of the terms in the model will have a p -value describing the probability of that term not being statistically significant. The equation is refined by removing terms with high p -values until the equation as a whole is acceptable. The engineering statistics e-handbook⁹⁵ provides an excellent source of information on the statistical DoE and modelling methods used and discussed here. Chapter 3 provides more information on how the process model was refined and tested.

1.6 Project aims and objectives

This project aimed to generate a scalable predictive modelling methodology suitable for the optimisation of range of industrially-relevant recombinant proteins and host systems. The yeast *P. pastoris* (expressing GFP) was used for developing the methods. The approach was then tested using a mammalian cell-line (hybridoma) expressing a monoclonal antibody to confirm the transferability of the modelling process. Additional factors (medium composition, antifoam agent use, induction strategy and cell surface hydrophobicity) which may impact target protein yield were identified and investigated for further optimisation of the process *i.e.* maximising product yield.

In order to accomplish this, the following objectives were identified:-

- Identify appropriate target proteins
- Create strains/transformants for each host system to express the target proteins
- Identify suitable factors/variables for process optimisation (applicable to small and large-scale)
- Identify suitable DoE/modelling processes
- Implement proposed modelling process
- Confirm the predictive capacity of the model at the small-scale

- Confirm the predictive capacity of the model at the large(r)-scale
- Confirm the scalable modelling process with an alternative target protein and expression system
- Identify and investigate further factors applicable to process optimisation

“Oh, well, the night is long the beads of time pass slow
Tired eyes on the sunrise, waiting for the eastern glow”

(Battle of evermore, Led Zeppelin)

2.1 Microbial strains used

The organisms used during the construction of expression vectors (*Escherichia coli*) and recombinant protein expression experiments (*Saccharomyces cerevisiae*, *Pichia pastoris* and hybridoma cell line) are shown below.

2.1.1 *Escherichia coli*

E. coli was used to generate quantities of plasmid DNA during expression vector construction and to provide sufficient expression vector for transformation of yeast cells. Different strains were used due to availability and purpose e.g. TOPO10 strain was used during expression vector construction as it had a high transformation efficiency. Details of the *E. coli* strains used here are available at http://openwetware.org/wiki/E._coli_genotypes

2.1.1.1 XL1-Blue super-competent cells

Wild type strain, Genotype: endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB⁺ lacI^q (lacZ)M15] hsdR17(r_K⁻ m_K⁺). Nalidixic acid resistant, tetracycline resistant (carried on the F plasmid) Available from Stratagene (Cat #200236). (<http://www.stratagene.com>)

2.1.1.2 DH5 α ^{97 98 99 100}

Genotype: F' endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ 80dlacZ M15 (lacZYA-argF)U169, hsdR17(r_K⁻ m_K⁺), λ - An Hoffman-Berling 1100 strain derivative (Meselson68)

Promega also lists DH5 α as phoA nalidixic acid resistant

2.1.1.3 BL21-DE3^{101 102}

Genotype: F⁻ ompT gal dem lon hsdS_B(r_B⁻ m_B⁻) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) an *E. coli* B strain with DE3, a λ prophage carrying the T7 RNA polymerase gene and lacI^q

Transformed plasmids containing T7 promoter driven expression are repressed until IPTG induction of T7 RNA polymerase from a lac promoter. Derived from B834 by transducing to Met⁺.

2.1.1.4 TOPO10^{99 103}

Genotype: F⁻ mcrA (mrr-hsdRMS-mcrBC) φ80lacZ M15 lacX74 nupG recA1 araD139 (ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ⁻

Streptomycin resistant

2.1.2 *Pichia pastoris*

Different *P. pastoris* strains were used due to initial availability, GS115 was immediately available at the Alpha Biologics laboratories, X33 was later supplied by the laboratory at Aston.

2.1.2.1 GS115

Genotype *his4*

The *Pichia* host strain GS115 and has a mutation in the histidinol dehydrogenase gene (*his4*) which prevents it from synthesising histidine. The expression plasmid carries the *HIS4* gene that complements *his4* in the host, so transformants are selected for their ability to grow on histidine-deficient medium.¹⁰⁴

2.1.2.2 X33

Wild type strain suitable for zeocin selection, available from Invitrogen (Cat# C180-00)

2.1.3 *Saccharomyces cerevisiae*

Different strains of *S. cerevisiae* were required for this project due to the different aspects of protein expression being examined. W303-1A was used in instances where a wild type strain was required. The Alcofree uracil auxotroph was used for protein expression with the pYX212 expression vector.

2.1.3.1 W303-1A ^{105 106}

Genotype: MATa/MAT α (leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15) [phi⁺]

Notes: W303 also contains a *bud4* mutation that causes haploids to bud with a mixture of axial and bipolar budding patterns. In addition, the original W303 strain contains the *rad5-535* allele. Further information is available at <http://wiki.yeastgenome.org/index.php/CommunityW303.html>

2.1.3.2 Alcofree (TM6* prototrophic)

Saccharomyces cerevisiae KOY.TM6*P was derived from the CEN.PK2-1C¹⁰⁷ strain.

Described in Elbing (2004) ¹⁰⁸

Further information on the commercialisation of Alcofree strain can be obtained from Gothia yeast solutions (<http://www.gothiayeast.com/>).

2.1.3.3 Alcofree (TM6* URA)

Uracil auxotroph of *Saccharomyces cerevisiae* KOY.TM6*P was derived from the CEN.PK2-1C¹⁰⁷ strain. Described in Elbing (2004) ¹⁰⁸

2.2 Mammalian cell lines used

2.2.1 Hybridoma cell line HD1

The HD1 hybridoma cell line was supplied by a customer company of Alpha Biologics Ltd and is subject to a confidentiality agreement.

2.3 Molecular biology

The following techniques were used during the manipulation of cDNA and the construction of expression vectors used with yeast cells.

2.3.1 DNA manipulation

The DNA manipulation techniques used for the construction of expression vectors, including cutting (restriction digests), joining (ligations) mutating and multiplying (PCR) are described below.

2.3.1.1 Restriction Digests

(All steps were carried out on ice unless otherwise stated)

To a pre-chilled 1.7 mL Eppendorf the following were added:

Restriction enzyme	1 μ L
Enzyme buffer	2.5 μ L
DNA	4 μ L
milliQ dH ₂ O	to 20 μ L

This was then incubated at the optimum reaction temperature (enzyme dependant) for 1 hour. To deactivate the enzyme, this was heated to 70 °C for 10 minutes and then returned to ice or stored at -20 °C.

Note: All timings, buffers and incubation conditions used were as per manufacturers' data sheet. Double digests, where possible, were performed as above using the most appropriate enzyme buffer.

2.3.1.2 Ligations

(All steps were performed on ice unless otherwise stated)

Ligation mixture contained:

Vector DNA (<i>e.g.</i> pYX212, pPICZ-C, pYES2)	100 ng
Insert DNA (2-3 times vector DNA)	<i>e.g.</i> 300 ng
dH ₂ O	to 18 μ L

This was heated to 45 °C for 5 minutes (to melt any annealed cohesive ends) and returned to ice. The following were then added:

5X ligase reaction buffer	4 μ L
T4 DNA ligase	0.5 μ L

This was then incubated at room temperature for 2-4 hours or overnight at 4 °C. The T4 DNA ligase was deactivated by heating to 65 °C for 10 minutes. 5-10 µL of the ligation mixture was then used to transform competent cells.

2.3.1.3 Polymerase chain reaction (PCR)

(All steps were carried out on ice unless otherwise stated)

A reaction mix was made of:

Reaction buffer	10 µL
25 mM MgCl	4 µL
2 mM dNTPs	10 µL
50 µM Forward primer	1 µL
50 µM Reverse primer	1 µL
milliQ dH ₂ O	72 µL
Polymerase	1 µL
Template DNA (0.5 ng)	1 µL

(The negative control reaction omitted the template DNA)

PCR annealing temperatures and extension times were optimised for each reaction by using a gradient PCR for each new reaction and adjusting extension time depending on fragment size.

Typical PCR conditions were:

95 °C	30 s (denaturing step)
45-65 °C	20 s (annealing step)
72 °C	40 s (extension step)

This was repeated for 35 cycles then followed by a final extension step of 3 minutes.

2.3.2 Agarose gel electrophoresis

Agarose gel electrophoresis was used to visualise DNA during the manipulation of cDNA and construction of expression vectors.

2.3.2.1 Pouring agarose gels

0.9 % agarose gel (50 mL):

To a suitable container (*e.g.* 125 mL conical flask) the following were added:

Agarose	0.45 g
1 x TAE Buffer	50 mL

The agarose and buffer were heated in a microwave (800 W) on full power until the agarose dissolved, typically this took 20-40 s. Care was taken on removal from the microwave as the contents occasionally boiled over. The dissolved agarose was allowed to cool slightly for 2-5min, 1 μ L of ethidium bromide was then added. The solution was agitated until the ethidium bromide was fully mixed. A gel cast was set up with an appropriately-sized comb, the agarose solution was poured into the gel cast and allowed to set for approximately 30 min at room temperature.

2.3.2.2 Sample preparation

Before DNA samples could be loaded onto agarose gels 5 μ L 2X loading buffer was added to 10 μ L DNA sample and mixed thoroughly.

For the 1kB ladder (reference bands at 10, 8, 6, 5, 4, 3, 2.5, 2, 1.5, 1, 0.75, 0.5, 0.25 kB) 3 μ L ladder buffer was added to 1 μ L 1kB DNA ladder.

2.3.2.3 Running conditions

A power pack was set to 70 V and run for 30 min, checking regularly to prevent bands running off the bottom of the gel.

2.3.2.4 Visualisation

Ethidium bromide-stained DNA was visualised under a UV source, *e.g.* UV light box or a geldoc system and the image saved as a jpeg file.

Appropriate personal protective equipment was used to protect against UV radiation

2.3.3 Plasmid preparation (mini/maxi preps)

Purified plasmid DNA was generated by using Qiagen (Cat #27104) mini-prep (20 µg DNA yield) or maxi-prep (500 µg DNA yield) Qiagen (Cat #12163) kits according to the manufacturer's instructions.

2.3.4 Sequence analysis

DNA sequence analysis was performed by the proteomics laboratory at Birmingham University. The DNA sample (e.g. 8 µL expression cassette of plasmid DNA) was mixed with (1 µL) of 5' and 3' sequencing primers (below) which had recognition sites ~100 bp from the sequence of interest. The results were received as ab1 files and viewed using Chromas v1.45 software. The sequences were then exported into Microsoft Word format. The forward, reverse and consensus sequences were aligned with the reference sequence and deviations identified.

5' sequencing primer 5'-GACTGGTTCCAATTGACAAGC-3'

3' sequencing primer 5'-GCAAATGGCATTCTGACATCC-3'

2.4 Transformation of cells

In order to obtain target protein expression from the cells, the appropriate DNA sequence (expression vector) was introduced into the cells. The DNA was either stably integrated into the genome for *P. pastoris* or held episomally as a plasmid in the case of the *S. cerevisiae* transformants.

The plasmid vectors (e.g. pPICZα-MPI(F)) were linearised by enzymatic digestion at a recognition site (e.g. *SacI*) allowing integration into the *P. pastoris* genome. The recombination event occurs away from the *AOX1* structural gene with 95-75 % frequency, resulting in a strain with the His⁺ Mut⁺ phenotype.¹⁰⁴ The method used for transformation was electroporation; passing a high voltage across the cells disrupts the cell membranes so forming pores through which the linearised DNA can pass into the cell and translocation to the nucleus. Positive transformants were selected on Zeocin.

Transformation of the Alcofree strain used the lithium acetate (LiAc) method to make the cell membrane porous, so allowing plasmid DNA (e.g. pYX212-α-MPI(F)) to enter the cell. The

vector is maintained as a circular structure and not integrated into the genome. Therefore, constant selection pressure was applied to prevent plasmid loss. The Alcofree strain has a disrupted gene (*URA3*) which codes for orotidine 5-phosphate decarboxylase (ODCase). ODCase has a role in pyrimidine ribonucleotide synthesis (*e.g.* uracil). If *URA3* is non-functional, uracil must be added to the culture medium for growth of a *URA3* strain. When used as a selection marker no uracil is added to the medium, but *URA3* is complemented by an intact copy of *URA3* on the plasmid which permits growth in uracil-deficient media.

Cells that were positively selected by the marker had variable expression of the target protein. Therefore, it was necessary to perform expression screening using small scale cultures. The cells that produced the highest levels of the target protein were banked and taken forward for further work.

2.4.1 Preparation of competent cells

For successful electroporations to be possible the cells being transformed must be competent, essentially being able to survive the electroporation process. This involved chilling mid-log phase cells on ice and the removal of salts from the cell suspension (medium) in order to avoid arcing when an electric pulse is passed over the cells.

2.4.1.1 *E. coli*

Cultures were set up in 250 mL shake flasks 50mL LB medium that were inoculated with 200 μ L glycerol stock *E. coli* BL21 (DE3). Cells were grown at 37 °C with elliptical agitation at 200 rpm to an OD₅₉₅ 0.5-0.6. The cells were chilled on ice for 15 min and transferred to a pre-chilled centrifuge bottle. The cells were harvested by centrifugation (20 min 5000g 2-4 °C). The cells were kept cold for the rest of the procedure. The cells were washed twice with the original culture volume of ice-cold sterile water and then centrifuged again as above. The final cell pellet was re-suspended with the residue from wash.

2.4.1.2 *P. pastoris*

The method used for preparing electrocompetent *P. pastoris* cells is essentially as described by Cereghino *et al* (2005)¹⁰⁹
A 5 mL YPD culture of *P. pastoris* cells was grown overnight in a 30 °C shaking incubator with an agitation rate of 250 rpm. The overnight culture was diluted to an OD₆₀₀ of 0.15-0.20

in a volume of 50 mL YPD in a flask large enough to provide good aeration. The culture was grown to an OD₆₀₀ of 0.8–1.0 in a 30 °C shaking incubator. Based on a generation time of 100–120 min, it was estimated that yeast should reach 0.8–1.0 in 4 to 5 h. the culture was harvested by centrifugation at 500× g for 5 min at room temperature and the supernatant discarded. The cell pellet was re-suspended in 9 mL of ice-cold BEDS solution (see below) supplemented with 1 mL 1.0 M dithiothreitol (DTT) and then incubated for 5 min at 100 rpm in a 30 °C shaking incubator. The culture was centrifuged at 500× g for 5 minutes at room temperature and the cells re-suspended in 1 mL (0.02 volumes) of BEDS solution without DTT and split into 40 µL aliquots. The competent cells were then ready for transformation. Alternatively, the cells were frozen slowly in small aliquots at -80 °C by placing the aliquots inside a styrofoam box to reduce the rate of the medium freezing. Competent cells were stored for up to 6 months at this temperature.

BEDS solution contained:

- 10 mM bicine-NaOH, pH 8.3
- 3 % (v/v) ethylene glycol
- 5 % (v/v) (dimethyl sulfoxide) DMSO
- 1 M sorbitol

Autoclave (121 °C 20 min) to sterilise, normally a 100 mL volume was prepared and stored at RT for 6 months.

2.4.2 Heat shock transformation of *E. coli* DH5α cells

Cells were thawed on ice (from glycerol stocks at -80 °C) prior to addition of transformation ligation mix (three transformations; with insert, self ligation and a positive control as in 2.3.1.2) and then incubated on ice for 30 min. The transformation mixes were transferred to 42 °C for 40 s and then incubated in ice for 1 min. The transformations were plated on LB agar with appropriate antibiotic selection and incubate at 37 °C until colonies appeared.

2.4.3 Electroporation

2.4.3.1 Electroporation of *E. coli* cells

This method was adapted from the Eppendorf Multiporator^R protocol # 4308 915.512 – 04/2002. An electroporation cuvette (2 mm gap width) was pre-chilled and had 40 µL of electrocompetent *E. coli* cells and 1 µL of DNA added which were gently mixed with a

pipette. Any moisture was wiped from the cuvette and then inserted into the electroporation device. The cells were electroporated using the prokaryote “o” mode with 1700 V and 5 ms time constant. 1 mL SOC medium was immediately added and then incubated at 37 °C with elliptical agitation at 200 rpm for 60 minutes. The cells were then plated on selective (e.g. zeocin / ampicillin) LB agar plates and incubated at 37 °C until colonies were easily distinguishable.

2.4.3.2 Electroporation of *P. pastoris* cells

The electroporation method that was used during the work was essentially as that of Cereghino *et al* (2005).¹⁰⁹

Approximately 4 µL (50–100 ng) of linearized plasmid DNA was mixed with 40 µL of competent cells in an electroporation cuvette and incubated for 2 min on ice. The samples were then electroporated using the following parameters in an Eppendorf multiporator set to “eukaryotic” mode. Pulse voltage set to 1200 V, with a 20 µs time constant. Immediately after electroporation the samples were re-suspended in 1 mL cold 1.0 M sorbitol and then plated on selective media (YNB, 2 % dextrose + 1.0 M sorbitol) for auxotrophic strains. Alternatively, if using zeocin-based plasmids, the samples were re-suspended in 0.5 mL 1.0 M sorbitol and 0.5 mL YPD then incubated in a 30 °C shaker for 1 h. Post this incubation period the sample was plated onto media containing increasing concentrations of zeocin (100, 250, 500, or 1000 µg/mL) for the selection of multicopy integrants. Note that increased numbers of transformants could be achieved for both types of selectable markers by incubating the re-suspended cells in a 30 °C shaker for longer periods of time (1–3 h). However, this is partly due to replication of transformants. This adapted method worked well and was used for all electroporations prepared.

2.4.4 Lithium acetate transformation of *S. cerevisiae* cells

The following method for the transformation of cells with lithium acetate was followed. 10 mL of YPD was inoculated with a loopfull of freshly-plated cells. The culture was incubated for approximately 24 h at 30 °C with 230 rpm agitation, TM6* taking longer. Post incubation 0.5 ml of culture was harvested by centrifugation at 7000 rpm for 2 min and the supernatant discarded. A 2 – 5 µg quantity of plasmid DNA was added together with 10 µl carrier DNA and added to the cells. 0.5 ml plate solution (below) was added and then vortexed briefly. The mixture was then incubated at room temperature overnight (approx 14 h). The bottom 100 µl was withdrawn from the tube and plated on SD^{-URA} agar.

The plating solution comprised of the following ingredients. Normally a 100mL volume was prepared and stored at RT for 6 months.

Plate Solution 40 % (w/v) PEG-3350

1 mM EDTA

10 mM Tris-HCl (pH 7.5)

0.1 M Lithium acetate

Autoclave (121 °C 20 min) to sterilise

2.4.4.1 Carrier DNA

The carrier DNA used for the transformation of *S. cerevisiae* was 7 mg/ml sonicated salmon sperm DNA. This reagent was prepared by sonicating (3x 20 sec pulses), 1 mL of 7 mg/mL salmon sperm DNA this was stored at -20 °C.

2.4.5 Modified Heat shock lithium acetate transformation of *S. cerevisiae* cells

10 mL of YPD was inoculated with a loopfull of freshly-plated cells. The culture was incubated for approximately 24 h at 30 °C with 230 rpm agitation. 0.5 mL culture was harvested by centrifugation at 7000 rpm for 2 min and the supernatant discarded. 2 – 5 µg of plasmid DNA and 10 µL carrier DNA was added to the cells. 0.5 mL plate solution was added and then vortexed briefly before being incubated on ice for 30 min. The transformation mixes were transferred to 42 °C for 40 s and then incubated in ice for 1 min. The mixture was then incubated at room temperature overnight (approx 14 h). The bottom 100 µL was withdrawn from the tube and plated on SD^{-URA} agar and incubated at 30 °C until colonies appeared.

2.5 Clone selection

Clone selection was used to isolate transformed cells (*E. coli*, *S. cerevisiae* or *P. pastoris*) which contained the intended plasmid vector (shuttle or expression) or had an integrated copy of the vector.

2.5.1 Confirmation of DNA uptake / integration (colonies on selective agar)

Following transformation 100 μ L cells were gently pipetted onto agar plates containing the appropriate selection marker e.g. 100 μ g/mL Zeocin and incubated at the optimum temperature (30 °C for *P. pastoris* and *S. cerevisiae* 37 °C for *E. coli*) until colonies were visible, typically 1-3 days. 20-30 colonies were then picked and re-streaked on a fresh selective plate and returned to the incubator until the colonies re-grew.

2.5.2 Expression confirmation (screening small cultures)

10-20 colonies were picked from re-streaked selection plates (as described in section 2.4.1.) and used to inoculate small (125 or 250 mL) shake flasks with the appropriate media (e.g. BMGY/BMMY, 2xCSM, SD-URA). Cultures were incubated at 30 °C with 250 rpm agitation. Supernatant and cell pellet samples were taken during the time course for expression confirmation by SDS-PAGE, western blotting and mass spectrometry.

2.6 Storage of strains (wild-type and transformed) by glycerol stock

Actively growing cells, early-mid log phase, were harvested by centrifugation 6,500 rpm in a microfuge and re-suspended in initial volume of fresh media containing 10 % w/v glycerol and stored at -80 °C.

2.7 Analytical methods

The methods used for the detection, identification and quantification of proteins, cell growth, physiological properties and carbon source utilisation of cultured cells are described below.

2.7.1 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Culture supernatant (12 μ L) was added to 4 μ L NuPAGE LDS sample buffer (Invitrogen Ltd. Cat # NP0007) heated to 95 °C and loaded into a NuPAGE 4-12% Bis-Tris polyacrylamide gel (Invitrogen Ltd. Cat # NP0322BOX). 5 μ L of a molecular weight marker (MWM) SeeBlue Plus2 (Invitrogen Ltd. Cat # LC5925) were also loaded for sample band weight determination. A 1x solution of MES SDS running buffer (20x) (Invitrogen Ltd. Cat # NP0002) was made using dH₂O. Proteins were separated by SDS-PAGE at 180 V for 0.75 h.

2.7.2 SDS-PAGE protein band visualisation by InstantBlue stain

Following separation the gel was washed briefly in 30 mL dH₂O and then bands were visualised by adding 5 mL of InstantBlue stain (Novexin Cat # ISB01L) and incubated with gentle agitation for a minimum of 20 min. The gel was then de-stained with multiple washes of dH₂O and an image scanned and saved as a jpeg file using a PC with a flat-bed scanner.

2.7.3 SDS-PAGE protein band visualisation by silver stain

Following separation, the gel was washed briefly in 30 mL dH₂O and the bands were then visualised using a SilverXpress silver staining kit (Invitrogen Ltd Cat # LC6100) allowing detection of sub ng quantities of protein.

2.7.4 Western blotting

An SDS-PAGE gel was run as per 2.7.1 but not stained. The protein bands were transferred to a nitrocellulose membrane (Invitrogen Ltd. Cat # LC2000) following Invitrogen instructions. The membrane was then blocked using 20 mL PBS-T (phosphate buffered saline with 0.005 % v/v Tween 20) with 5 % w/v non-fat skimmed milk powder and incubated at room temperature with gentle agitation for 45 minutes. The membrane was washed with 20 mL PBS-T and then probed using a 1:5,000 dilution of rabbit polyclonal anti-His₆ antibody conjugated to HRP (Abcam Cat# ab1187-100) in PBS-T with 1 % w/v non-fat skimmed milk powder and incubated for 1 h at room temperature with gentle agitation or overnight at 4 °C with no agitation. The membrane was then washed three times with 20 mL PBS-T before visualisation using EZ-ECL reagent (Biological Industries. Beit Haemek, Israel. Cat# 20-500-120) and a gel-doc imaging system with 15 minute cumulative exposure time. The resulting image was saved as a jpeg file.

2.7.5 Mass spectrometry

An SDS-PAGE gel was run as per 2.7.1 and visualised as per 2.7.2. The protein band of interest was excised using a clean sharp scalpel and de-stained following the Novexin InstantBlue handbook protocol. The de-stained protein bands were then sent for ESI-QUAD-TOF at the protein science group at Birmingham University. Analysis was carried out using the online program MASCOT MS/MS ion search, accessible at http://www.matrixscience.com/search_form_select.html

2.7.6 GFP_{uv} fluorescence analysis

Culture supernatants (100 μ L) and BMMY blanks (100 μ L) were loaded onto black 96-well plates (Greiner Maxisorp Cat # M9685). All samples and blanks were buffered using 50 μ L 1 M potassium phosphate pH 8.0. The plates were assayed using a fluorescent plate reader (SpectraMax Gemini or PerkinElmer Victor3 with GFP specific filter set) with an excitation wavelength of 397 nm and emission wavelength of 506 nm. Each sample was loaded in triplicate and each well was read three times. Results were exported as excel spread sheets and used for yield determination.

2.7.7 Antibody quantification by Protein-A purification

The HD1 antibody was quantified using product peak area integration following elution from a protein-A chromatography column. A standard purification protocol was devised based on the use of a 1 mL HiTrap MabSelect SuRe (GE Healthcare Cat # 11-0034-93) and an AKTA purifier purification system (GE Healthcare). Clarified (0.2 μ m filtered) HD1 culture supernatant (5 mL) was loaded into a 5 mL sample loop attached to the AKTA sample valve. The MabSelect column had a column volume (CV) of 1 mL. The column was equilibrated with 5 CV 1x phosphate buffered saline (PBS) pH 7.4, the sample (5 CV) was then loaded onto the column and washed with 2 CV 1x PBS. An isocratic elution was then performed over 5 CV 0.1 M glycine (pH 2.7) with 0.5 CV fractions. A Frac100 (GE Healthcare) fraction collector was used to automatically collect the fractions. The collection tubes contained 50 μ L 1 M Tris-HCl (pH 9.0) to neutralise the pH of the eluate. The column was then cleaned with 2 CV 0.1 M glycine to remove any unbound material. The fractions containing the product peak were identified from the chromatogram, pooled and stored at 4 °C for future analysis. A high yielding sample pool was quantified by $A_{280\text{nm}}$ and used to calculate the HD1 product concentration in the supernatant samples from the associated peak areas (from individual chromatograms).

2.7.8 Optical density at 595 and 600 nm wavelength ($OD_{595/600}$)

Two wavelengths were used for cell density estimation because of the different photospectrometer filter sets in place at Aston University and at Alpha Biologics. Culture samples were diluted with dH₂O to be in the linear range; an approximate OD_{595} of 0.15 - 0.3 units. The maximum single dilution was 1 in 100. If further dilution was require serial dilution steps were used. The spectrophotometer was set to a wavelength of 595 nm or 600

nm and allowed to warm-up for 15 minutes before blanking a 1 cm path length disposable cuvette with the major diluent. The diluted sample was then read and the absorbance multiplied by the dilution factor.

2.7.9 Cell surface hydrophobicity (CSH) assays

2.7.9.1 Microbial Attachment to Hydrocarbons (MATH) assay

Method adapted from Smith *et al* (1998).¹¹⁰

Cell growth was monitored by OD₅₉₅. Replicate samples of cultures were taken, with sufficient volume for 4 mL of an OD₅₉₅-adjusted sample of 0.5 units. Cells were harvested by centrifugation 6000 rpm 10 min, the supernatant was removed and the cell pellet re-suspended in the original sample volume of PUM buffer (below). The cells were pelleted by centrifugation at 6000 rpm for 10 min, the supernatant was removed and cell pellet re-suspended in 4 mL PUM buffer (OD₅₉₅ 0.5). 1 mL was removed for the OD₅₉₅ control reading and 1.5 mL of re-suspended cells dispensed into acid-washed glass tubes (12x75 mm) and 300 µL of n-hexadecane added. This was incubated for 10 min at room temperature and then vortexed for 3x 30 seconds and let to stand for 10 min. The n-hexadecane layer was removed by pipette and incubated at 4 °C for 15 min. The remaining solidified n-hexadecane was removed with a plastic inoculation loop. The samples were brought to room temperature and read at OD₅₉₅ of treated sample.

The hydrophobic index was calculated by:

$$\text{Relative cell surface hydrophobicity (\%)} = \frac{\text{Initial OD}_{595} - \text{final OD}_{595}}{\text{Initial OD}_{595}} \times 100$$

PUM buffer contained:

127 mM K₂HPO₄

53 mM KH₂PO₄

30 mM urea

0.8 mM magnesium sulphate (heptahydrate)

This was made up with dH₂O and filter sterilised (0.22 µm).

2.7.9.2 Hydrophobic interaction chromatography (HIC) assay

Adapted from the PhD Thesis of Alexander John Foster, Aston University, May 2004, Cell surface analysis of the basidiomycete yeast *Cryptococcus neoformans*.

Columns were prepared (1 column volume = 1 mL) using gravity flow. A glass Pasteur pipette was plugged with silanised glass wool and then packed with 1 mL of either sepharose CL-4B (Sigma Cat # CL4B200) or phenyl-sepharose CL-4B (Sigma Cat # P-7892) and then washed (3x) with 4 mL PUM buffer

Cell growth was monitored by OD₅₉₅. Replicate samples were taken from X33 cultures, of sufficient volume to give 2 mL of an OD₅₉₅-adjusted sample of 0.5 units. Cells were harvested by centrifugation at 6,000 rpm for 10 min, the supernatant was removed and cell pellet re-suspended in the original sample volume of PUM buffer. Cells were pelleted by centrifugation at 6,000 rpm for 10 min, the supernatant removed and cell pellet re-suspended in 2 mL PUM buffer (OD₅₉₅ 0.5). 1 mL of sample was loaded onto each column (control and phenyl). The sample was eluted from the column with 4 mL of PUM buffer and fractions collected. The fractions were pooled and centrifuged at 6,000 rpm for 10 min, the cell pellet was re-suspended cells to 1 mL and OD₅₉₅ measured for control and phenyl eluates. The relative CSH was calculated:

$$\text{Relative cell hydrophobicity (\%)} = \frac{\text{control OD}_{595} - \text{experimental OD}_{595}}{\text{Control OD}_{595}} \times 100$$

2.7.10 Ethanol concentration assay (UV-method)

Culture samples were assessed for ethanol concentration using an ethanol assay kit (Roche Cat # 10 176 290 035) which used an enzymatic reaction to provide a colour change detected using a spectrophotometer. The assay was carried out as per manufacturer's instructions.

2.8 Shake flask (Erlenmeyer flask) cultures

Shake flasks of various sizes (125, 250 500, 1000, 2000 mL), closure types (0.2 µm vented screw cap and foil) and with or without baffles were used. Unless otherwise stated, culture volume did not exceed 20 % total volume of each flask. Flasks were sterilised by autoclave

(121°C, 20 min). Cultures were incubated in bench or floor-mounted shaking incubators with agitation 200-250 rpm.

2.9 Media

The media used for culture (solid and or liquid) of *E. coli* (LB), *S. cerevisiae* (YPD, 2x CBS, CSM, BSM and SD^{URA}), *P. pastoris* (YPD, BMGY, BMMY and SOC) and liquid culture of hybridoma cells (ExCell 302) are described below.

2.9.1 LB

(per litre)

Bacto-tryptone	10 g
Yeast extract	5 g
NaCl	10 g
dH ₂ O	800 mL
pH was adjusted with 1M NaOH	
dH ₂ O	1000 mL

The LB was autoclaved using a 121 °C, 20 min cycle to sterilise.

2.9.2 YPD

(per L)

Yeast extract	10 g
Peptone	20 g
Agar (plates only)	20 g
Glucose (40 % stock solution)	50 mL

The YPD was autoclaved using a 121 °C, 20 min cycle to sterilise.

Addition of glucose after autoclaving was routine.

2.9.3 BMGY (Buffered Glycerol-complex Medium)

(1% yeast extract, 2% peptone, 100mM potassium phosphate pH 6.0, 1.34 % YNB, 4×10^{-5} % biotin, 1 % glycerol). 10 g yeast extract and 20 g peptone were dissolved in 700 mL dH₂O and

autoclaved for 20 min at 121 °C. This was then cooled to room temperature and the following added:

1 M potassium phosphate buffer pH 6.0	100 mL
10x YNB	100 mL
500x biotin	2 mL
10x glycerol	100 mL

The medium was stored at 4 °C for up to 2 months.

Stock solutions:

10x YNB (13.4 % Yeast nitrogen base with ammonium sulphate with out amino acids)

YNB	134 g
dH ₂ O	1000 mL
Filter sterilise (0.2 µm)	
Store at 4 °C	

500x biotin (0.02 % biotin)

biotin	20 mg
dH ₂ O	100 mL
Filter sterilise (0.2 µm)	
Store at 4 °C	

10x GY (10% glycerol)

Glycerol	100 mL
dH ₂ O	900 mL
Filter sterilise (0.2 µm)	

1 M potassium phosphate buffer pH 6.0

1 M K ₂ HPO ₄	132 mL
1 M KH ₂ PO ₄	868 mL

pH was adjusted with phosphoric acid or KOH

Autoclave (121 °C 20 min) cycle

2.9.4 BMMY (Buffered Methanol-complex Medium)

(1 % yeast extract, 2 % peptone, 100 mM potassium phosphate pH 6.0, 1.34 % YNB, 4×10^{-5} % biotin, 0.5 % methanol). 10 g yeast extract and 20 g peptone were dissolved in 700 mL dH₂O and autoclaved for 20 minutes at 121 °C. This was cooled to room temperature and the following was added:

1 M potassium phosphate buffer pH 6.0	100 mL
10x YNB	100 mL
500x biotin	2 mL
10x methanol	100 mL

Media was stored at 4 °C for up to 2 months.

Stock solutions:

10x YNB (13.4 % Yeast Nitrogen Base with Ammonium Sulphate with out amino acids)

YNB	134 g
dH ₂ O	1000 mL
Filter sterilise (0.2 µm)	
Store at 4 °C	

500x biotin (0.02 % biotin)

biotin	20 mg
dH ₂ O	100 mL
Filter sterilise (0.2 µm)	
Store at 4 °C	

10x M (5 % methanol)

methanol	50 mL
dH ₂ O	950 mL
Filter sterilise (0.2 µm)	

1 M potassium phosphate buffer pH 6.0

1 M K ₂ HPO ₄	132 mL
1 M KH ₂ PO ₄	868 mL

Adjust pH with phosphoric acid or KOH

Autoclave (121 °C 20 min) cycle

2.9.5 2xCBS

(Per L)

Ammonium sulphate	10 g
Potassium phosphate	6 g
Magnesium sulphate	1 g
Glucose	10 g

This was autoclaved (121 °C, 20 min) and once cooled to room temperature the following was added:

Trace element solution	1 mL
Vitamin solution	1 mL

Trace element solution for CBS (per L)

EDTA (di-sodium-)	15.0 g
Zinc sulphate heptahydrate	4.5 g
Manganese chloride tetrahydrate	1.0 g
Cobalt (II)-chloride hexahydrate	0.3 g
Copper (II)-sulphate penthydrate	0.3 g
Di-sodium molybdenu dihydrate	0.4 g
Calcium chloride dihydrate	4.5 g
Iron sulphate-heptahydrate	3.0 g
Boric acid	1.0 g
Potassium iodide	0.1 g

Filter sterilise (0.2 µm).

Vitamin solution for CBS (per L)

D-biotin	0.05 g
Ca D(+) panthothenate	1.0 g
Nicotinic acid	1.0 g
Myo-inisitol	25.0 g
Thiamine hydrochloride	1.0 g
Pyridoxol hydrochloride	1.0 g
p-amino benzoic acid	1.0 g

Filter sterilise (0.2 µm).

2.9.6 CSM (Complete synthetic medium)

(Per L)

Yeast nitrogen base with out amino acids	1.7 g
Ammonium sulphate	5.0 g
Glucose	20.0 g
1M MES pH 6.0	2.5 mL
10x complete Drop Out (DO) solution	50 mL

10x DO Solution (complete) (per L)

L-Adenine hemisulphate salt	200 mg
L-Arginine HCl	200 mg
L-Histidine HCl monohydrate	200 mg
L-Isoleucine	300 mg
L-Leucine	1000 mg
L-Lysine HCl	300 mg
L-Methionine	200 mg
L-Phenylalanine	500 mg
L-Threonine	2000 mg
L-Tryptophan	200 mg
L-Tyrosine	300 mg
L-Valine	1500 mg
Uracil	1500 mg

2.9.7 BSM (Basal salts medium)

(Per L)

85 % H_3PO_4	26.7 mL
$\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$	0.93 g
K_2SO_4	18.2 g
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	14.9 g
KOH	4.13 g
Glycerol	40 g

pH (25 °C) was adjusted to pH 5.0 with 30 % NH_4OH

The BSM was autoclaved using a 121 °C, 20 min cycle to sterilise.

2.9.8 SD^{-URA}

(per L)

Yeast nitrogen base with out amino acids	6.7 g
Agar (plates only)	20 g

The medium was autoclaved using a sterilisation cycle (121 °C, 20 minutes), brought to room temperature and then the following was added:

DO solution (10x Stock)	100 ml
Glucose	20 g

10x DO Solution (– uracil)

(per L)

L-Adenine hemisulphate salt	200 mg
L-Arginine HCl	200 mg
L-Histidine HCl monohydrate	200 mg
L-Isoleucine	300 mg
L-Leucine	1000 mg
L-Lysine HCl	300 mg
L-Methionine	200 mg
L-Phenylalanine	500 mg
L-Threonine	2000 mg
L-Tryptophan	200 mg
L-Tyrosine	300 mg
L-Valine	1500 mg

The 10x DO Solution (– uracil) was autoclaved using a 121 °C, 20 min cycle to sterilise.

The medium was stored at 4 °C for up to 1 year.

2.9.9 SOC

(Per L)

Bacto-tryptone	20 g
yeast extract	5 g

NaCl	0.5 g
1M KCl	2.5 mL
dH ₂ O	to 1 L

pH was adjusted to 7.0 with NaOH.

The medium was autoclaved using a sterilisation cycle (121 °C 20 minutes), brought to room temperature and 10 mL 20 % glucose was added.

2.9.10 ExCell 302

Ex-Cell 302 is a serum-free medium without L-glutamine for CHO Cells and is available from Sigma Aldrich (Cat # 14324C). The medium (per litre) was supplemented with 10 mL foetal bovine serum, 10 mL L-glutamine and 10 mL 100x non-essential amino acid solution (Sigma Aldrich, Cat # M7145)

2.10 Antibiotics

Antibiotics were used to create selection pressure on transformed cells to select for positive transformants *i.e.* those that contained genes encoding antibiotic resistance from expression vectors. Ampicillin sodium salt (Melford Cat # 69-52-3) and Zeocin (Invitrogen Ltd. Cat # R250-01) were the two antibiotics used to select clones containing the pYX212 and pPICZαA plasmids. These were used at 100 µg/mL unless otherwise stated. 100 mg/mL working stocks of each were stored at -20 °C and thawed in the dark prior to use.

2.11 Applikon Micro 24 bioreactor

The Applikon Micro 24 bioreactor is based on a modified 24 well plate incubation cassette and a control unit which both monitors and controls the culture conditions of each well independently (see Figures 2.1 and 2.2). The incubation cassettes are supplied as sterile single-use units, with a 10 mL total volume and a 3-7 mL working volume per well. The closures for the wells used here were type D, which are for use with high gas flow rates; cultures with DO set points above 40 %. The incubation cassettes have two optical sensor pads and a gas sparging port in the bottom of each well. Culture pH and DO were measured using the optical sensors and controlled by sparging CO₂, ammonium hydroxide vapour and O₂ through the culture. CO₂ and O₂ were supplied from (minimum 1.5 Bar) gas cylinders and the ammonium hydroxide from a pressure vessel containing 15 % ammonium hydroxide

solution. Culture temperature was controlled and monitored using a thermocycler type heating block. Adjacent wells could have set points varying by up to 4 °C. The lowest temperature set point was limited to 2 °C above ambient room temperature. The Micro 24 requires a sterile 6 Bar clean dry air (CDA) supply to run its vacuum generator, achieved using a suitably sized compressor supplied by Applikon Biotechnology. The Micro 24 was capable of agitation 500-800 rpm with a circular orbit of 5 mm. When the Micro 24 was first used it was found that the solenoid valve controlling O₂ supply to well A6 was stuck in the open position. This removed DO control from well A6; experiments were therefore designed and carried out to avoid any problems related to this lack of control. The solenoid valve was fixed when the unit was used for performing DoE2 and DoE3 (Chapters 3 and 4) Both the Micro 24 and laptop computer were run using an uninterruptible power supply (UPS) to prevent equipment failure and data loss in the event of a power outage.

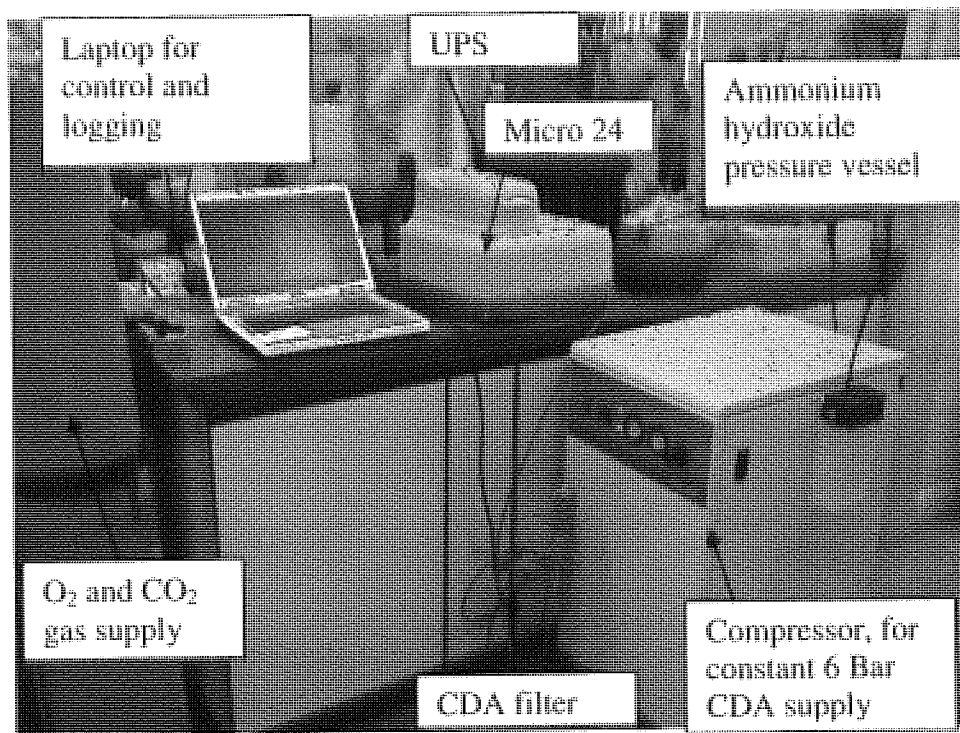


Figure 2.1. Micro 24 micro-bioreactor (Applikon Biotechnology Ltd.) in use at Aston University. Labels highlight the required ancillary equipment for running the Micro 24. CDA is clean dry air and UPS is uninterruptable power supply.

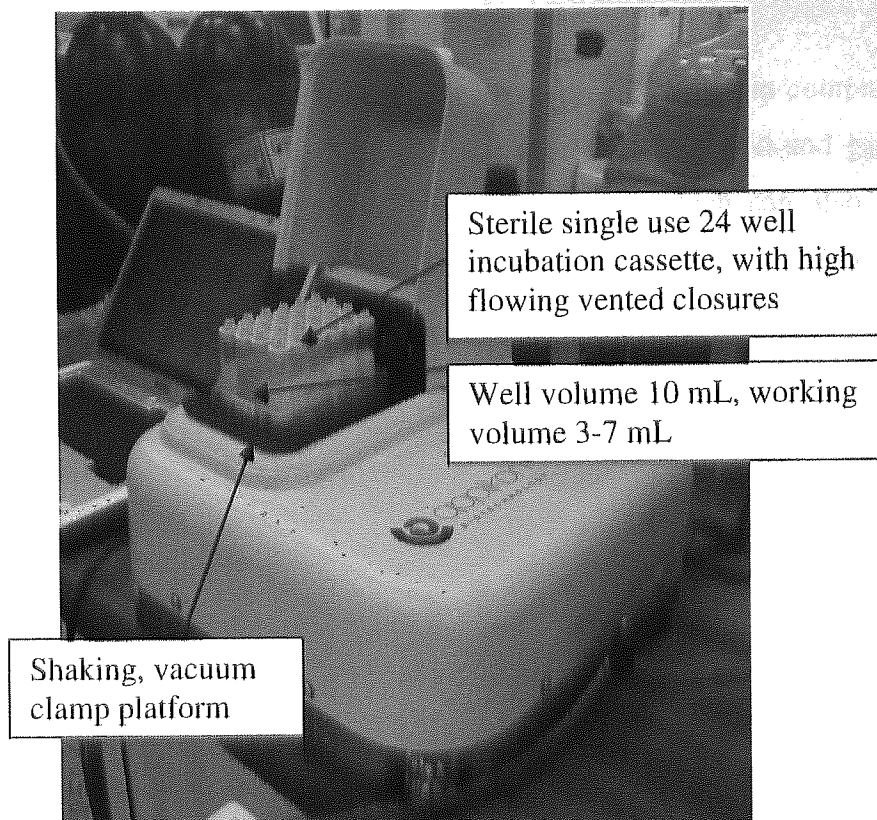


Figure 2.2. Micro 24 micro-bioreactor (Applikon Biotechnology Ltd.) in use at Aston University. The incubation cassette, well closures and shaking platform are highlighted.

2.11.1 Installation

Installation of the Micro 24 was performed by Applikon Biotechnology Ltd personnel. Upon installation a test run using a simple plate layout with multiple replicate conditions across the plate was performed, referred to as the installation run.

As the Micro 24 uses gas sparging to control the pH and DO of each well, culture foaming was expected and was controlled using an antifoam agent. Antifoam suitability was confirmed by a foam test. 700 ppm (4.2 μL) of four antifoam agents (Fluka P2000 (Sigma-Aldrich Cat #81380, Sigma Antifoam-C (Sigma-Aldrich Cat #A8011), Struktol SB2121 (Schill & Seilacher, Germany) and Struktol J673A (Schill & Seilacher, Germany) were added to 5 mL YPD medium in wells of the incubation cassette. Gas was then sparged through the medium to identify antifoam agents that prevented foaming. The antifoam P2000 at 700 ppm was selected for use in the Micro 24 as it prevented foaming and was commonly used by others in the laboratory at Aston University.

2.11.2 Programming/ run set-up

The run conditions for each of the wells were programmed using a laptop computer and the Micro 24 control software. The gas flow method (standard) was selected and gas cylinders and CDA attached to the appropriate ports. The PID settings which can also be altered, remained as supplied unless otherwise stated: pH (P = 8, I = 4), DO (P = 4, I = 15). The logging rate was set to 20 minutes as this was suitable for the intended culture duration.

2.12 7 L vessel Fermentor/ ADI1030 control unit (Scale-up experiments)

2.12.1 Fermentor set-up/ operation

A jacketed 7 L (total volume) glass vessel (as shown in Figure 2.3) was used for all large-scale bioreactor runs. All seals, o-rings and tubing were inspected and replaced as required prior to use. The head plate had ports for the impellor drive shaft, septum, 2 x sample tubes, 1x top tri-port for addition tubes, 3x single top "drip" addition port, 1x single bottom addition tube port, sparger bar (gas in), condenser (gas out), pH probe, DO probe, temperature probe and a universal sampling device.

The choice of impellor design was dependant on the organism being cultured. For mammalian cell culture (hybridoma HD1 cell-line) a single marine blade impellor was positioned one blade diameter from the bottom of the vessel. For *P. pastoris*, three 6-blade Rushton turbine impellors were used, positioned so that the bottom impellor was one diameter from the bottom of the vessel, the top impellor was positioned one diameter from the top of the (expected) culture volume. The third impellor was located equidistant to the bottom and top impellors.

The heat exchanger used to control the jacket temperature, was not controlled by the ADI1030 unit but has its own internal feed-back loop. Therefore, control of culture temperature was dependant on an operator defining a heat exchanger set-point that was appropriate for the culture set point. This was typically 2 °C higher than the culture set-point.

Aeration was primarily provided by a sparger bar supplying 0.2 µm filtered laboratory air from an air pump at 0.2-2 vvm. Pure O₂ was supplemented on-demand as required to maintain set-point. Gas flow regulators were used to step down the flow rates, depending on organism being cultured and process conditions.

The vessel was decontaminated, thoroughly cleaned and double autoclaved (121 °C, 20 min) prior to each run. When switching from one organism to another all seals were checked, silicone tubing was replaced and the vessel was autoclaved three times with at least one day between runs to ensure no cross contamination occurred.

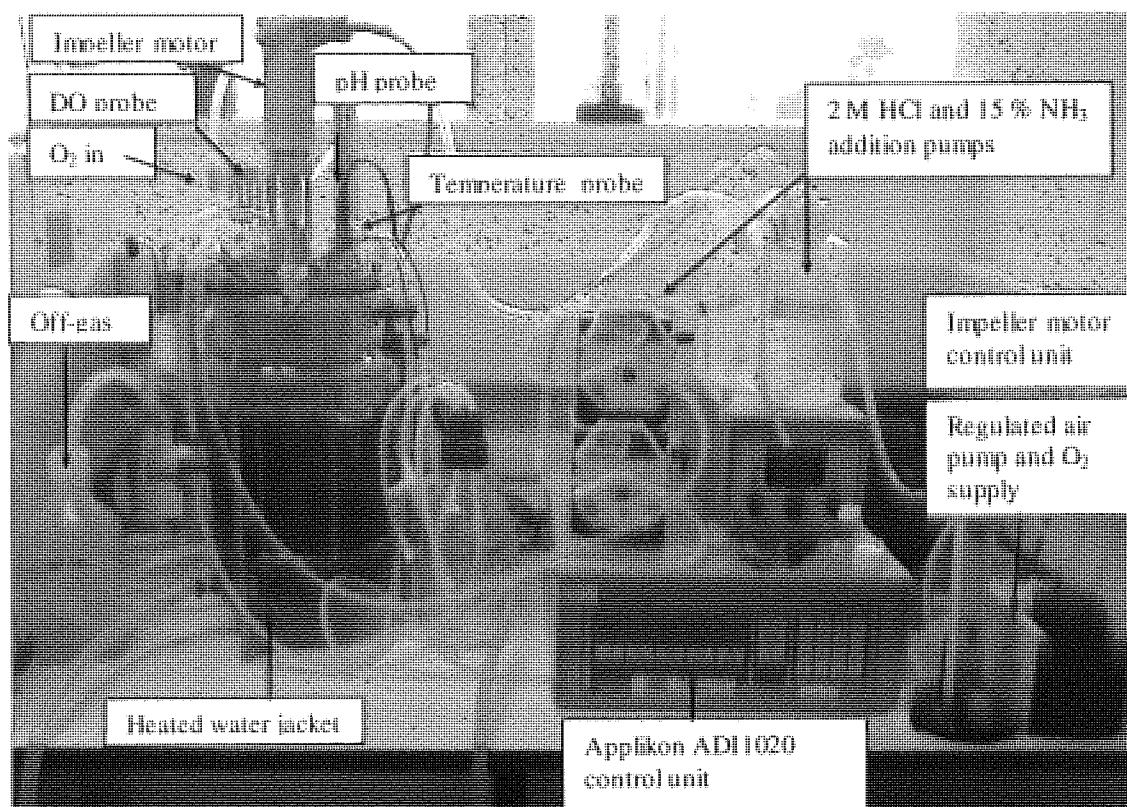


Figure 2.3. 7 L bioreactor with Applikon ADI1030 control unit and ancillary equipment, in use at the Alpha Biologics laboratory.

2.12.2 Applikon ADI1030 set-up and operation

An Applikon ADI1030 proportional–integral–derivative (PID) control unit was used for control of process conditions (pH, DO) and used to control the peristaltic feed pumps for inducer or glucose addition. Devices such as solenoid valves, peristaltic pumps and probes (pH, DO and temperature) were connected to the ADI1030 as required for each run.

2.12.3 Applikon BioXpert Lite software

BioXpert Lite software was used for data-logging of on-line measurements (pH, DO and temperature) and plotting of off-line measurements (e.g. cell density, cell viability, glucose concentration, heat exchanger set point, agitation rate, gas flow rates). The BioXpert software was run on a PC connected to the ADI1030 control unit with an RS232 cable. At the end of a run the data were exported in Microsoft Excel format and trace images as bitmap files. The

software removed the data-holes that occurred during overnight periods where there was not operator present to manually log the on-line measurements from the ADI1030 control unit, allowing a fuller picture of the culture profiles to be generated.

2.13 MiniTab statistical software

MiniTab software (version 15.1.30.0) was used extensively during the DoE planning and model building processes. The DoE response surface design creation wizard was used to generate the experimental matrix for the model building experiments in the Micro 24. Once the experimentation was completed the responses (product yields) were analysed using the “analyse response surface design” function of the software. This produced an initial regression model and corresponding statistical analysis (ANOVA). The “contour/surface plot” function was used to generate all graphical representations of the models. Determination of optimal factor settings was achieved using the “response optimiser” function of the software. Further information on the use of this software can be found on the MiniTab website, at <http://www.minitab.com/products/minitab/>.

2.14 Model building and refinement

The models generated by the Minitab software contained all of the factor terms including interactions (e.g. pH*DO). Each of these terms has a p -value associated with it, indicating the probability of that term not being statistically significant to the model. It was possible to refine the model, as a whole, by removing those terms with high p -values (e.g. > 0.300). After each round of refinement a new model was generated and the refinement process continued until the higher p -value terms were removed or the R^2_{adj} value of the model was maximised. This process is more fully discussed in Chapter 3.

“Hand against hand, foot against foot,
There is no unstoppable technique”

(Wing Chung proverb)

CHAPTER 3

3 OPTIMISATION OF RECOMBINANT PROTEIN YIELD USING PARALLEL MINI-BIOREACTORS AND A DESIGN-OF-EXPERIMENTS MODELLING APPROACH

Proteins lie at the heart of biology: understanding their structures and mechanisms of action gives direct insight into cellular function as well as providing targets for the investigation of disease. Since the vast majority of proteins are not readily available from natural sources, recombinant protein overproduction is a universally-recognised solution to obtaining the milligram quantities required for applications as diverse as structural genomics and biopharmaceutical manufacture.¹ Suitable host cell factories for producing recombinant proteins include microbes such as *Escherichia coli*, *Saccharomyces cerevisiae* and *Pichia pastoris*, as well as mammalian cell-lines (e.g. CHO, NS0 and BHK) and insect cells (e.g. Sf9, Sf21, S2 and Tn5B1-4 cell-lines) transfected with viral vectors. Each system has certain benefits and potential drawbacks in their use (see Chapter 1), but yeast, with its well-established genetic and molecular biological resources, combines the ease and speed of use of bacterial systems with its ability as a eukaryote to secrete post-translationally-modified proteins. As a consequence it is an increasingly popular choice in both academic and commercial laboratories.² Nonetheless the routine achievement of high production yields in yeast or indeed any of the alternative systems mentioned above continues to be a substantial bottleneck to further progress.

Following the initial identification and selection of positively expressing clones, the methods to optimise and scale-up promising small-scale results are still largely based on an empirical trial-and-error basis. In order to devise a strategy for the optimisation and scale-up process, several groups have recently adopted the 'design of experiments' (DoE) type approach that is frequently used in the bioindustry. Studies have focused on how critical elements of the experimental set-up in production systems as bioreactors or shake flasks, such as the medium composition,^{70,71,87} nutrient feed rates,⁸⁹ induction of expression⁶³ and cell biomass,⁹⁴ can affect the total product yield. Recently Islam and colleagues¹¹¹ used DoE in a micro-well plate format to examine the optimisation of firefly luciferase production in *E. coli* by applying an experimental design that included liquid fill level and culture agitation rates as input parameters. These two parameters both strongly influence oxygen transfer rates in the wells. The authors of this study found that while oxygen transfer was indeed important, they were only able to look at it indirectly in their set-up via agitation rate and liquid fill volume. In addition, pH which routinely varies in growth conditions unless controlled by a pH feed-back

loop, was not included as a factor in their DoE design. While the model generated predicted yield values in good agreement with those obtained ($R^2 = 0.817$), the protein yields were unexpectedly low and appeared to result from lysed cells. Subsequently, the same authors examined how oxygenation efficiency (as measured by the oxygen mass transfer coefficient, k_La) is a parameter that can be used to guide scale-up from small scale processes such as micro-well plate experiments up to larger, more conventional bioreactor production systems.¹¹²

In this study we wanted to determine a simpler, more direct method, for the rapid scale-up of initial and potentially promising small-scale screening methods, thereby minimising empirical, time-consuming trial-and-error practices. Here, we describe how a DoE-derived model generated using a parallel mini-bioreactor (that can control pH, temperature and dissolved oxygen tension (DO)) is predictive not only of protein yield normalised for culture density in the same system, but is directly scalable to cultures grown in a 7 L bioreactor. By further optimising both the accumulation of cell density in batch and improving the fed-batch induction regime, we then demonstrate that additional yield improvement can be achieved. By essentially investigating the cell growth and productivity phases separately, our approach has the benefit of reducing process development time by allowing the direct scale-up of the best production conditions from bench to bioreactor.

Figure 3.1 illustrates a typical fed-batch *P. pastoris* fermentation profile: Region A is the glycerol batch phase where biomass is accumulated; it is also common to include a glycerol feeding stage to further increase achievable biomass; Region B is the carbon-source starvation phase where all remaining glycerol is utilised for cell growth; Region C is the adaptation phase where the process conditions (temperature, pH and DO) are gradually changed from those for optimal cell growth to those for optimal protein expression; Region D is the induction phase where methanol is supplied to the cells at a defined feeding rate (methanol activates the *AOX1* promoter frequently used with the *Pichia sp.* system).

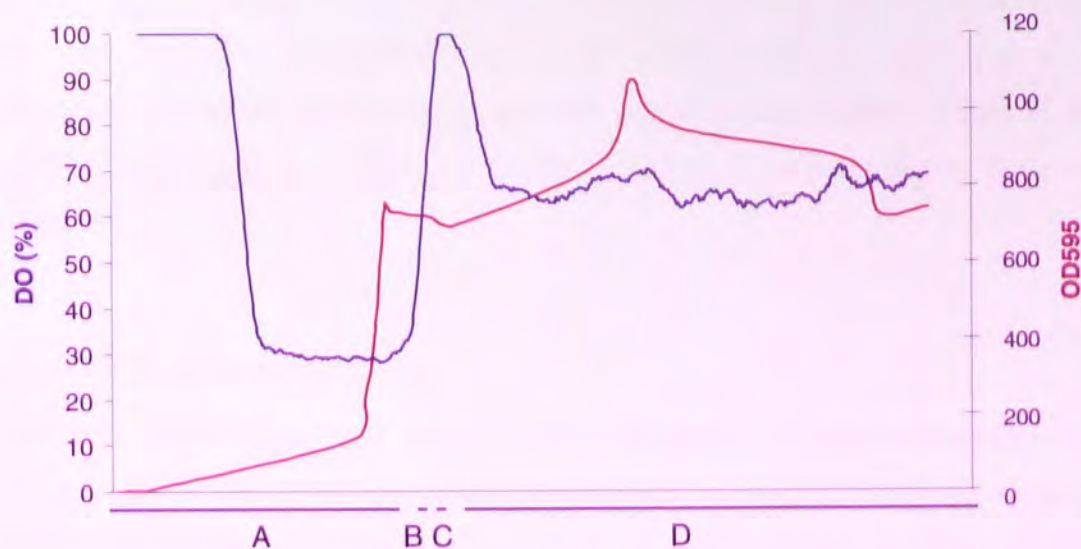


Figure 3.1. Typical biomass (OD₅₉₅) and dissolved oxygen (DO; %) profiles for a fed-batch fermentation of *P. pastoris*. Process phases are indicated by: A glycerol batch, B starvation, C adaptation and D induction (methanol feed).

3.1 Optimisation of GFP expression in *P. pastoris*

As discussed in Chapter 1, green fluorescent protein (GFP) is a widely used tool in modern bioscience. It was originally isolated from *Aequorea victoria* commonly known as the crystal jellyfish. To date recombinant GFP (rGFP) has been expressed in micro-organisms,¹¹³ plants²² and animals.¹¹⁴ In the majority of these rGFP expression experiments, the production of GFP is not an end in itself, but more likely an indicator of cellular activities, such as gene regulation,²⁵ intra-cellular pH,²⁴ viral sub-cellular localisation¹⁹ or as a “stress-probe”.¹¹⁵ The aim of our work was to generate rapid methods for maximising recombinant protein expression based on the control of the process conditions, namely culture temperature, pH and aeration. As such, GFP fitted the needs of the method generation procedure as it enables simple and quick detection of protein yield by fluorescence, with an excitation wavelength (λ_{ex}) of 397 nm and emission wavelength (λ_{em}) of 506 nm.

The optimisation method required the construction and characterisation of a GFP expressing strain of *P. pastoris*. As described in section 1.4.3.2, of the many GFP variants available, we used cycle 3 mutant GFP¹¹⁶ referred to as GFP_{uv}. The GFP_{uv} expressing cell-line was designated X33GFP_{uv}. The strain creation involved the design and construction of an expression vector (pPICZaA-GFP_{uv}) containing the DNA sequence coding for GFP_{uv}. The vector was constructed using standard molecular biology techniques such as PCR, restriction digests and subsequent ligation. The vector was then used to transform parental *P. pastoris* cells (strain X33), producing a novel strain X33GFP_{uv}.

The new strain was then characterised for growth and productivity. Such profiling provides the operator with an expectation of how the organism will behave in culture. Productivity profiling confirmed the expression of GFP_{uv} and at what point during the culture it was detectable.

3.1.1 GFP as the target protein

GFP was chosen as a target protein, due to its ease of detection coupled with numerous examples in the literature of successful recombinant expression. The main attraction of using GFP for developing our optimisation methods was the rapid assay techniques that could be applied to determine secreted product yield in the culture supernatant *i.e.* fluorescence analysis of culture supernatant. Campbell and Choy¹¹³ indicated the steps required to gain successful expression of two secreted GFP variants from *P. pastoris*. This showed that generating a GFP-expressing strain of *P. pastoris* was feasible.

3.1.2 Strain creation

The parental *P. pastoris* strain, wild-type X33, was transformed with a linearised vector containing the GFP_{uv}-coding cDNA generating the strain X33GFP_{uv}. The GFP_{uv} cDNA was sourced from Clontech and supplied in the plasmid holding vector pGFP_{uv} which also contained an antibiotic resistance (ampicillin) gene and an origin of replication for *E. coli* (pUC). The holding vector was propagated in *E. coli* to provide sufficient material (GFP_{uv} cDNA) for manipulation prior to transformation of X33 cells. Clone Manager v5.02 software was used to design the plasmid and identify suitable restriction enzymes for each manipulation of the DNA sequence.

The GFP_{uv} coding sequence was isolated and amplified from pGFP_{uv} by PCR. The forward and reverse primers were designed to introduce restriction sites at either end of the GFP_{uv} sequence: 5' *EcoRI* (GAA TTC) and 3' *XbaI* (TCT AGA). The restriction sites allowed the insertion of GFP_{uv} into the *P. pastoris* expression vector pPICZαA as the multiple cloning site (MCS) contains both *EcoRI* and *XbaI* recognition sites. The PCR product (*EcoRI*-GFP_{uv}-*XbaI*) and expression vector (pPICZαA) were sequentially digested with *EcoRI* and *XbaI* to provide fragments with "sticky ends". The digested fragments were ligated to form the plasmid pPICZαA-GFP_{uv}. Figure 3.2 shows an annotated plasmid map for pPICZαA-GFP_{uv} generated using Clone Manager v5.02. Before the vector could be used to transform

competent X33 cells, the expression cassette was sequenced to confirm the fidelity of the completed vector. The attached data CD contains sequence analysis files showing that the correct sequence had been obtained.

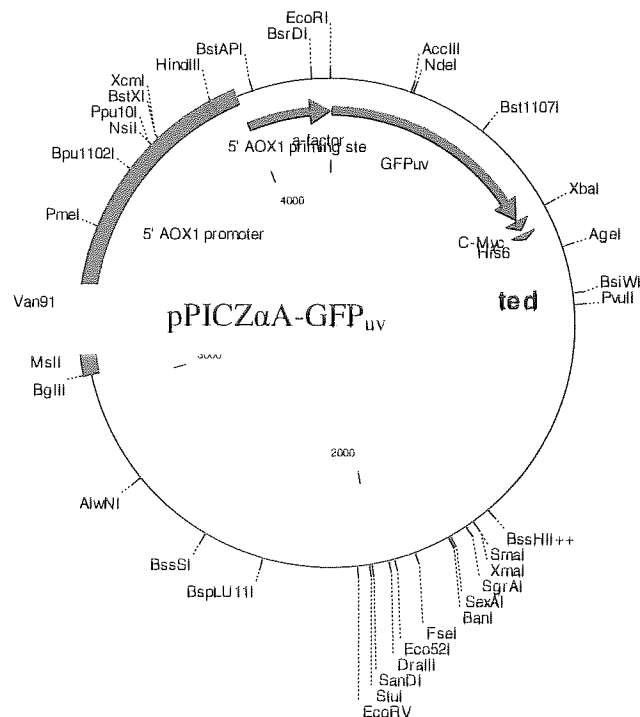


Figure 3.2. Plasmid map of pPICZαA-GFP_{uv} generated in Clone Manager v5.02. Notable features are the AOXI promoter, α-factor secretion signal, GFP_{uv} coding sequence and the N-terminal tags, c-myc and His₆. The location of the PmeI recognition site is also visible in the AOXI promoter region.

The pPICZαA-GFP_{uv} vector was incorporated into the host genome thus providing a stable integration. To achieve this, the plasmid was linearised by enzymatic digestion, the selected enzyme having only one recognition sequence in the plasmid so fragments were not generated. pPICZαA-GFP_{uv} was linearised with *PmeI*, as the alternative *SacI* had recognition sites within the GFP_{uv} sequence. Parental *P. pastoris* X33 cells were transformed by electroporation, which subjects the cells to a large (>1200 V) electrical pulse forming pores in the membrane through which linearised vector DNA can enter the cell. To prevent damage to the cells during the electrical pulse, the cells were made electrocompetent. The preparation involved extensive cell washing to remove any remaining medium salts and chilling of the cells to 2-4 °C. Any medium remaining in the cell slurry allows arching of the current and a characteristic ‘pop’ can be heard, which reduces transformation efficiency.

For the cell transformation, 5 μ L of *Pme*I-linearised pPICZ α A-GFP_{uv} (approximately 20 μ g from a Qiagen Maxiprep) were mixed with 4 μ L electrocompetent X33 cells, transferred to an electroporation cuvette and pulsed at 1800 V. The cells were recovered at 30 °C with 200 rpm agitation for 1 h. The cells were then plated on selective agar (100 μ g/mL zeocin) and incubated for 72 h. From this process a total of five colonies were obtained. These colonies were individually re-streaked onto fresh selection plates and all five colonies re-grew indicating that the zeocin resistance gene from pPICZ α A-GFP_{uv} had been integrated. Glycerol stocks of each X33GFP_{uv} clone were generated and stored at – 80 °C, as described in Chapter 2.

3.1.3 Strain characterisation

The X33GFP_{uv} clones that re-grew on selective agar, were able to survive on account of having integrated the zeocin antibiotic resistance portion of the linearised plasmid into the cellular genome. It was theoretically possible that the clones did not contain the GFP_{uv} expression cassette and so to confirm full integration, expression screening was carried out as part of the strain characterisation. Shake flask cultures were carried out to identify positively-expressing clones. Five 500 mL baffled shake flasks containing 50 mL BMGY medium were inoculated with a single colony of each of the five clones and grown overnight. Clone 2 did not grow in liquid medium despite growing well on agar and was therefore not used in any further work. In contrast, clones 1, 3, 4 and 5 grew well and were used for expression studies. Overnight BMGY cultures were used to inoculate 50 mL BMMY medium (1 L baffled shake flasks) to OD₅₉₅ 1.0. The flasks were incubated at 30 °C with 200 rpm agitation for 48 h with 100 % methanol added to 1 % v/v at 24 h to maintain induction. The cultures looked visibly green when exposed to UV light and supernatant samples were assayed by SDS-PAGE and anti-His₆ Western blotting. All the clones showed two bands at approximately 27 kDa by anti-His₆ Western blot in agreement with the expected size for native GFP. Two bands were excised from the corresponding SDS-PAGE gel in this region and sent for analysis by ESI-QUAD-TOF mass spectrometry. The raw results were subjected to a MS/MS ion search using the web-based Mascot program (available at www.matrixscience.com). The full mass spectrometry results can be found in the attached data CD. The upper band (see Figure 3.3) was found to be cycle 3 mutant GFP_{uv} and Figure 3.4 shows the mass-spec fragment results on a snake plot of the GFP_{uv} amino acid sequence.

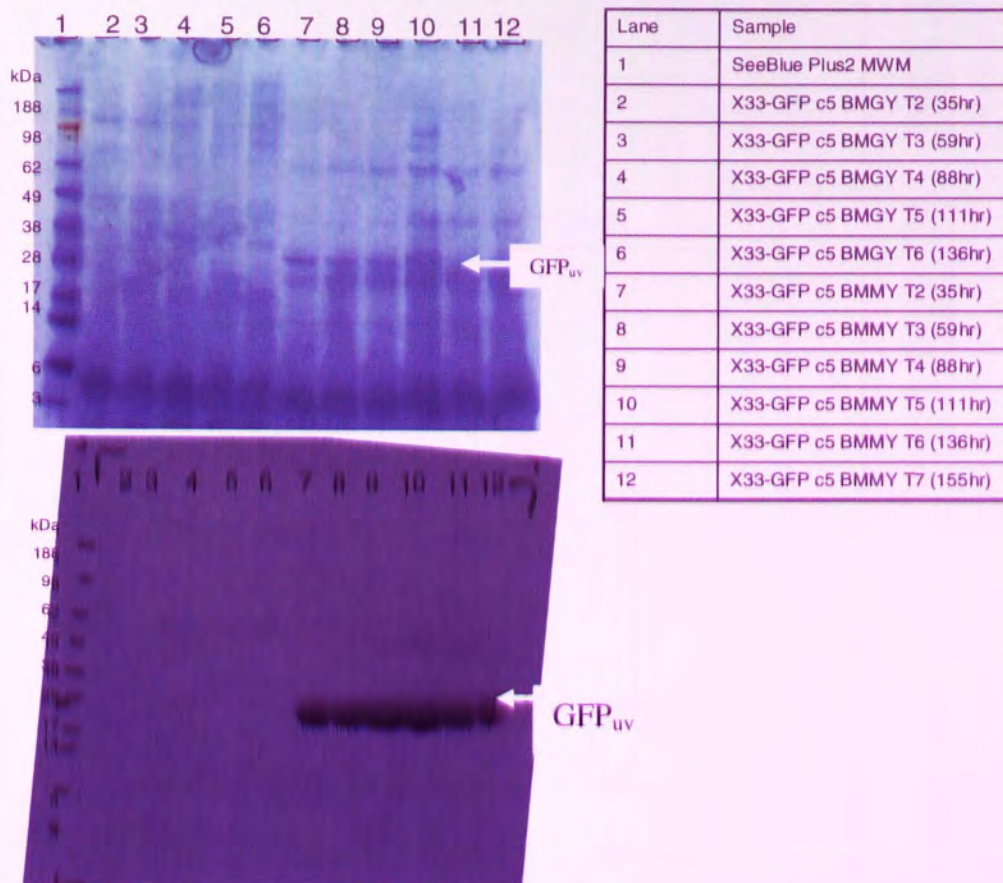


Figure 3.3. SDS-PAGE and anti-His₆ Western blot analysis of *P. pastoris* X33GFP_{uv} culture supernatants. Samples were taken from extended time course shake flask cultures with sampling points as shown in Figure 4. Arrows indicate the position of the GFP_{uv} product band. The upper panel shows an SDS-PAGE gel visualised with Novexin InstantBlue (a Coomassie-type stain) and the lower panel shows a Western blot detection by chemiluminescent probe and photographic film with approximate molecular weight positions transposed from the blot membrane to film.

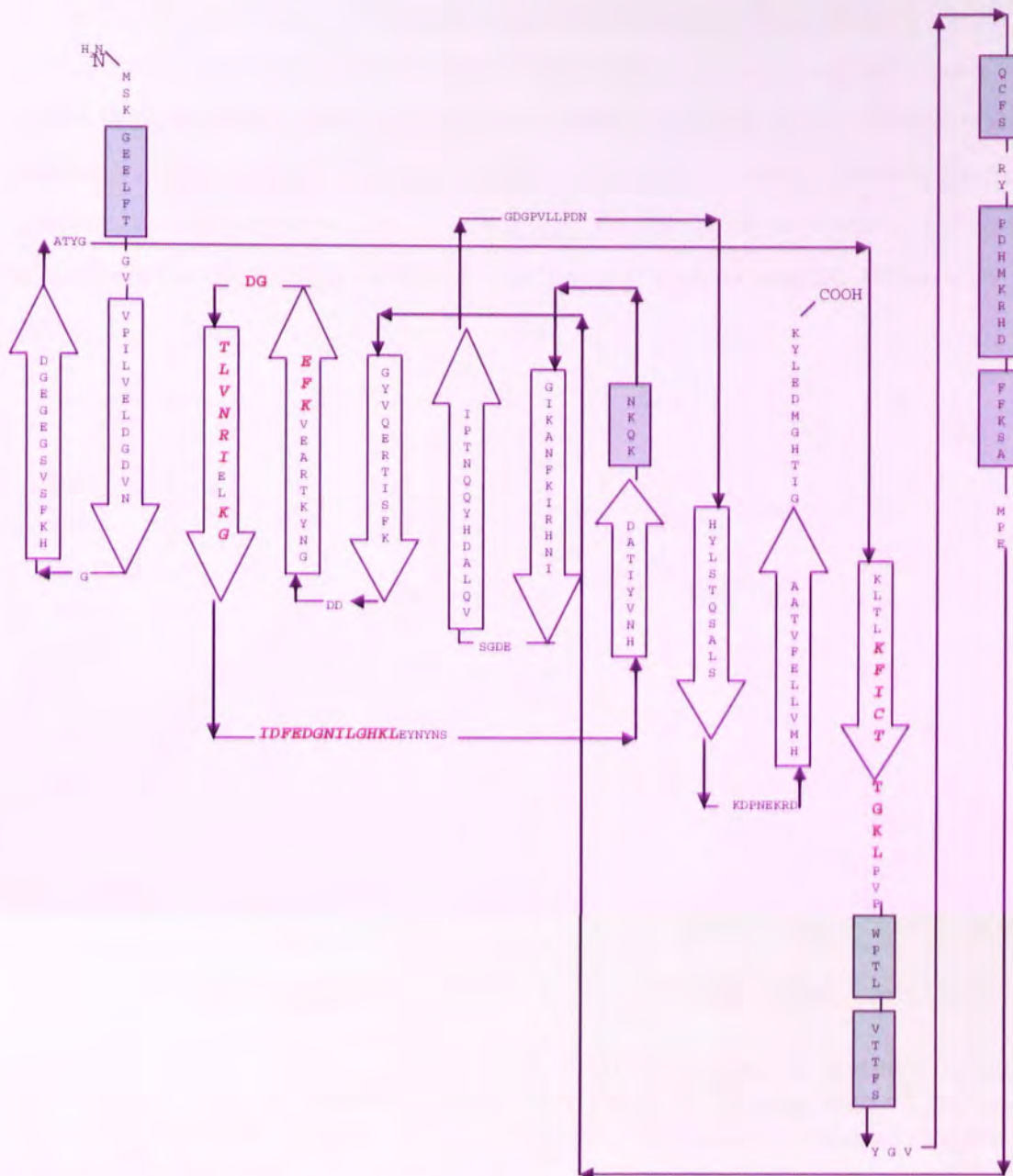


Figure 3.4. Snake-plot of the GFP_{uv} amino-acid sequence showing the mass spectrometry result fragments highlighted in red and larger font. The sample was the upper 27 kDa band from the SDS-PAGE analysis of *P. pastoris* X33GFP_{uv} culture supernatant.

3.1.3.1 Growth curves

In order to gain more detailed growth profiles for the X33GFP_{uv} clones 1, 3, 4 and 5 eight shake flask cultures were set-up to run over an extended time course of 155 h. A BMGY seed culture of each clone was used to inoculate 500 mL shake flasks containing 100 mL BMGY or BMMY to OD₅₉₅ 0.7. Therefore each clone was cultured in both non-induced (BMGY) and induced (BMMY) conditions. Samples were taken at 24 h intervals for biomass estimation (OD₅₉₅) and SDS-PAGE protein analysis. Induction was maintained at 24 h intervals by addition of 100 % methanol to 1 % (v/v), with the non-induced cultures having an equivalent

glycerol addition. Figure 3.5 shows the growth profiles and sampling points for these extended time course shake flask cultures. The data for clone 1 are not shown because the culture became contaminated. It can be seen that the non-induced cultures reach a higher biomass than the induced cultures. It is also noticeable that despite a carbon source fed-batch approach, biomass accumulation slows towards plateaux in all the cultures. This is typical and indicative of another culture factor being limiting, such as additional nutrients or oxygen supply.

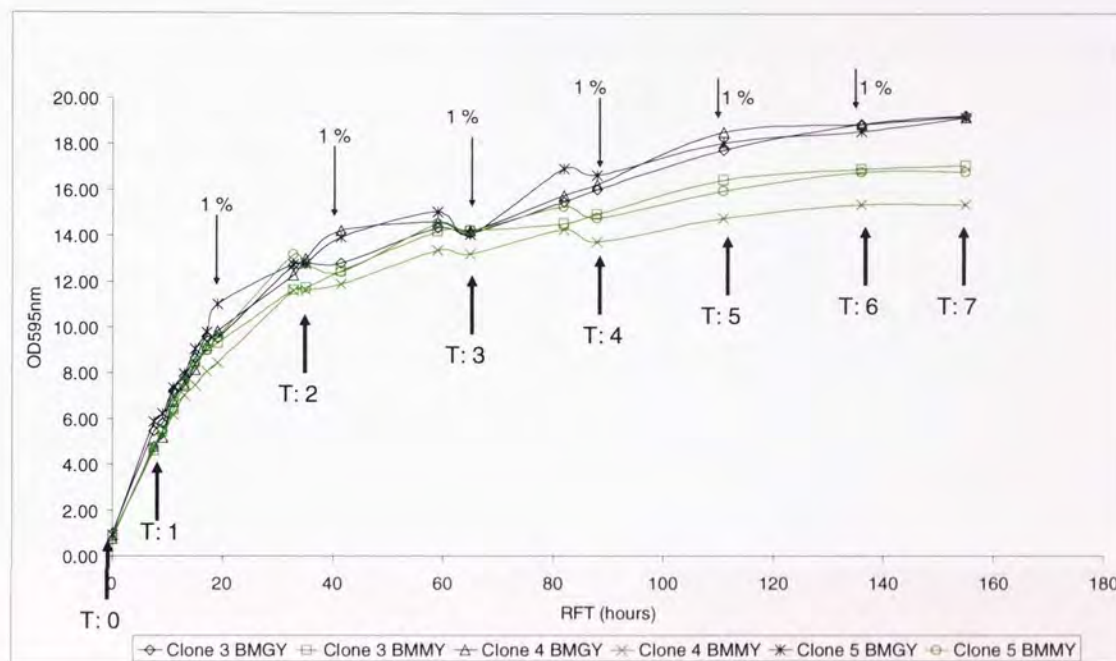


Figure 3.5. Growth profiles of *P. pastoris* X33GFP_{uv} clones 3, 4 and 5 in non-induced (BMGY) and induced (BMMY) shake flask cultures. Arrows with “1 %” indicate the point of methanol or glycerol addition to 1 % of the culture volume. Arrows (T: 0 – 7) indicate sampling points.

Figure 3.3 shows the SDS-PAGE and anti-His₆ Western blot analysis of supernatant samples taken from both non-induced (BMGY) and induced (BMMY) X33GFP_{uv} clone 5 cultures. The ~27 kDa upper band is indicated by an arrow. It is possible to see the accumulation and later degradation of the GFP_{uv} band in the induced (BMMY) cultures, which is detectible at 35 h post induction. No GFP_{uv} bands are seen in the non-induced samples. Sample analysis of clones 3 and 4 showed similar trends of GFP_{uv} accumulation and degradation but had lower density product bands. Clone 5 was therefore selected for use in future work as it showed a higher yield, which would therefore be easier to detect.

3.1.4 Fluorescence analysis

One of the main motivations for using GFP_{uv} as a model target protein is the ease of its detection by fluorescence. The GFP_{uv} fluorophore is excited by light at a wavelength of 395 nm and emits at 509 nm (green light). To determine if the fluorescence signal in relative fluorescence units (RFU) possessed a linear relationship to product (GFP_{uv}) band density on SDS-PAGE gels, a supernatant sample containing GFP_{uv} was serially diluted and assayed by SDS-PAGE and fluorescence. Figure 3.6 shows a plot of the SDS-PAGE band densitometry and corresponding RFU values. A linear relationship can be seen between fluorescence and band density, which shows that RFU can be used to rapidly estimate GFP_{uv} concentration in a supernatant sample.

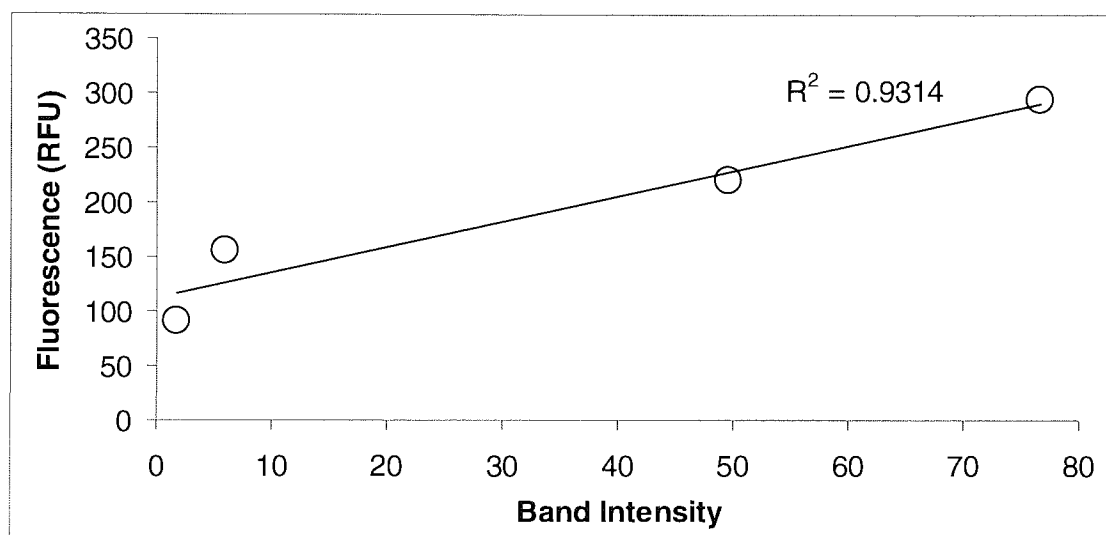


Figure 3.6. Scatter plot showing the relationship between *P. pastoris* X33GFP_{uv} culture supernatant fluorescence (RFU) and SDS-PAGE band densitometry. Serial dilutions of a GFP_{uv}-containing culture supernatant were assayed by fluorescence and SDS-PAGE.

In order to quote GFP_{uv} protein concentrations and not just the associated RFU values, GFP_{uv} was further quantified using a Bradford assay. Serial dilutions of X33GFP_{uv} supernatant samples were compared to a bovine serum albumin (BSA) standard curve to determine the total protein concentration of the X33GFP_{uv} samples. The X33GFP_{uv} supernatant samples were also assayed by fluorescence and the RFU values plotted against the Bradford assay results. The equation for the line of fit provides a conversion for fluorescence (RFU) to mg/mL GFP_{uv}. Figure 3.7 shows the scatter plot of GFP_{uv} fluorescence and protein concentration. The resulting conversion factor for RFU to mg/mL is 3.101×10^{-7} . This was used in all yield calculations where protein concentration was required.

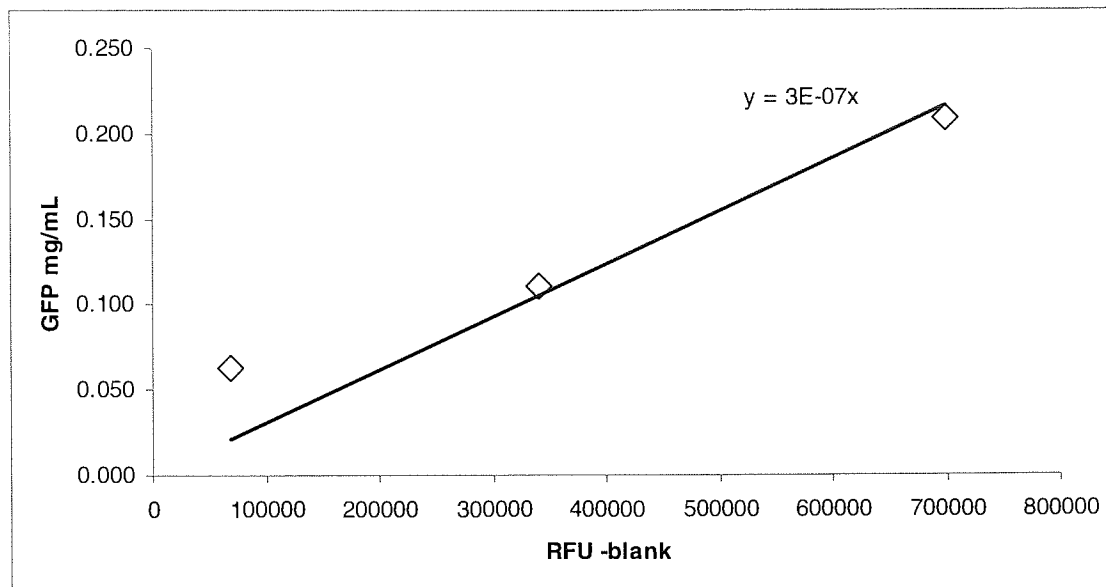


Figure 3.7. Calibration plot for GFP_{uv} fluorescence (RFU) to GFP_{uv} protein concentration (mg/mL). The line of fit provides the conversion factor, protein concentration (mg/mL) = $3.101 \times 10^{-7} \times \text{RFU}$.

3.2 Initial optimisation of GFP yield using a design-of-experiments approach

If an optimisation process were based around varying one factor at a time for a large number of variables, the time required to carry out the experimental runs would be considerable and costly. If it were instead possible to extract the same information from fewer experiments in parallel this would dramatically reduce the costs and time associated with optimising the process. We therefore explored the use of statistical design of experiments (DoE) in combination with a parallel mini bioreactor in process optimisation to maximise the yield of GFP from clone 5. DoE uses statistical techniques to model the process based on a number of “snap-shots” or treatment combinations and then infer what occurs elsewhere in the design space within the process limits. The role of DoE in process optimisation is more fully discussed in Chapter 1.

3.2.1 The Applikon Micro 24 Bioreactor

Generating a scalable process model can greatly reduce the time and cost of developing an optimised production process. A key requirement of any small-scale system used for model building is equivalence: the process must run in the same manner at both small and production scale. Here we used the Micro 24 Bioreactor from Applikon Biotechnology. This is a bench-top, computer-controlled micro-bioreactor system. The Micro 24 uses a modified deep 24-well plate as an incubation cassette to provide 24 individually-controlled “vessels” which are run in parallel. Figures 3.8 and 3.9 show annotated photographs of the Micro 24

system, whilst in use at Aston University. The Micro 24 is a relatively new system with few examples in the literature describing its use. It is of particular use to those in industrial process development settings where a rapid development timeline is important.

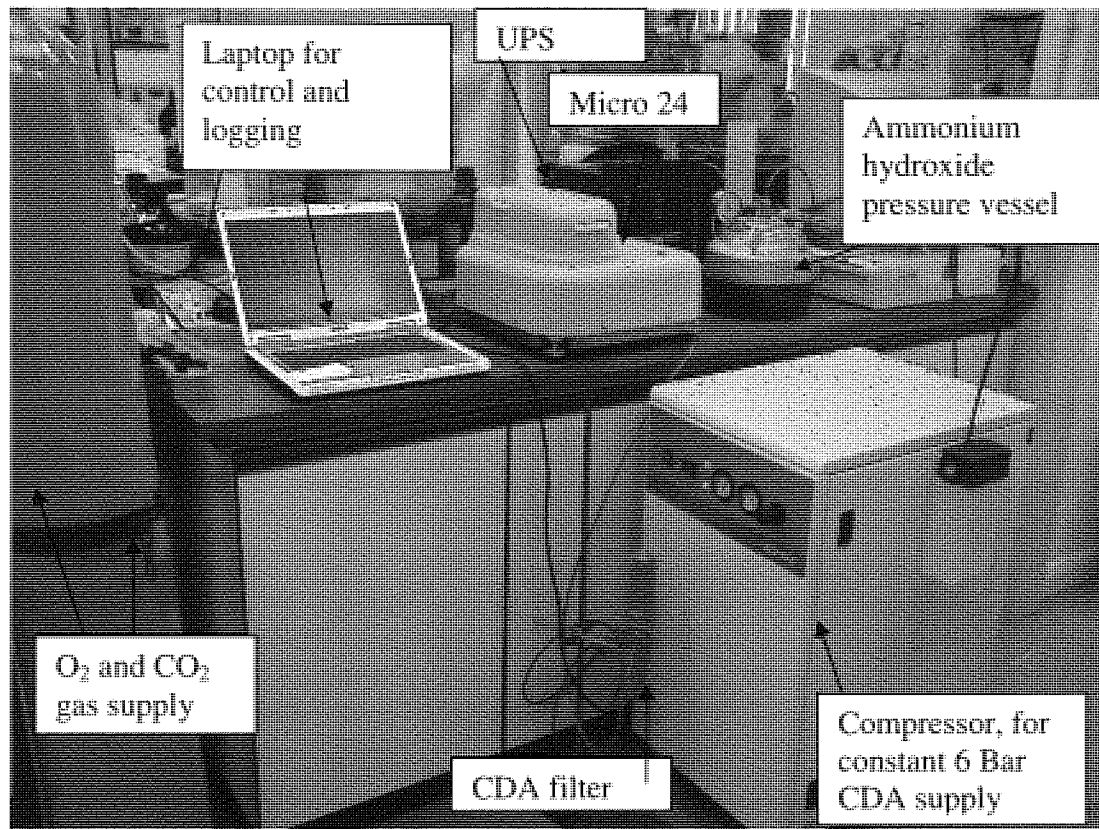


Figure 3.8. Applikon's Micro 24 bioreactor system in use at the Aston laboratory, highlighting the required ancillary equipment and bioreactor unit size. The uninterruptable power supply (UPS) and clean dry air (CDA) supply are also shown.

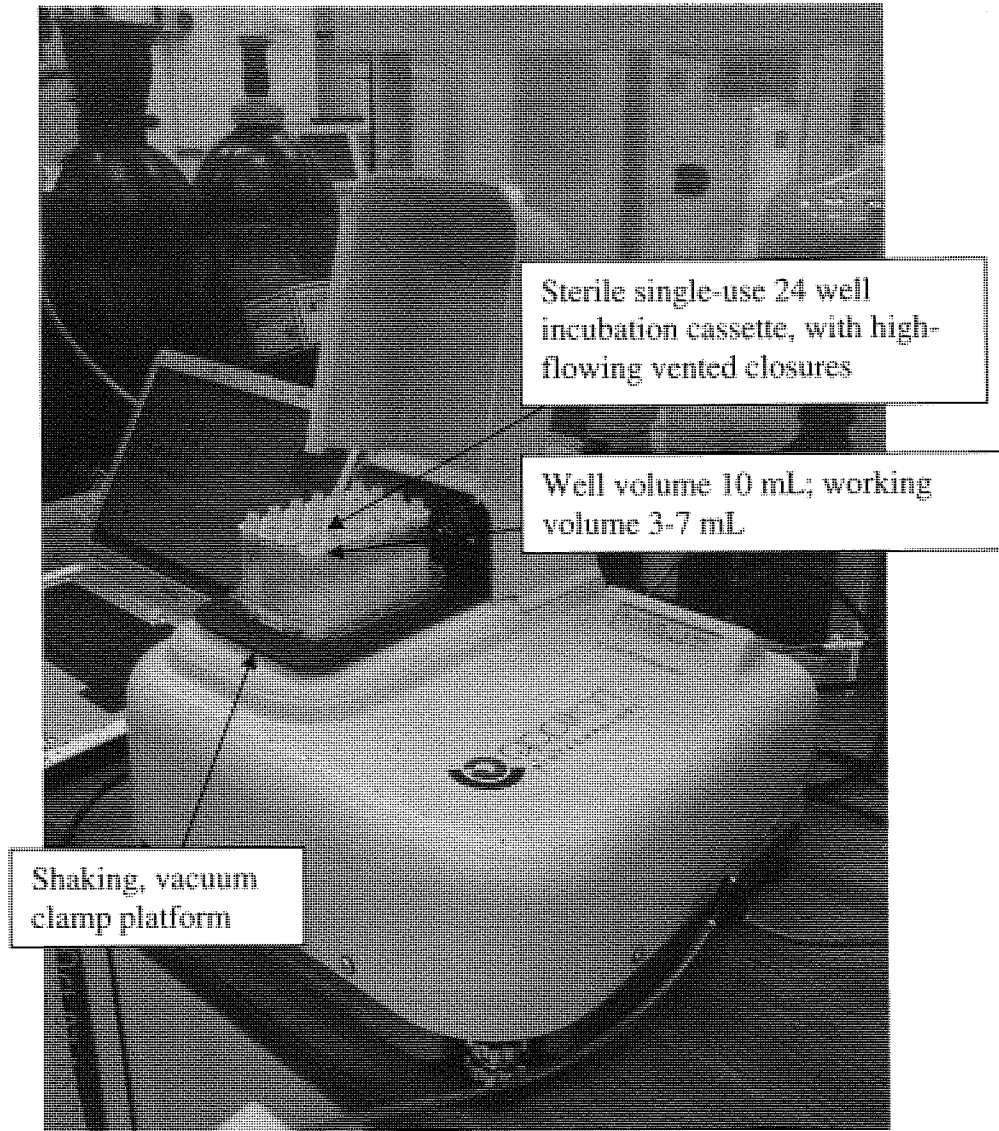


Figure 3.9. Applikon's Micro 24 bioreactor system in use at the Aston laboratory. The lid of the incubation chamber is open showing the incubation cassette, well closures and cultures during an experimental run.

The incubation cassette has dissolved oxygen tension (DO), pH optical sensors, a sparging port and two thermal conduction pads in the bottom of each well for monitoring and control. One thermal pad is used for monitoring the culture temperature and the other to transfer heat into the well from the heating element in the base unit. The culture pH and DO are controlled to set point by sparging gas through the medium e.g. O₂ for DO and CO₂ or ammonium hydroxide vapour for pH. The actual gassing strategy is determined by the requirements of the organism. The settings used here may not be suitable for other organisms or even experiments, depending on the run conditions such as a requirement for micro-aerobic conditions.

The wells were sealed using Type A closures, which are 0.2 μm vented caps with a non-return ball valve suited to high gas flow rates. Two other types are available for differing gas

flow/evaporation rates. The incubation cassette was secured to the shaker platform by a vacuum generator running from a 6 bar clean dry air (CDA) supply. The laptop and Micro 24 unit were powered via an uninterruptable power supply (UPS) to avoid down time due to power fluctuations. The software provided both control of run conditions and on-line logging of the condition in each well. The software's display shows graphs of gas flow rates for pH and aeration, heater drive (temperature) and actual well conditions. Figure 3.10 shows an example of the Micro 24's control software display. The Applikon Biotechnology website (www.applikon-bio.com) contains additional specifications and promotional information, such as organisms already used in the system.

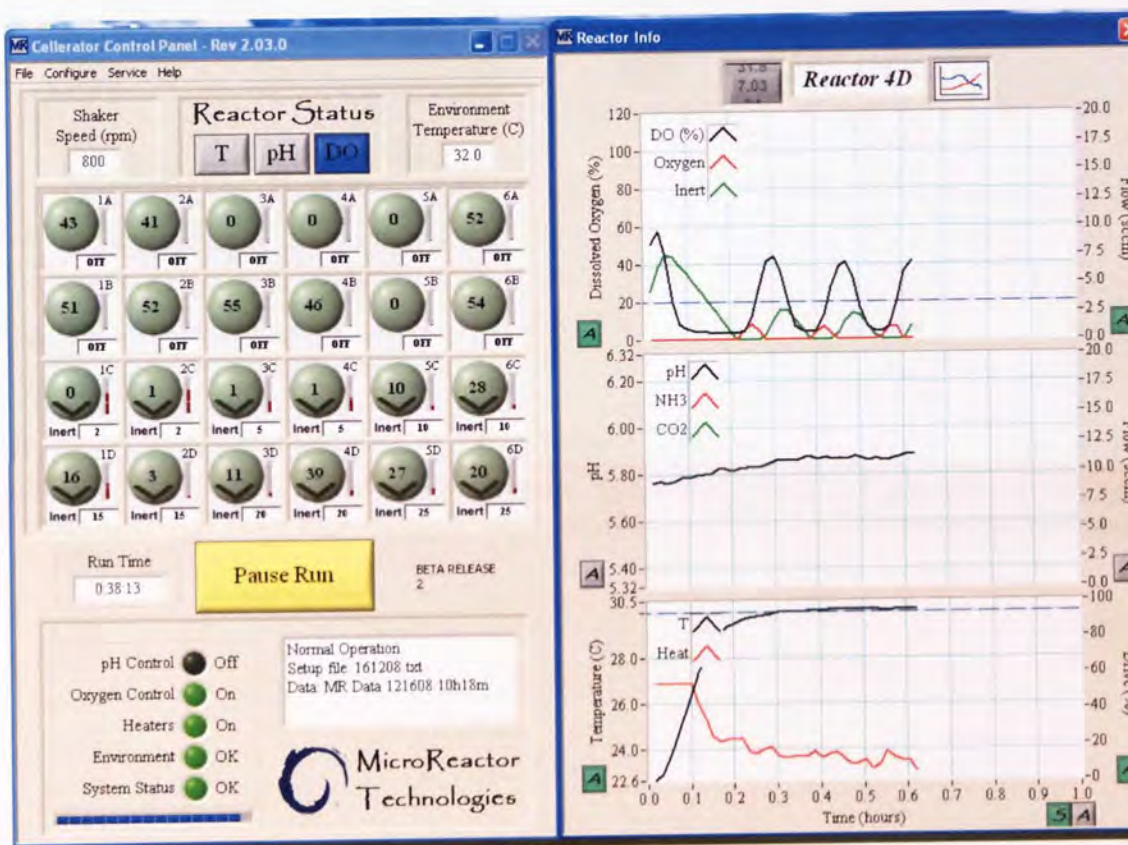


Figure 3.10. Example of the Micro 24 control software display. Left hand panel shows current well conditions, set points and system information. Right hand panel shows logging data (temperature, pH and DO) for an individual well, with actual conditions, level and method of control.

3.2.2 Why use Design of Experiments?

The question asked of DoE here is whether altering the process conditions (culture temperature, pH and DO) increases GFP_{uv} yield from *P. pastoris* X33GFP_{uv}? The culture temperature, pH and DO were therefore each varied at three levels -1, 0 and +1 (low extreme, mid point and high extreme) within a Box-Behnken design. Figure 3.11 shows a representation of the design space, which is rotatable, but contains regions of poor prediction

quality. These "missing corners" can be useful when the experimenter wishes to avoid combined factor extremes. There are also no combinations requiring all 3 factors to be maximal. This property prevents a potential loss of data in those cases. As the design requires 3 levels for each factor being investigated, these were based on previous experience of the organism (*P. pastoris* X33GFP_{uv}) and process settings used in the literature. Table 3.1 shows the DoE matrix of required factor settings for the initial predictive model.

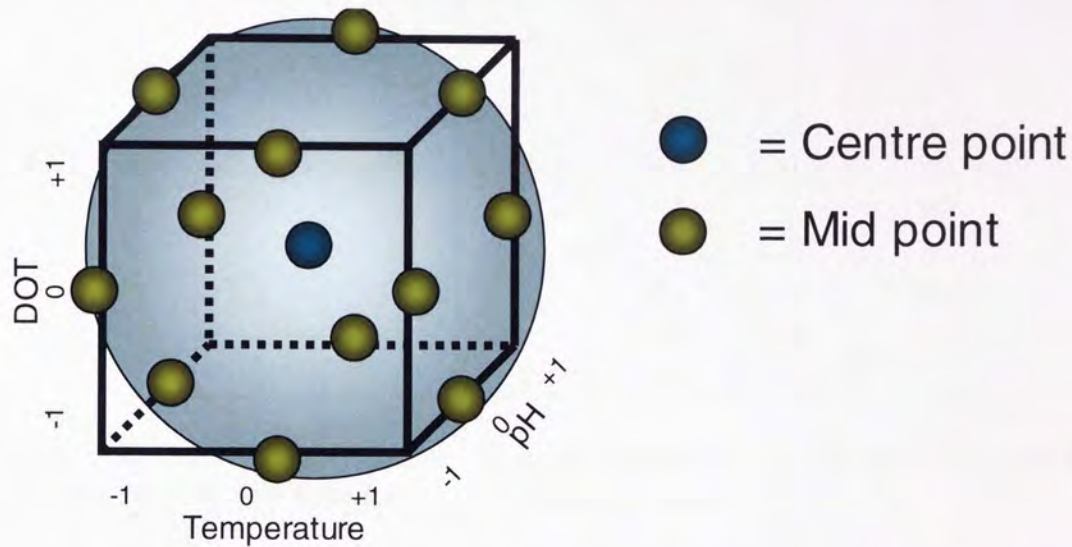


Figure 3.11. Representation of the Box-Behnken design space. The geometry of this design suggests a sphere within the process space such that the surface of the sphere protrudes through each face with the surface of the sphere tangential to the midpoint of each edge of the space.

CODED FACTORS			INPUT FACTORS		
T	pH	DO	T (°C)	pH	DO (%)
-1	1	0	20	6.7	45
-1	-1	0	20	5.3	45
-1	0	1	20	6	80
-1	0	-1	20	6	10
0	1	-1	25	6.7	10
0	0	1	25	6	80
0	0	-1	25	6	10
0	1	1	25	6.7	80
0	0	0	25	6	45
0	-1	-1	25	5.3	10
0	-1	1	25	5.3	80
1	1	0	30	6.7	45
1	0	-1	30	6	10
1	0	1	30	6	80
1	-1	0	30	5.3	45

Table 3.1. Experimental matrix showing coded (-1, 0, 1) and non-coded set-points required for the Box-Behnken DoE1 predictive model.

3.2.3 Normalised yield allows additional flexibility in process optimisation

The modelling procedure was built around the use of normalised yield data, which is determined by dividing yield (RFU) by biomass (OD₅₉₅) at the point of harvest. This is a measure of productivity on a per cell basis and so is more likely to be predictive over a range of different culture set-ups that typically generate different culture densities. Additionally, a model based on normalised yield would be consistent with a two step optimisation, allowing for specific accumulation of biomass prior to induction using the predicted optimal 'per cell' production conditions of the model, thus leading to further improvement in total yield.

3.2.4 Data collection for predictive model

A classic quadratic Box-Behnken design⁹¹ was employed for generating the predictive model. This allowed a reduced number of treatment combinations to be used in building the model and gave no bias towards any potential optimum region of the process space. Minitab v15.1.30 software was used to generate an experimental matrix of factor combinations required to build the predictive model. Table 3.2 shows the input variables and measured outputs for the model building experiments which were performed in the Micro 24 parallel multi-well bioreactor. The factor combinations were carried out in at least triplicate to provide

statistical confidence in the model building data. Biomass (OD_{595}) and yield (RFU) measurements were taken at 48 h post induction. The input values for the model were based on the normalised yield (RFU/ OD_{595}) values with outliers removed from the full data set. The rationale for this was to increase the confidence in the data used to produce the predictive model. The full data set was examined and outliers identified. The culture data for each outlying data point was interrogated and if process deviations were encountered e.g. factor set point was not maintained, the removal of the associated outlying data point was seen as justifiable.

Input Factors (Controlled on-line)			Yield (RFU)			Achieved Biomass (OD ₅₉₅ nm)			Normalised Yield (RFU/OD ₅₉₅ nm)			Full data set			Data set with outliers removed		
T (°C)	pH	DO	y1	y2	y3	OD1	OD2	OD3	norm.y1	norm.y2	norm.y3	mean norm.y	SD	SD % of mean	mean norm.y	SD	SD % of mean
20	6.7	45	1532.3	2006.2	1461.2	16.5	21.7	14.8	93.1	92.7	99.1	95.0	3.6	3.8	95.0	3.6	3.8
20	5.3	45	-32.7	-89.8	-42.4	2.5	2.9	1.6	-13.1	-31.0	-26.5						
20	5.3	45	-141.6	0.0	-57.3	4.9	7.8	5.8	-29.2	0.0	-10.0	-18.3	12.5	-68.2	-28.9	2.2	-7.8
20	6	80	675.2	707.0	668.8	12.7	16.1	16.1	53.2	43.9	41.7	46.2	6.1	13.2	42.8	1.6	3.7
20	6	10	261.6	250.3	382.7	16.2	17.1	14.0	16.1	14.7	27.3						
20	6	10	836.9	858.9	746.5	25.6	26.6	28.7	32.8	32.4	26.0	24.9	7.8	31.4	29.6	3.4	11.6
22.5	6.7	80	1280.3			25.9			49.5			49.5	0.0	0.0	49.5		
25	6.7	10	2196.4	2331.4	2318.6	19.4	22.7	21.7	113.2	102.9	107.1	107.7	5.2	4.8	107.7	5.2	4.8
25	6	80	514.4	321.5	-47.7	13.0	17.7	9.0	39.7	18.2	-5.3						
25	6	80	1146.1	336.3	1334.2	25.6	23.6	26.4	44.8	14.3	50.6	27.0	21.5	79.6	45.0	5.5	12.1
25	6	10	-55.6	-48.3	-20.6	6.9	7.3	7.7	-8.1	-6.6	-2.7	-5.8	2.8	-48.1	-5.8	2.8	-48.1
25	6.7	80	1353.4	1999.2	2880.1	15.2	21.2	20.4	89.0	94.5	141.5	108.4	28.9	26.6	91.8	3.9	4.2
25	6	45	586.0	713.7	569.6	16.7	16.9	16.3	35.2	42.4	34.9						
25	6	45	38.7	672.0	501.2	11.2	17.5	18.0	3.5	38.4	27.9						
25	6	45	516.9	191.3	57.2	17.3	17.4	20.2	29.9	11.0	2.8						
25	6	45	940.4	1084.0	281.8	25.6	26.6	27.4	36.7	40.8	10.3						
25	6	45	1073.9	1224.2	942.4	26.4	30.2	27.6	40.8	40.6	34.1	28.6	14.3	49.8	34.4	8.6	25.0
25	6	50	871.4	510.4	540.9	19.8	17.5	20.7	44.0	29.2	26.2	33.1	9.5	28.8	27.7	2.1	7.6
25	5.3	10	-127.2	-42.8	-156.3	7.8	7.3	9.2	-16.3	-5.9	-17.1						
25	5.3	10	66.4	0.0	164.3	10.8	12.6	12.8	6.1	0.0	12.9	-3.4	12.1	-357.2	9.5	4.8	50.1
25	5.3	80	358.4	364.3	363.0	15.0	13.6	14.4	24.0	26.7	25.2	25.3	1.4	5.5	25.3	1.4	5.5
27.5	6.7	80	2584.7	2805.0		22.7	21.4		113.9	131.4		122.6	12.4	10.1	122.6	12.4	10.1
27.5	6.5	60	1811.7	1652.3	1596.5	19.8	17.2	16.9	91.5	96.1	94.6	94.0	2.3	2.5	94.0	2.3	2.5
27.5	6.3	60	1346.9	1321.8	1428.0	19.2	18.6	18.1	70.3	71.1	79.1	73.5	4.9	6.6	73.5	4.9	6.6
30	6.7	45	1854.6	1944.1	1909.7	19.6	17.1	18.9	94.5	114.0	101.3	103.3	9.9	9.6	103.3	9.9	9.6
30	6	10	86.7	109.7	80.5	15.0	15.3	17.5	5.8	7.2	4.6	5.9	1.3	22.0	5.9	1.3	22.0
30	6	80	869.3	-91.6	798.3	15.9	11.7	15.7	54.7	-7.9	50.8						
30	6	80	908.6	1035.5	950.4	22.3	21.3	20.0	40.7	48.7	47.6	39.1	23.5	60.0	48.5	5.1	10.5
30	5.3	45	321.4	332.2	345.2	13.9	14.2	14.3	23.2	23.4	24.1	23.6	0.5	2.1	23.6	0.5	2.1

Table 3.2. Specification of the input factors and measurable outputs for the initial model building experiments. The input factors were temperature (T), pH and dissolved oxygen tension (DO). Yield (RFU) and achieved biomass (OD₅₉₅), were measured in at least triplicate 48 h post induction. The standard deviation (SD; n 3) is given for the yield normalised to OD₅₉₅.

3.2.5 Building a predictive model for yield

The normalised yield data where outliers had been removed from the Micro 24 cultures, as shown in Table 3.2, was used to build a predictive model of normalised yield as a function of process conditions (temperature, pH and DO). The model building process used Minitab (v15.1.30) software to generate coefficients for each of the factor terms as shown in Table 3.3. The model was refined by removing those terms with high p -values associated with them, see Chapter 2 section 2.14. The resulting refined model was named “DoE1”. Equation 3.1 shows the terms and coefficients for the DoE1 model.

Term	Coefficient
Constant	1027.92
T	18.7193
pH	-491.271
DO	0.822675
T*T	-0.01201
pH*pH	54.205
DO*DO	0.004139
T*pH	-3.15409
T*DO	0.042111
pH*DO	-0.32396

Table 3.3. Initial predictive model terms and coefficients based on the model-building data shown in Table 3.2, leading to Equation 3.1 for normalised yield (RFU) where T = temperature (°C), pH = pH and DO = dissolved oxygen tension (%).

$$\text{Normalised yield (RFU/OD}_{595}) = 853.114 + (19.9937 * T) - (447.489 * \text{pH}) + (2.248 * \text{DO}) + (50.5565 * \text{pH} * \text{pH}) - (3.15409 * T * \text{pH}) - (0.32396 * \text{pH} * \text{DO})$$

Equation 3.1. Refined DoE1 predictive model, where; T = temperature (°C), DO = dissolved oxygen tension (%)

The model was examined statistically by analysis of variance (ANOVA; Table 3.4). It can be seen that the p -value associated with the regression F -statistic (11.09 %) is 0.002, this is strong evidence (99.8 %) against the null hypothesis (no linear relationship between any of the predictors and Y)¹¹⁷ therefore, there is evidence of a relationship between predicted and experimental GFP_{uv} yields.

Source	DF	SS	MS	F statistic	p value
Regression	6	21343.9	3557.32	11.09	0.002
Linear	3	18313.4	6104.46	19.03	0.001
Square	1	2291.1	2291.09	7.14	0.028
Interaction	2	739.5	369.73	1.15	0.363
Residual	8	2566.4	320.80		
Total	14	23910.3			

Table 3.4. ANOVA analysis of DoE1 model. Where; DF = degrees of freedom, SS = sum of squares, MS = mean square.

The model described by Equation 3.1 can be used to predict the coordinates on the design space of the optimal conditions. Minitab has a ‘response optimiser’ function which generates an output as shown in Figure 3.12. From this the optimal conditions on the design space were identified as: 20.0 °C, pH 6.7 and 80 % DO.

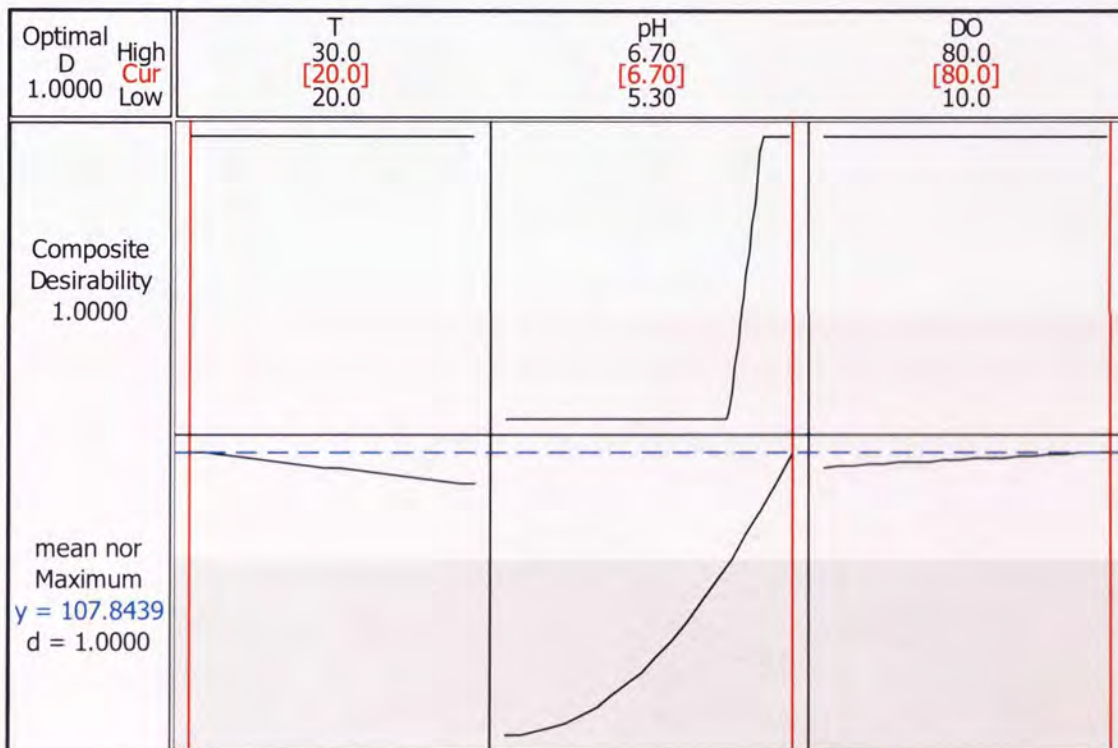


Figure 3.12. Response optimisation for DoE1 model as predicted by Minitab software. The predicted optimal conditions are 20 °C, pH 6.7 and 80 % DO.

The model was plotted visually using response surface regression to demonstrate how the factor interactions affected yield. Figure 3.13 shows contour plots of each of the factors: the hold values used were the “0” mid-points. Figure 3.14 shows a surface wire frame plot of yield against culture temperature (T) and pH, with DO held at 80 % (the “+1” factor setting). It can be seen that pH has a major influence on GFP_{uv} yield while DO has a much smaller impact. Figures 3.13 and 3.14 showed, however, that DoE1 does not capture the optimised set of conditions for maximal protein yield as the response surface plot does not achieve a

maximum value. However, they did provide a very useful indication of how to move the design space to find the optimal conditions and specifically to examine higher pH levels in a revised DoE.

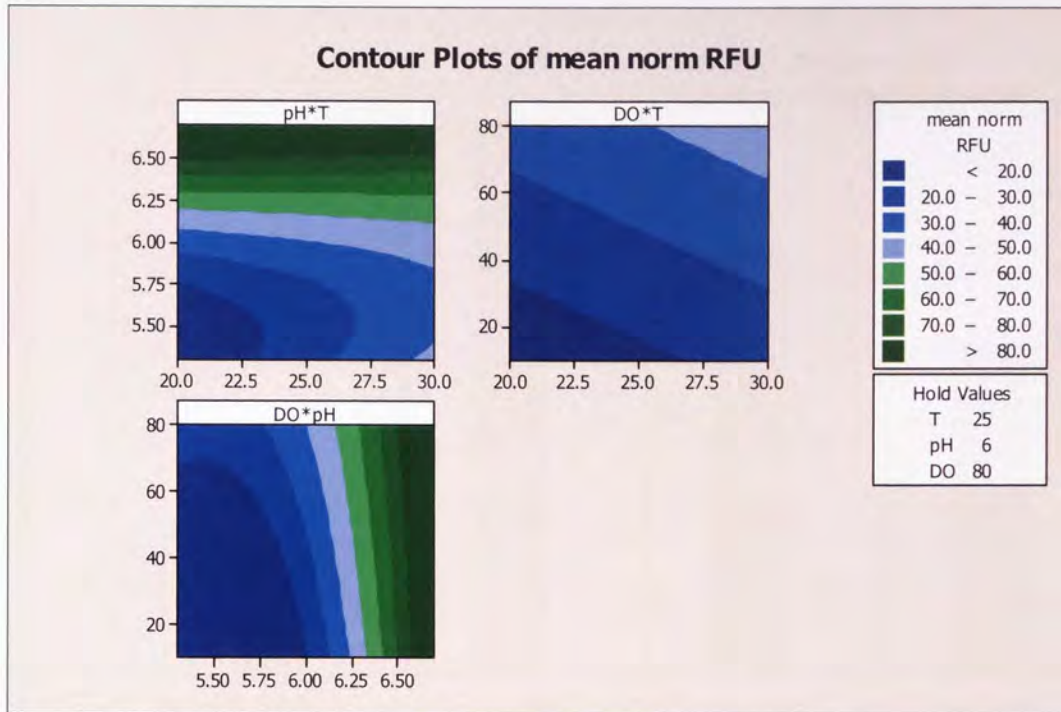


Figure 3.13. Response surface contour plot for DoE1 model showing normalised yield (RFU/OD₅₉₅) changing with each of the factors. Where T = temperature (°C), pH = pH and DO = dissolved oxygen tension (%). All hold values are the “0” mid-point values.

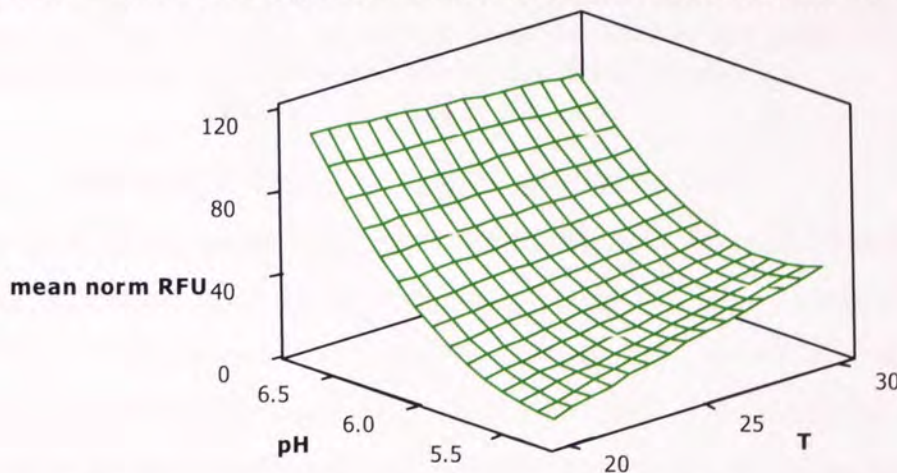


Figure 3.14. Response surface wire-frame plot for DoE1 model showing normalised yield (RFU/OD₅₉₅) changing with temperature and pH. Where T = temperature (°C), pH = pH. DO = dissolved oxygen tension (%). DO hold value is the “+1” value of 80 %.

3.2.6 Validating the model on a small scale

Although the predictive capability of the DoE1 model had been tested statistically (p -value = 0.002), experimental validation was required to show the predicted yields were achievable. The Micro 24 bioreactor was therefore used to run “check point” conditions to demonstrate

how well the model predicted real experimental yields. Figure 3.15 shows the resulting GFP_{uv} yields normalised to cell density. The data show that the model does predict experimental yield and the dramatic affects of varying pH are confirmed Figure 3.15.

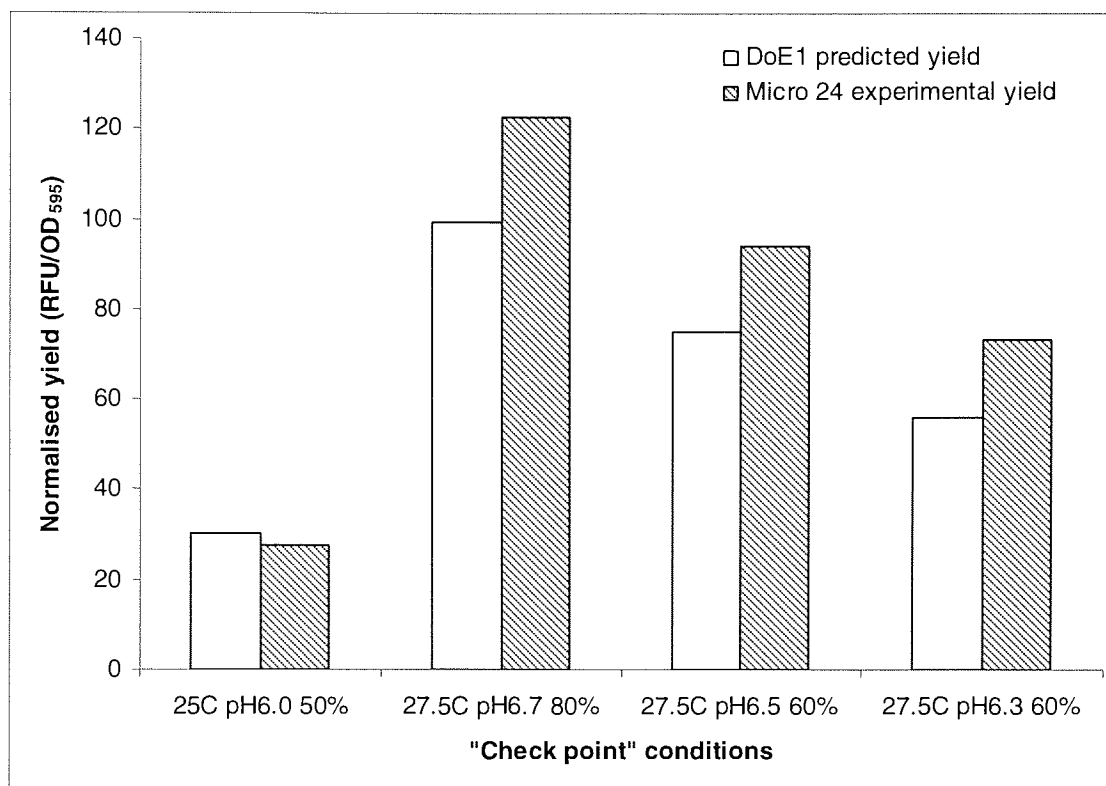


Figure 3.15. DoE1-predicted normalised yield compared with experimental normalised yield in Micro 24 cultures. Check point conditions are from within the model design space with yield normalised to cell density (RFU/OD₅₉₅).

3.2.7 Improving the predictive model: generating DoE2

The limits of the design space used to construct the DoE1 model were shown not to capture a truly optimised set of culture conditions (T, pH and DO). The DoE1 model building and validation process had shown that the methods devised would work provided the correct initial question or hypothesis was asked of the DoE. A second round of model building was therefore devised based on the Box-Behnken design with the design space moved towards the summit of the DoE1 yield 'mountain': towards lower temperature and higher pH and DO values. The -1 temperature was set at 19 °C as it was not feasible to go lower since room temperature must be at least 2 °C lower than the culture set-point. Reducing room temperature below 17 °C may have interfered with other ongoing experiments in the laboratory and it was not possible to house the apparatus in a cold room. On balance, however, it was concluded that the yield benefits at even lower temperatures would be outweighed by slow growth. Table 3.5 shows the coded and non-coded factor settings used for the second round of DoE (DoE2).

CODED FACTORS			INPUT FACTORS		
T	pH	DO	T (°C)	pH	DO (%)
-1	-1	0	19	6	60
1	-1	0	29	6	60
-1	1	0	19	8	60
1	1	0	29	8	60
-1	0	-1	19	7	30
1	0	-1	29	7	30
-1	0	1	19	7	90
1	0	1	29	7	90
0	-1	-1	24	6	30
0	1	-1	24	8	30
0	-1	1	24	6	90
0	1	1	24	8	90
0	0	0	24	7	60
0	0	0	24	7	60
0	0	0	24	7	60

Table 3.5. Experimental matrix showing coded (-1, 0, 1) and non-coded set-points required for the Box-Behnken DoE2 predictive model.

3.2.7.1 Data collection for predictive model DoE2

As with the DoE1 model, Minitab software (v15.1.30) was used to generate an experimental matrix of factor combinations required to build a predictive model to a Box-Behnken design. Table 3.6 shows the input variables and detailed measured outputs for the model building experiments which were performed in the Micro 24 parallel multi-well bioreactor. The factor combinations were carried out in at least triplicate to provide statistical confidence in the model building data. Biomass (OD_{595}) and yield (RFU) measurements were taken at 48 h post induction. The centre point conditions (24 °C, pH 7.0, 60 % DO) had 9 replicate runs (cultures) to ensure good quality data for the centre point of the model. Table 3.7 shows the RFU yield data from Table 3.6 also converted to μg . It can be seen that while culture pH had a major influence on cell density, with the highest values being achieved at pH 6.0 (mean OD_{595} 20.6, Table 3.7) and the lowest at pH 8.0 (mean OD_{595} 2.8), the effect of temperature and DO on growth were less marked. The overall relationship between yield and OD_{595} was found to fit to the line $y = 0.1305 x$ ($R^2 = 0.53$), indicating a tendency for higher yields to be associated with denser cultures (Figure 3.16).

Input Factors (Controlled on-line)			Yield (RFU)			Achieved biomass (OD ₅₉₅)			Normalised yield (RFU/OD ₅₉₅)			Full data set			Data set with outliers removed		
T (°C)	pH	DO (%)	y1	y2	y3	OD1	OD2	OD3	Norm 1	Norm 2	Norm 3	Mean norm RFU	SD	SD % of mean	Mean Norm RFU	SD	SD % of mean
19	6	60	899.70	946.73	846.06	8.90	17.19	20.30	101.09	55.08	41.68	47.24	28.71	60.77	43.71	10.51	24.04
19	6	60	456.92	802.41	523.08	20.85	23.35	17.85	21.91	34.36	29.30	109.08	26.25	24.06	123.29	12.93	10.49
19	8	60	105.95	97.03	137.14	0.80	0.85	1.70	132.43	114.15	80.67	84.46	4.30	5.09	84.46	4.30	5.09
19	7	30	923.10	1069.84	1302.21	10.85	12.10	16.30	85.08	88.42	79.89	73.03	6.88	9.42	73.03	6.88	9.42
19	7	90	1201.63	505.76	1070.43	15.15	7.70	14.45	79.32	65.68	74.08						
24	6	30	899.68	900.15	618.41	25.35	24.75	22.95	35.49	36.37	26.95						
24	6	30	781.60	941.22	1032.11	13.60	9.90	13.40	57.47	95.07	77.02	54.73	26.87	49.11	32.94	5.21	15.81
24	6	90	1235.85	1137.80	1111.65	12.40	14.35	12.25	99.67	79.29	90.75						
24	6	90	1336.48	713.51	1153.53	21.25	23.70	20.80	62.89	30.11	55.46	69.69	25.49	36.58	77.61	18.50	23.84
24	8	30	182.77	147.25	156.73	4.80	4.80	4.80	38.08	30.68	32.65						
24	8	30	124.15	171.19	487.66	4.28	4.84	6.56	29.01	35.37	74.34						
24	8	90	191.56	140.28	199.27	1.30	1.30	1.30	147.35	107.91	153.28	40.02	17.12	42.79	33.16	3.63	10.95
24	7	60	720.98	2756.23	3332.21	6.60	14.35	14.65	109.24	192.07	227.45	136.18	24.66	18.11	150.32	4.19	2.79
24	7	60	2284.01	369.61	2366.47	14.50	4.30	13.85	157.52	85.96	170.86						
24	7	60	2298.43	2397.34	2138.59	17.25	16.35	11.30	133.24	146.63	189.26						
24	7	60	2217.92	2282.81	2292.11	24.10	23.50	23.25	92.03	97.14	98.59						
24	7	60	2312.09	2209.96		23.10	20.25		100.09	109.13		136.37	44.91	32.93	127.61	33.81	26.50
29	7	30	2502.21	2576.22	2183.73	24.60	25.05	25.20	101.72	102.84	86.66	97.07	9.04	9.31	102.28	0.80	0.78
29	8	60	145.18	245.01	154.99	3.98	4.00	3.90	36.48	61.25	39.74						
29	8	60	135.59	140.25	160.60	4.38	4.32	5.36	30.96	32.47	29.96						
29	8	60	111.34			4.26			26.14			36.71	11.69	31.84	32.62	4.85	14.86
29	6	60	959.74	1109.98	270.20	19.75	23.60	19.00	48.59	47.03	14.22	36.62	19.41	53.01	47.81	1.10	2.31
29	7	90	1505.73	1219.82	2522.97	13.50	11.80	19.95	111.54	103.37	126.46	113.79	11.71	10.29	113.79	11.71	10.29

Table 3.6. Specification of the input factors and measurable outputs for the DoE2 model building experiments. The input factors were temperature (T), pH and dissolved oxygen tension (DO). Normalised yield (RFU/OD₅₉₅) and achieved biomass (OD₅₉₅), were measured in at least triplicate, 48 h post induction. The standard deviation (SD; n 3) is given for the yield normalised to OD₅₉₅.

INPUT FACTORS (controlled on-line)				MEASURABLE OUTPUTS (measured off-line)							
T (°C)	pH	DO (%)	Mean OD595nm	Mean RFU	Mean Norm RFU	Mean Norm Yield (µg/mL·OD595nm)	STDEV (µg/mL·OD595nm)	Total GFP yield (µg)	Mean RFU/mL	Norm RFU/mL·OD	Norm yield (ng/mL·OD)
19.0	6.0	60.0	20.28	865.06	43.71	0.013	0.003	1.595	8650.64	426.58	127.97
19.0	8.0	60.0	0.83	101.49	123.29	0.037	0.004	0.183	1014.85	1230.12	369.04
19.0	7.0	30.0	13.08	1098.39	84.46	0.025	0.001	1.989	10983.86	839.53	251.86
19.0	7.0	90.0	12.43	925.94	73.03	0.022	0.002	1.634	9259.39	744.72	223.42
24.0	6.0	30.0	24.35	806.08	32.94	0.010	0.002	1.444	8060.80	331.04	99.31
24.0	6.0	90.0	16.21	1195.06	77.61	0.023	0.006	2.265	11950.63	737.24	221.17
24.0	8.0	30.0	4.70	156.42	33.16	0.010	0.001	0.281	1564.18	332.52	99.76
24.0	8.0	90.0	1.30	195.41	150.32	0.045	0.001	0.352	1954.12	1503.17	450.95
24.0	7.0	60.0	17.64	2138.25	127.61	0.038	0.010	4.052	21382.45	1212.09	363.63
29.0	7.0	30.0	24.83	2539.22	102.28	0.031	0.000	4.570	25392.16	1022.85	306.85
29.0	8.0	60.0	4.37	141.33	32.62	0.010	0.001	0.256	1413.26	323.65	97.09
29.0	6.0	60.0	21.68	1034.86	47.81	0.014	0.000	1.865	10348.59	477.44	143.23
29.0	7.0	90.0	15.08	1749.51	113.79	0.034	0.004	3.089	17495.07	1159.89	347.97

Table 3.7. Input factors for the DoE2 model with resulting normalised GFP_{uv} yields converted from fluorescence values (RFU) to amount of GFP_{uv} protein (µg).

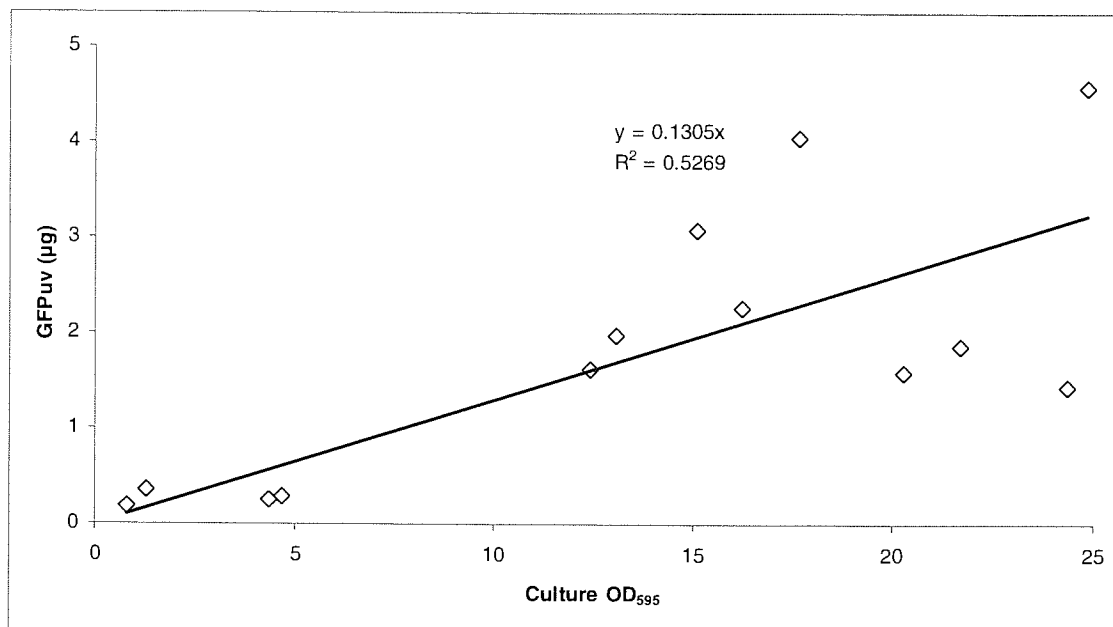


Figure 3.16. Relationship between biomass (OD₅₉₅) and GFP_{uv} yield (µg) for the cultures in Table 3.7. The line of fit is described by the equation $y = 0.1305x$ with an R^2 of 0.53.

3.2.7.2 Building a predictive model for yield (DoE2)

The normalised yield data where outliers had been removed, as shown in Table 3.6, was used to build a predictive model of normalised yield as a function of process conditions (T, pH and DO). The model building process was exactly as for DoE1. The model refinement process for DoE2 was more extensive than that used for DoE1, as the DoE2 design space was found to include optimal factor settings. Chapter 2 describes the initial refinement process which was then extended to improve the quality of the DoE2 model. The additional refinement steps were based around how the adjusted R^2 value (R^2_{adj}) for the regression changed as terms were removed. R^2_{adj} is a modification of R^2 that adjusts for the number of terms in the model, R^2_{adj} increases if the new model terms improve the model. Terms were removed from the model based on their p -values (highest first) giving R^2_{adj} values of -0.16 (full model), 0.115 (1 term removed), 0.274 (2 terms removed), 0.324 (3 terms removed) and 0.292 (4 terms removed). It was found that when 3 terms were removed from the full model, the R^2_{adj} was maximised. The resulting refined model was named “DoE2” and is described by Equation 3.2, which shows the terms and coefficients for this model. In this model, the yield was expressed in µg/OD₅₉₅ rather than RFU/OD₅₉₅ because it was felt that this was a more widely useful unit.

$$\text{Yield } (\mu\text{g}/\text{OD}_{595}) = (-21814.9 + (328.6 \times T) + (5502.1 \times \text{pH}) - (37.8 \times \text{DO}) - (325.6 \times \text{pH}^2) - (47.9 \times T \times \text{pH}) + (6.4 \times \text{pH} \times \text{DO})) \times \gamma,$$

Equation 3.2. Refined DoE2 predictive model, where; T = temperature (°C), DO = dissolved oxygen tension (%) and $\gamma = 3.0 \times 10^{-7}$ and is the conversion factor from total RFU to μg of protein.

The model was examined statistically by ANOVA (Table 3.8). It can be seen that the *p*-value associated with the regression *F*-statistic (1.96 %) is 0.217 this shows some evidence (78.3 %) against the null hypothesis (no linear relationship between any of the predictors and Y)¹¹⁷ therefore, there is some evidence of a relationship between predicted and experimental GFP_{uv} yields. The *p*-values for the DoE2 model were poorer than those seen from the DoE1 model. It was expected that the impact of this on the predictive capability of the DoE2 model would be negative, but this required testing experimentally with “check point” conditions, to confirm this.

During the model building process we noted that terms including DO had higher *p*-values (increased probability of not being significant) than those terms without. This suggested that of the factors investigated here, DO had the least influence on normalised yield.

Source	DF	SS	MS	<i>F</i> statistic	<i>p</i> value
regression	6	1288405	214734	1.96	0.217
linear	3	586988	223083	2.04	0.21
square	1	326196	326196	2.98	0.135
interaction	2	375221	187610	1.71	0.258
residual	6	657203	109534		
total	12	1945608			

DF = degree of freedom
SS = sum of squares
MS = mean square

Table 3.8. ANOVA results for the DoE2 model. Where; DF = degrees of freedom, SS = sum of squares, MS = mean square.

The response optimiser function of the Minitab software was used again to predict the optimal conditions for GFP yield on a per cell basis (yield normalised to OD₅₉₅). Figure 3.17 shows the output from the response optimiser indicating that the factor settings 20.2 °C, pH 7.76 and 90 % DO should be the highest yielding set if process conditions within the DoE2 design space. The lower panels in Figure 3.17 show how the individual factors influence yield over the design space: higher yields are obtained at lower temperatures, with 20.2 °C being the

optimal temperature. Yield also increases with culture pH from pH 6.0 to pH 7.76 (optimal setting) then gradually decreases to the high extreme “+1” (pH 8.0).

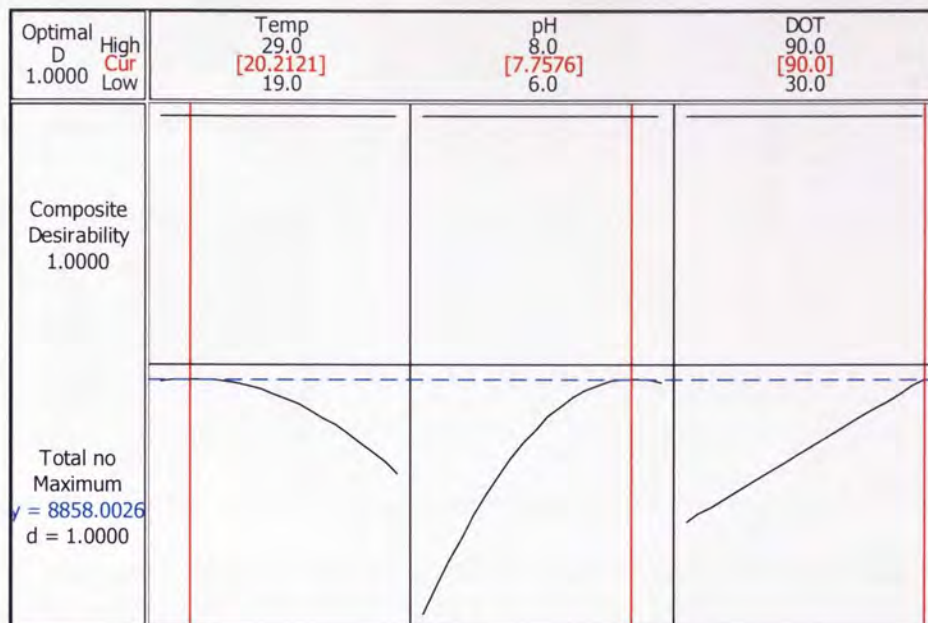


Figure 3.17. Response optimisation for DoE2 as predicted by Minitab software. Suggested optimal conditions are 20.2 °C, pH 7.76 and 90 % DO.

The DoE2 predicted optimal conditions used (21.5 °C, pH 7.6 and 90 % DO) differ from those in Figure 3.17 (20.2 °C, pH 7.76 and 90 % DO). This is due to further refinement of the DoE2 model since the experiments were performed. A rapid turn-around was required during the experimental period and time was not available for extensive model refinement. The curves in the bottom panel of Figure 3.17 suggest further increases in GFP_{uv} expression could be achieved using the optimal set-points described by the fully refined DoE2 model (20.2 °C, pH 7.76 and 90 % DO).

The model was plotted visually using response surface regression to demonstrate how the factor interactions affected yield. Figure 3.18 shows contour plots of each of the factors, with the hold values set at the “0” mid-points. Figure 3.19 shows a wire frame plot of yield against culture temperature (T) and pH with DO held at 90 % (the “+1” value). It can be seen that the interaction of temperature and pH has a major influence on GFP yield.

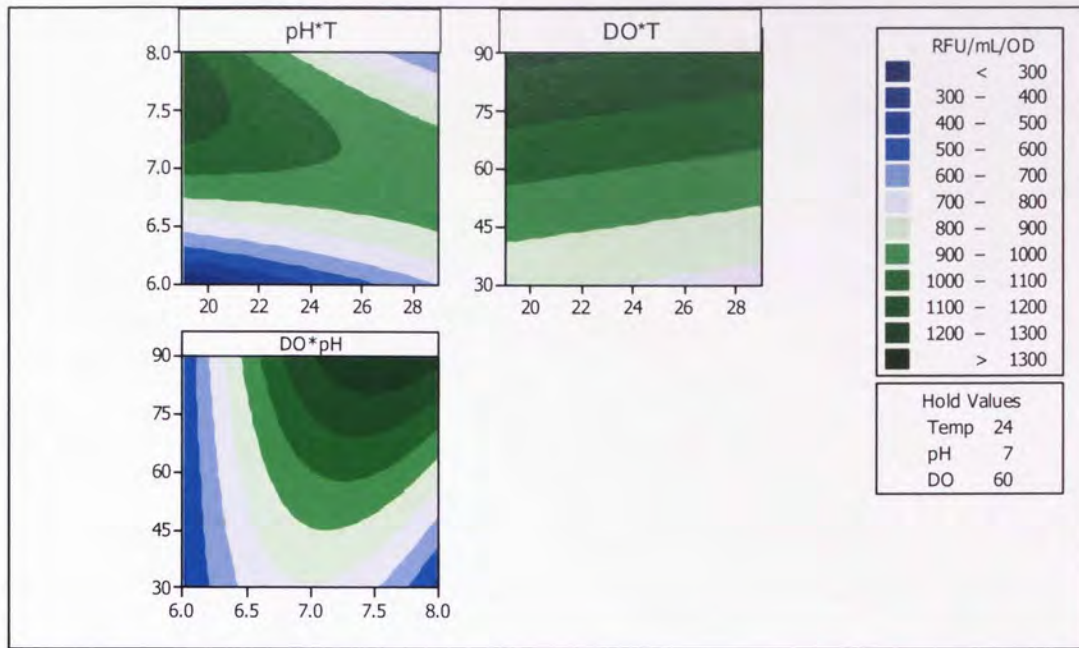


Figure 3.18. Response surface contour plot for DoE2 model showing normalised yield (RFU) changing with each of the factors. Where T = temperature ($^{\circ}\text{C}$), pH = pH and DO = dissolved oxygen tension (%). All hold values are “0” mid-point values.

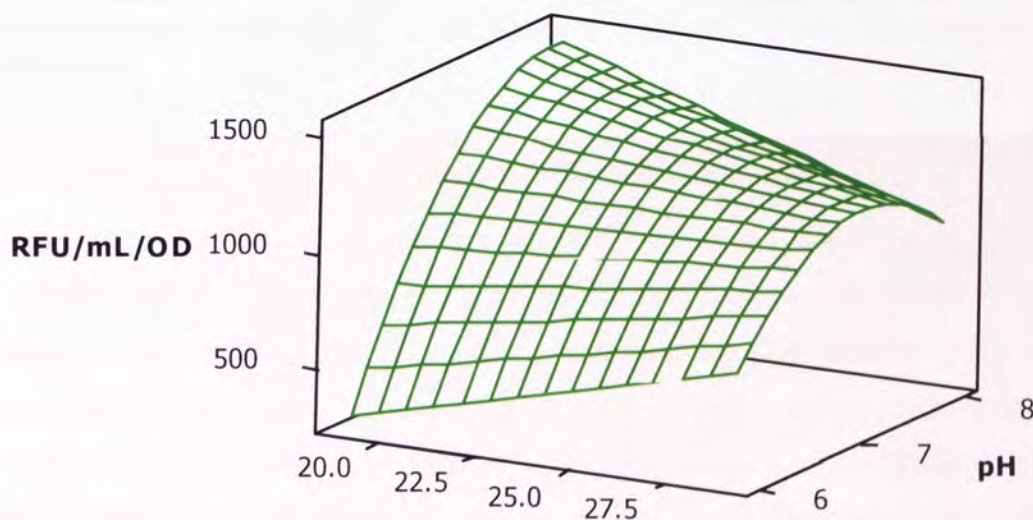


Figure 3.19. Response surface wire-frame plot for DoE2 model showing normalised yield (RFU) changing with temperature and pH. Where T = temperature ($^{\circ}\text{C}$), pH = pH. DO = dissolved oxygen tension (%). The DO hold value is the “+1” value of 90 %.

Figures 3.17 and 3.18 show that the DoE2 design space captures the optimised set of conditions for maximal protein yield. Figure 3.19 shows how DoE2 and DoE1 fit together to provide a composite view of how yield changes with culture temperature and pH. The optimised conditions can be seen on these response surface plots.

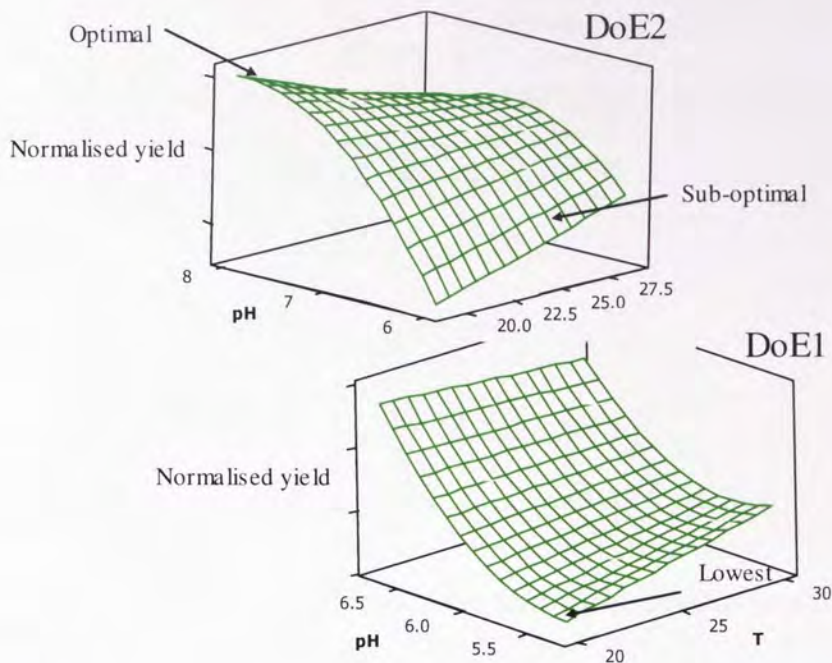


Figure 3.19. Response surface wire-frame plots for DoE1 and DoE2 showing how the two design spaces fit together demonstrating the full range of yield conditions. Arrows indicate features of the design space: optimal yielding, sub-optimal yield and lowest yielding conditions. Where T = temperature (°C), pH = pH, DO = dissolved oxygen tension (%).

3.2.7.3 Testing the DoE2 model at small scale (check points)

The predictive capability of the DoE2 model had been tested statistically (p -value = 0.217), but experimental validation was required, as for DoE1. Table 3.9 shows the check point conditions and resulting GFP_{uv} yields (normalised to cell density) from small scale experiments in the Micro 24. Figure 3.20 shows the DoE2 check point data as a bar graph, with error bars representing standard deviation (SD) where $n = 3$. Within error, the DoE2-predicted yields give a fair indication of expected yield on a small scale. The DoE2-predicted optimal conditions (20.2 °C, pH 7.76 and 90 % DO) were run as 20 °C, pH 7.7 and 80 % DO and gave the highest experimental yield of those conditions tested.

T (°C)	pH	DO (%)	DoE2 predicted yield (ng/OD ₅₉₅)	Micro 24 experimental yield (ng/OD ₅₉₅)
20	7.5	60	2058.39	1529.33
20	7.7	80	2408.61	2223.45
27	8	50	1071.36	865.10
28	7.5	90	2155.23	1610.99
28	6	80	1203.12	1042.06
23.6	7.25	60	2120.00	2062.35
27.5	6.7	80	1990.96	2207.22
27.5	6.5	60	1711.26	1692.87
27.5	6.3	60	1546.07	1323.05
21.5	7.6	20	1356.07	1082.17
21.5	7.6	40	1719.67	1327.46
21.5	7.6	60	2083.27	1683.97

Table 3.9. Factor settings (T, pH and DO) and yield results (ng/OD₅₉₅) for the check point conditions used to validate the DoE2 model in the Micro 24. These factor settings were not used in the model building process.

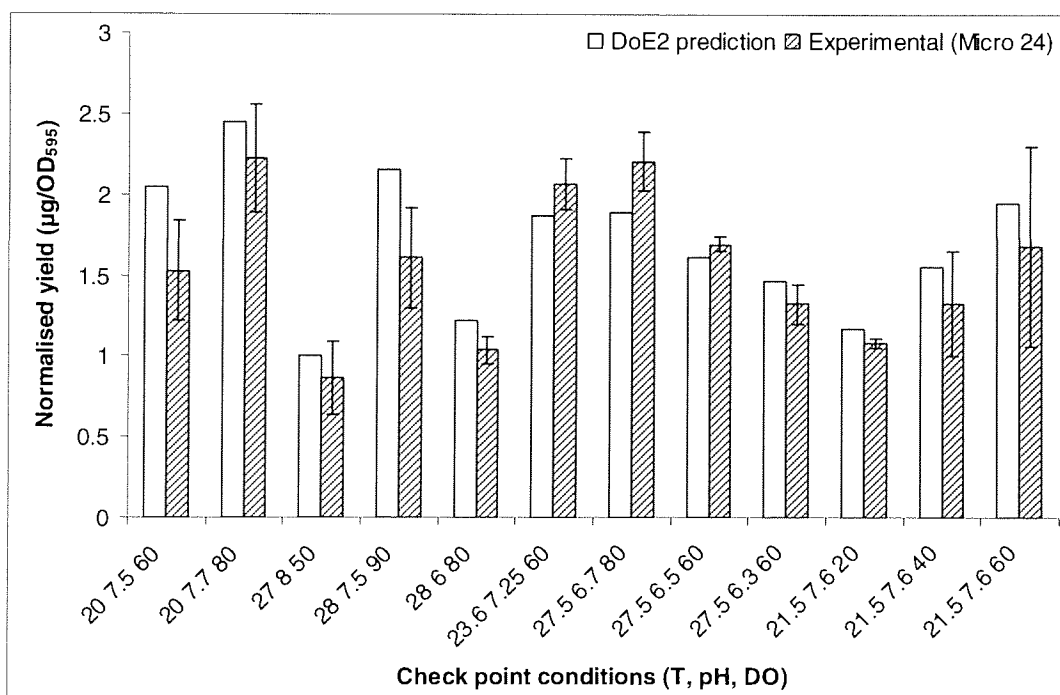


Figure 3.20. Normalised GFP yield (µg/OD₅₉₅) as predicted by the DoE2 model. Error bars represent standard deviation where n = 3.

Figure 3.21 shows a scatter plot of DoE2-predicted and Micro 24 experimental yields.

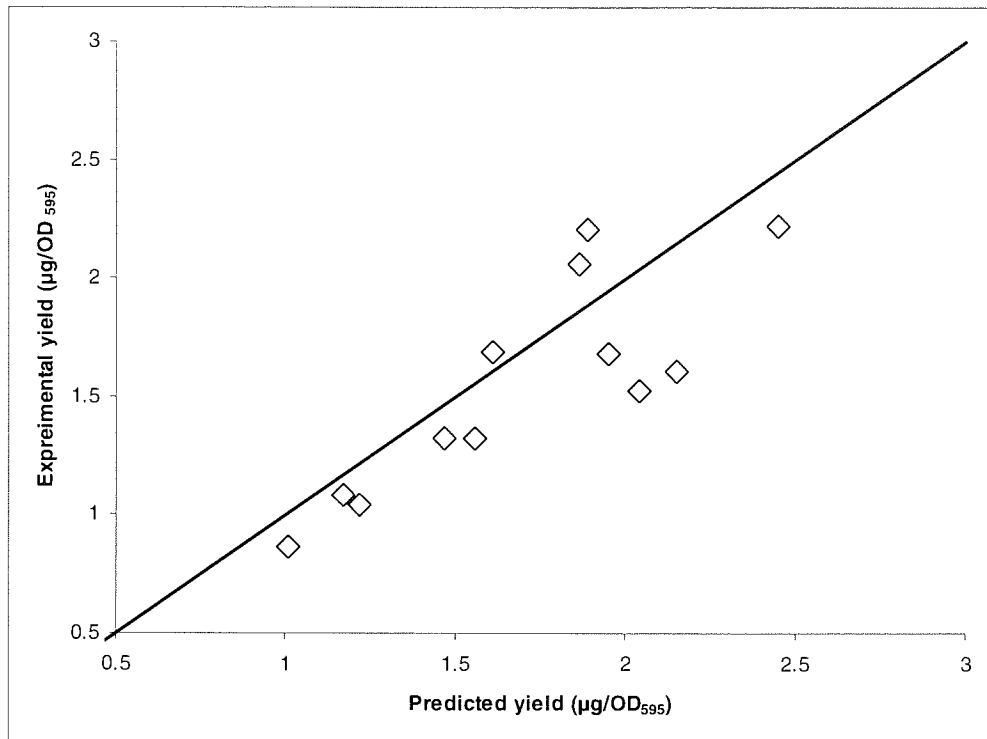


Figure 3.21. Relationship between predicted and experimental normalised yields from Micro 24 cultures from Table 3.9 and Figure 3.20. Check point conditions are from within the model design space with yield normalised to cell density (ng/OD₅₉₅). The line $y=x$ illustrates the spread of the predicted and experimental check points compared to the ideal situation where predicted yield = experimental yield.

As Figure 3.20 demonstrates, the DoE2 model provided a reasonable prediction of normalised GFP yield based on varying process conditions (T, pH and DO). It was now necessary to validate the model on a larger scale to demonstrate that it was indeed scalable and hence a useful tool for large scale process optimisation.

3.3 Is the predictive model scalable?

Having shown the DoE2 model was predictive on a small-scale in the Micro 24, the model was tested on a larger scale to determine if a large-scale process could be optimised directly from a model built using small-scale data. To show if this were possible, a number of check point conditions were selected to be run in a 7 L stirred tank bioreactor (STB). The conditions selected were DoE2-predicted optimal conditions (21.5 °C, pH 7.6, 90 % DO), DoE2 centre point (24 °C, pH 7.0, 40 % DO), DoE2 best small-scale experimental fit to the model (27.5 °C, pH 6.5, 60 % DO) and standard *P. pastoris* conditions (30 °C, pH 6.0, 20 % DO).

Figure 3.22 shows an overview of a typical scale-up STB experiment. A BMGY shake flask seed culture was inoculated using a freshly-plated single colony of X33GFP and grown to an OD₅₉₅ 2-6. This was used to inoculate the 7 L vessel containing 3 L BMGY medium

equilibrated to standard growth conditions (30°C, pH 6.0, 20% DO). The culture was grown until the glycerol carbon source was exhausted and the culture had entered the starvation phase. This ensured all glycerol had been consumed by the cells as it represses protein expression by the *AOXI* promoter.¹¹⁸ The culture conditions were then gradually altered to the chosen DoE2-defined expression conditions during an adaptation phase, which was carried out over 1-3 h so the cells were not suddenly exposed to different conditions. The methanol fed-batch phase used three methanol feed rates starting with a slow feed (0.167 mL/min to a 3 L culture) to allow the cells to adapt their metabolism to methanol utilisation.¹¹⁹ The methanol addition rate was then increased to 0.417 mL/min (to a 3 L culture) and finally to 0.583 mL/minute. The final feed rate was maintained until the point of harvest at 48 h post-induction as with the Micro 24 cultures.

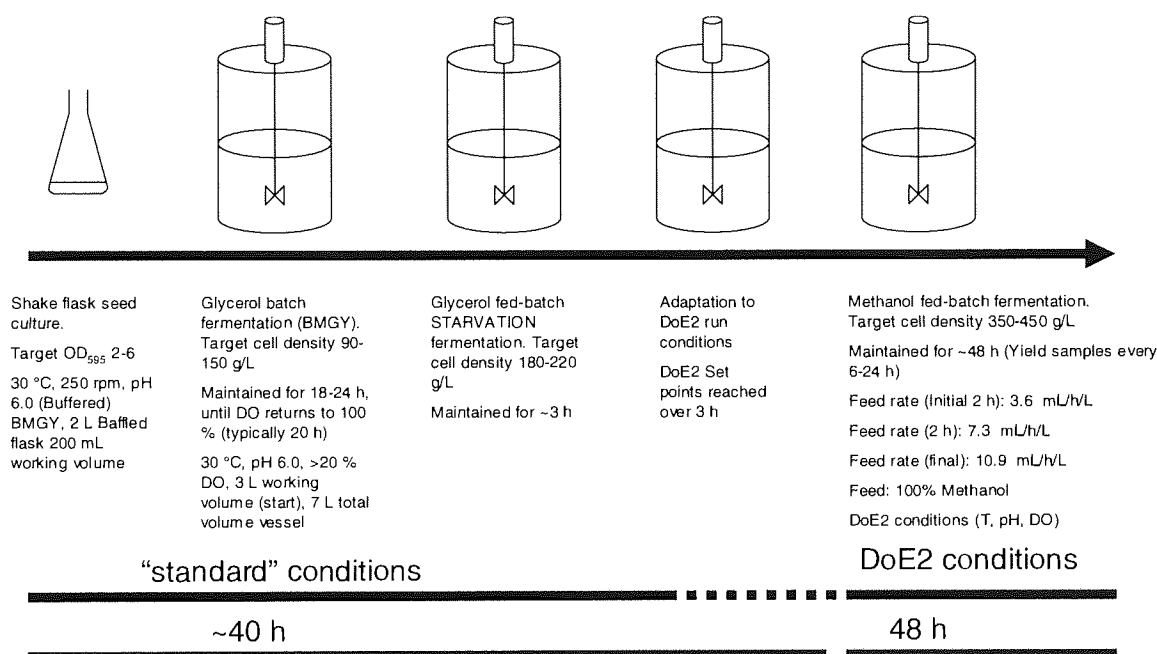


Figure 3.22. Schematic overview of a typical scale-up bioreactor experiment. The timings and expected biomass values are based on the Invitrogen *Pichia pastoris* fermentation handbook and previous experience of the X33 strain.

3.3.1 Bench-scale (7L vessel) experimental set-up

A jacketed 7 L (total volume) glass vessel with an Applikon ADI1030 proportional–integral–derivative (PID) control unit was used, as shown in Figure 3.23. The heat exchanger used to control the jacket temperature, not visible in Figure 3.23, was not controlled by the ADI1030 unit but had its own internal feed-back loop. Therefore, control of culture temperature was dependant on an operator defining a heat exchanger set-point that was appropriate for the culture set point; which was typically 2 °C higher than the culture set-point. Aeration was primarily provided by a sparger bar supplying 0.2 µm filtered laboratory air from an air pump

at 0.5-2 vvm. Pure O₂ was supplemented on-demand as required to maintain set-point under the control of the ADI1030 control unit and a feed-back loop. Control of culture pH was by the ADI1030 unit using a PID feed-back loop with addition of either 2 M HCl or 15 % NH₃ as required to maintain set-point. During methanol induction, the methanol feed was administered using a peristaltic pump (not shown in Figure 3.23) with calibrated tubing. The feed rate was achieved by “pump on time” *i.e.* number of seconds per minute that the pump would run, which was controlled by the ADI1030 unit.

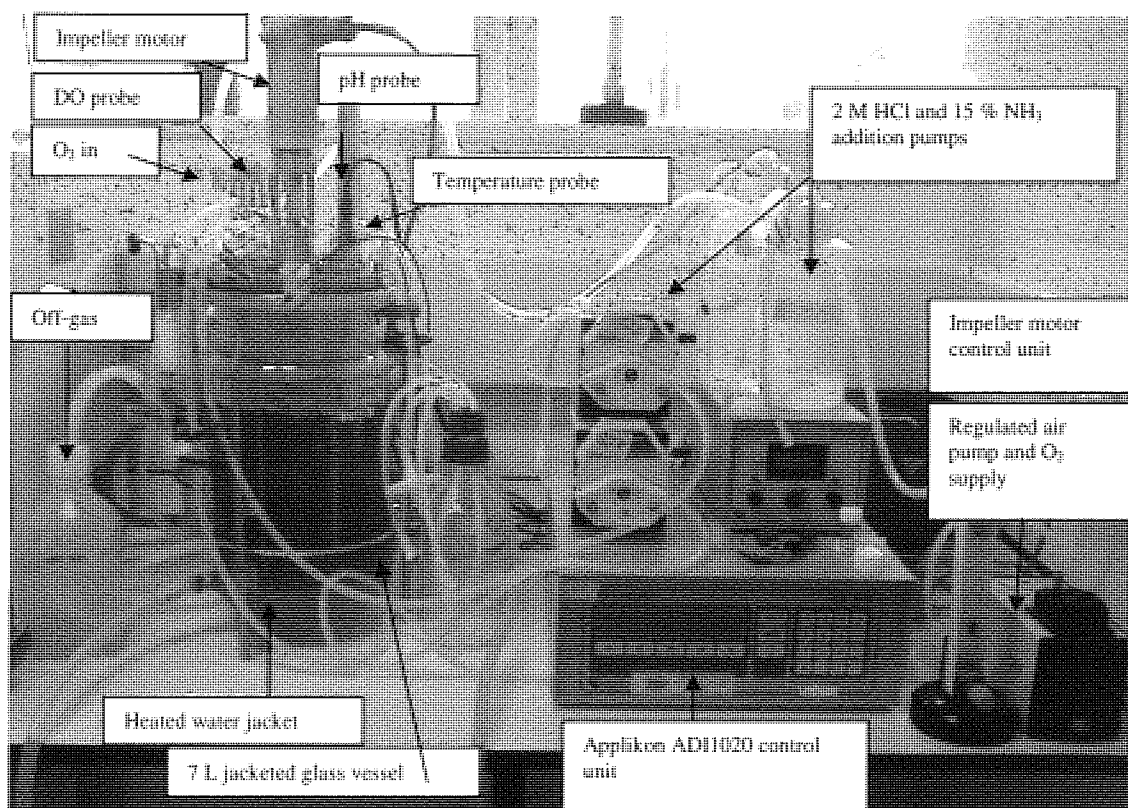


Figure 3.23. Set-up of 7 L (total volume) stirred tank bioreactor used for the *P. pastoris* X33GFP_{uv} scale-up experiments.

The medium used for the initial scale-up bioreactor runs was BMGY for the glycerol batch phase (biomass generation) to mirror the Micro 24 set-up. BSM medium, which is typically used in bioreactors in place of BMGY, also suffers from heavy precipitation when pH > 5.5¹²⁰ meaning it would have been unsuitable for these experiments. The use of a STB allows additional improvements to be made to the process not possible in the Micro 24, such as switching to a fed-batch strategy using a feed pump for nutrient or inducer additions. The induction phase was initially run using methanol induction as per the Invitrogen *Pichia pastoris* fermentation handbook using three methanol feed rates of 0.167, 0.417 and 0.583 mL/min. The sterile 100 % methanol feed was supplemented with PTM₁ trace salts to 12 mL/L.

3.3.2 Initial DoE2 scale-up bioreactor experiments

Demonstrating that the DoE2 model predicts the per cell yield from the *P. pastoris* X33GFP_{uv} strain in bioreactors was approached in the same manner as for the Micro 24. A number of check point conditions were selected: DoE2 optimised conditions (21.5 °C, pH 7.6, 90 % DO), DoE2 centre point conditions (24 °C, pH 7.0, 40 % DO), DoE2 best small-scale experimental fit conditions (27.5 °C, pH 6.5, 60 % DO) and “standard” conditions (30 °C, pH 6.0, 20 % DO). The number of check point conditions used at the large scale was lower than when testing at the small-scale in the Micro 24. As the model had already been validated at the small-scale it was thought that any failure in predictive capacity at the large scale would be identified by the widely spaced scale-up check points which spanned the DoE2 design space. The scale-up experimental runs are detailed in Table 3.10.

Each of the initial “direct transfer” experiments (f002-006) were monitored to provide culture profiles for each experimental run, with biomass (OD₅₉₅) measured off-line whilst culture temperature (°C), DO (%) and pH were on-line measurements but logged manually. Following run f005, data logging software (BioXpert Lite, Applikon Biotechnology) was available and so the overnight “data gaps” could be avoided, providing a more complete picture of the process. Figure 3.24 shows the culture profiles for each of the runs detailed in Table 3.10.

In all the runs it is possible to identify the culture phases as outlined in Figure 3.1. With the glycerol batch phase typically running from 0 to 24 hours effective fermentation time (EFT), ending when the DO spikes back to 100%. The glycerol batch phase is when biomass is generated, typically reaching OD₅₉₅ ~30. The subsequent starvation period of at least 1 h ensured all glycerol was utilised prior to induction. Starvation is followed by an adaptation period when the process conditions (temperature, pH and DO) are gradually altered from growth to expression conditions. The adaptation was generally performed over 1.5 h. Once expression process conditions were met, induction was initiated by methanol addition (indicated arrows in Figure 3.24)

Run name	Run description	Data logging (Y/N)	Glycerol batch phase conditions	Glycerol batch phase duration (hours)	Adaptation phase (Y/N)	Induction phase conditions	Induction phase harvest biomass (OD ₅₉₅)	Apparent decrease in biomass (Y/N)	Effective induction period (hours)	Maximum total GFP yield (mg/L)	Total yield at harvest (mg/L)	Normalised yield at harvest ($\mu\text{g/mL}/\text{OD}_{595}$)
f002	Standard handbook conditions (BMGY medium)	N	30 °C, pH 6.0, 20 % DO	22.7	N	30 °C, pH 6.0, 20 % DO	38.3	Y	8	34.75	29.33	0.77
f003	DoE2 centre point conditions	N	30 °C, pH 6.0, 20 % DO	24.1	Y	24 °C, pH 7.0, 40 % DO	28.6	N	30	187.86	97.38	3.40
f004	DoE2 good small scale experimental match conditions	N	30 °C, pH 6.0, 20 % DO	23.6	Y	27.5 °C, pH 6.5, 60 % DO	39.7	N	20	262.23	219.94	5.54
f005	DoE2 optimised conditions	Y	30 °C, pH 6.0, 20 % DO	22.3	Y	21.5 °C, pH 7.6, 90 % DO	23.2	N	8	90.32	46.67	2.01
f006	DoE2 good small scale experimental match conditions (repeat)	Y	30 °C, pH 6.0, 20 % DO	22.6	Y	27.5 °C, pH 6.5, 60 % DO	29.2	Y	22	75.66	56.84	1.95

Table 3.10. Initial DoE2 scale up experimental run descriptions and yield data. Note: Run f001 was terminated following the glycerol batch phase due to media component precipitation making biomass estimation problematic and DO not being under PID control.

During the glycerol batch phase of f004 biomass accumulation was two fold higher (OD₅₉₅ ~40) than expected (OD₅₉₅ ~20). This run was therefore repeated (f006) resulting in glycerol

batch phase biomass accumulation ($OD_{595} \sim 20$), being more similar to the other runs. This indicates that f004 was subject to non-standard conditions most likely medium composition allowing higher biomass accumulation.

Following the initial induction all cultures showed an increase in biomass as additional carbon source was supplied to the culture. In some cases (f002 and f006) this was followed by an apparent decrease in biomass at approximately 50 h (EFT; Figure 3.24). The apparent decrease in biomass could be due to the cells swelling rather than dividing leading to an increase in culture transmittance, therefore a decrease in optical density. Alternative explanations would be culture dilution due to the methanol feed or the toxic affects of methanol accumulation leading to cell death and subsequent lysis, which is supported by corresponding increases in DO.

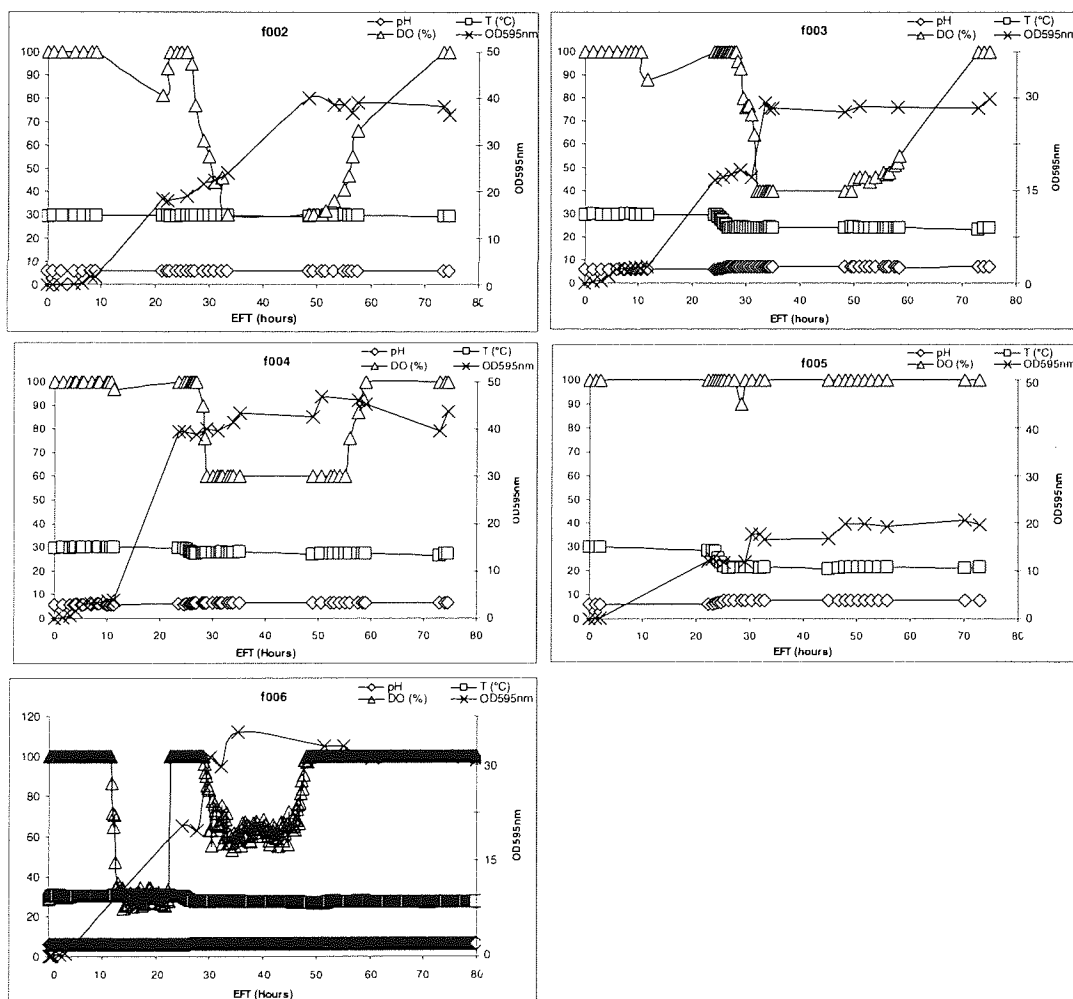


Figure 3.24. Initial DoE2 scale up culture profiles for fed-batch *P. pastoris* X33GFP_{uv} experiments, using the conditions indicated by Table 3.10. Where culture temperature = T (°C), dissolved oxygen = DO (%) and biomass = OD₅₉₅. Culture biomass is shown on the secondary y-axis. Point of induction is indicated by an arrow.

Figure 3.25 shows the GFP_{uv} yields (accumulated GFP_{uv}) from the bioreactor runs (f002 – f006). The GFP_{uv} yield increases from 0 to typically 10 h post-induction for all runs. The post induction period where GFP_{uv} yield is increasing can be thought as the effective induction period. Table 3.10 shows that the effective induction period lasts for 8-30 hours, depending on the expression conditions.

The GFP_{uv} yield profiles (Figure 3.25) can be linked to the effects of methanol accumulation seen in the culture profiles (Figure 3.24). The effective induction periods end approximately when the DO trace increases to 100 %. Following the effective induction period the GFP_{uv} detectable in culture supernatants decreases, this can be linked to methanol accumulation, cell lysis and the release of proteases from the dead cells.

There is no major decrease in yield in f004, as seen in the other scale-up runs; this is most likely associated with the higher pre-induction biomass allowing the cells in culture to cope with the methanol feeding rates used. This increase in pre-induction biomass could be used as an option for further optimising the process.

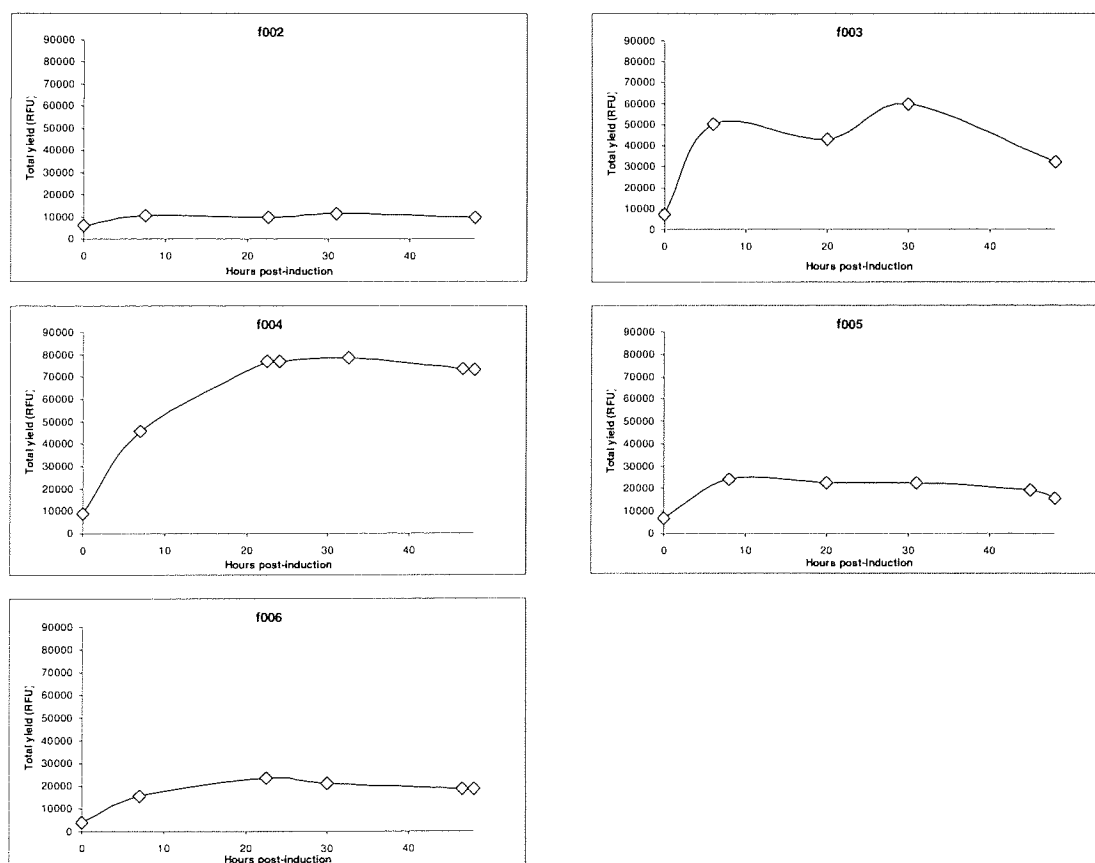


Figure 3.25. GFP yield data from fed-batch *P. pastoris* X33GFP_{uv} DoE2 scale-up experiments using conditions described in Table 3.10. Culture supernatant samples were assayed for GFP_{uv} fluorescence (RFU) at time points during the induction phase of the experiments. Total GFP yield (RFU) profile is shown for each fermentation.

3.3.2.1 DoE2 scale-up conclusions from initial runs

The DoE2 scale-up process aimed to prove that a model built using data from small-scale (6 mL) cultures could be directly used to optimise culture conditions with a five hundred fold increase in volume (3 L). Figure 3.26 demonstrates the predictive capability of the DoE2 model for normalised GFP_{uv} yield (ng/mL/OD₅₉₅) at both small (Micro 24) and large (7 L Bioreactor) scales. Where conditions used in the Micro 24 differ from those used in the bioreactor, the most similar Micro 24 conditions are substituted.

The yields obtained from the large scale experiments are higher than predicted by the DoE2 model but do fit the trend indicated by the small scale experimental data. As the large scale experiments were not repeated, with the exception of f004 (as f006), it is possible that these results may not be representative of the culture conditions (set points). Another possibility is that the small scale system is not completely equivalent to the large scale system. Differences do exist between the two systems (Micro 24 and 7 L bioreactor) notably mode of agitation and pH control, which could account for the poor predictive power of the model. The biggest difference, however, was the induction regime. It was noted that there was scope for further optimisation of the large-scale process, which was not possible at the small-scale due to limitations of the equipment *e.g.* no provision for on-line liquid additions to the cultures.

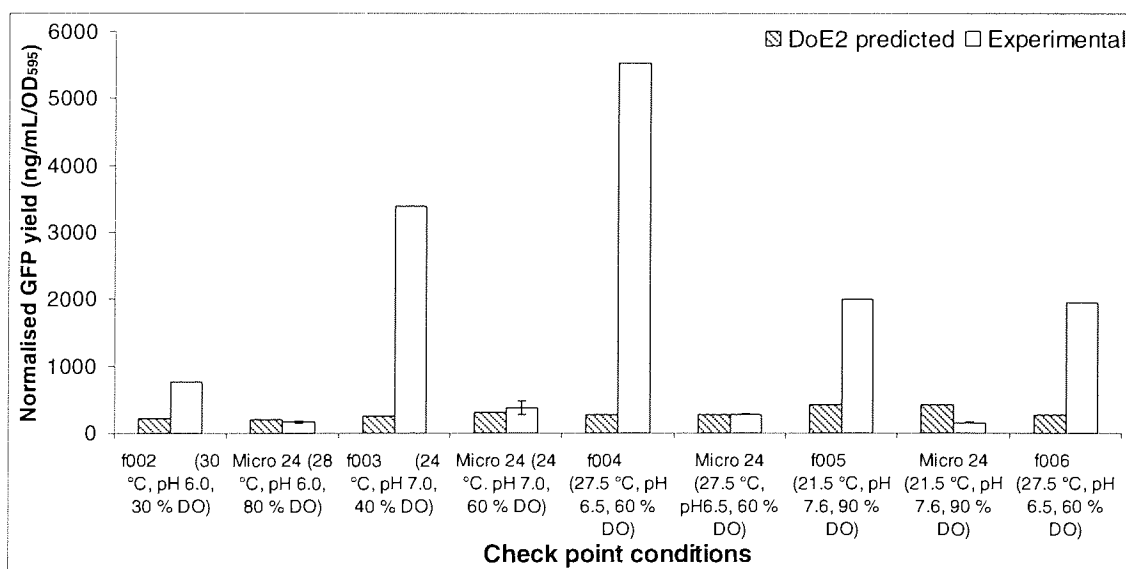


Figure 3.26. Comparison of normalised yield (ng/mL/OD₅₉₅) data at 48 h post-induction for DoE2 model predictions, small-scale Micro 24 experimental runs and large-scale (7 L Bioreactor) experimental runs. Error bars represent standard deviation where n = 3.

The large scale runs highlighted the importance of optimising the rate of methanol addition used during the induction phase of the cultures. The feed rates were not well matched to the metabolic activity of the X33GFP_{uv} cells at non-standard (DoE2) induction conditions and lead to methanol accumulation at ~8-30 h post-induction. This resulted in sub-optimal protein yields as the “useful” induction/expression period was limited to 8-30 h and not the full 48 h. This was especially noticeable when using the DoE2 predicted optimal conditions (21.5 °C, pH 7.6, 90 % DO). However, despite the short period of “useful” expression in f005 (DoE2 predicted optimal conditions) the total yield was higher than that achieved using the standard Invitrogen handbook conditions in f002 (See Figure 3.27), which had a longer effective induction period.

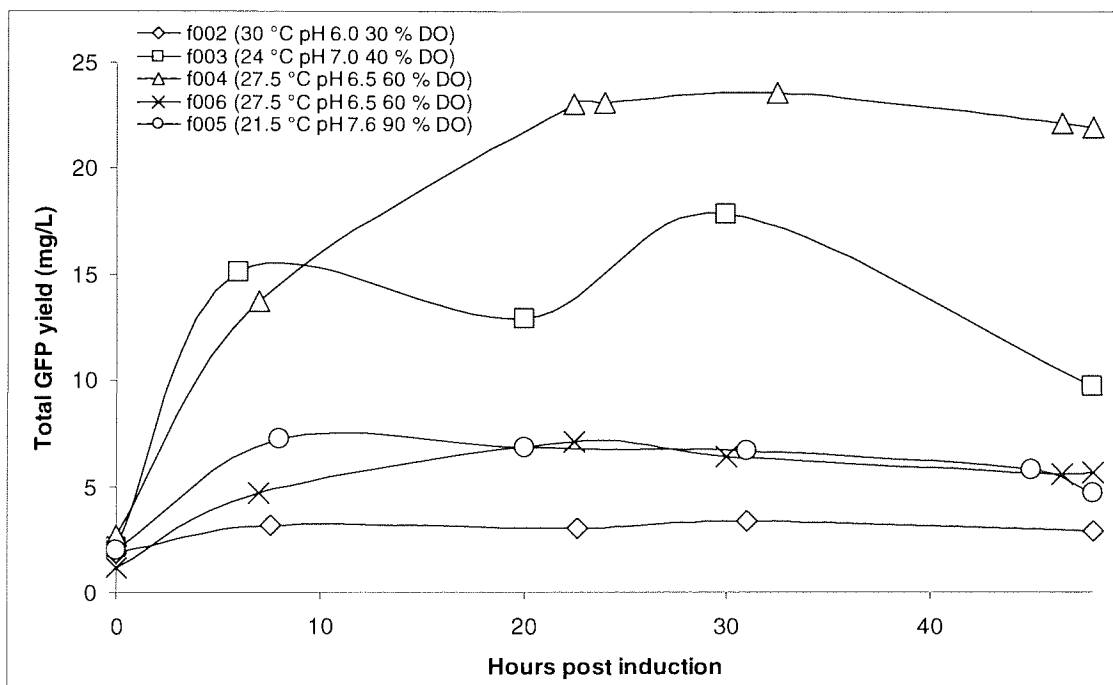


Figure 3.27. Total GFP_{uv} yields (RFU) profiles for the initial DoE2 scale-up bioreactor runs. Culture supernatant samples were assayed for GFP_{uv} fluorescence (RFU) during the 48 h induction phase of the cultures.

Figure 3.27 also demonstrates the influence of increasing pre-induction biomass on GFP_{uv} yield, as experienced in f004 (DoE2 best small scale fit conditions) where GFP_{uv} yield was higher than expected. In this case the methanol feeding rate during induction did not lead to rapid methanol accumulation and toxic effects. It is thought that the methanol utilisation rate of the higher density culture was better matched to the methanol addition (feed) rate.

As all large-scale experiments suffered from methanol accumulation, to some extent, the “standard” large-scale process requires altering to avoid methanol toxicity so increasing yield. Two improvements were suggested by the initial experiments; change the induction protocol

(*i.e.* methanol feed) and increase the pre-induction biomass. These alterations can be used in addition to the optimised process conditions (temperature, pH and DO) predicted by the small-scale modelling *e.g.* DoE2 optimised conditions (21.5 °C, pH 7.6, 90 % DO).

3.4 Further improvements are possible to the DoE modelled process

The DoE model building process provided an optimised set of process conditions (21.5 °C, pH 7.6, 90 % DO) for maximising per cell GFP yield from *P. pastoris* X33GFP. During the initial scale-up bioreactor runs it was noted that the fermentation protocol used in the small-scale Micro 24 experiments could be improved upon for use in a traditional bioreactor. The toxic effects of methanol accumulation during the induction phase of the runs limited the duration of effective induction period; typically 8-30 h of GFP_{uv} expression achieved from a possible 48 h. Two approaches were identified to avoid methanol accumulation; increase pre-induction biomass and use a mixed feed for the induction. Increasing the pre-induction biomass provides more cells to utilise the methanol feed (at the original addition rate), although ideally the utilisation rate should be balanced by the addition rate. Using a mixed induction feed of a non-repressive carbon source (sorbitol) and the inducer (methanol) might be effective in two ways. Firstly the concentration of methanol supplied to the culture is reduced due to dilution with sorbitol. Secondly the sorbitol provides an additional/alternative carbon source for the cells to utilise.

3.4.1 Increasing pre-induction biomass with DoE2 optimised process conditions

The DoE2 predicted optimal conditions for per cell (normalised) GFP_{uv} yield had been shown to fit small and large scale experimental data. The fermentation run f004 had indicated that higher total or “volumetric” yield was obtainable by increasing the number of cells (biomass) present prior to induction with methanol. To harness this for increasing total yield, medium composition was identified as an area for further investigation. The medium used in the Micro 24 experiments and initial bioreactor runs was BMGY; this medium was used as the standard fermentation medium, basal salts medium (BSM), suffers from precipitation at pH >5.5. The pH set-points defined by the model design space exceeded pH 5.5 and would result in heavy precipitation.¹²⁰ The carbon source used in both formulations is glycerol (represses the *AOXI* promoter) but the glycerol concentrations differ at 10 g/L (BMGY) and 40 g/L (BSM). The high concentration found in BSM can support higher biomass accumulation than BMGY; such cell densities require a high oxygen supply, which is feasible in a bioreactor but not in shake-flask culture. Once the process had been moved to larger-scale (7 L) bioreactors it

would be possible to sustain high pre-induction cell densities using a modified BMGY medium. The glycerol concentration of BSM (40 g/L) was adopted for use in BMGY; f007 utilised the modified BMGY (40 g/L glycerol) medium for the glycerol batch-phase to see if the “useful” induction period could be extended by increasing pre-induction biomass.

Figure 3.28 shows the f007 “DoE2 predicted optimal induction conditions” culture profile with biomass (OD_{595}) measured off-line whilst culture temperature ($^{\circ}C$), DO (%) and pH were on-line measurements with data-logging by Applikon’s BioXpert Lite software. It is possible to identify the culture phases as outlined in Figure 1. The glycerol batch phase runs from 0 to 25.6 h EFT ending when the DO spikes back to 100 %, achieving $OD_{595} \sim 40$. This was an approximate two fold increase in biomass achieved during the glycerol batch phase over standard BMGY with 10 g/L glycerol. As the expression conditions differed from “standard” conditions used for the growth phase there was a requirement for an adaptation phase $\sim 26.5 - 28$ h (EFT). During the induction phase, methanol was fed at $\sim 0.2 - 0.4$ mL/min and the biomass remained fairly constant. This plateau was maintained until harvest at 48 h post-induction (78 hours EFT).

During induction, DO was seen to fluctuate during early-mid induction (< 60 h EFT). The drop seen at ~ 35 h (EFT) was attributed to cellular utilisation of methanol and the subsequent increase (to 100 % DO) due to methanol accumulation. During induction, the methanol feed rate was varied in an attempt to match the rate of addition to utilisation; the DO drops seen at ~ 50 and 55 h (EFT) were due to reduction of the feed rate. When the DO reached 100 % (49 h EFT), the feed was stopped until the DO was seen to drop to 60-80 %. This was seen as evidence that the ~ 0.4 mL/minute feed rate was supplying methanol at a rate above that of utilisation. The feed was therefore set at ~ 0.2 mL/min for the remainder of the run. The DO remained below 100 % for ~ 10 h, but it returned to 100 % during the over-night period from 56 h (EFT) when there was no operator monitoring. This indicates that the combined methanol addition had surpassed methanol utilisation resulting in methanol accumulation.

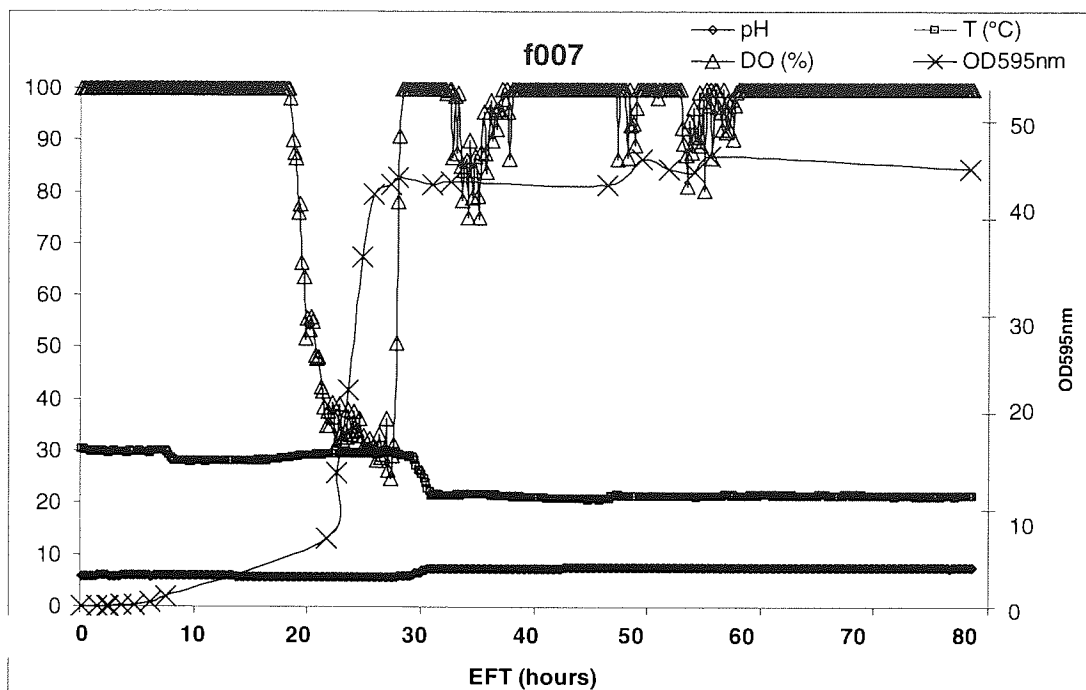


Figure 3.28 Culture profile for fed-batch *P. pastoris* X33GFP_{uv} DoE2 scale-up experiment (f007) using “DoE2 predicted optimal” conditions (21.5 °C, pH 7.6, 90 % DO) with 40 g/L glycerol in batch phase, where culture temperature = T (°C), dissolved oxygen = DO (%) and biomass = OD₅₉₅. Culture biomass is shown on the secondary y-axis.

Figure 3.29 shows the GFP_{uv} yield (accumulated GFP_{uv}) from the “DoE2 predicted optimal” condition bioreactor run with 40 g/L glycerol batch phase (f007). The total GFP_{uv} yield increased from 0 to ~15 h post-induction followed by a slight decline until harvest (48 h post-induction). This yield profile can be linked to the effects of methanol accumulation leading to a halt in GFP_{uv} expression and increased cell death/lysis leading to proteolysis. Total yield reached a maximum of ~55000 RFU at ~ 15 h post-induction and then decreased until harvest at 48 h post-induction, suggesting the early part of induction was the only productive stage and no further GFP_{uv} was expressed by the culture after ~15 h post-induction.

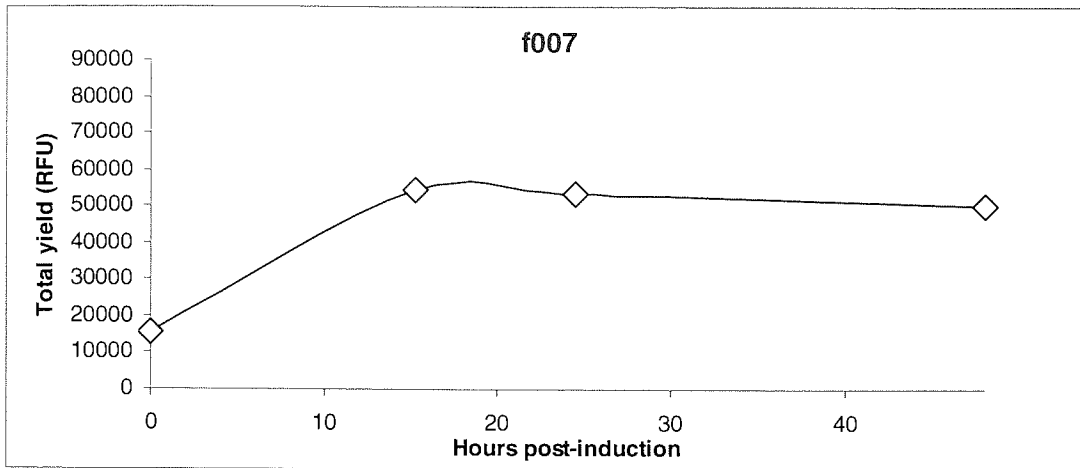


Figure 3.29 GFP_{uv} yield data from (f007) fed-batch *P. pastoris* X33GFP_{uv} DoE2 scale-up experiment using “DoE2 predicted optimal” conditions (21.5 °C, pH 7.6, 90 % DO) with 40 g/L glycerol in batch phase. Culture supernatant samples were assayed for GFP fluorescence (RFU) at 0, 15.3, 24.5 and 48 h post-induction. Total GFP yield (RFU) accumulation is shown over the time course for the fermentation.

Figure 3.30 shows how using the modified BMGY medium lead to an approximate two fold increase in total GFP yield when using the DoE2 predicted optimal conditions with standard (10 g/L) glycerol BMGY and five fold increase when compared to standard process conditions and media. The effective induction period was still curtailed by cell death/stress due to methanol accumulation, probably a poorly-balanced methanol addition rate. It was therefore concluded that further development of the induction strategy would be required to eradicate the limited expression period.

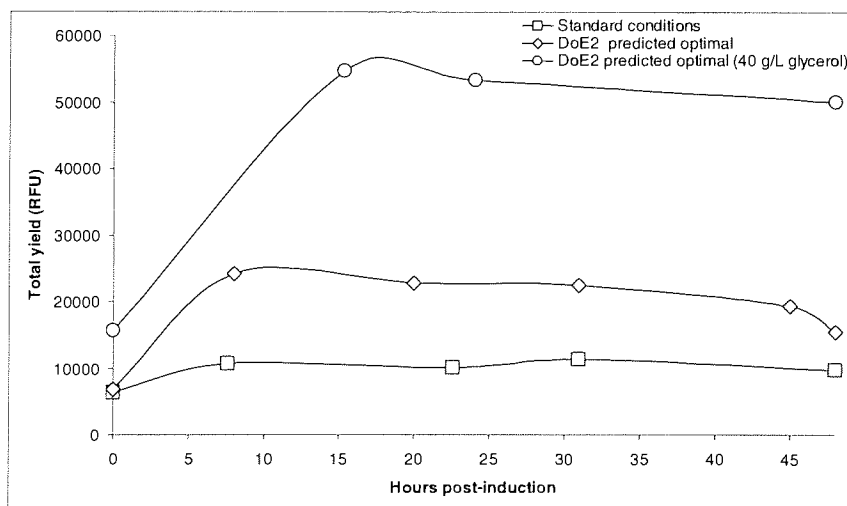


Figure 3.30. Total GFP_{uv} yields (RFU) profiles for the initial DoE2 scale-up and increased pre-induction biomass bioreactor runs. Culture supernatant samples were assayed for GFP_{uv} fluorescence (RFU) during the 48 h induction phase of the cultures.

3.4.2 Increasing pre-induction biomass and using a mixed feed induction with DoE2 optimised process conditions

Increasing pre-induction biomass improved the volumetric yield for DoE2-predicted optimal conditions, as seen in Figure 3.30. The use of a mixed feed induction was identified in the literature as an additional possible method of increasing specific product formation rates in *P. pastoris* fermentations.¹²¹ It was suspected that an additional benefit of providing a mixed feed during induction was the elimination of methanol accumulation. The mixed feed selected for use was 60 % w/v sorbitol 40 % (v/v) methanol with PTM₁ trace salts at 12 mL/L. As the mixed feed contained a lower methanol concentration, the feed rates would supply 60 % less methanol whilst maintaining the same addition of 'carbon source', which might allow the cells to utilise the supplied methanol thus avoiding accumulation and extending the period of useful induction to 48 h.

To show the DoE2 model remained predictive for yields following the introduction of the new large-scale process, check point conditions were used. These were performed in a similar manner to the previous model confirmation experiments. The details of these runs are shown in Table 3.11; f007 (modified BMGY medium only) is also shown for comparison between induction strategies.

The culture profiles for each of the runs described in Table 3.11 are shown in Figure 3.31. As with the initial scale-up runs it is possible to see the culture phases illustrated in Figure 1: glycerol batch, starvation, adaptation and induction. The glycerol batch phase typically lasted for 26 h reaching OD₅₉₅ ~80. It should be noted that the biomass achieved during f011 was OD₅₉₅ ~50. This is more comparable to the biomass achieved during the glycerol batch phase with standard BMGY (10 g/L glycerol), suggesting the BMGY medium used for the batch phase of f011 may not have contained 40 g/L glycerol.

Following the start of induction in f008, the DO level remained around set-point under control of oxygen supplementation *via* the ADI1030 control unit and did not climb to 100 %. However, with cultures f009 and f011 the DO was very unstable, fluctuating from set-point to 100 %. This was attributed to the carbon source being supplied at a limiting rate. The carbon feed limited cellular activity so the cells were not able to consume all the available oxygen. If these runs were not being used for model confirmation, the rate of mixed feed addition would have been increased from 0.16 mL/minute to 0.58 mL/minute (standard handbook rate).

In all three runs (f008, f009 and f011) the biomass began to increase following the start of induction. In f009 and f011 this continued to increase until harvest at 48 h post induction. In f008 there was an apparent drop in biomass following ~55 h (EFT) which was attributed to changes in cellular morphology, as previously discussed. Both the increases in post induction biomass and DO profiles indicate that methanol accumulation was not occurring when the mixed feed induction was used.

Run name	Run description	Data logging (Y/N)	Glycerol batch phase conditions	Glycerol batch phase duration (hours)	Adaptation phase (Y/N)	Induction phase conditions	Induction phase harvest biomass (OD ₅₉₅)	Apparent decrease in biomass (Y/N)	Effective induction period (hours)	Maximum total GFP yield (mg/L)	Total yield at harvest (mg/L)	Normalised yield at harvest ($\mu\text{g/mL/OD}_{595}$)
f007	DoE2 optimised conditions with improved medium composition (40 g/L glycerol)	Y	30 °C, pH 6.0, 20 % DO	25.6	Y	21.5 °C, pH 7.6, 90 % DO	42.66	N	15	175	150	3.52
f008	DoE2 optimised conditions with improved medium composition (40 g/L glycerol) and mixed feed induction (60 % sorbitol 40 % methanol)	Y	30 °C, pH 6.0, 20 % DO	27.8	Y	21.5 °C, pH 7.6, 90 % DO	101.25	Y	48	1801	1246	12.31
f009	Standard conditions with improved medium composition (40 g/L glycerol) and mixed feed induction (60 % sorbitol 40 % methanol)	Y	30 °C, pH 6.0, 20 % DO	28.2	N	30 °C, pH 6.0, 20 % DO	145.53	N	48	330	328	2.26
f011	DoE2 good small scale experimental match conditions with improved medium composition (40 g/L glycerol) and mixed feed induction (60 % sorbitol 40 % methanol)	Y	30 °C, pH 6.0, 20 % DO	21	Y	27.5 °C, pH 6.5, 60 % DO	153.9	N	48	2006	2006	13.03

Table 3.11 DoE2 scale up experimental run descriptions and yield data for cultures using improved medium composition (40 g/L glycerol) and mixed feed (60 % w/v sorbitol, 40 % v/v methanol) induction.

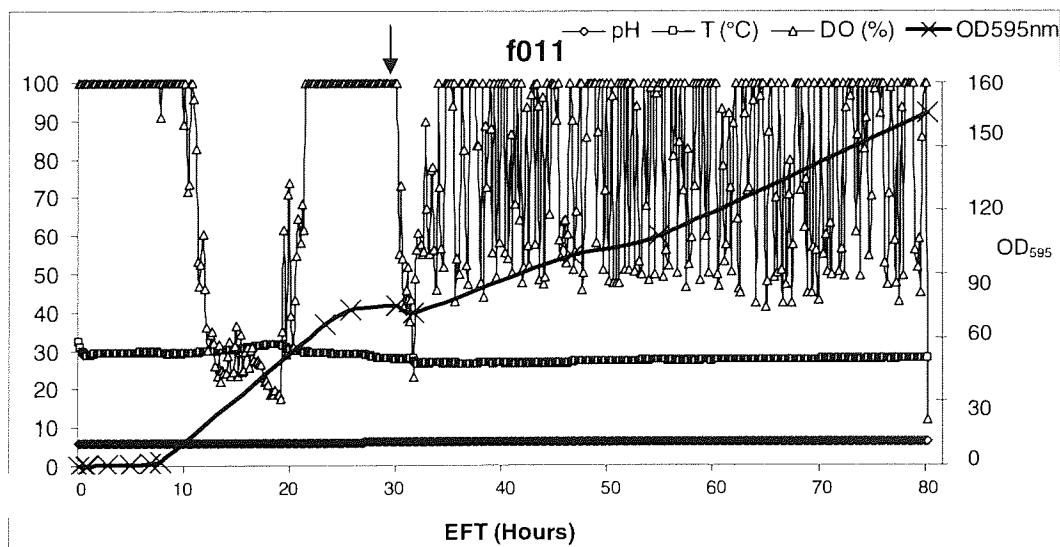
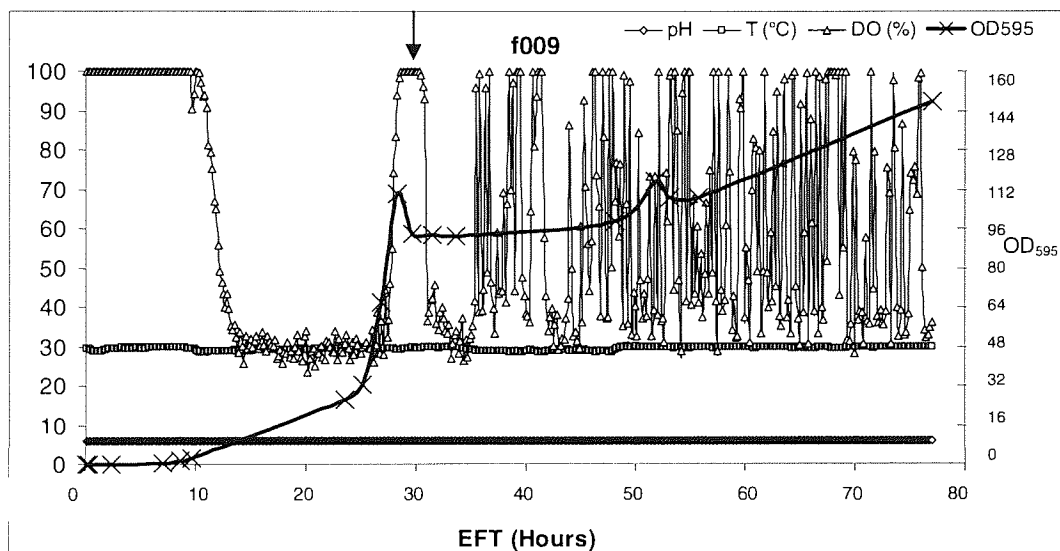
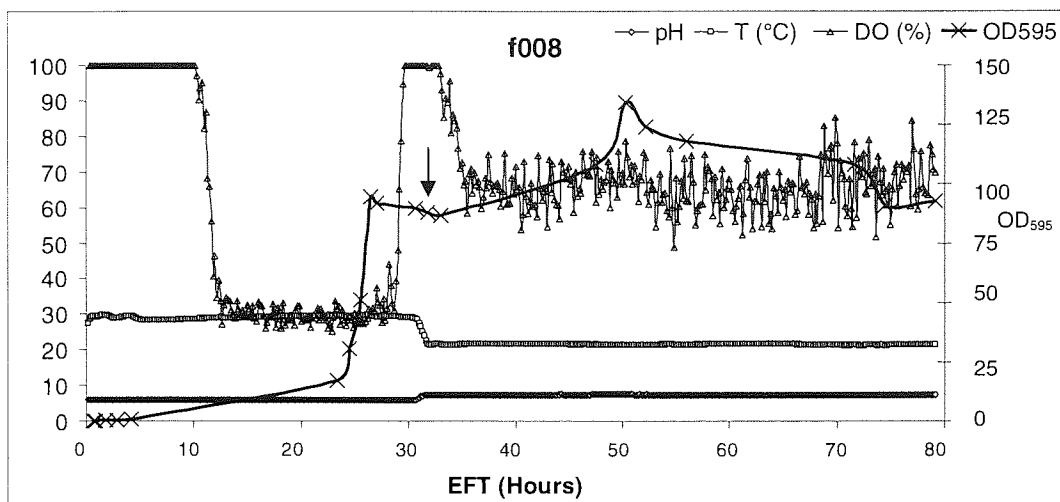


Figure 3.31 DoE2 scale up culture profiles for fed-batch *P. pastoris* X33GFP_{uv} experiments, using improved medium composition (40 g/L glycerol) and mixed feed (60 % sorbitol, 40 % methanol) induction. Run conditions are indicated by Table 3.11. Where culture temperature = T (°C), dissolved oxygen = DO (%) and biomass = OD₅₉₅. Culture biomass is shown on the secondary y-axis. Point of induction is indicated by an arrow.

Figure 3.32 shows the GFP_{uv} yields from the bioreactor runs f008, f009 and f011 as total GFP_{uv} (RFU). The GFP_{uv} yield increases from 0 to 48 h post-induction for all runs, giving an effective induction period of 48 h. The accumulation of product (total GFP yield) is representative of the result originally expected for a fed-batch culture.

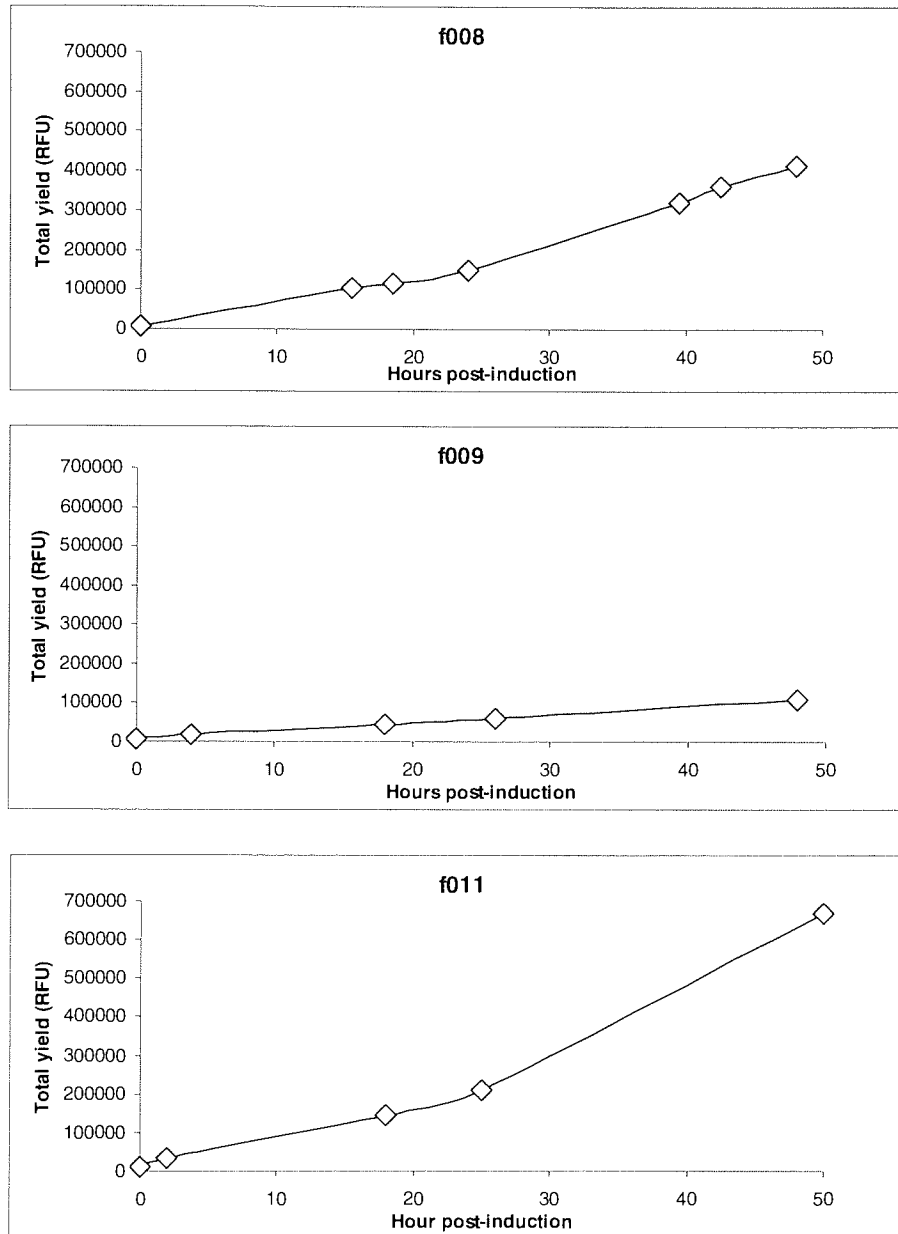


Figure 3.32 GFP_{uv} yield data from fed-batch *P. pastoris* X33GFP_{uv} DoE2 scale-up experiments using improved BMGY medium, mixed feed induction and process conditions described in Table 3.11. Culture supernatant samples were assayed for GFP_{uv} fluorescence (RFU) at time points during the induction phase of the experiments.

The total yield of f011 was expected to be between those of the “standard” (f009) and DoE2 optimised conditions (f008), but this is not the case (Figure 3.33). The total yield of f011 at harvest is 50 % higher than that of the DoE2 predicted optimal conditions. The normalised yields of f007, f008, f009 and f011 at harvest are higher than expected; approximately ten fold

higher than the yields from the small scale (Micro 24), Figure 3.34 shows these results on the secondary axis. Additionally the normalised yield of f011 is higher than expected ($13.03 \mu\text{g/mL/OD}_{595}$) exceeding that of the DoE2 optimised conditions (f008) $12.31 \mu\text{g/mL/OD}_{595}$.

The high yields from the improved process (40 g/L glycerol batch and mixed feed induction) are not predicted well by the DoE2 model. As the DoE2 model was based on data from the standard process (10 g/L glycerol batch and methanol induction), the lack of predictability is not surprising and would require a new model to be built using data from experiments using the improved process.

The high yield from f011 is possibly due to a better matching of pre-induction biomass, methanol feed rate and process conditions and hence cell metabolism and methanol induction strategy. The unexpected yield data is a strong indication that future optimisation experiments should include induction feed rates in the initial DoE plan or in a further round of DoE and modelling building on existing models *e.g.* DoE2.

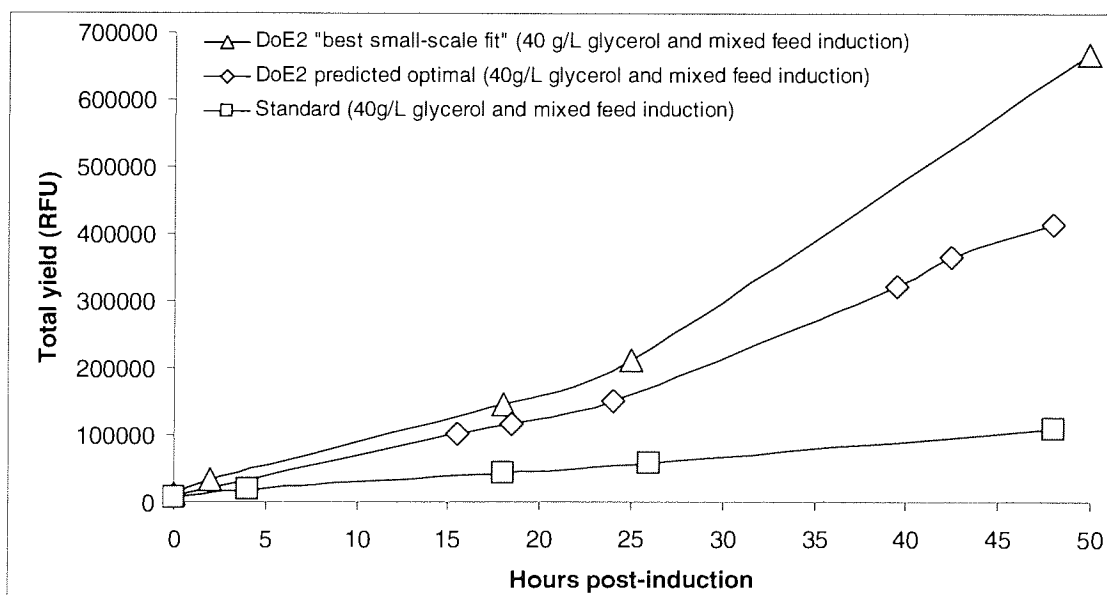


Figure 3.33. Total GFP_{uv} yields (RFU) profiles for the DoE2 scale-up increased pre-induction biomass and mixed feed induction bioreactor runs. Culture supernatant samples were assayed for GFP_{uv} fluorescence (RFU) during the 48 h induction phase of the cultures.

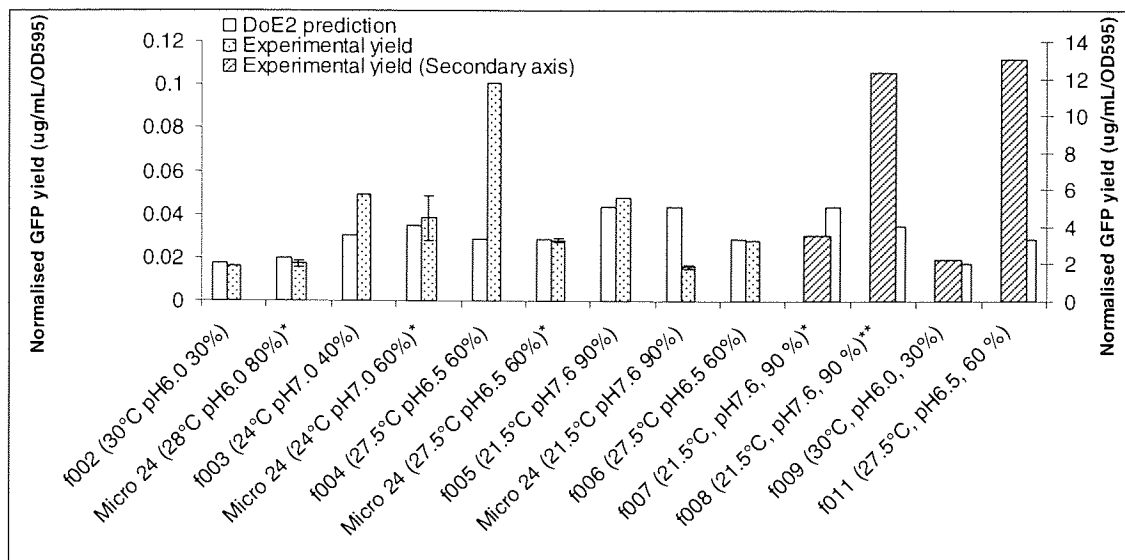


Figure 3.34 Comparison of normalised yield ($\mu\text{g/mL/OD}_{595}$) data at harvest (48-50 h post-induction) for DoE2 model predictions, small-scale Micro 24 experimental runs and large-scale (f002-f011) experimental runs. Error bars represent standard deviation where $n = 3$. *f007 used DoE2 optimised conditions with modified BMGY (40 g/L glycerol) and low methanol feed rate. **f008 used DoE2 optimised conditions with modified BMGY (40 g/L glycerol) and mixed feed induction. f009 used “standard” conditions with modified BMGY (40 g/L glycerol) and mixed feed induction. f011 used “DoE2 best small-scale experimental fit” conditions with modified BMGY (40 g/L glycerol) and mixed feed induction. Experimental yields for f007, f008, f009 and f011 are shown on the secondary axis.

3.5 Use of predictive scale-down/up modelling and micro-bioreactors for process development

The work described in this chapter outlines a series of steps that can be used for rapid and scalable process development of recombinant protein expression. Figure 3.35 summarises the methodologies employed in optimising process conditions for maximising secreted GFP_{uv} yield (the yield normalised to cell density) from bioreactor cultures of *Pichia pastoris* X33 GFP_{uv} . These steps should be applicable to other target proteins and alternative expression systems. The key consideration in using this work scheme is posing an appropriate question of the DoE which can be fulfilled by the small-scale equipment. For example, here temperature pH and DO were selected as factors in the DoE and then shown to be important in maximising yield at both small and large-scales. The induction strategy and medium composition were later recognised as factors that should be included in the DoE question. However, there was no provision for on-line-controlled liquid additions of methanol to the small-scale bioreactor so this could not be included in the small-scale model-building experiments. Medium composition and specifically BMGY glycerol concentration, was not initially envisaged as being critical during the small-scale work so should be included in an additional round of DoE (e.g. DoE3) or a factor in the initial design space for future

processes, as this was seen to be important on a large scale. Where the small-scale equipment prevents a factor being included in the model building it may be necessary to perform optimisation at the large-scale. As recombinant protein expression is complex involving many factors, a truly optimal process may be difficult to achieve. The cyclical nature of optimisation methods (Figure 3.35), however, can provide substantial improvements to the process.

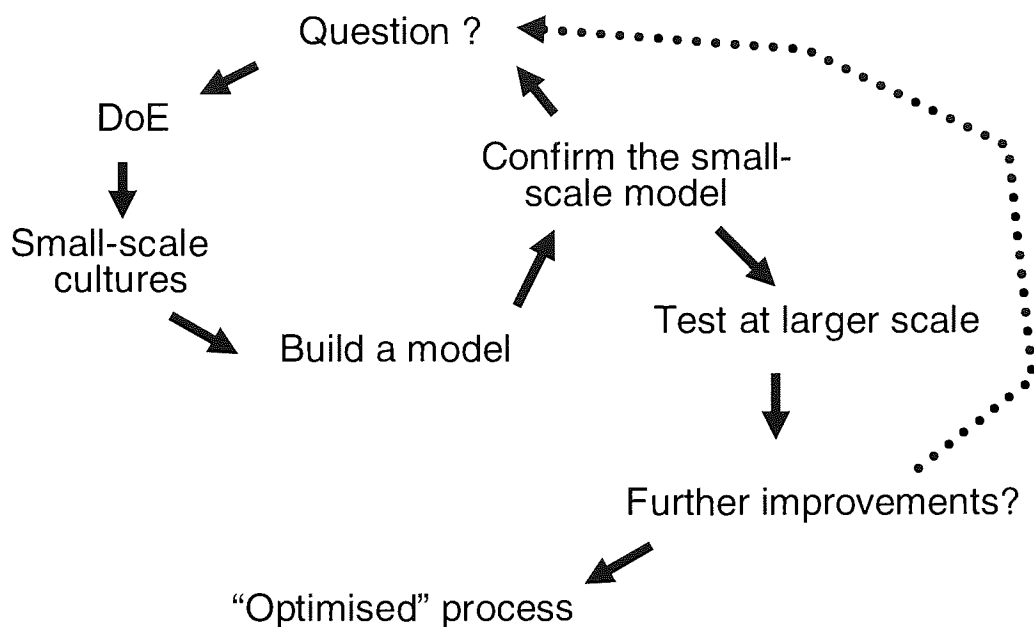


Figure 3.35. Overview of the optimisation process used for maximising GFP_{uv} expression from *P. pastoris* X33GFP_{uv} bioreactor cultures as well as other target proteins and expression systems.

“If I could have my wasted days back
Would I use them to get back on track?”

(Frantic, Metallica)

CHAPTER 4

4 PRODUCTION OF INDUSTRIALLY RELEVANT PROTEINS

In Chapter 3, GFP_{uv} was used as a tool to demonstrate the power of DoE in optimising process parameters in a recombinant protein production experiment. In this chapter the approach was extended to a hybridoma cell-line. In addition other industrially-relevant targets, mini-pro-insulin (MPI), hPGH and a monoclonal antibody-based influenza vaccine, were trialled for expression to establish if they could also be used in a DoE approach.

As part of the partnership between Aston University and Alpha Biologics a hybridoma cell-line was made available for use in this work. The cell-line expressed a monoclonal antibody raised against a tumour-inducing factor. The possible uses for such an antibody include attaching chemotherapy agents for direct targeting of tumour cells.³³ The hybridoma cell-line (referred to here as hybridoma HD1) did not require any further cell-line development work, so could be used immediately for evaluating the scalable process optimisation methods developed in Chapter 3. These methods, although designed using *Pichia pastoris* expressing GFP_{uv}, were intended to be transferable for use with other target proteins and expression systems. The hybridoma cell-line is markedly different to the methylotrophic yeast *P. pastoris*, notably the cells are mammalian requiring different handling and culture techniques and expression is constitutive not inducible. These major attribute changes would fully challenge the robustness of the derived optimisation methodologies.

The creation of yeast strains expressing MPI and hPGH were chosen as additional targets as they have been produced on an industrial scale in yeast. These strains were intended to be used for comparing expression levels between the *S. cerevisiae* Alcofree strain and *P. pastoris* and to further test the DoE scalable modelling methodology.

4.1 Extension of the DoE approach to a hybridoma cell-line

The DoE scalable process optimisation methodologies developed using the GFP_{uv} expressing strain of *P. pastoris* (X33GFP_{uv}) were applied to the hybridoma HD1 cell-line. The product was truly industrially relevant being an anti-tumour factor antibody. The protein expression also differed from the methylotrophic yeast *P. pastoris* being a mammalian cell-line (hybridoma) that constitutively expressed the antibody.

The question asked of the DoE was whether altering the process conditions (culture temperature, pH, DO and glucose concentration) would increase the product antibody yield from the hybridoma HD1 cell-line. Temperature, pH, DO and medium glucose concentration were therefore varied at three levels -1, 0 and +1 (low extreme, mid point and high extreme). The Box-Behnken experimental design was used in conjunction with response surface methodologies (RSM) to build a predictive model and determine the conditions for maximising antibody yield.

The methods set-out in Chapter 3 were followed closely only deviating where culture technique or product yield determination required *e.g.* gas supply strategy, pH control and yield determination. The fluorescence assay could not be used for yield determination as the antibody does not contain a fluorophore. A robust method was therefore required for yield determination from the small-scale cultures. A Protein-A chromatography protocol was devised for use with a GE Healthcare AKTA purification system where the elution peaks could be directly compared and quantified by peak area integration. A standard of known product concentration was assayed and the resultant peak area was used to generate a conversion from peak area to protein concentration for each sample.

4.1.1 Cell-line characterisation

Determining the growth (cell number and viability) and antibody productivity profiles is an important first step in working with a new cell-line. To gain an understanding of how the hybridoma HD1 grew some preliminary flat and shake-flask cultures were set-up in triplicate and monitored over 12 days (288 h). These extended time course cultures were monitored daily for cell number, cell viability, culture pH, glucose concentration of the medium and product determined by SDS-PAGE. This provided information to form an understanding of the growth characteristics and productivity of the cell-line enabling a robust expression strategy to be put in place. The productivity profile allowed identification of earliest point of harvest, allowing the process optimisation runs to be performed over a short duration, therefore reducing development time.

The medium used with hybridoma HD1 was ExCell 302 supplemented with 10 mL/L glutamine, 10 % bovine serum albumin (BSA) and 10 mL/L non-essential amino acids. The flasks used were single use pre-sterilised T 75 cm² flat flasks (T75) and 250 mL shake flasks, both with 0.2 µm vented closures. All cultures were incubated at 37 °C with 5 % CO₂, flat

flask cultures were incubated without agitation and shake flask cultures with 75 rpm agitation. The established standard seeding density for the hybridoma HD1 cell-line was 2×10^5 cells/mL.

4.1.1.1 Growth profiling

The growth profiles of hybridoma HD1 cells cultured in flat (T75) and shake (250 mL) flasks were generated from triplicate cultures set-up in parallel using the same seed culture to ensure the reproducibility. Table 4.1 describes the cultures used for hybridoma HD1 cell-line characterisation. Figures 4.1 and 4.2 show representative growth profiles for T75 culture (15 mL medium) and 250 mL shake flask (50 mL medium) cultures over 288 hours (12 days).

Culture type	Culture volume (mL)	Culture duration (h)	VCN ($\times 10^5$ cells/mL)			pH		Glucose concentration (mg/mL)	
			Inoculation	Harvest	maximum	Inoculation	Harvest	Inoculation	Harvest
T75 Flat flask (75 cm ²) (A)	15	288	2.0	18.1	23.9	8.22	7.09	22.2	3.3
T75 Flat flask (75 cm ²) (B)	15	288	2.0	13.5	20.3	8.22	7.08	22.2	3.0
T75 Flat flask (75 cm ²) (C)	15	288	2.0	18.4	18.7	8.22	7.10	22.2	3.0
Shake flask (250 mL) (D)	50	288	2.0	16.5	16.6	8.22	7.23	22.2	3.0
Shake flask (250 mL) (E)	50	288	2.0	13.1	26.9	8.22	7.22	22.2	3.0
Shake flask (250 mL) (F)	50	288	2.0	15.1	17.6	8.22	7.28	22.2	3.0

Table 4.1. Culture descriptions for hybridoma HD1 cell-line characterisation experiments. Growth profiles of cultures B (T75 flat flask) and D (250 mL shake flask) are shown in Figures 4.1 and 4.2.

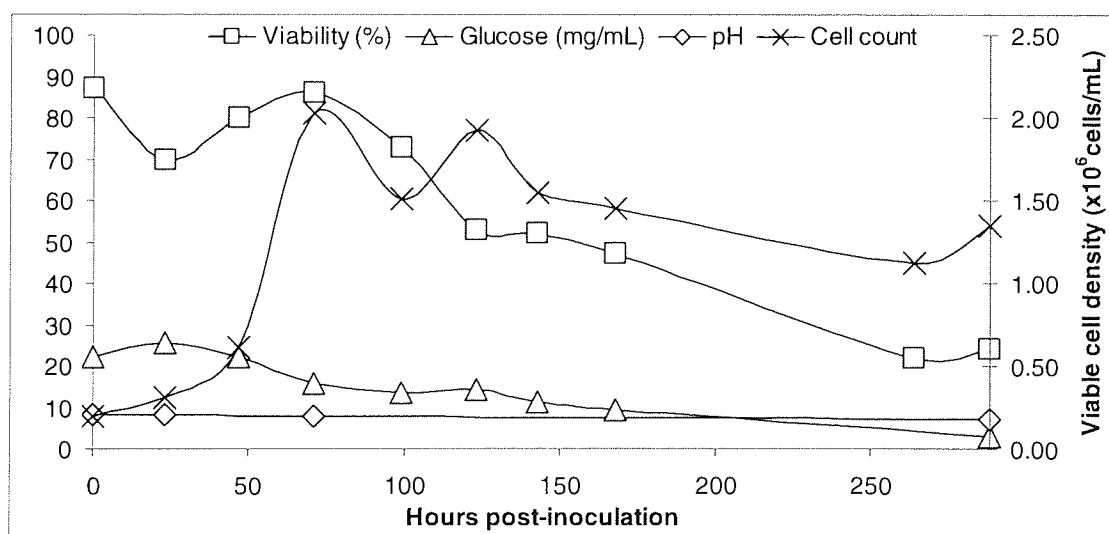


Figure 4.1. Example of the growth profiles for T75 flat flask (culture B) batch culture of hybridoma HD1 cells over 12 days. Values for cell viability (%), medium glucose concentration (mg/mL) and culture pH are shown on the left hand y-axis, viable cell count (cells/mL) is on the secondary y-axis.

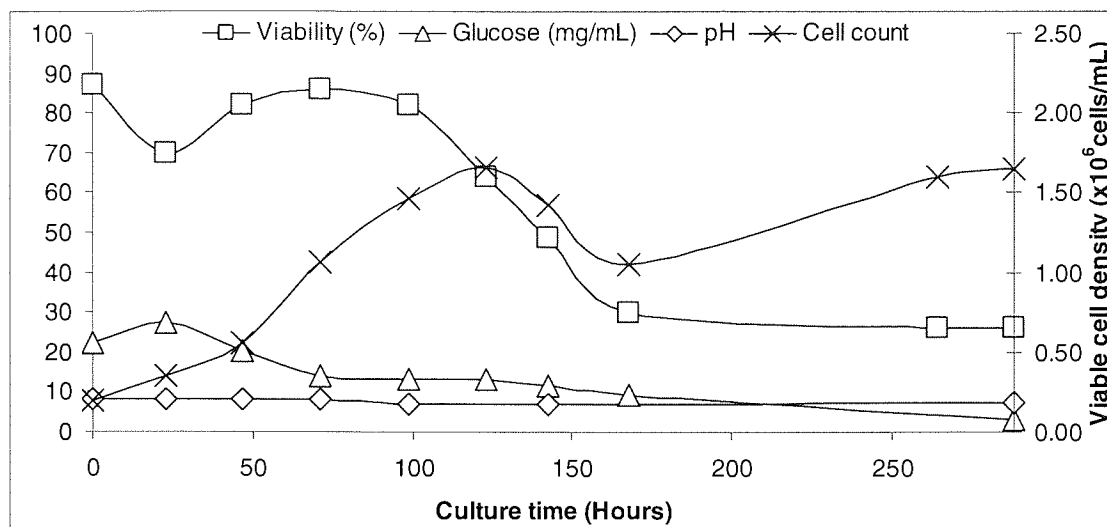


Figure 4.2. Example of the growth profiles for 250 mL shake flask (culture D) batch culture of hybridoma HD1 cells over 12 days. Values for cell viability (%), medium glucose concentration (mg/mL) and culture pH are shown on the left hand y-axis, viable cell count (cells/mL) is on the secondary y-axis.

The replicates of each flask type showed very similar profiles and were further confirmed by the growth profiles of the additional “back-up” cultures, using both T75 cm² flat and 250 mL shake flasks. A batch culture of hybridoma HD1 using standard media could be expected to reach a cell density of $\sim 1.5 \times 10^6$ cells/mL in 3 days when seeded at 2×10^5 cell/mL.

The growth characterisation demonstrated the link between the measured factors: culture pH drops as the cells metabolise the glucose forming lactic acid and as cell number increases the glucose concentration decreases as the cells utilise the available carbon source. Cell viability initially decreased as the cells were introduced at a low cell density. Viability increased as the cell density increased (more secreted growth factors present in the culture medium) and then decreased as the available nutrients were exhausted.

4.1.1.2 Product detection

During the growth profile characterisation of the cultures described above (4.1.1.1), samples were taken daily for productivity assays. Supernatant samples were taken daily at 0, 23, 46.5, 71, 99, 123, 142.5, 168, 264 and 288 h post inoculation. The aim of this analysis was to determine the time point of earliest product detection. Cell free supernatants were run on SDS-PAGE in both non-reduced and reduced (β -mercaptoethanol addition) states. The reduced state resulted in the antibody heavy and light chains separating. Figure 4.3 shows that the intact hybridoma HD1 antibody has an approximate mass of 160 kDa, also existing as a dimer

(~300 kDa) and there are a number of break-down products including a readily identifiable band at ~ 16 kDa.

The hybridoma HD1 product dimer can be seen as a distinct band above the 188 kDa marker. The antibody is detectable from 71 hours (3 days) post-inoculation as indicated by the accumulation of the 16 kDa breakdown product (arrowed). Figure 4.4 shows the Coomassie stained image of the reduced SDS-PAGE gel. The hybridoma HD1 product can be seen as a band at approximately 160 kDa. The reduced state of the samples has removed the product dimer (~300 kDa) seen in the non-reduced gel (Figure 4.7). The ~16 kDa break-down product is detectable from 71 h (3 days) post-inoculation. These gels indicate that the hybridoma HD1 antibody can be detected by SDS-PAGE at 3 days post-inoculation. It was concluded that future experiments should have a minimum duration of 3 days. To ensure reliable detection from cultures with lower yields than those shown in Figures 4.3 and 4.4, the standard culture duration was 96 hours (4 days).

The hybridoma HD1 antibody was detected in all samples including the time 0 (point of inoculation). This is a result of product carry over from the inoculum culture. The cells used to inoculate the experimental cultures were transferred directly from a seed culture and were not subjected to a full media exchange. The use of the ~16 kDa break-down product as an indicator for point of earliest detection was intended to give an approximation for shortest culture duration. This was exceeded by the 4 day culture period selected for further experiments. Additionally SDS-PAGE band densitometry could not be used for yield analysis in the DoE experiments. The chosen method was product peak integration from Protein-A chromatograms with “blanking” using the time 0 result to remove carry over effects.

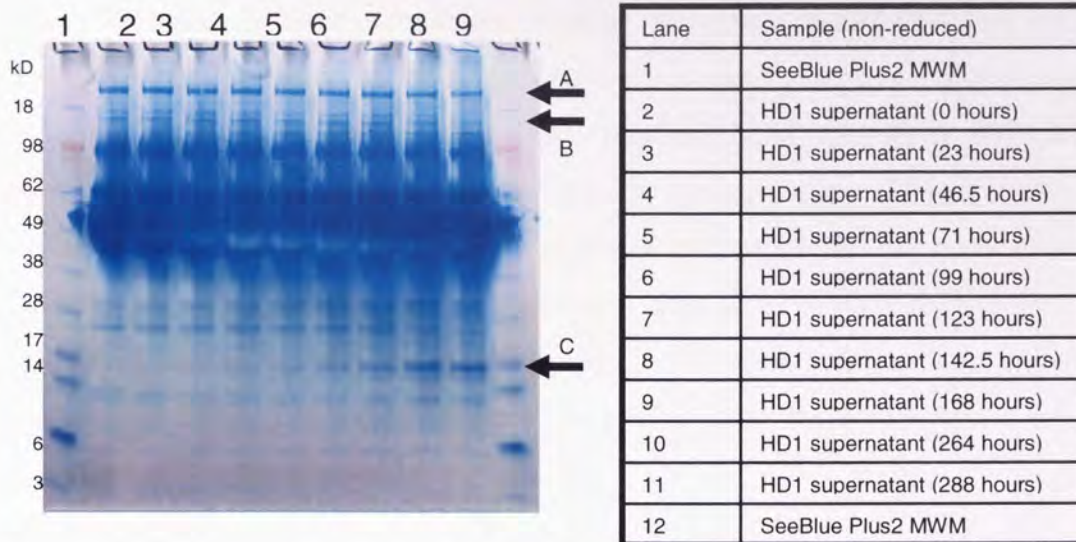


Figure 4.3. Coomassie stain visualised SDS-PAGE image of non-reduced supernatant sample from 250 mL shake flask (culture D) batch culture of hybridoma HD1 cells over 12 days (0-288 hours). Arrows indicate position of hybridoma HD1 expressed antibody dimer (A), monomer (B) and breakdown product (C). Product accumulation is detectable following 3 days (71 hours) post inoculation.

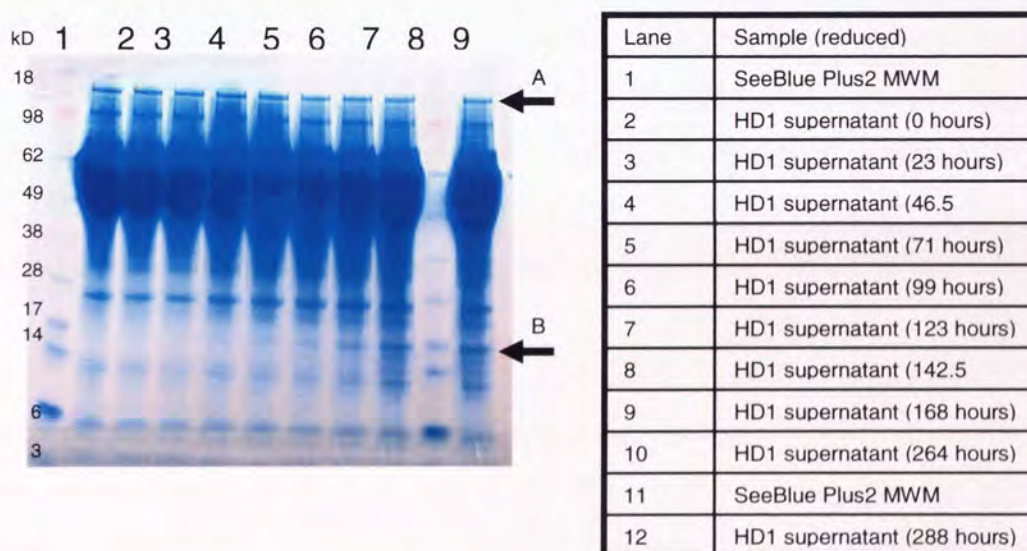


Figure 4.4. Coomassie stain visualised SDS-PAGE image of reduced supernatant sample from 250 mL shake flask (culture D) batch culture of hybridoma HD1 cells over 12 days (0-288 hours). Arrows indicate position of hybridoma HD1 expressed antibody (A) and breakdown product (B). Product accumulation is detectable following 3 days (71 hours) post inoculation.

4.1.2 The Applikon Micro 24 Bioreactor for mammalian cell culture

The Micro 24 is a bench-top, computer-controlled micro-bioreactor system, which uses a modified deep 24-well plate (incubation cassette) to provide 24 individually-controlled “vessels” which are run in parallel. The system is more fully described in Chapter 3 including annotated diagrams of the equipment. Further information is available from the Applikon Biotechnology internet site (<http://www.applikon-bio.com>).

Using the Micro 24 for mammalian cell culture requires a slightly different equipment setup and operational procedure to those used for yeast culture (as described in Chapter 3). The rate of agitation for the incubation cassette was reduced to 500 rpm (from 800 rpm): this is the lowest setting possible which still achieves adequate mixing. This is much higher than agitation rates used in STB which are typically <100 rpm, but the wells of the incubation cassette provide a low(er) shear environment than a STB with an impellor. There were initially concerns over how the hybridoma HD1 cells would cope with the agitation rate (500 rpm) in the Micro 24, but data from other groups using the Micro 24 for a CHO cell-line showed this was not an issue.⁸²

Concerns over the method of pH control using carbon dioxide and ammonium hydroxide vapour were also shown to be unfounded for mammalian cell culture (with the hybridoma HD1 cell-line). It was thought that the ammonium hydroxide might be too harsh, resulting in cell death and that a pH buffer addition strategy should be employed. This method of pH control had been used previously for CHO cells in the Micro 24.⁸² The buffer addition method required additional liquid to be supplied to the wells, with each culture having differing set points (including pH) that would require a different buffer type/volume to be added. This liquid addition method was not seen as ideal for a “high-throughput” methodology. It was decided that a gaseous pH control by CO₂ and ammonium hydroxide vapour was trialled.

The Micro 24 has 5 gas-in ports; CDA, 1, 2, 3 and an auxiliary (AUX) port. The AUX port directs gas into the incubation chamber and not into the culture. When using the Micro 24 for yeast culture the AUX port was not used. The control strategy for the hybridoma HD1 cultures required both up and down control of pH which meant it was not possible to have both up and down control of DO as there are insufficient gas ports leading to the cultures. Therefore, upward control of DO remained as with the yeast strategy using O₂ addition to the culture but downward DO control was attempted by having nitrogen “overlay” in the incubation chamber using the AUX port. This was hoped to reduce the DO in the cultures. The well closures used were the free flowing 0.2 µm vented type so allowing nitrogen in the chamber to enter the culture wells and strip the oxygen from the cultures therefore reducing DO. Culture DO was controlled upward with O₂ addition, allowing full control of DO to set-point.

4.1.3 DoE plan for hybridoma HD1 yield optimisation (DoE3)

The non-coded levels for each factor were based on previous experience of the standard hybridoma HD1 cell-line process, discussions with those working on the hybridoma HD1 project and examples in the literature of hybridoma culture. The lower “-1” limit for media glucose concentration was set by the medium composition at 2 mg/mL as it was not possible to obtain a lower glucose concentration with ExCell 302. Table 4.2 shows the DoE matrix of required factor settings for the hybridoma HD1 (DoE3) predictive model in both coded (-1, 0, 1) and the non-coded input factors used as set-point for the experimental Micro 24 runs. The run order of the model building conditions allowed each run to generate a model, allowing earlier indication of trends. Once the full data set was available a model could be built using data from replicate runs.

CODED FACTORS				INPUT FACTORS				Micro 24 experimental run order			
Temp	pH	DO	Glucose	Temp (°C)	pH	DO (%)	Glucose (ng/mL)				
-1	-1	0	0	34	6.6	40	6		2	3	4
1	-1	0	0	38	6.6	40	6		2	3	4
-1	1	0	0	34	7.4	40	6		2	3	4
1	1	0	0	38	7.4	40	6	1	2	3	
0	0	-1	-1	36	7	20	2	1	2	3	
0	0	1	-1	36	7	60	2	1	2	3	
0	0	-1	1	36	7	20	10	1	2	3	
0	0	1	1	36	7	60	10	1	2	3	
-1	0	0	-1	34	7	40	2	1	2	3	
1	0	0	-1	38	7	40	2	1	2	3	
-1	0	0	1	34	7	40	10	1	2	3	
1	0	0	1	38	7	40	10	1	2	3	
0	-1	-1	0	36	6.6	20	6	1	2	3	
0	1	-1	0	36	7.4	20	6	1	2	3	
0	-1	1	0	36	6.6	60	6	1	2	3	
0	1	1	0	36	7.4	60	6	1	2	3	
-1	0	-1	0	34	7	20	6	1	2	3	
1	0	-1	0	38	7	20	6	1	2	3	
-1	0	1	0	34	7	60	6	1	2	3	
1	0	1	0	38	7	60	6	1	2	3	
0	-1	0	-1	36	6.6	40	2	1	2	3	
0	1	0	-1	36	7.4	40	2	1	2	3	
0	-1	0	1	36	6.6	40	10	1	2	3	
0	1	0	1	36	7.4	40	10	1	2	3	
0	0	0	0	36	7	40	6	1			4
0	0	0	0	36	7	40	6	1			4
0	0	0	0	36	7	40	6	1			4

Table 4.2. Experimental matrix showing coded (-1, 0, 1) and non-coded set-points required for the Box-Behnken DoE3 predictive model for optimising antibody yield from hybridoma HD1 hybridoma cell culture. The experimental run order is also indicated demonstrating that each run provides sufficient data for an independent model to be generated.

4.1.4 Data collection for predictive model building

The classic quadratic Box-Behnken design⁹¹ was employed for generating the predictive model. This allowed a reduced number of treatment combinations to be used in building the model and gave no bias towards any potential optimum region of the process space. Minitab v15.1.30 software was used to generate an experimental matrix of factor combinations

required to build a predictive model to a Box-Behnken design. Table 4.3 shows the input variables and measured outputs for the model building experiments which were performed in the Micro 24. The factor combinations were carried out in at least triplicate to provide statistical confidence in the model building data. Cell density, viability and glucose concentration were monitored daily by off-line methods. Antibody yield was measured at point of harvest (4 days post inoculation) by Protein-A purification of 5 mL cell-free culture supernatant.

A conversion factor was generated by purification of a high yielding sample of hybridoma HD1 culture supernatant, taken from a 7 L bioreactor culture. The sample was purified using the hybridoma HD1 purification protocol on the GE Healthcare AKTA system, the resulting eluted product fractions being pooled and quantified by A_{280} . The conversion factor was determined by dividing the product concentration (from A_{280}) by the associated peak area and sample volume (5 mL), giving the conversion antibody yield ($\mu\text{g/mL}$) = $0.1564 \times \text{peak area}$.

INPUT FACTORS				RUN DATA										COMBINED DATA	
T (°C)	pH	DO (%)	Glucose (mg/mL)	Installation Run	Yield (µg/mL)	Run 1	Yield (µg/mL)	Run 2	Yield (µg/mL)	Run 3	Yield (µg/mL)	Run 4	Yield (µg/mL)	Mean yield (µg/mL)	SD
34	6.6	40	6					A1	0.06	A1	0.12	A1	0.29	0.16	0.12
38	6.6	40	6					B6	0.30	A6	0.41	A6	0.27	0.33	0.07
34	7.4	40	6					B1	0.05	B1	0.03	B1	0.01	0.03	0.02
38	7.4	40	6					A6	0.22	B6	0.07			0.18	0.10
36	7	20	2					A3	0.22	A3	0.34			0.23	0.11
36	7	60	2					B3	0.31	B3	0.39			0.31	0.08
36	7	20	10					C3	0.44	C3	0.45			0.42	0.04
36	7	60	10					D3	0.20	D3	0.26			0.26	0.07
34	7	40	2					A1	0.26	C1	0.23			0.25	0.02
38	7	40	2					B6	0.14	C6	0.18			0.22	0.10
34	7	40	10					B1	0.14	D1	0.15	C1	0.15	0.15	0.01
38	7	40	10					C6	0.24	D6	0.35			0.38	0.15
36	6.6	20	6					A2	0.27	A3	*			0.22	0.07
36	7.4	20	6					A5	*	A4	0.17			0.22	0.07
36	6.6	60	6					B2	0.19	B4	0.04	B3	0.13	0.12	0.07
36	7.4	60	6					B5	0.19	C4	0.18			0.22	0.07
34	7	20	6					C1	0.29	D4	0.21			0.20	0.01
38	7	20	6					D5	*	A2	0.13			0.19	0.09
34	7	60	6					D1	0.21	A5	0.30	A5	0.12	0.28	0.15
38	7	60	6					D6	*	B2	0.13			0.19	0.05
36	6.6	40	2					D2	0.11	D5	0.28	B5	0.21	0.29	0.08
36	7.4	40	2					C5	0.30	C2	0.12			0.17	0.09
36	6.6	40	10					C2	0.33	D2	0.13			0.23	0.09
36	7.4	40	10					A4	0.11	A5	0.19			0.27	0.07
36	7	40	6	A1				B4	0.16	B5	0.11			0.12	0.02
36	7	40	6	A4				C4	0.19			A2	0.14	0.15	0.04
36	7	40	6	C6				D4	0.13			B2	0.17		
36	7	40	6	D2					0.13			C2	0.09		
36	7	40	6						0.24						
36	7	40	6						0.24						

Table 4.3. Specification of the input factors and measurable outputs for the hybridoma HD1 model building experiments. The input factors were temperature (T), pH and dissolved oxygen tension (DO). Antibody yield (µg/mL) was measured 48 h post induction in at least triplicate. The standard deviation (SD; n 3) is given for the yield. ^ indicates a contaminated culture and * an error in purification resulting in no yield data.

4.1.5 Building the DoE3 predictive model

The yield data from the Micro 24 cultures, as shown in Table 4.3, were used to build a predictive model of yield and process conditions (T, pH, DO and glucose concentration). The model building process used Minitab (v15.1.30) software to generate coefficients for each of the factor terms as shown in Table 4.4. The model was refined as with the DoE1 and DoE2 models in Chapter 3; terms were removed from the full model, based on their associated *p*-values, to maximise the R^2_{adj} value of the model. The R^2_{adj} values were 83.09 (full model), 84.26 (1 term removed), 85.21 (2 terms removed), 84.77 (3 terms removed) and 84.21 (4 terms removed). The model with 2 terms removed (R^2_{adj} 85.21) was selected as the refined model and was named “DoE3”.

Term	Coefficient	<i>p</i> -value
Constant	5.87195	0.389
T	-0.57227	0.029
pH	1.48234	0.218
DO	-0.02624	0.088
Glucose	-0.09185	0.382
T * T	0.00765	0.034
pH * pH	-0.10615	0.216
DO * DO	0.00017	0.000
Glucose * Glucose	0.00485	0.000
T * Glucose	0.00833	0.001
pH * DO	0.00245	0.242
pH * Glucose	-0.03346	0.005
DO * Glucose	-0.00075	0.002

Table 4.4. Refined DoE3 predictive model terms, coefficients and associated *p*-values based on the model building data shown in Table 4.3.

$$\text{Antibody yield } (\mu\text{g/mL}) = 5.87195 - (0.57227 * T) + (1.48234 * \text{pH}) - (0.02624 * \text{DO}) - (0.09185 * \text{Glucose}) + (0.00765 * T * T) - (0.10614 * \text{pH} * \text{pH}) + (0.00017 * \text{DO} * \text{DO}) + (0.00485 * \text{Glucose} * \text{Glucose}) + (0.00833 * T * \text{Glucose}) + (0.00245 * \text{pH} * \text{DO}) - (0.03346 * \text{pH} * \text{Glucose}) - (0.00075 * \text{DO} * \text{Glucose})$$

Equation 4.1. Refined DoE3 predictive model for Yield ($\mu\text{g/mL}$) where T = temperature ($^{\circ}\text{C}$), pH = pH, DO = dissolved oxygen tension (%) and Glucose = glucose concentration mg/mL .

The DoE3 model was examined statistically by ANOVA (Table 4.5). It can be seen that the *p*-value associated with the regression *F*-statistic (13.97 %) is 0.000. This shows strong evidence (> 99.9 %) against the null hypothesis (no linear relationship between any of the predictors and Y)¹¹⁷ therefore, there is strong evidence of a relationship between predicted and experimental antibody yields. The *p*-values for the DoE3 model were lower than those seen from the DoE1 and DoE2 models. It was expected that the impact of this on the

predictive capability of the DoE3 model would be positive, but this required testing experimentally with “check point” conditions, to confirm this.

Source	DF	SS	MS	F statistic	p-value
Regression	12	0.173971	0.014498	13.97	0.000
Linear	4	0.063080	0.003647	3.51	0.033
Square	4	0.066020	0.016505	15.90	0.000
Interaction	4	0.044871	0.011218	10.81	0.000
Residual	15	0.015570	0.001038		
Total	27	0.189540			

Table 4.5. ANOVA analysis of DoE3 model where: DF = Degrees of freedom, SS = Sum of squares, MS = Mean square.

The model was plotted visually using response surface regression to demonstrate how the factor interactions affected yield; Figure 4.5 shows contour plots of each of the factors, the hold values being “0” mid-points. Figure 4.6 shows a surface wire frame plot of yield against culture temperature (Temp) and pH with glucose concentration held at 10 mg/mL “+1”. It can be seen that temperature and pH have a major influence on antibody yield. Figures 4.5 and 4.6 show the DoE3 design space does not appear to capture the optimised set of conditions (factors) for maximal antibody yield. However, extending the design space in the direction of the optimal yielding conditions may not provide higher yields. This is due to the process limits (-1, 0, 1) being set to the extremes of expected hybridoma HD1 cell tolerance. The use of check point conditions exceeding the design space limits would show if this were the case.

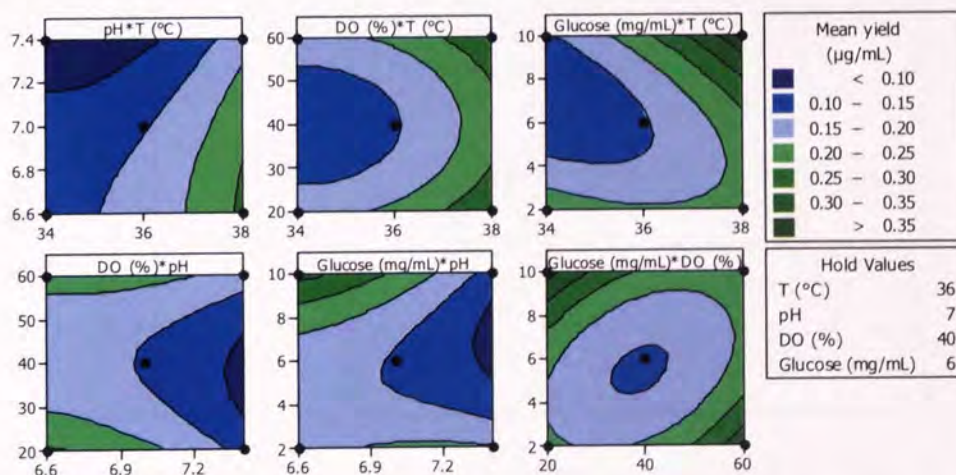


Figure 4.5. Response surface contour plot for DoE3 model showing yield ($\mu\text{g/mL}$) changing with each of the factors. Where T = temperature ($^{\circ}\text{C}$), pH = pH and Glucose = glucose concentration (mg/mL). Black circles represent design space corner and centre point locations.

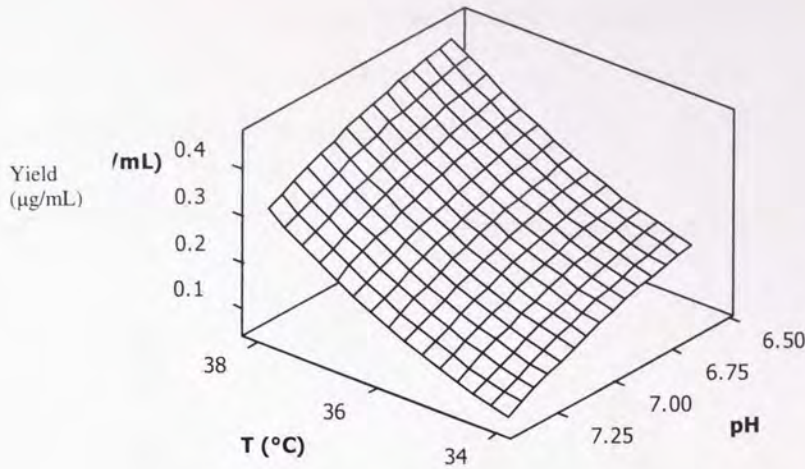


Figure 4.6. Response surface wire-frame plot for DoE3 model showing yield ($\mu\text{g/mL}$) changing with temperature and pH factors. Where T = temperature ($^{\circ}\text{C}$), pH = pH. Glucose =glucose concentration (mg/mL). Glucose hold value is “+1” 10 mg/mL.

The model described by Equation 4.1 can be used to predict the position or “coordinates” on the design space of the optimal conditions. The ‘response optimiser’ function of Minitab software was used to identify the “optimised” conditions on the design space to be 38.0 $^{\circ}\text{C}$, pH 6.6 and 10 mg/mL glucose, as shown in Figure 4.7.

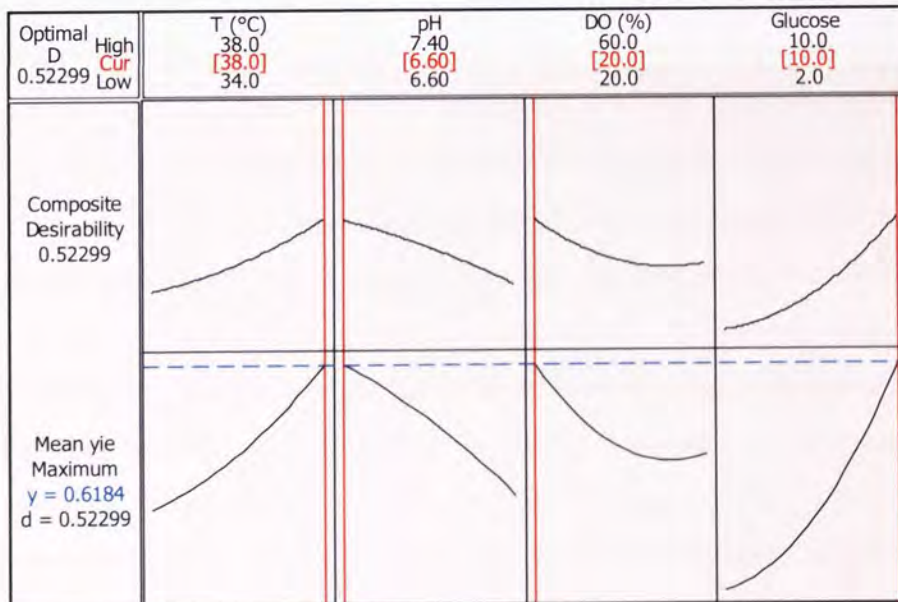


Figure 4.7. Response optimisation for DoE3 model as predicted by Minitab software. Suggested optimal conditions (Factor settings) are 38 $^{\circ}\text{C}$, pH 6.6, 20 % DO and 10 mg/mL glucose.

4.1.6 Testing the model at small scale

The predictive capability of the DoE3 model had been tested statistically (by ANOVA). This capability was tested experimentally to show the model predicted yields were achievable. The

Micro 24 bioreactor was used to run “check point” conditions to demonstrate how well the model predicted experimental conditions. Figure 4.8 shows that the model does predict optimal experimental yield, from those tested.

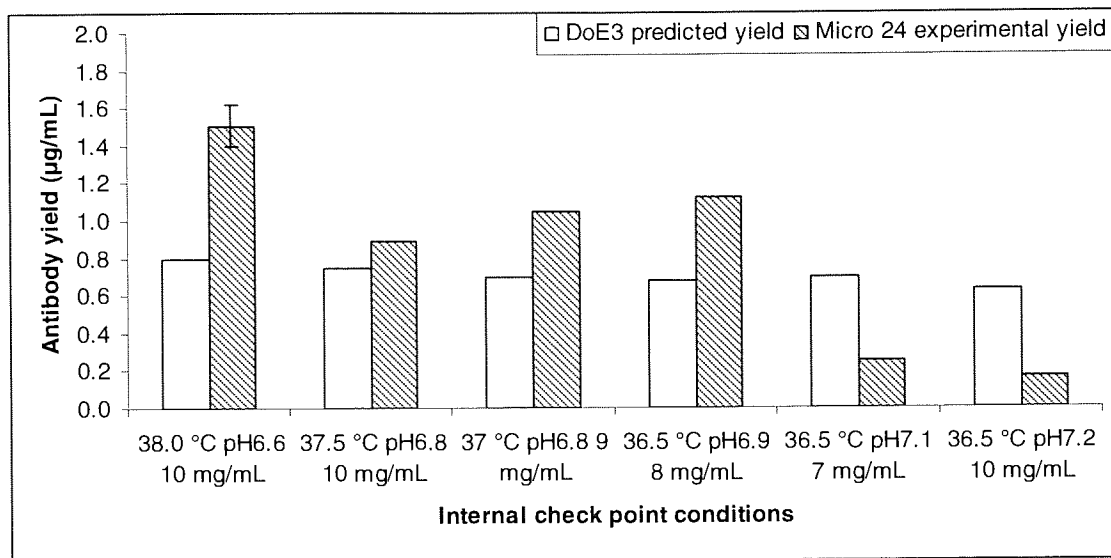


Figure 4.8. DoE3 model predicted and experimental yield from model confirmation Micro 24 cultures. Check point conditions are from within the model design space with yield in µg/mL.

The p -values given by the DoE3 model for each term are much lower than those from the DoE2 (*P. pastoris* GFP expression) model indicating a higher quality model, statistically. However, the predictive capability for experimental values is poorer than expected. Figure 4.9 shows an XY scatter plot of DoE3 predicted versus Micro 24 experimental yield (µg/mL) for the internal check point conditions. It can be seen that the data points lie either side the line $y=x$, indicating that the model both over and under-estimates the antibody yield values. This is most likely due to the variability in the experimental model building data and no repeat data from the model confirmation run. Despite this the predicted conditions did provide the highest yields. If possible an increased number of experimental repeats and should be performed to identify and remove outlying results to reduce the standard deviation for each set of model building conditions.

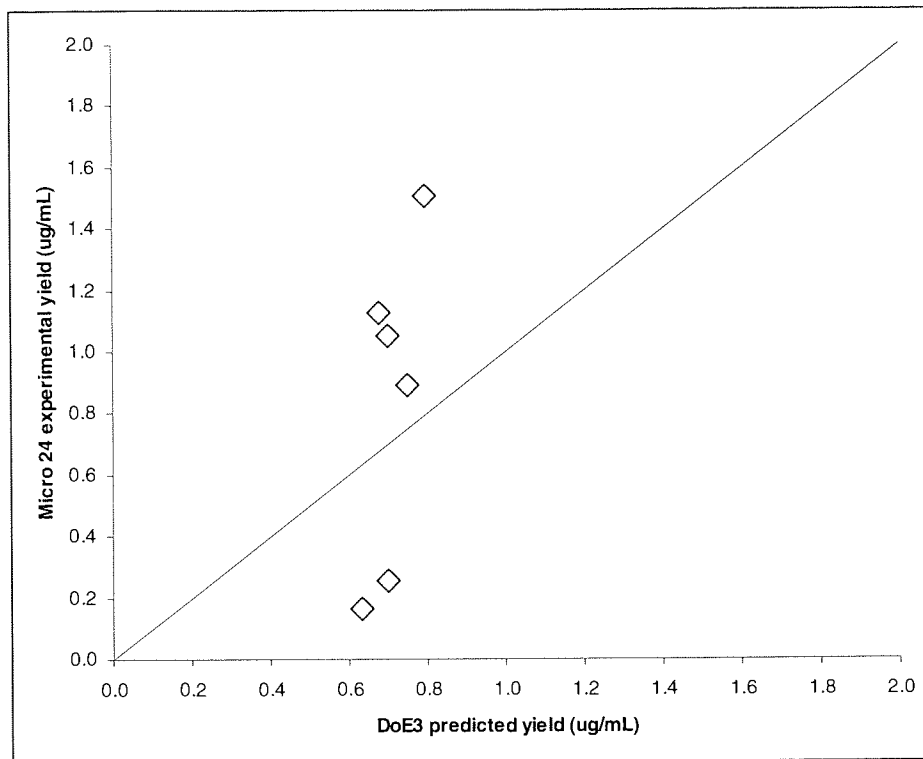


Figure 4.9. XY scatter plot of DoE3 model predicted and experimental yield ($\mu\text{g/mL}$) from model confirmation Micro 24 cultures (as shown in Figure 4.12). Check point conditions are from within the model design space. The line $y=x$ is shown to demonstrate the ideal relationship between model predicted and experimental yields.

It was previously noted that the RSM plots for DoE3 indicate that the design space does not encompass the optimal factor setting for maximal product yield. During the experimental planning stage the upper and lower process limits for each factor were selected based on previous experience and expectations of hybridoma HD1 behaviour. The model suggests that increasing both temperature and glucose concentration beyond the “+1” limits will result in further increased yield. To demonstrate if this were the case and the process limits were not limits, factor settings outside the design space were run as check point conditions. Figure 4.10 shows the antibody yield ($\mu\text{g/mL}$) for the check point conditions outside of the design space. The conditions tested lower ($33\text{ }^{\circ}\text{C}$) and higher ($39, 40\text{ }^{\circ}\text{C}$) temperatures and higher glucose concentrations (15 mg/mL).

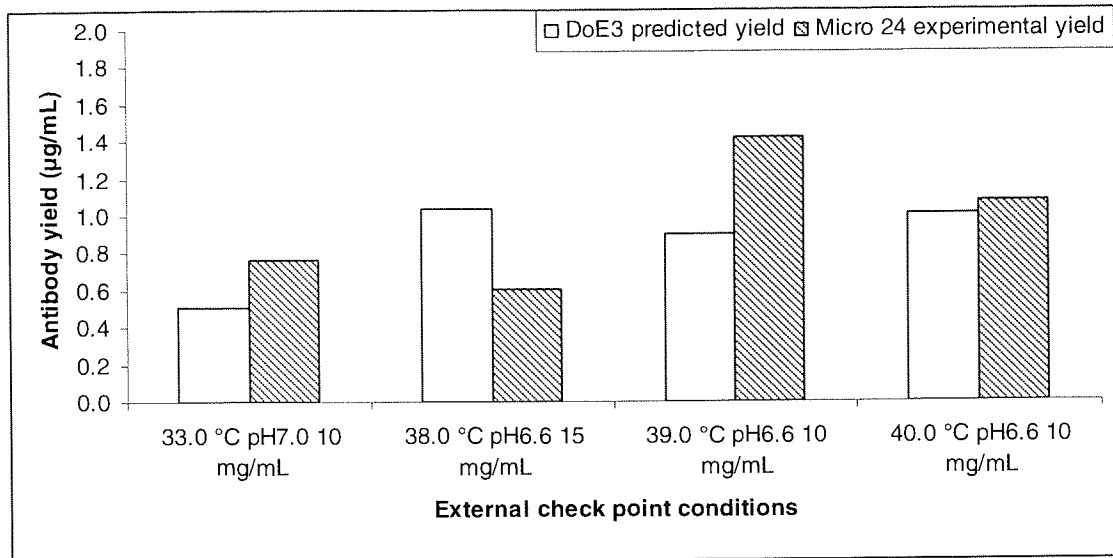


Figure 4.10. DoE3 model predicted and experimental yield ($\mu\text{g/mL}$) from model confirmation Micro 24 cultures. Check point conditions are outside the model design space.

Figure 4.11 shows the yields ($\mu\text{g/mL}$) for all check point and “optimised” conditions. The increases in yield suggested by the contour plots of the DoE3 model, when the design space limits are exceeded, are not met by the experimental data. This shows the design space limits were correctly set and the further yield increases are not possible whilst investigating these (temperature, pH and glucose concentration) factors alone.

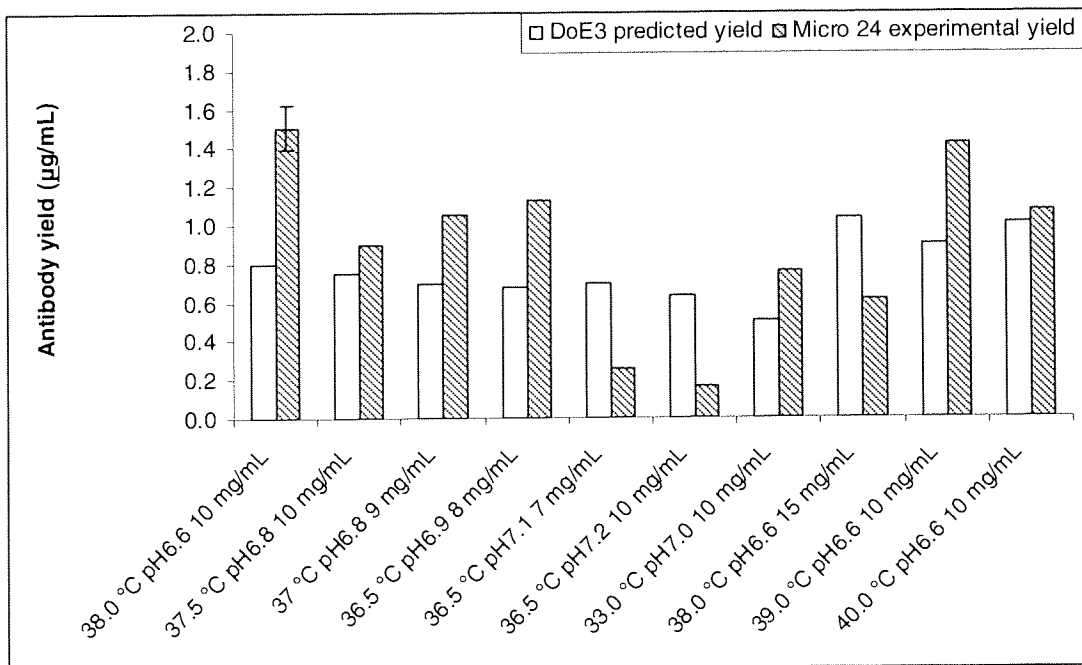


Figure 4.11 DoE3 model predicted and experimental yield from model confirmation Micro 24 cultures. Showing antibody yield ($\mu\text{g/mL}$) from optimised conditions and both internal and external design space check point conditions. (Error bars show SD, where $n=3$).

4.1.7 Model scalability

Having shown the DoE3 model was able to predict optimal conditions at the small-scale in the Micro 24 system, the model required testing at a larger scale as with the DoE2 model. To show if this were possible check point conditions were selected to be run in a 7 L stirred tank bioreactor (STB) as with the small-scale model testing. The conditions selected were “standard” (37.0 °C, pH 7.0, 2.5 g/L glucose and 40 % DO).

DoE3 optimised (38.0 °C, pH 6.7, 10 g/L glucose and 40 % DO) and DoE3 best small-scale experimental fit (36.5 °C, pH 6.9, 8 g/L glucose and 40 % DO). Each of the scale-up runs was named as shown in Table 4.6.

Name	Run conditions
BR01	Bi-phase Day 0-4 “standard” (37.0 °C, pH 7.0, 2.5 g/L glucose and 40 % DO) Day 4-5 DoE3 “optimised” (38.0 °C, pH 6.7, 10 g/L glucose and 40 % DO)
BR02	Day 0-2 DoE3 “optimised” (38.0 °C, pH 6.7, 10 g/L glucose and 40 % DO)
BR03	Bi-phase Day 0-4 DoE3 best small-scale experimental fit (36.5 °C, pH 6.9, 8 g/L glucose and 40 % DO) Day 4-7 DoE3 “optimised” (38.0 °C, pH 6.7, 10 g/L glucose and 40 % DO)

Table 4.6. DoE3 scale-up experimental run conditions. Note: BR02 was not run in a bi-phase manner as it was already using DoE3 optimised conditions.

4.1.7.1 Bench-scale (7L vessel) experimental set-up

The vessel used for the DoE3 scale-up experiments is the same as was used for the DoE2 scale-up work. As DoE3 is based on mammalian cell culture the vessel required some changes in set-up to allow culture of hybridoma cells instead of yeast. Figure 4.12 shows an annotated photograph of the vessel set-up and required ancillary equipment. The main differences between the microbial and mammalian set-ups are impellor design, agitation rate, gas type and flow rate (vvm) and pH control strategy.

The six bladed Rushton turbine impellers used in DoE2 scale-up produce a high shear environment capable of maintaining high oxygen transfer. However, mammalian cells have a lower tolerance to shear than yeast cells and do not require such high oxygen transfer rates. For these reasons a single marine blade impellor was used in place of the three Rushton turbine impellers. The agitation rates used with the marine blade impellor were much lower

(80-90 rpm) than those used in DoE2 scale-up (600-900 rpm) to reduce the shear of the system and was sufficient to meet the lower oxygen demand of the hybridoma HD1 cells.

The gas flow rates were reduced to 0.2 vvm for hybridoma HD1 culture; those used in DoE2 scale-up (0.5-2.0 vvm) would produce deleterious effect to the hybridoma cells. Shear forces are exerted by bubbles popping, which would be exacerbated by high gas flow rates. In addition the ExCell 302 medium used with the hybridoma HD1 cells is supplemented with 10 % BSA; this increases the protein concentration of the medium for product stability but also foaming. Low gas flow rates help to reduce the volume of foam generated, but to further reduce foaming, an antifoam agent (Sigma Antifoam-C) was added to 0.02 % v/v which had been shown to have no detrimental effect to hybridoma HD1 cell growth, yet be effective in foam control.

The pH control strategy also differed from that used for DoE2 scale-up: CO₂ replaced HCl and sodium hydroxide replaced ammonium hydroxide. These control agents are more suited for use in mammalian cell culture, as they are seen as being less “harsh” to the cells.

Glucose “slug” addition to the culture was via a stand alone peristaltic pump, not controlled by the ADI1030 unit. Glucose (200 g/L stock solution) was added to culture set-point using calibrated pump tubing. The peristaltic pump was used at its maximum rate (0.917 mL/minute) for “X” minutes. The time “X” was calculated from culture g/L glucose and pump rate.

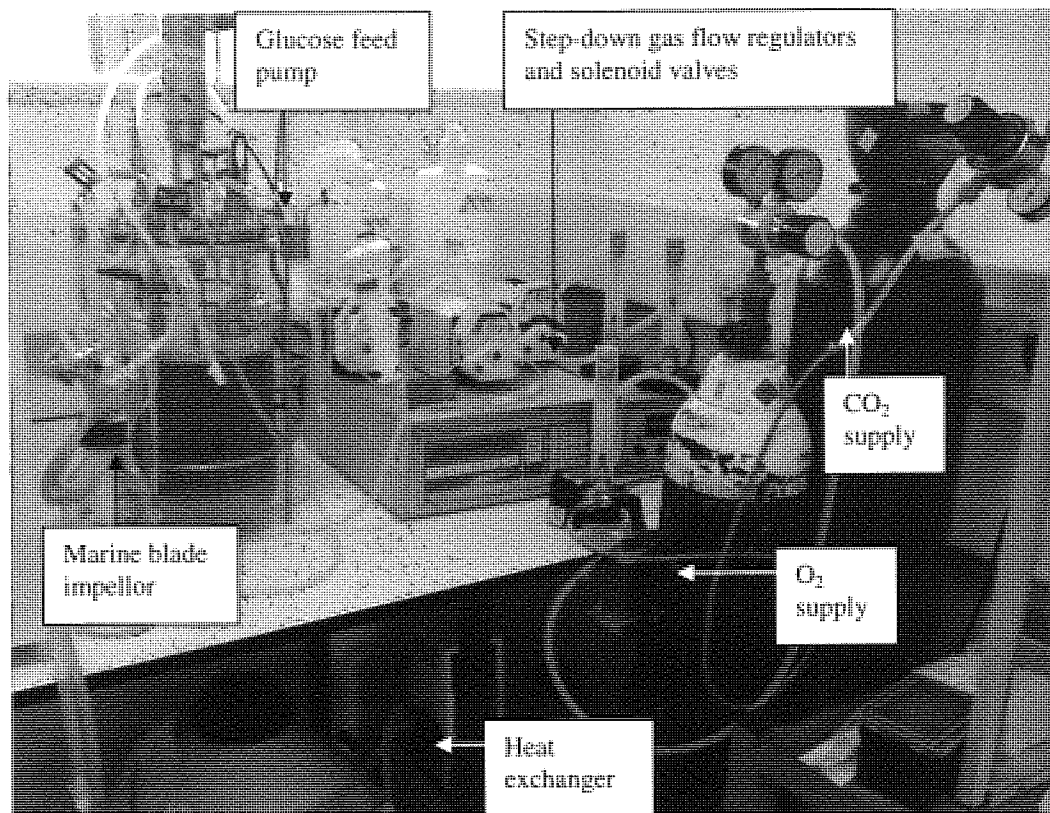


Figure 4.12 Set-up of 7 L (total volume) stirred tank bioreactor used for the mammalian cell culture of the hybridoma HD1 cell-line for DoE3 scale-up experiments.

4.1.7.2 DoE3 scale-up experiments

Demonstrating that the DoE3 model predicts the per cell yield from the hybridoma cell-line HD1 in bioreactor cultures was approached in the same manner as for the Micro 24. A number of check point conditions were selected: standard mammalian cell culture conditions (37 °C, pH 7.0, 2.5 g/L glucose and 40 % DO), DoE3 derived optimised conditions (38 °C, pH 6.7, 10.0 g/L glucose and 40 % DO) and DoE3 best small-scale experimental fit (36.5 °C, pH 6.9, 8 g/L glucose and 40 % DO). The number of check point conditions used at the large scale was lower than when testing at the small-scale in the Micro 24. As the model had already been validated at the small-scale it was proposed that any failure in predictive capacity at the large scale would be identified by the widely spaced scale-up check points which spanned the DoE3 design space. The scale-up experimental runs are detailed in Table 4.7.

The DoE3 scale-up experiments used a bi-phasic process with the check point conditions run for the first 4 days to provide comparative data for model scalability. The process conditions were then switched to the DoE3 predicted optimal conditions for the second phase of the experiment. This bi-phasic process is more representative of a production experiment where

cell biomass is generated using the optimal growth conditions (standard conditions) and then optimal expression conditions used to gain maximal product from the high biomass. In such a process, normalised (per cell) and volumetric yields can be maximised giving a higher total yield than possible by either optimisation process alone. As BR02 had DoE3 optimised conditions as the check point, these were to be maintained during the second phase of the experiment.

Each of the DoE3 scale-up experiments (BR01-BR03) were monitored to provide culture profiles for each experimental run, with viable cell density ($\times 10^5$ cells/mL) and glucose concentration (g/L) measured off-line whilst culture temperature ($^{\circ}\text{C}$), DO (%) and pH were on-line measurements but logged manually. Data logging software (BioXpert Lite, Applikon Biotechnology) was used so the overnight “data gaps” could be avoided, providing a more complete picture of the process. Figure 4.13 shows the culture profiles for each of the runs detailed in Table 4.7.

Culture samples were taken at the point of inoculation, end of Phase 1 and at harvest (end of Phase 2) for antibody yield determination by Protein-A analysis. As expression is constitutive, there was product carry-over from the inoculum culture. The time zero sample therefore served as a blank for each protein expression experiment.

Each experiment was inoculated at a viable cell density of $1.1\text{-}1.8 \times 10^5$ cells/mL. A higher seeding density would ideally be used e.g. 5×10^5 cells/mL, but this was not possible due to the volume of seed culture available and the minimum operating volume of the bioreactor.

During BR01 the Phase 2 optimised conditions were maintained for only 2 days as the culture became contaminated. This can be seen in Figure 4.13 as a drop in DO at approximately 135 h. A supernatant sample was taken at harvest (culture termination) to determine antibody production from the second phase of BR01.

In BR02 a drop in culture temperature can be seen prior to inoculation (during the overnight equilibration). This was due to a leak from a tube connecting the heat exchanger and the vessels water jacket. This was fixed and the culture temperature returned to set-point before the vessel was inoculated.

The cell density at point of inoculation of BR02 was 1.1×10^5 cells/mL. This appears to be too low to allow cell growth under the optimised expression conditions, indicated by the post inoculation drop of cell density (cells/mL) and cell viability (%). Cells were 85 % viable at inoculation which decreased to 24 % at 23 h post inoculation and 6.8 % at 48 h. The decision was taken to end the run prematurely as the cells were dead and the process was not reproducing the small-scale growth curves.

The BR03 culture was grown for 4 days with supernatant samples taken at the point of inoculation and at the end of the Phase 1. The process conditions were then altered to the DoE3 optimised set-points and the run continued for 3 days. The culture was harvested after ~218 h, 3 days post-shift to DoE3 optimised conditions as a failure with the feed-back loop had lead to a loss of pH control, reaching pH 7.2 during an unmonitored overnight period. Cell viability had decreased to 28.7 %, with 30 % being the standard cut-off limit for maintaining a culture.

Run name	Run description	Phase 1 conditions	Phase 1 duration (days)	Phase 1 seeding density ($\times 10^5$ cells/mL)	Phase 1 maximum density ($\times 10^5$ cells/mL)	Phase 1 end density ($\times 10^5$ cells/mL)	Phase 1 total antibody yield ($\mu\text{g/mL}$)	Phase 1 normalised antibody yield ($\text{ng/mL}/10^5$ cells)	Adaptation phase (Y/N)	Phase 2 conditions	Phase 2 duration (days)	Phase 2 starting density ($\times 10^5$ cells/mL)	Phase 2 maximum density ($\times 10^5$ cells/mL)	Phase 2 end density ($\times 10^5$ cells/mL)	Phase 2 total antibody yield ($\mu\text{g/mL}$)	Phase 2 normalised antibody yield ($\text{ng/mL}/10^5$ cells)
BR01	Bi-phase Day 0-4 standard conditions Day 4-5 DoE3 optimised conditions	37.0 °C, pH 7.0, 2.5 g/L glucose, 40 % DO	4	1.8	13.8	10.4	0.99275	95.45663462	Y	38.0 °C, pH 6.7, 10 g/L glucose, 40 % DO	2	11.6	11.6	9.7	1.7	177.0
BR02	Day 0-2 DoE3 "optimised"	38.0 °C, pH 6.7, 10 g/L glucose, 40 % DO	2	1.1	1.1	0.4	N/A	N/A	N	N/A	N/A	N/A	N/A	N/A	N/A	N/A
BR03	Bi-phase Day 0-4 DoE3 best small-scale experimental fit conditions Day 4-7 DoE3 optimised conditions	36.5 °C, pH 6.9, 8 g/L glucose, 40 % DO	4	1.8	16.1	14.1	1.50066	106.4296454	Y	38.0 °C, pH 6.7, 10 g/L glucose, 40 % DO	3	14.1	14.1	8.6	-1.1	-130.9

Table 4.7. DoE3 scale up experimental run descriptions and yield data. Note that run BR02 was terminated following 48 h due to cell viability dropping to 6.8 %. Negative yield is seen in BR03 Phase 2 possibly due to protease release from dying cells.

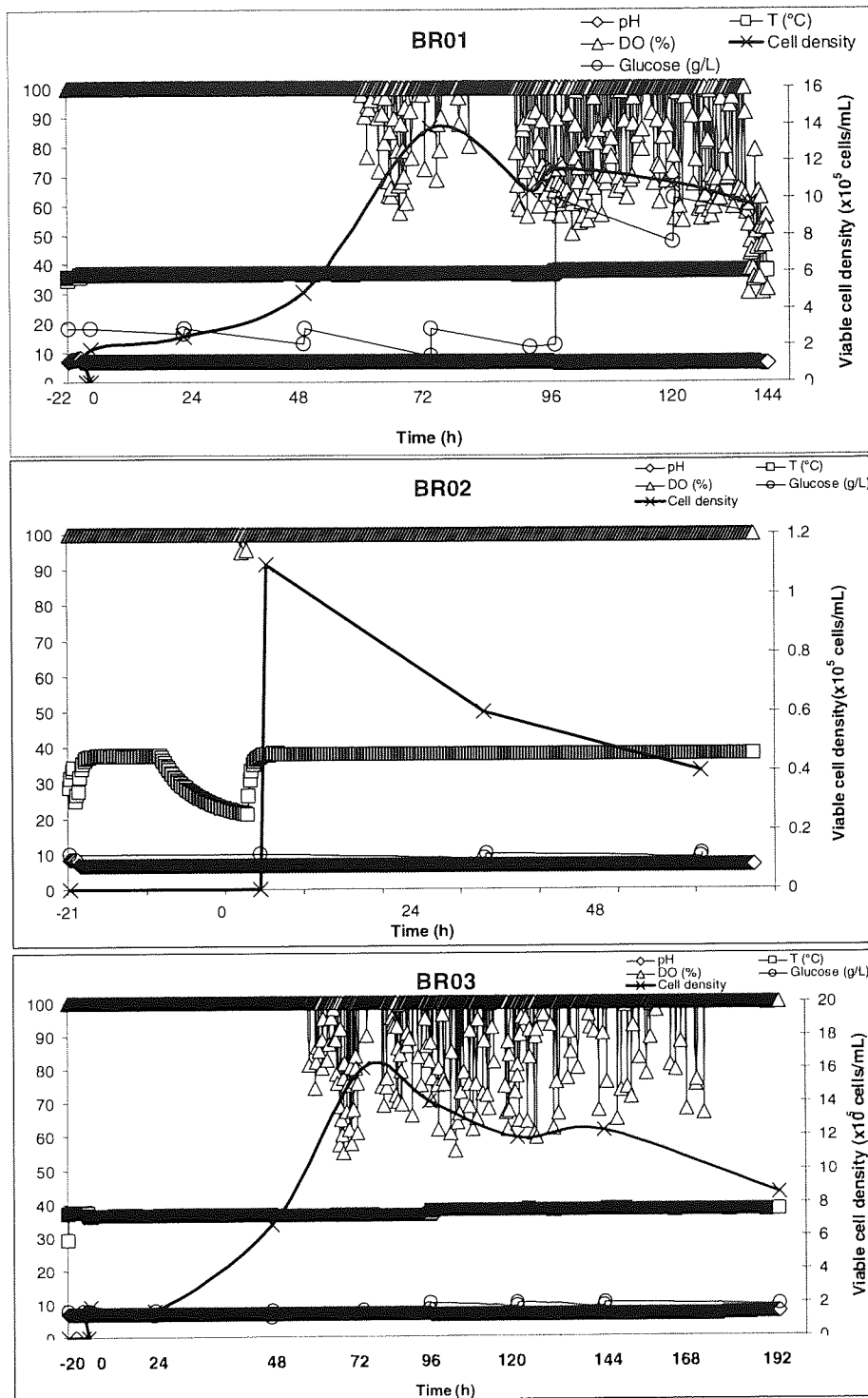


Figure 4.13. DoE3 scale up culture profiles for fed-batch hybridoma (hybridoma HD1) cell culture experiments, using the conditions indicated by Table 4.7. Where culture temperature = T ($^{\circ}$ C), dissolved oxygen = DO (%), glucose concentration = glucose (g/L) and biomass = viable cell density ($\times 10^5$ cells/mL). Culture biomass is shown on the secondary y-axis.

4.1.7.3 DoE3 scale-up antibody yield results

The supernatant samples taken during the DoE3 scale-up experiments (BR01, BR02 and BR03) were assayed for antibody concentration (yield) at both day 4 (end of scale-up phase of the run) and at point of harvest (end of process development phase). The BR02 experiment did not show any expression. It is thought the low cell density at inoculation coupled with conditions not well suited to cell growth (DoE3 “optimised” expression conditions) resulted in the cell death and lysis leading to proteolysis. Figure 4.14 shows the antibody yield data from BR01 and BR03. These experiments were bi-phase processes, therefore have two yields associated with them. The yield from the BR01 standard conditions (37 °C, pH 7.0, 2.5 g/L glucose and 40 % DO) was 0.993 µg/mL, which is similar to the small-scale (1.051 µg/mL) but higher than the DoE3 (0.150 µg/mL) predicted yields. BR03 “best small-scale experimental conditions” gave an antibody yield of 1.500 µg/mL, higher than either the small-scale (1.126 µg/mL) and DoE3 predicted (0.204 µg/mL) yields.

BR01 demonstrates that yield can be increased by using a bi-phase process, the 2 day period using DoE3 optimised conditions gave a higher yield (1.717 µg/mL) than either “standard” (BR01) or “best small-scale experimental fit” (BR03) conditions over 4 days. Similar increases in yield were not seen during the DoE3 optimised conditions phase of BR03. It is thought that the loss of pH control towards the end of the culture resulted in both cell death and product protein degradation. As shown by Figure 4.14 the yield from the second phase of BR03 (alternative conditions) was lower than the sample used as a blank, therefore giving a negative yield. This decrease in detectable product is likely to be due to proteolytic degradation by enzymes released following cell death.

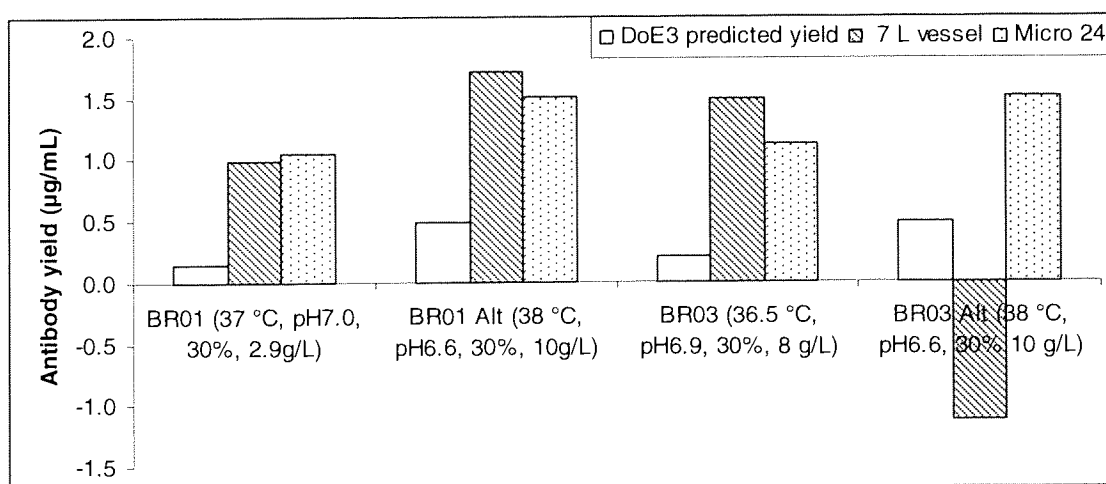


Figure 4.14. Antibody yields from the bi-phase DoE3 scale-up experiments BR01 and BR03.

4.1.8 Conclusions on DoE3 model scalability

In this Chapter the methods of process optimisation (designed in Chapter 3) were tested using an alternative target protein and expression system. The methods were shown to be transferable and capable of predicting antibody yield and optimal expression conditions from hybridoma HD1 (hybridoma) cultures at the small-scale. The DoE3 scale-up process highlighted some deficiencies in the process, most notably seeding density (as seen in BR02). These scale-up experiments also demonstrated the benefit of bi-phase conditions, where optimal conditions for growth are used to generate biomass followed by optimal expression conditions to provide maximal per cell yield. The bi-phase conditions used in BR01 could be further developed to give a higher yielding process.

4.2 Examination of additional industrially-relevant target proteins to the DoE process in yeast

The initial choice of target proteins was dependent on the protein having a proven record of expression at an industrial scale and having examples in the literature using *P. pastoris* and *S. cerevisiae* expression systems. The target proteins selected were insulin (MPI), human placental growth hormone (hPGH), and a monoclonal antibody-based influenza vaccine. The strategy is outlined in Figure 4.15. Detailed sequence information relating to the construction of each of the expression vectors can be found on the attached data CD.

The quantity of each complete expression vector was increased by transformation of competent *E. coli* cells. This is possible as the plasmids contain origins of replication and selection markers (*e.g.* ampicillin resistance) compatible with *E. coli*, and are hence referred to as shuttle vectors. Subsequent shake flask culture of positive transformants and DNA purification were used to generate sufficient plasmid DNA for analysis and yeast transformations. Glycerol stocks of each *E. coli* transformant were stored at -80 °C for long term storage.

Prior to use in yeast transformations, samples of each plasmid were taken and subjected to DNA sequence analysis to confirm the fidelity of the expression cassette. Those that showed deviations from the reference sequence were not used in any further work. The plasmids which had the correct sequence were used for transforming yeast cells.

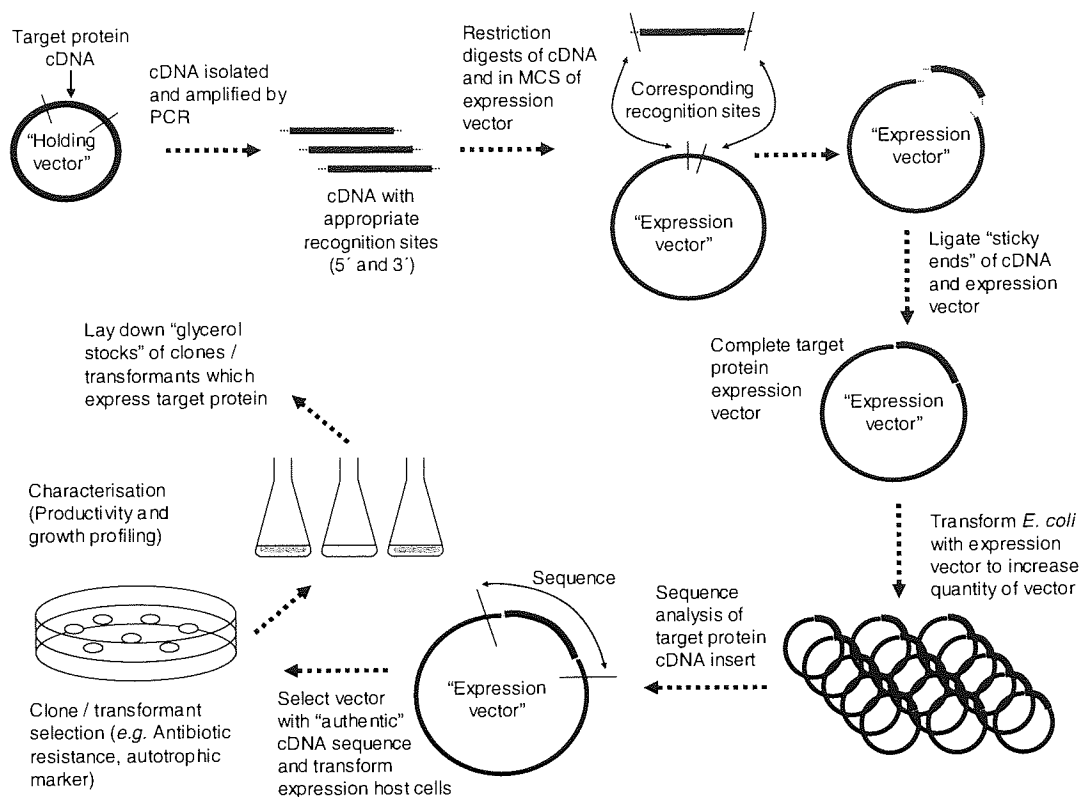


Figure 4.15. Schematic of the clone creation strategy. The cDNAs for each of the target proteins were provided in “holding” vectors suitable for cDNA propagation in *E. coli*. The coding sequence was removed by PCR using primers designed to introduce recognition sites 5′ and 3′ for restriction enzymes found in the multiple cloning site (MCS) of the expression vector. The cDNA and expression vector were then ligated together and used to transform *E. coli* cells allowing the production of increased quantities of the complete expression vector. The target protein insert was sequenced to ensure authenticity and then used to transform expression host cells. Transformants/clones were screened for vector up-take and product expression. Positive transformants/clones had glycerol stocks generated and kept at $-80\text{ }^{\circ}\text{C}$ for long term storage.

4.2.1 Insulin

Insulin is a suitable target for optimising protein production in yeast as it is a soluble, secreted protein produced by industry in yeast expression systems. However, full length insulin cDNA from *Homo sapiens* is not efficiently expressed in yeast (*S. cerevisiae* or *Pichia pastoris*).⁴² Several modifications are required, most importantly the removal of the C-peptide sequence and either direct fusion of the B and A chains or the insertion of a mini C-peptide (AAK). This gives a mini-proinsulin (MPI) construct which is efficiently expressed in yeast.⁴⁰ MPI is the form of insulin which is used in this project.

The attachment of a secretion factor, the α -mating factor, aids efficient secretion of the MPI into the culture medium. The expression of insulin analogues in *P. pastoris* and *S. cerevisiae* was documented by Kjeldsen *et al.*⁴² They investigated how pro-leader and C-peptide

replacement sequences impact on expression levels. By using the *S. cerevisiae* α -mating factor they obtained secreted proinsulin from both yeast strains. The α -mating factor is used in the pPICZ α (Invitrogen) vector series to confer secretory expression to the “expression cassette”.

It has been shown that the C-peptide of the insulin precursor (A chain, C-peptide, B chain) can be replaced with a short linker sequence or completely removed and still result in correct disulphide bond formation between the A and B chains.³⁷ In a later paper Kjeldsen *et al*¹²² demonstrated that a tri-peptide linker can be used to successfully express proinsulin in *S. cerevisiae*. These shortened proinsulin proteins are referred to as mini-proinsulin (MPI).

As an extensive modification of the original sequence was required (C-peptide substitution); two tags (c-myc and His₆) were attached to aid detection and purification of the expressed protein.

Construction of an expression cassette containing MPI with a secretion factor and tags required several steps; manipulating the native human sequence using a combination of polymerase chain reactions (PCR), specific restriction enzyme digests and ligations. Construction of the three expression vectors pPICZ α A-MPI(F), pYES2- α -MPI(F) and pYX212- α -MPI(F) follows the strategy outlined in T Kjeldsen,⁴² who expressed insulin precursors in a *S. cerevisiae* system.

The molecular biology strategy used is outlined below:

- Removal of pre-B chain and post-A chain sequences and introduction of *EcoRI* (5') and *XbaI* (3') sites to insulin sequence (manipulations carried out by Round 1 PCR)
- Ligation of PCR Round 1 product (*EcoRI*-Insulin-*XbaI*) with pUC19 (2686 bp, Fermentas)
- C-peptide replacement by reverse PCR (PCR Round 2) to give MPI
- Attachment of α -factor, c-myc epitope, His₆ tag and stop codon to MPI by ligation into pPICZ α A, producing the α -MPI expression cassette and the expression vector pPICZ α A-MPI(F)
- Addition of *HindIII* and *SacI* restriction sites to 5' and 3' ends of α -MPI (PCR Round 3)

- Ligation of α -MPI into *S. cerevisiae* expression vectors; pYX212 and pYES2 using the *Hind*III and *Sac*I sites producing expression vectors pYX212- α -MPI(F) and pYES2- α -MPI(F).

Using Clone Manager 5 software, diagrammatic illustrations of the three expression vectors were created, as shown in Figure 4.16.

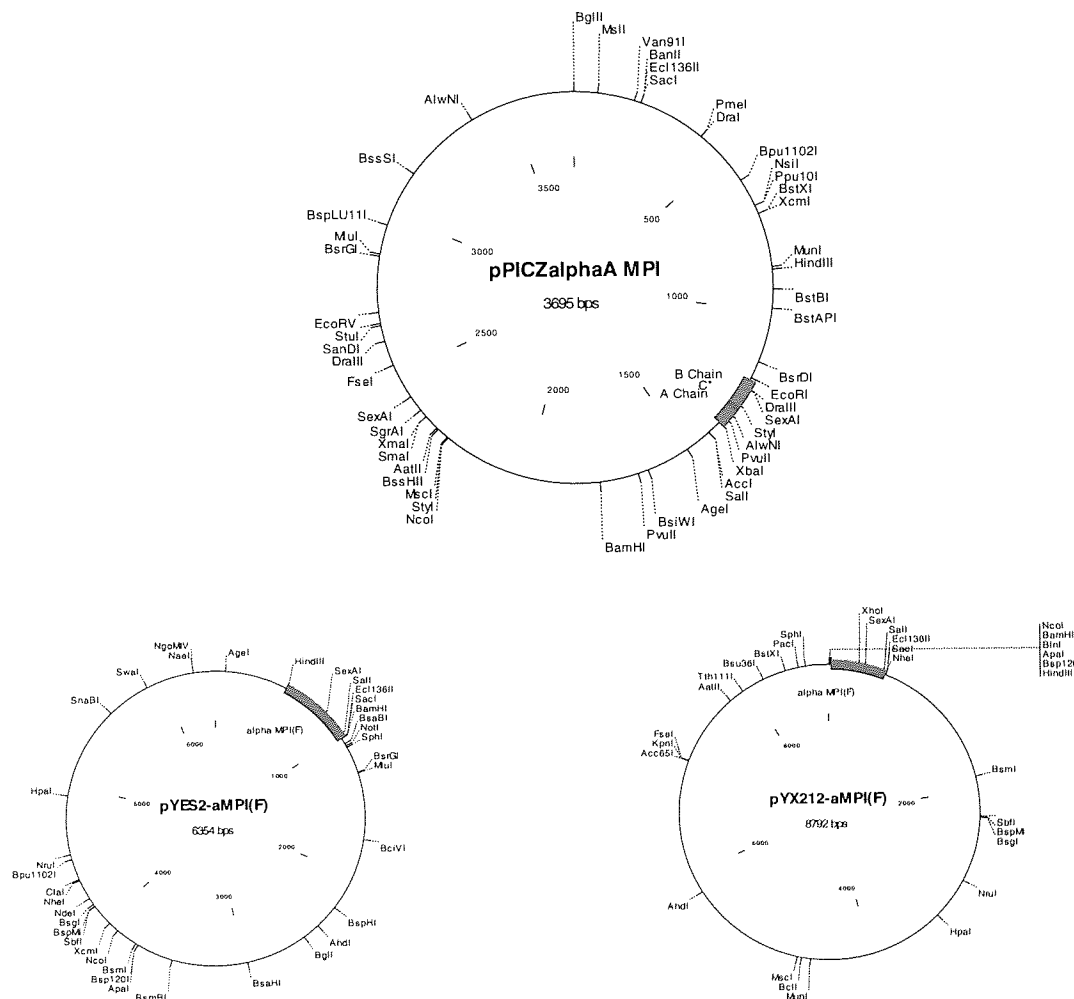


Figure 4.16. Representations of the three expression vectors used for the production of secreted MPI in *P. pastoris*, *S. cerevisiae* Alcofree and *S. cerevisiae* W303-1A

After electroporation the cells were plated on YPDS agar with Zeocin and incubated until colonies were visible; 10-20 colonies were re-streaked on fresh Zeocin plates to ensure cultures for production screening were from clonal cells.

A transformation was carried out using competent *P. pastoris* GS115 and X33 cells with 5 μ L and 10 μ L *Sac*I linearised pPICZ α -MPI(F). This yielded no colonies on selective agar after incubation for 6 days. This was repeated using *Bst*XI-linearised pPICZ α -MPI(F); no

colonies formed with the GS115 transformations, while the X33 selection plates grew to “lawn coverage”. Cells on the X33 selection plate were resuspended in 2 mL YPD and diluted 1:10 in YDP; 20, 40 80 and 160 μ L of the suspension streaked out on fresh selection plates. No colonies formed after 8 days of incubation.

A further transformation produced 20 colonies, which were re-streaked. Ten colonies (2-11) re-grew and were used to inoculate production screening cultures. SDS-PAGE analysis of the supernatants 90 hours post induction showed that clones 6 and 10 had strong 4 kDa bands on Quick stain (Coomassie type) visualisation. Western Blot analysis using an anti-His₆ HRP conjugated antibody only detected the His₆ tag positive control. It was thought the His tag may have been cleaved from the MPI so an anti-insulin antibody was used to repeat the analysis. This failed to highlight any product bands

It was noted that β -mecaptoethanol had not been added to samples prior to SDS-PAGE and Western Blot analysis, which might explain the negative results as few of the disulphide bonds may have been broken by the heat treatment leading to poor display of the His₆ tag or insulin antibody recognition sites, especially in the case of the samples with 4 kDa bands visualised by Coomassie stain. This highlighted the His₆ positive control but no product bands when using the anti-His₆ antibody and only the insulin B chain in the molecular weight marker (MWM) with an anti-insulin antibody

LiAc transformation of TM6 *ura3* cells with pYX212- α -MPI(F) resulted in >20 colonies; 10 of these were re-streaked and all re-grew on fresh selective agar plates (SD^{-ura} agar). Single colonies were picked and used to inoculate SD^{-ura} shake flask cultures and grown for 72 hours. Cell free supernatant was concentrated by acetone/ethanol precipitation resulting in a one hundred fold concentrate which was analysed by SDS-PAGE and anti-His₆ Western Blotting. Figure 4.17 shows images from the analysis. The low molecular weight (~4 kDa) band seen in panel A was initially thought to be product however it was not detected by Western Blot (Panel B). The Western Blot (panel B) shows a band at ~17 kDa and a number of higher weight bands (multiples of 17 kDa). This is thought to be dimer (17 kDa) and higher order aggregates, or perhaps the result of hyper-glycosylation. Clones 1 and 6 were selected for further study, as they showed higher yields than the other transformants.

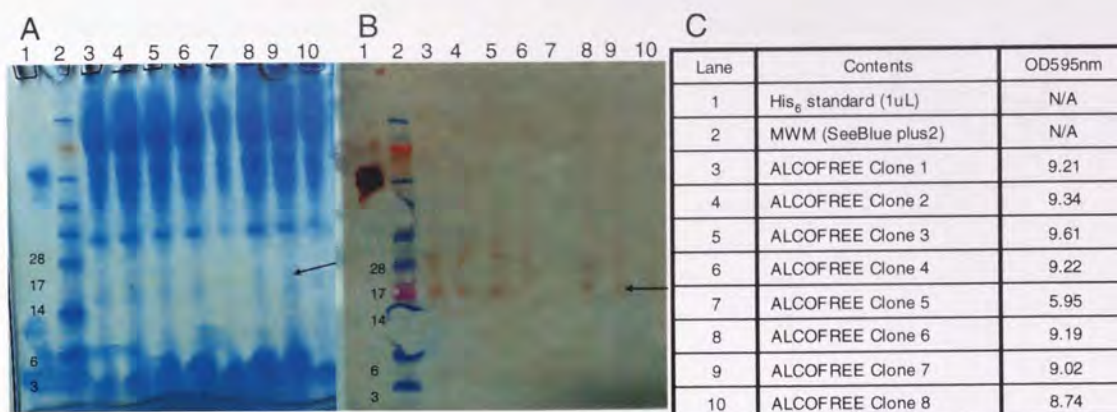


Figure 4.17. Alcofree MPI clone selection analysis of 100x concentrated supernatant. (A) shows a Coomassie stained SDS-PAGE image. (B) shows AEC stained Western Blot using anti-His₆ antibody. (C) describes the samples loaded in A and B. Also shown is the culture cell density (OD₅₉₅).

A time course experiment was set up to evaluate an alternative to SD^{-ura} medium; 2xCBS. The 2xCBS medium had been shown to give higher Alcofree biomass than SD^{-ura} (Chapter 5), the time course would also allow culture characterisation, growth profiling and product accumulation. Shake flask cultures were set up using baffled 500 mL shake flasks with 50 mL 2xCBS, OD₅₉₅ was monitored and supernatant samples taken at 24 h intervals over 4 days. In order to aid detection of secreted MPI, Silver staining was used as it is more sensitive than Coomassie staining and would remove the need for yield determination by Western Blot. Figure 4.18 shows the growth profiles (OD₅₉₅) for clones 1 and 6 during this time course experiment. Figure 4.19 shows an enhanced colour image of the coomassie stained SDS-PAGE gel with the supernatant samples taken during the time course. Figure 4.20 shows the Silver stain image of the SDS-PAGE gel. The ~17 kDa protein band is detectable in both cultures from 24 h and has an apparent maximal density at 72 h post inoculation in the clone 6 culture. The ~17 kDa band intensity in the samples from clone 1 are noticeably lower than those from clone 6. Therefore clone 6 was used in further work with cultures harvested around 72 h post inoculation under these conditions.

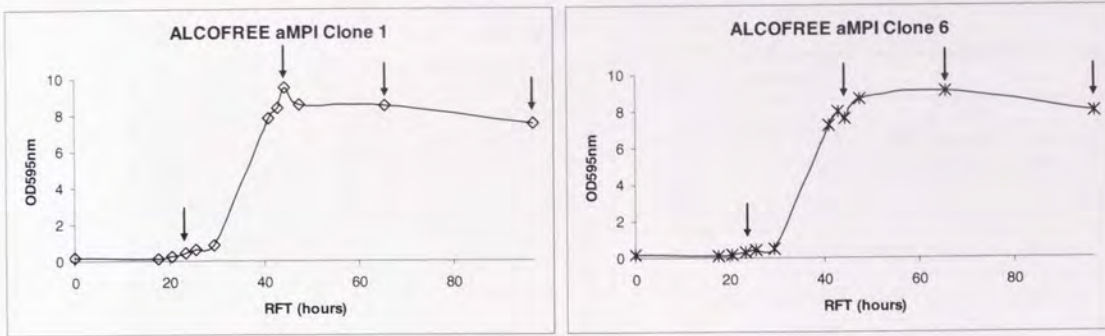


Figure 4.18. Growth curves of *S. cerevisiae* Alcofree pYX212- α -MPI clone 1 and 6 in baffled 500 mL shake flasks with 50 mL 2xCBS medium over 97 h (Relative fermentation time; RFT), arrows indicate supernatant sampling point.

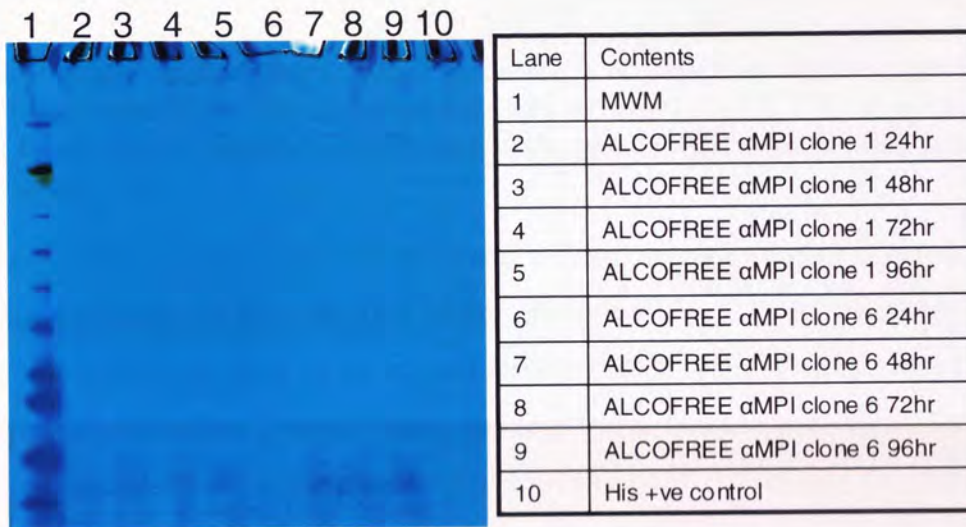


Figure 4.19. Enhance colour image of coomassie stained SDS-PAGE gel. Samples were taken at 24 h intervals during shake flask cultures (See Figure 6.18) of Alcofree MPI clones 1 and 6 to evaluate MPI expression. The supernatants were not concentrated as in Figure 6.17. The only visible band is the His₆ positive control (Lane 10)

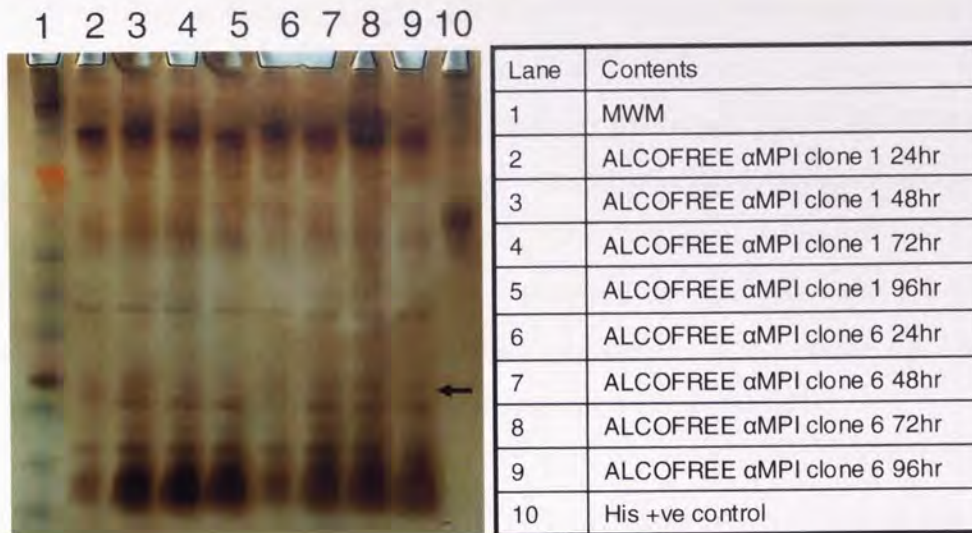


Figure 4.20. Silver stained SDS-PAGE gel. Samples were taken at 24 h intervals during time course shake flask culture (See Figure 4.18) of Alcofree MPI clones 1 and 6 to evaluate MPI expression. The supernatants are as Figure 4.19. The band at ~17 kDa, indicated by the arrow, increases in density with culture time in both clone 1 and clone 6 supernatants.

The MPI (~17 kDa) protein band was detectable without acetone/ethanol concentration if the silver staining method was used in place of Coomassie staining. The use of silver staining for yield determination is not ideal as an internal standard must be included in every SDS-PAGE gel used for quantification. Despite silver staining being quicker to perform than Western Blotting it is not a high throughput method, also it becomes challenging for more than four silver stains at once. Care must be taken when “developing” the gel as over-development results in a decrease in colour intensity of the protein band. This would complicate or even negate protein quantification by band densitometry.

As a positive *P. pastoris* MPI clone was not produced, it was not possible to carry out a comparison of MPI expression between the Alcofree strain and *P. pastoris*. The Alcofree pYX212-α-MPI(F) work detailed here indicates that increased biomass would be needed to easily detect a secreted recombinant protein when constitutive expression is utilised, as with the pYX2121 plasmid system. The failure to generate a positively expressing MPI *Pichia* strain, indicates that strain creation is not as simple as the literature, especially Invitrogen’s *Pichia* expression handbook, would suggest.

4.2.2 hPGH

Placental specific human growth hormone was regarded as a suitable choice as a target protein as it is a secreted soluble protein with a history of recombinant expression in the both

yeasts^{55, 123, 124} and at an industrial scale in *E. coli*.^{43, 125} Unlike insulin, hPGH did not require any modifications to the cDNA from *Homo sapiens* for efficient expression in either yeast system used here.¹²⁶ A similar approach was taken to that employed in the creation of MPI expression vectors. The hPGH cDNA was removed from the holding vector and inserted into the pPIZ α A plasmid, so creating pPIZ α A-hGH2. This was used to generate a clonal *Pichia Sp.* cell-line expressing secreted hPGH. The expression cassette was copied from pPIZ α A-hGH2 and inserted into the Alcofree expression plasmid pYX212 creating pYX212-hGH2 and used to generate Alcofree transformants expressing secreted hPGH.

The sequencing analysis of the completed hPGH expression vectors showed that pPICZ α A-HGH2 clone 1 and pYX212-HGH2 clone 11 contained an authentic hPGH coding sequence, which were used for all yeast transformations where appropriate. After electroporation the cells were plated on YPDS agar with Zeocin and incubated until colonies were visible; 10-20 colonies were re-streaked on fresh Zeocin plates to ensure cultures for production screening were from clonal cells.

The initial electroporation with X33 cells used *Pme*I linearised “clone 1” pPICZ α A-HGH2 and produced 20 colonies on 100 μ g/mL Zeocin YPDS agar. No colonies came through on the higher (500, 1000 and 2000 μ g/mL) Zeocin concentration plates, suggesting low copy number integration events. All 20 colonies were re-streaked and re-grew on 100 μ g/mL Zeocin YPDS agar plates.

The X33-HGH2 colonies which re-grew may only have contained the Zeocin resistance gene from the expression vector and not the entire sequence. To verify if the full vector, including expression cassette, was integrated the clones were subjected to a selection screen. A single colony of X33-HGH2 clone 1 was used to inoculate 50 mL BMGY in 500 mL baffled shake flask and grown overnight. This seed culture was then used to inoculate 50 mL BMMY medium in 1 L baffled shake flask and incubated with methanol induction for 48 h. Samples were taken at 24 and 48 h post induction and analysed by SDS-PAGE and anti-His₆ Western Blotting. No product bands were visible or detected in either the Coomassie stained gel or by Western Blotting.

Other X33-HGH2 clones (6, 7, and 8) were isolated and screened in a similar manner but over an extended time course induction of 91 h. No product (hPGH) was detected by either SDS-PAGE or Western Blotting. Only the His₆ positive control was detected. Further screening

was not pursued due to success with creating a GFP expressing cell-line. The remaining X33-HGH2 clones were therefore banked for future use.

The standard LiAc transfection method was used with Alcofree (TM6 ura3) cells and pYX212-HGH2 clone 11 plasmid DNA. This generated 20 colonies on the 2xCBS agar selection plate. All of these colonies were picked and re-streaked onto fresh 2xCBS agar plates; all grew. Single colonies from each re-streak were used to inoculate small shake flask expression screening cultures. Transformants 1 and 7 grew poorly so were not used in further work. At 72 h post inoculation supernatant samples from the remaining 18 cultures were assayed by SDS-PAGE and anti-His₆ Western Blot. No product or non-specific protein bands were visible on the Coomassie stained gel or by ECL detection of the Western Blot. The cultures were grown to 144 h post inoculation and again supernatant samples were assayed by SDS-PAGE and anti-His₆ Western Blot; no ECL signal was detected. To determine if hPGH had not been secreted into the culture medium, samples of 144 h cell pellets were subjected to three freeze (-20 °C) thaw (95 °C) cycles and the soluble fraction assayed by SDS-PAGE and anti-His₆ Western Blot; there were many non-specific protein bands but no product was detected.

Further attempts to create hPGH expressing *Pichia sp.* and Alcofree clones/transformants were not pursued. It was noted that the cell densities achieved in shake flask cultures of Alcofree were approximately half those obtained with *Pichia sp.* This is not ideal as the pYX212 vector used with the Alcofree strain confers constitutive recombinant protein expression; therefore, product titre is proportional to cell density. In an attempt to address this; optimisation of media composition was identified as a route to increasing Alcofree biomass and thus product yield and is discussed in Chapter 6.

4.2.3 Monoclonal antibody based influenza vaccine

The aim of this “proof of concept” project was to investigate an alternative expression system for a novel vaccine product being developed by a client of the industrial sponsor of the studentship, Alpha Biologics. If expression were feasible in yeast, then this production system would reduce the time and overall cost for production when compared to the baculovirus mediated insect cell culture system already in use. This work is subject to a confidentiality agreement (CDA) between the customer company, Alpha Biologics and Aston University. Due to this, aspects of the work have been anonymised for publication here with the customer

company's permission. The product protein was based on a monoclonal antibody with an undisclosed specificity.

The project aimed to generate *P. pastoris* clones with confirmed secretory expression of the target monoclonal anti-body. The following deliverables were to be provided from the project:

- 2 ampoules (glycerol stocks) of up to ten clones generated from transformations
- Agar plate stocks for the clones derived from the transformations
- Samples from small scale shake flask cultures from each clone for product screening
- Samples of cell paste and culture supernatant from larger (400 mL) cultures for each clone
- Western Blot data from the samples generated in the 400 mL shake flask cultures
- Technical report detailing the above work

The expression cassette for the monoclonal antibody was supplied in an appropriate plasmid, based on pPICZ α A, for transforming competent *P. pastoris* cells. The expression vector had been analysed for authenticity by the customer prior to delivery.

The supplied plasmid DNA was linearised with *Bst*XI, as there was only one recognition sites for *Bst*XI on the plasmid. This was then used to transform electro-competent *P. pastoris* cells and plated on selective agar containing Zeocin. 10 colonies were visible after 48 h which were re-streaked on fresh selective agar plates and incubated for 72 h. 6 of the 10 clones re-grew and had glycerol stocks layed down.

Single colonies of the 6 clones were picked from the re-streaked selection plates and used to inoculate small shake flask expression screening cultures. Induction was maintained for 72 h with samples taken for biomass determination by OD₅₉₅ and product analysis by Western Blot. Product analysis was carried out by the customer company; two clones were selected for further study.

The supplied plasmid DNA was used to transform Alcofree cells by the standard LiAc method. After 72 h incubation on selective agar plates 35 colonies were visible, these were picked and re-streaked on fresh agar plates. All colonies plated out were able to be cultured on the agar plates.

Single colonies from 10 re-streaked transformants were used to inoculate small shake flask expression screening cultures. These were incubated for 72 h; OD_{595nm} was measured at the point of harvest. Samples of supernatant and cell pellet were analysed by SDS-PAGE and Western Blot using appropriate antibody (supplied by the customer) detection and visualisation.

The first SDS-PAGE gel and Western Blot of the samples showed possible product in the supernatant and no visible product in the cell pellet. A repeat of this gel/blot was conducted but with a ten-fold sample concentration using an ethanol/acetone method. In addition a lower standard concentration and blank lanes to identify spill over from the standard. Visualisation was by staining using an AEC kit (Invitrogen) but only the applied standards were visible. The blot was repeated and visualised using an ECL method. Product was able to be visualised but only very faintly in the supernatants of clones 2, 3, 6 and 7.

The supernatant samples for clones 2, 3, 6 and 7 were concentrated one hundred-fold and re-run for Western Blotting. ECL detection showed a clear product band in all clones but visibly denser banding in clones 3 and 7. See Figure 4.21. Clones 3 and 7 were selected for 400 mL cultures.

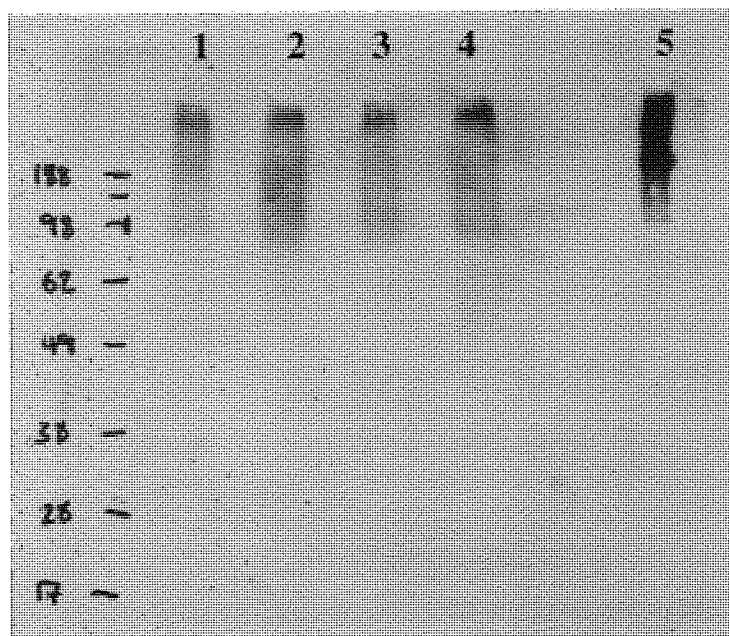


Figure 4.21. Image of a photographic film developed from ECL detection of a Western Blot. Samples are of 100x concentrated supernatants of *S. cerevisiae* Alcofree clones. Lane 1; Clone 2 supernatant, lane 2; Clone 3 supernatant, Lane 3; Clone 6 supernatant, Lane 4; Clone 7 supernatant, Lane 5; 0.1 μ L H5 Standard with blank lanes between each sample and 2 blank lanes between the samples and standard. The positions of the molecular weight markers have been added for reference.

It was noted that the apparent molecular weight of the product band expressed in the Alcofree system was higher than that of the positive control supplied by the customer. It was thought this may be the result of increased glycosylation over that of the control expressed in a mammalian cell-line. Another explanation may be the aggregate formation during the actone/ethanol concentration step. The use of high cell density cultures (fermentations) to express the protein at levels detectable without concentration would show if this were the case.

The customer company required additional material for characterisation of the product identified during the clone/transformant selection process. Larger shake flask cultures were used to express product from clones/transformants selected by the customer.

Pichia pastoris “production” shake flask cultures were set up in baffled 2 L shake flasks with 400 mL BMMY and induced for the two clones selected by the customer company. The cultures were induced for 75.5 h. Samples (5 mL) were taken during the time course of the cultures; these and all supernatants and cell pellets at harvest were sent to the customer for analysis.

Alcofree “production” shake flask cultures were set up in baffled 2 L shake flasks with 400 mL SD^{ura} for the two clones selected by the customer company. The cultures were grown for 72 h. Samples (5 mL) were taken during the time course; these and all the culture supernatant and cell pellets at harvest were sent to the customer for analysis.

The growth profiles of the 400 mL production cultures (*P. pastoris* and Alcofree) were comparable with previous cultures. The use of baffled flasks allowed higher biomass accumulation to be reached than with un-baffled flasks, this is due to improved mixing, and so, increased oxygen transfer to the medium.

A detailed technical report was written on the above work and submitted along with all required deliverables to the customer company. A second project was secured, on the basis of the work here, to evaluate an alternative vector construct for expression using *P. pastoris*. Transformation by electroporation was used but failed to produce any colonies on selective (zeocin) agar plates, this was repeated but was again unsuccessful. The customer is currently reviewing the use of yeast as an alternative expression system.

4.3 Conclusions from yeast strain creation

This work highlights two main points for the using *P. pastoris* and the Alcofree strain for the expression of industrial proteins: generating positively expressing clones/transformants is not a straightforward process and can be highly time consuming. When using systems with constitutive expression it may be important to have high biomass levels to produce the required amount of target protein.

Despite problems with generating positively expressing clones, the methods used were deemed by colleagues to be appropriate, especially as attempts by others were also unsuccessful. The cause of the poor transformation efficiencies is currently unknown. Areas for further investigation are plasmid/vector fidelity and parental cell-line authenticity. This would initially require further vector sequencing analysis and cell-line phenotyping.

The problem of low biomass (low expression) with the Alcofree system was identified as an area for further investigation. It was decided to investigate this by media design techniques as the SD^{-URA} and 2xCSM media were suspected to be sub-optimal for use with this strain of *S. cerevisiae*. Moving from shake flask cultures to a fully controlled bioreactor vessel could also aid in increasing achievable biomass, due to improved aeration and culture homogeneity. This is discussed in Chapter 6.

“There's a darkened sky before me
There's no time to prepare
Salvage a last horizon
But no regrets from me”

(No more lies, Iron Maiden)

CHAPTER 5

5 OPTIMISING ADDITIONAL FACTORS: INVESTIGATIONS OF THE INFLUENCE OF CELL SURFACE HYDROPHOBICITY AND ANTIFOAMS ON PRODUCTIVITY IN *PICHIA PASTORIS* CULTURES.

Many optimisation experiments described in the literature investigate obvious factors influencing protein expression such as culture temperature¹²⁷, pH⁶¹, aeration^{128, 129} and medium composition⁹⁰. However, the optimisation of many additional factors offers the possibility for further improvements to the overall process. Here we examined how cell surface hydrophobicity and antifoam agent use affect the characteristics of recombinant protein expression in yeast systems. It could be envisaged that if cell physiology changes (*e.g.* as indicated by a change in cell surface hydrophobicity) the protein production process could, perhaps, be manipulated to exploit this. In a similar way if simply using an alternative antifoam agent or alternative concentration of the selected antifoam can improve cell growth or product yield, this could prove a rapid and convenient way for improving the target protein yield.

It was hypothesised that cell surface hydrophobicity might be subject to change during the high cell density (HCD) culture of *Pichia pastoris*. At HCD cells are in close proximity to each other and it is expected that they will interact in a different manner to that seen at low cell densities. If this were the case, it might be possible to exploit cell surface hydrophobicity to improve the process yield.

There is a wide variety of antifoam agents available for use in microbial and cell culture applications. The mode of action may be the same for each antifoam type, but the molecules vary dramatically in their physiochemistry *e.g.*: oils, polyalkylenglycol, siloxane polymers and polypropylene glycol. It would be very surprising if these different molecules influenced cell growth and subsequent protein expression in the same way. The effect of each of the different types of antifoams on the cells and protein expression were to be examined. The study was trying to determine if any had a beneficial effect on the overall process yields.

Two secreted target proteins were used to determine changes in protein yield; an Fc-fusion protein and green fluorescent protein (GFP_{uv}). The Fc-fusion protein was used as it was industrially important and real protein being developed for use in clinical trials. GFP_{uv} was used due to its ease of detection and higher expression levels allowing small changes in

productivity to be observed. In addition the methods for GFP_{uv} expression had already been established in Chapter 3.

5.1 Investigation of cell wall hydrophobicity in high cell density *Pichia pastoris* cultures

Previous chapters have dealt with various aspects of recombinant protein expression involving the yeasts *P. pastoris* and *S. cerevisiae* as host systems. Here one facet of the host's physiology is further investigated: how the cell surface hydrophobicity (CSH) of the yeast cell wall changes when grown as HCD cultures.

If the CSH of *P. pastoris* cells changes during a culture this could provide an opportunity for improving the process yield. An example of a similar situation is flocculation in *S. cerevisiae* during the brewing processes. Nutrient supply can be used to control the onset of flocculation which affects the fermentation process as a whole and therefore the quality of the resulting beer.¹³⁰ If CSH can be used as a control factor in the production of recombinant proteins in *P. pastoris* then the parameters to cause the CSH change could be incorporated into a process to provide a mechanism of optimisation.

As demonstrated in Chapter 4 and Chapter 6 there have been limitations on the overall achievable yeast biomass when working with the *S. cerevisiae* strains. Although progress was made in improving the biomass, the cell densities were still lower those attained with *P. pastoris* cultures. Therefore, the work described here uses data from *P. pastoris* X33GFP_{uv} shake flask and HCD bioreactor cultures. When the problem of low biomass accumulation can be overcome with the *S. cerevisiae* Alcofree strain, determination of CSH during its culture should also be examined.

5.1.1 Cell surface hydrophobicity; Important aspects

The cell wall of any microorganism provides structural support and protection between the cell membrane and the extra-cellular environment. The structural support offered by the cell wall prevents unwanted swelling as a result of water ingress. The major components of the yeast cell wall are mannoproteins, 1,6- β -glucan, 1,3- β -glucan and chitin: this makeup is tightly controlled by the cell cycle¹³¹. These components directly affect the CSH as they are the outer surface of the cell. If the ratio of the cell wall components is changed, the CSH will

reflect this. This is useful as simple techniques are available to determine CSH¹³² from which cell wall composition may be inferred.

CSH has important consequences in an industrial setting as the hydrophobicity of the cells could affect the dynamics of a HCD culture, influencing factors such as cell aggregation, biofilm generation, nutrient availability and possibly on-line data acquisition (i.e. Supervisory Control and Data Acquisition or SCADA).¹³² If cells increase the hydrophobicity of their outer surface, they will gain a tendency to repel water. Therefore, any dissolved medium components have a reduced opportunity for reaching the cells. When hydrophobic entities are dispersed in an aqueous environment they tend to aggregate. A notable example of this is the formation of lipid bilayers,¹³³ where lipids form a polarised bi-layer with the hydrophilic head groups displayed towards the aqueous cytosol and the hydrophobic tails “hidden” from the water.

There are few references in the literature on the dynamics of CSH in HCD cultures of *P. pastoris*; many articles relate to the pathogenic yeast *Candida albicans* and its mode of attachment in the colonisation of human tissue.¹³⁴ The role of CSH in flocculation of *S. cerevisiae* cells has also been investigated and it was shown that CSH is a major determinant in flocculation, as a result of nutrient limitation, with hydrophobicity significantly increasing shortly before initiation of flocculation.¹³²

5.1.2 Determination of CSH

The two methods (MATH and HIC)¹¹⁰ used here for the calculation of relative CSH rely on the propensity of the cells to avoid water, i.e. transfer to the non-aqueous phase. This is the hydrocarbon (n-hexadecane) layer in the microbial attachment to hydrocarbon (MATH) assay or the chromatography matrix in the hydrophobic interaction chromatography (HIC) assay. The values provided by both assay methods are relative percentages of the cells in the sample which transfer to the non-aqueous phase. The cell density of the sample is measured (e.g. OD₅₉₅ or viable cell count) prior to the assay and compared to the cell density after the cells have had been exposed to the non-aqueous phase.

In the work presented here, the CSH was determined in non-induced and induced cultures of *P. pastoris* X33GFP_{uv} in both shake flask and bioreactor (HCD) cultures. Samples were taken at time points throughout the course of the culture to capture early, mid and late log phases of

the growth profile and several points post induction. The MATH assay was primarily used due to its high(er)-throughput nature, when compared to the HIC assay.

5.1.3 Preliminary CSH investigation with shake flask cultures

To gain some understanding of how CSH changes during the course of a *P. pastoris* culture an experiment was setup using shake flasks and two medium types: BMGY and BMMY. This aimed to highlight any differences between non-induced (BMGY) and induced (BMMY) cultures expressing GFP_{uv} and to generate a preliminary data set prior to data acquisition from HCD fermentations.

Two 1 L shake flasks containing 100 mL of medium (either BMGY or BMMY) were equilibrated and then inoculated to OD₆₀₀ 0.4 using an overnight YPD seed culture (typically at OD₆₀₀ 8-13). Samples were taken at 0, 6, 8 and 25 h post inoculation and OD₆₀₀ measurements were taken at each time point to follow the growth profiles of the cultures. The growth curves in Figure 5.1 indicate when samples for CSH determination were required to capture early/mid and late log phases; ~8 and 25 h post inoculation respectively.

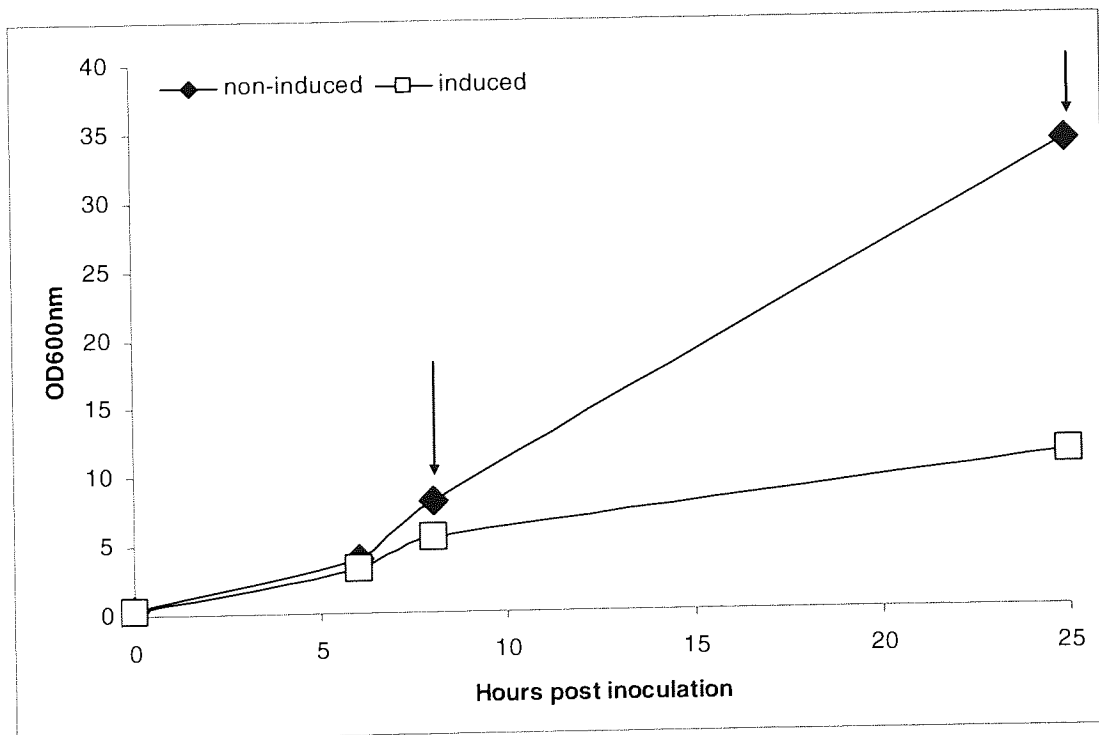


Figure 5.1. Growth curves (OD₆₀₀) for *P. pastoris* X33GFP_{uv} shake flask cultures grown in non-inducing (BMGY) and inducing (BMMY) media. Arrows indicate sampling points for CSH assays at 8 h (early/mid log phase) and 25 h (late log) post inoculation.

It is noticeable from Figure 5.1 that the non-induced culture reaches a higher biomass than the induced culture, as there is no metabolic load for recombinant protein expression. This is expected as in a protein production experiment the culture has two phases: growth phase using glycerol as the carbon source to generate biomass and then the production phase where methanol is used as both a carbon source and trigger for induction of expression.

Figure 5.2 shows the CSH results from the MATH assay for the non-induced *P. pastoris* X33GFP_{uv} shake flask culture grown in BMGY medium. The hydrophobicity measurement at 8 h (29.0 % relative CSH) is very similar to the measurement at 25 h (29.4 % relative CSH). The HIC assay was not used for these samples (non-induced) as the number of HIC columns was limited and we wanted to more-fully capture the effects of secreted protein (GFP_{uv}) expression from the induced culture. These results suggest that CSH is not affected by the growth phase of the culture over the two points sampled here. As this was a single culture, repeats would ideally have been performed to confirm this. However, as the culture did not reach a “high” cell density (for *P. pastoris*), subsequent experiments were performed in a controlled bioreactor where cell densities can be >400 (OD₅₉₅).

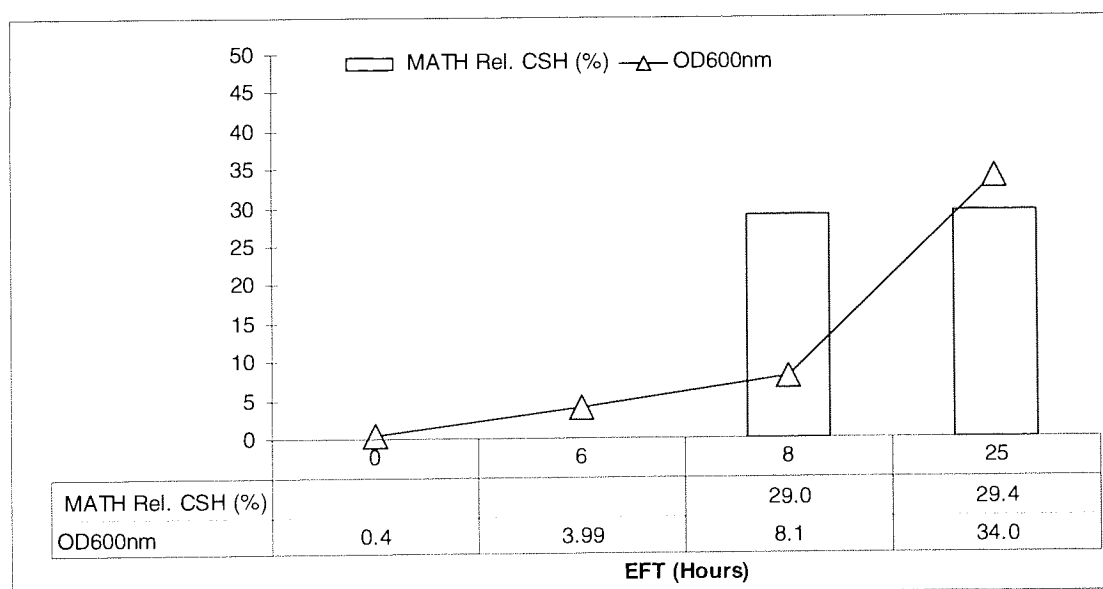


Figure 5.2. Non-induced (BMGY) *P. pastoris* X33GFP_{uv} shake flask culture. Relative CSH and cell density (OD₆₀₀) at early/mid and late log phases of growth curve are shown; 8 h and 25 h post inoculation respectively.

Figure 5.3 shows the CSH results from the MATH and HIC assays for the induced *P. pastoris* X33GFP_{uv} shake flask culture grown in BMMY medium. The hydrophobicity measurements at 8 h (40.4 % MATH, 0 % HIC) differ from the measurements at 25 h (14.6 % MATH, 7.2 % HIC). The two assays show very different trends over the time point sampled; CSH decreases according to the MATH results but increases according to the HIC assay.

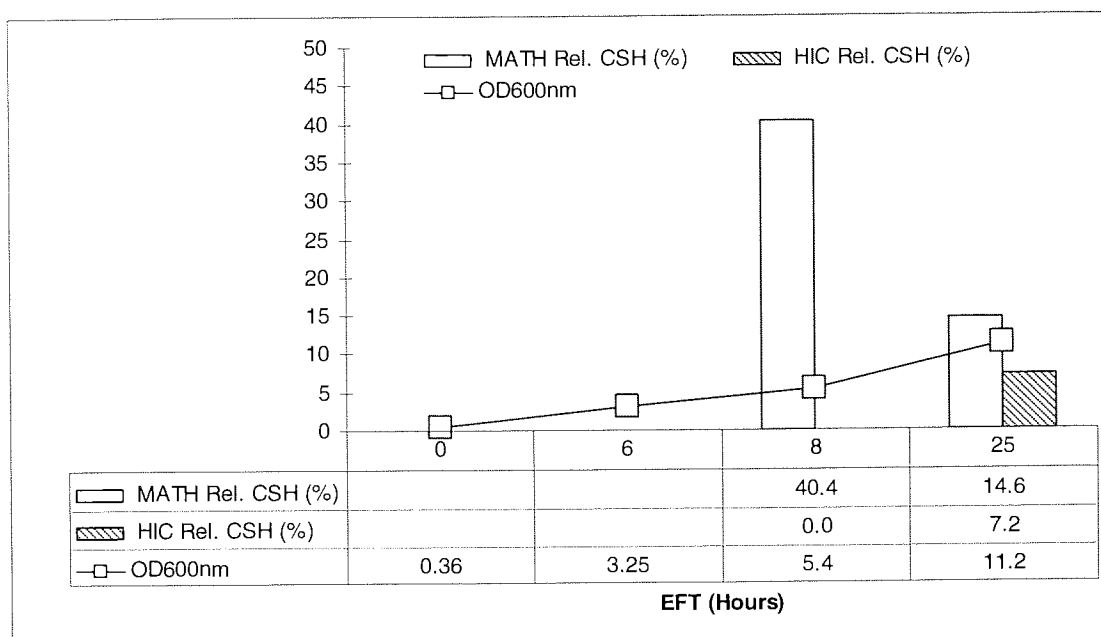


Figure 5.3. Induced (BMMY) *P. pastoris* X33GFP_{uv} shake flask culture. Relative CSH and cell density (OD₆₀₀) at early/mid and late log phases of growth curve are shown; 8 h and 25 h post inoculation respectively.

If the CSH of non-induced (Figure 5.2) and induced (Figure 5.3) cultures are compared, it can be seen that both non-induced and induced cultures are ~ 35 % (± 5) hydrophobic at 8 h post inoculation. The cell densities of the cultures are also similar: non-induced OD₆₀₀ 8.09, induced OD₆₀₀ 5.4. At 25 h post inoculation the non-induced CSH remains at 29 % whereas the induced culture drops to 14.6 % by MATH (7.2 % HIC). The cell densities at 25 h also show differences: the non-induced culture reaches OD₆₀₀ 34, induced culture OD₆₀₀ 11.2. The difference in cell density is as expected, with the non-induced culture reaching a higher density than the induced culture. The CSH change in the induced culture could be due to the protein expression (secreted protein trapped in the cell wall) or more directly linked to the methanol addition resulting in a physiological change to the cell surface.

As this was the first attempt at running these assays and no references were found in the literature on how CSH varies in *P. pastoris* cultures, repeats would be necessary to confirm whether the trends seen in the preliminary shake flask data are valid. However, further experiments were performed using a bioreactor as the cultures were more tightly controlled and capable of yielding higher cell densities (OD₅₉₅ > 100) than those possible in shake flasks (OD₅₉₅ 10-30). For all subsequent experiments relative CSH was calculated using the MATH assay as the HIC method required the time consuming preparation of chromatography columns and was therefore not pursued further.

5.1.4 High Cell Density 7 L bioreactor culture

The CSH values determined from *P. pastoris* X33GFP_{uv} shake flask cultures showed that CSH may vary with culture growth phase. However, the results were based on only one culture. The cell densities achieved in the shake flask cultures (see Section 1.2.1.1) could not attain the high levels reported in the literature of $OD_{600} > 500$ ¹ and from personal experience ($OD_{595} > 100$): these HCD cultures can only be achieved with the use of a controlled bioreactor, providing appropriate aeration, temperature and pH control. A 7 L jacketed glass vessel with an Applikon ADI1030 control unit and required ancillaries was therefore used to culture *P. pastoris* X33GFP_{uv} to cell densities (OD_{595}) > 100 . These bioreactor runs had fully controlled temperature and pH, but only upward DO control as the cell growth would use the dissolved oxygen and cause the DO to be decreased. The medium used for the initial batch phase was BMGY with non-standard glycerol concentrations (40 g/L not 10 g/L), which is equivalent to the glycerol concentration found in the Invitrogen fermentation BSM medium. The increase in carbon source (glycerol) used in the batch or growth phase had been found to yield higher achievable biomass than standard BMGY, whilst avoiding the problems of precipitation seen with BSM at pH's > 5 ,⁶⁴ which was discussed further in Chapter 3.

As in the shake flask experiments, samples were withdrawn from the cultures at time points covering early, mid and late log growth phases. In addition, samples were also taken during the fed-batch induction phase of the culture. The induced cells were expressing secreted GFP_{uv} under the control of the AOXI promoter, induced by methanol (or methanol/sorbitol mixed feed) addition.

Three fermentations were sampled for CSH determination. These were part of the design of experiments scale-up process (Chapter 3), which involved 10 fermentations in total and as such had different process set-points post-induction. The conditions during the batch growth phase, however, are the same for each of the fermentations (30 °C, pH 6.0, 30 % DO). For each run the vessel was equilibrated overnight to the growth phase set-points to determine sterility of the medium pre-induction. The starting volume for the fermentations was 3 L (BMGY medium) 40 g/L glycerol with P2000 antifoam added to 0.07 % (v/v). Samples were withdrawn aseptically using the sample port device to determine biomass (OD_{595}) and CSH (MATH assay).

Figure 5.4 shows the on-line and off-line measurements for the 7 L bioreactor run using model-predicted optimal induction/expression conditions (21.5 °C, pH 7.6, 90 % DO).

Induction was by a slow (0.167 mL/min) 100 % methanol feed with 12 mL/L PTM₁ trace salts. The sample points for the CSH assay are indicated by the arrows in the growth (batch) phase at early, mid and late log phases of the growth curve: 4, 19 and 26 h post inoculation respectively. The induced (fed-batch) phase was sampled at early, mid and late induction: 44, 52 and 79 h post inoculation (15, 23 and 50 h post-induction)

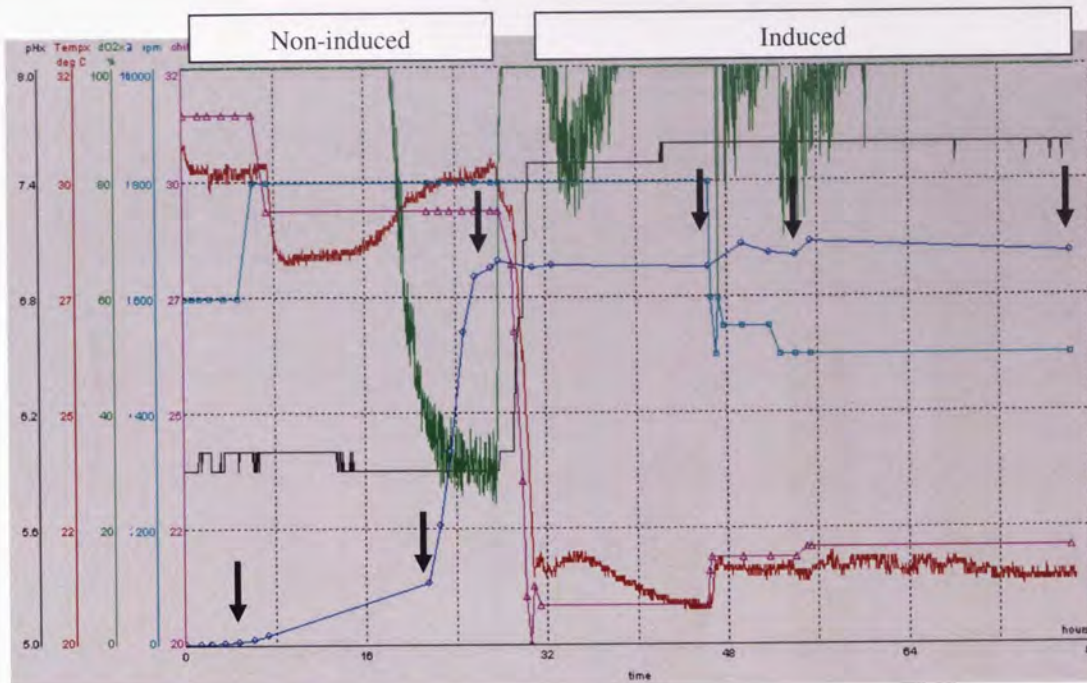


Figure 5.4. Trace from BioXpert Lite monitoring software of *P. pastoris* X33GFP_{uv} 7 L bioreactor culture with model predicted optimal expression conditions as described in Chapter 3. The initial set points for the non-induced batch phase were 30 °C, pH 6.0, 30 % DO. This was followed by starvation and adaptation phases prior to induction with methanol (fed-batch) with set-points 21.5 °C, pH 7.6, 90 % DO. Arrows indicate sample points for CSH assay.

Figure 5.5 shows the biomass (OD₅₉₅) and CSH (%) measurements from time points indicated in Figure 5.4. It can be seen that during the growth (batch) phase, relative CSH increases with biomass/progression through early (0.6 % CSH), mid (4.1 % CSH) and late log phases (20.2 % CSH). Once the culture is induced with the methanol feed, CSH decreases (9.2 % CSH at 15 h post induction) and then gradually increases with time post induction (13.5 % CSH at 23 h post induction, 15.8 % CSH at 50 h post induction). This could be due to either a reaction to methanol accumulation in the culture or a result of protein (GFP_{uv}) accumulation in the cell wall. To determine the cause of the gradual increase in CSH post induction an alternative induction strategy was used to avoid methanol accumulation. This was based on the initial work aimed at improving GFP_{uv} yield as described in Chapter 3. If methanol accumulation were occurring the cells would be likely to be affected, which could result in a change in cell CSH.

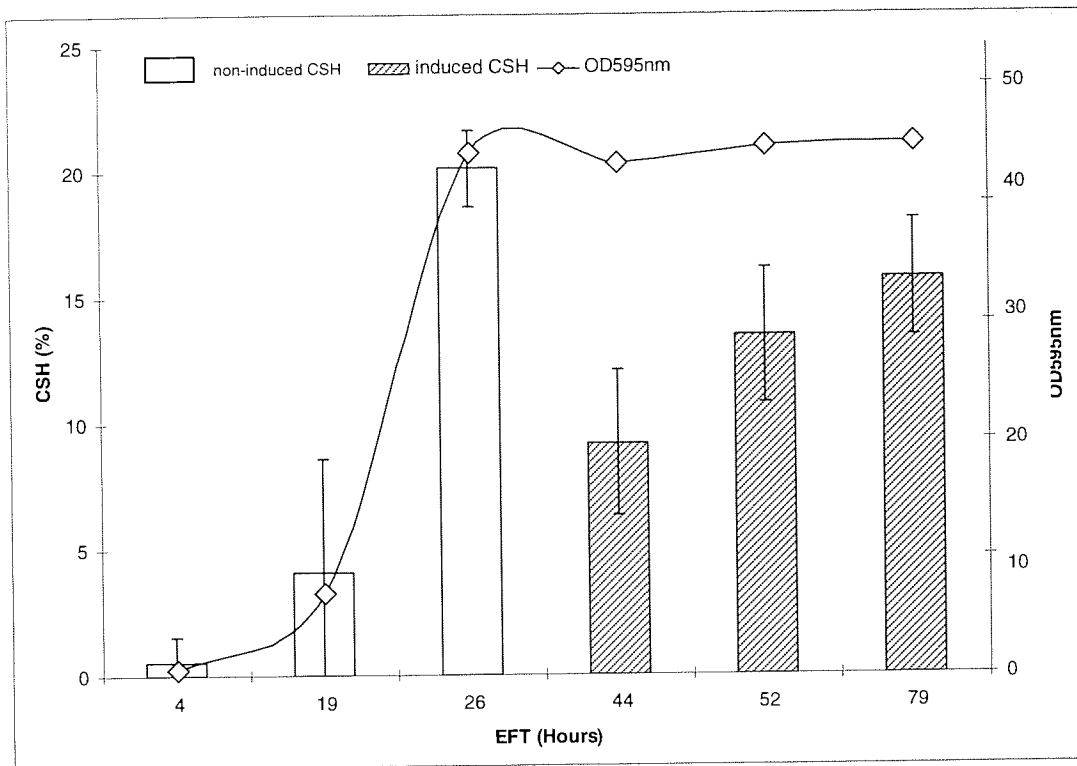


Figure 5.5. Non-induced (BMGY batch) and induced (methanol fed-batch) *P. pastoris* X33GFP_{uv} 7 L bioreactor culture, as shown in Figure 5.4. Relative CSH (%) and cell density (OD₅₉₅) of non-induced and induced phases of the culture are shown. The growth (batch) phase at early, mid and late log phases of growth curve are shown: 4, 19 and 26 h post inoculation respectively. The induced (fed-batch) phase was sampled at early, mid and late induction: 44, 52 and 79 h post inoculation (15, 23 and 50 h post-induction). Error bars show the standard deviation (n=3).

It was considered that the fermentation depicted in Figures 5.4 and 5.5 suffered from methanol accumulation during the induction phase due to poor matching of the standard methanol feeding rate and cellular metabolism at the non-standard induction conditions (21.5 °C, pH 7.6, 90 % DO). To address this, an alternative induction regime was devised using a mixed feed of 40 % methanol 60 % sorbitol. Thorpe *et al* (1999)¹²¹ demonstrated that a non-limiting mixed feed induction of *P. pastoris* can result in an increased specific protein production rate. This was a method of circumventing methanol accumulation at non-standard induction conditions (*e.g.* low temperature) by supplying a non-repressive carbon source (sorbitol) and a lower methanol concentration.

Figure 5.6 shows the on-line and off-line measurements for the 7 L bioreactor run from BioXpert Lite monitoring software. The run conditions use model-predicted optimal expression conditions (21.5 °C, pH 7.6, 60 % DO) as in the previous run. Induction was by a mixed feed of 40 % methanol, 60 % sorbitol with 12 mL/L PTM₁ trace salts at a rate of 0.167 mL/min. It is noticeable that the DO trace is maintained at the set-point (60 %) and biomass

continues to increase following induction with the mixed feed. This suggests that the cells in the previous run were subjected to methanol accumulation: the DO trace was not held at the set-point and biomass did not increase following induction with 100 % methanol. The apparent drop in biomass (OD_{595}) was not seen in off-line (wet) cell weight measurements and is possibly due to the cells swelling and increasing in size rather than dividing. This would lead to a increase in transmittance¹³⁵ and therefore a decrease in optical density as $OD_{\lambda} = -\log_{10}T$ where T is transmittance.

The sample points for the CSH assay are indicated by the arrows in the growth (batch) phase at early, mid and late log phases: 2, 23 and 26 h post inoculation respectively. The induced (fed-batch) phase was sampled at early and late induction: 48 and 89 h post inoculation (17 and 58 h post-induction).

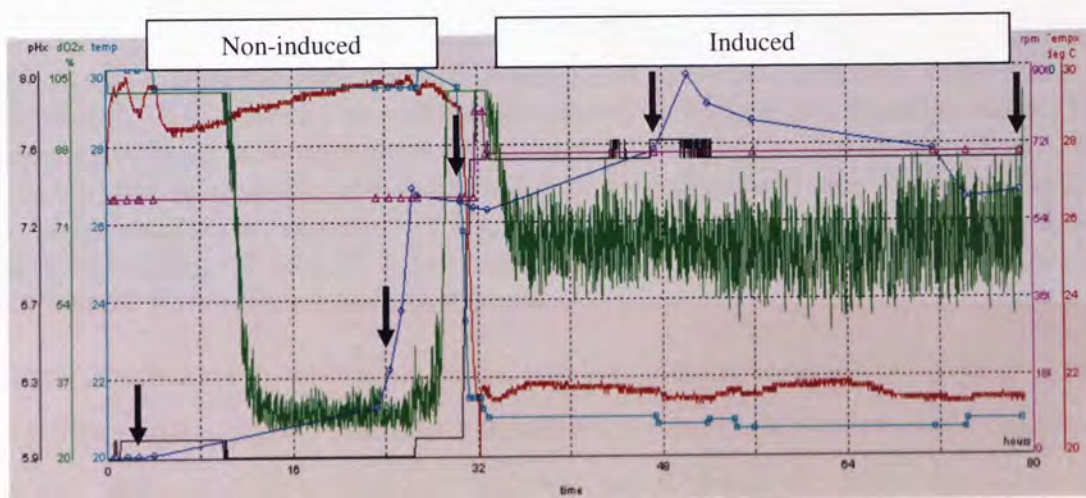


Figure 5.6. Trace from BioXpert Lite monitoring software of (f008) *P. pastoris* X33GFP_{uv} 7 L bioreactor culture with model-predicted optimal expression conditions and mixed feed induction as described in Chapter 3. The initial set points for the non-induced batch phase were 30 °C, pH 6.0, 30 % DO. This was followed by starvation and adaptation phases prior to induction with a methanol/sorbitol mixed feed (fed-batch) with set-points 21.5 °C, pH 7.6, 60 % DO. Arrows indicate sample points for the CSH assay.

Figure 5.7 shows the biomass (OD_{595}) and CSH (%) measurements from time points indicated in Figure 5.6: 2, 23, 26, 48 and 89 h post inoculation. It can be seen that during the growth (batch) phase, relative CSH increases with biomass/progression through early (12.1 % CSH) and mid (31.8 % CSH) log phases but drops in late log phase (23.27 % CSH). Once the culture is induced with the methanol/sorbitol feed CSH remains fairly constant within error with a possible decrease at late induction (19.3 % CSH at 58 h post induction).

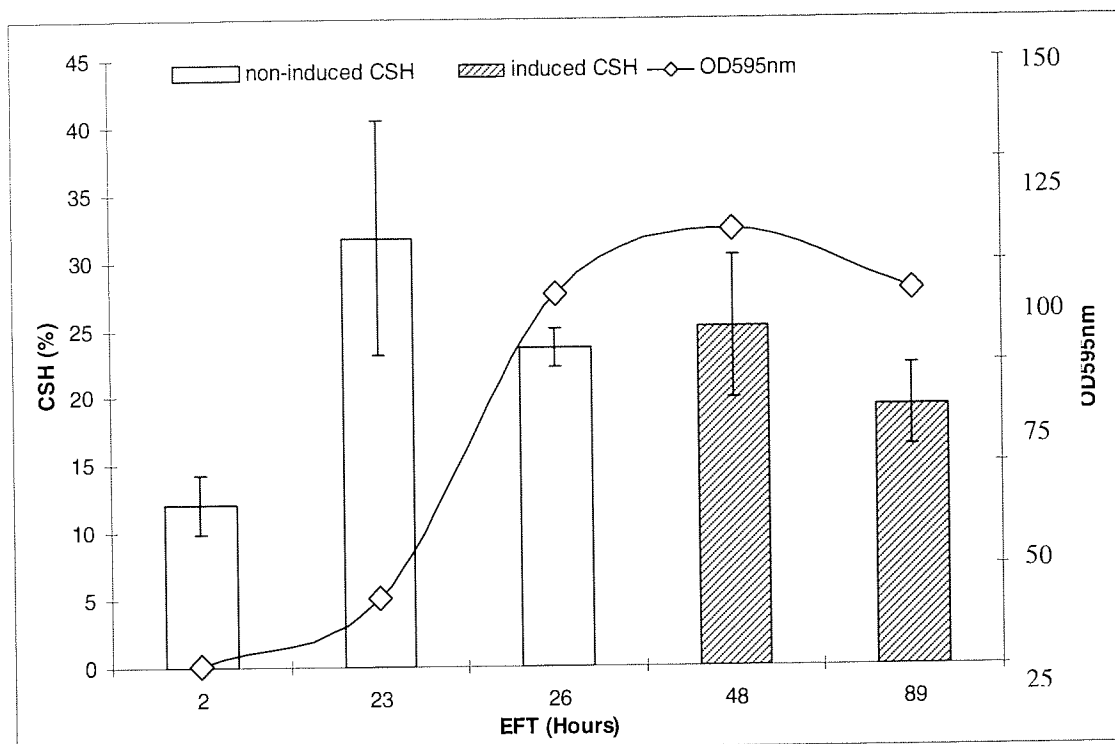


Figure 5.7. Non-induced (BMGY batch) and induced (methanol fed-batch) *P. pastoris* X33GFP_{uv} 7 L bioreactor culture, as shown in Figure 5.6. Relative CSH (%) and cell density (OD₅₉₅) of non-induced and induced phases of the culture are shown. The growth (batch) phase at early, mid and late log phases of growth curve are shown: 2, 23 and 26 h post inoculation respectively. The induced (fed-batch) phase was sampled at early and late induction: 48 and 89 h post inoculation (17 and 58 h post-induction). Error bars represent the standard deviation (n=3).

The data obtained from this run would suggest CSH increases during periods of rapid growth such as mid-log phase. This does not necessarily contradict the previous results, as the time points when sampling occurred are different, and therefore the cells may not have been in late log but perhaps in stationary phase. To confirm this, an additional repeat fermentation run would be required.

Figure 5.8 shows the on-line and off-line measurements for the 7 L bioreactor run from BioXpert Lite monitoring software. The run conditions for (f009) use the standard expression conditions (30 °C, pH 6.0, 30 % DO) as suggested by the Invitrogen handbook. Induction was by a mixed feed of 40 % methanol, 60 % sorbitol with 12 mL/L PTM₁ trace salts at a rate of 0.167 mL/min. It is noticeable that the DO trace is maintained above the set-point (30 %) and biomass continues to increase following induction with the mixed feed. Off-line biomass estimation by (wet) cell weight confirmed that biomass continued to increase post-induction. This, with the f008 data, suggests that the cells in the initial run (f007) were subjected to methanol accumulation: the DO trace was not held at the set-point and biomass did not increase following induction with 100 % methanol.

The sample points for the CSH assay are indicated by the arrows in the growth (batch) phase at early, mid and late log phases of growth curve: 8.5, 23.5 and 28.25 h post inoculation respectively. The induced (fed-batch) phase was sampled at early, mid and late induction: 31.25, 48, 53 and 77.5 h post inoculation (1.75, 18.5, 23.5 and 48 h post-induction).

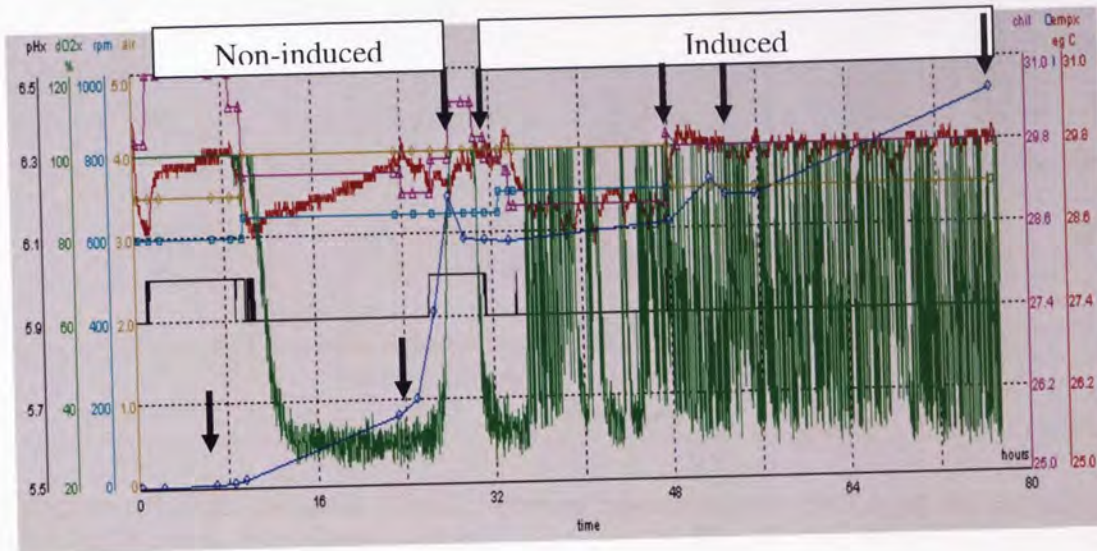


Figure 5.8. Trace from BioXpert Lite monitoring software of (f009) *P. pastoris* X33GFP_{uv} 7 L bioreactor culture with standard process conditions (Invitrogen handbook) and mixed feed induction. The initial set points for the non-induced batch phase were 30 °C, pH 6.0, 30 % DO. This was followed by a starvation phase prior to induction with a methanol/sorbitol mixed feed (fed-batch) with set-points 30 °C, pH 6.0, 30 % DO. Arrows indicate sample points for CSH assay.

Figure 5.9 shows the biomass (OD₅₉₅) and CSH (%) measurements from time points indicated in Figure 5.6: 8.5, 23.5, 28.25, 31.25, 48, 53 and 77.5 h post inoculation. It can be seen that during the growth (batch) phase relative CSH increases from 12.8 % at early log (8.5 h post inoculation) to 30.1 % during mid log phase (23.5 h post inoculation) then drops to 16 % at late-log/stationary (28.25 h post inoculation). Within error it appears that CSH is ~ 15 % when the cells are in the lag and late log/stationary phases of the growth cycle and become more hydrophobic when rapidly growing/dividing during the mid-log phase. This is in agreement with the data from the previous bioreactor run (Figure 5.7).

This trend of increased CSH during rapid growth is also seen during the induction phase. Shortly after induction starts (31.25 h EFT) CSH is 16.8 %, but once the cells resume growth (using the sorbitol carbon source supplied in the mixed feed induction) at 18.5 h post induction, CSH increases to 36.3 % but then drops to 27 % (23.5 h post induction) and 20.7 % (48 h post induction).

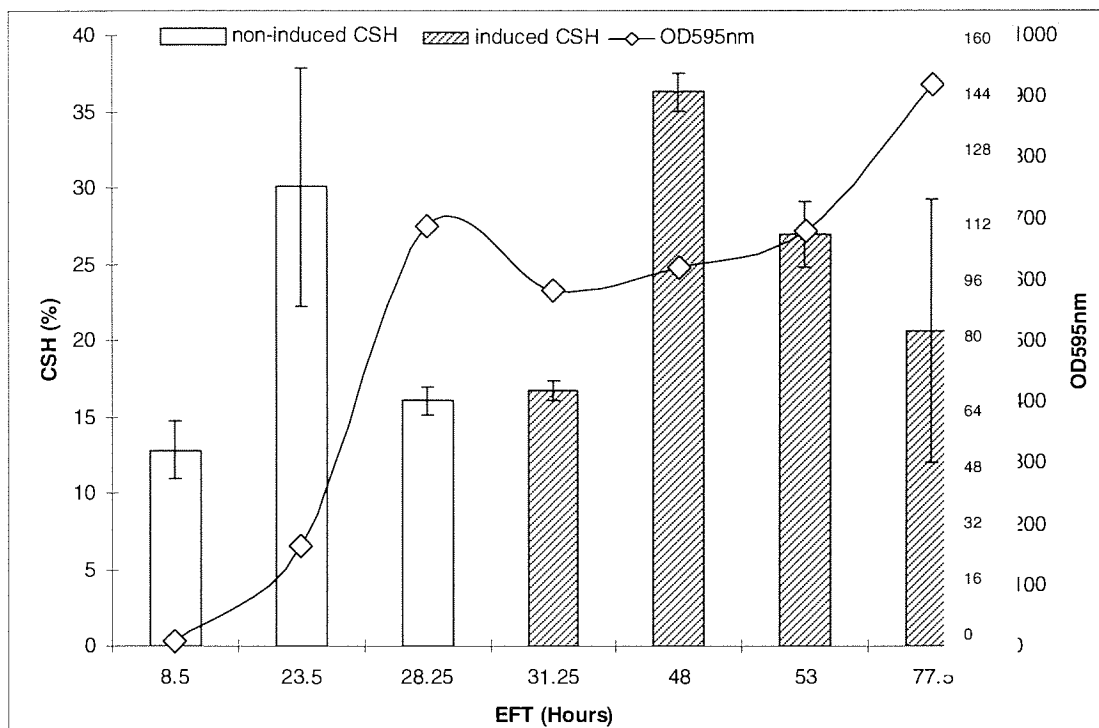


Figure 5.9. Non-induced (BMGY batch) and induced (methanol fed-batch) *P. pastoris* X33GFP_{uv} 7 L bioreactor culture, as shown in Figure 5.8. Relative CSH (%) and cell density (OD₅₉₅) of non-induced and induced phases of the culture are shown. The growth (batch) phase at early, mid and late log phases of the growth curve are shown: 8.5, 23.5 and 28 h post inoculation respectively. The induced (fed-batch) phase was sampled at early and late induction: 31.2, 48, 53 and 77.5 h post inoculation (1.75, 18.5, 23.5 and 48 h post-induction). Error bars represent the standard deviation.

The results from the latter bioreactor runs suggest that the late-log phase sample from the first bioreactor run was not in the intended portion of the growth curve (late-log) and were instead cells from the mid-log phase. The effects of methanol accumulation during the induction of initial bioreactor run can be seen in Figure 5.4 as a drop in temperature (generated by the cells) and an increase in DO (supplied oxygen not being utilised by the cells) following induction with the 100 % methanol feed. The increase in CSH associated with this culture (see Figure 5.5) post-induction is thought to be related to the effects of methanol accumulation and not cell growth as with the other HCD cultures.

The reference¹³² on *S. cerevisiae* CSH did not use relative hydrophobicity (%) as a measure of CSH, but rather used a scoring system (e.g. +++, ++, ±) so the results here can not be directly compared to those involved in *S. cerevisiae* flocculation. The example of CSH determination with *Candida albicans* used cell adhesion to styrene beads (hydrophobic) measured by flow cytometry, providing a relative (%) attachment to beads,¹³⁴ again not directly comparable to the methods used here. Smith *et al*¹¹⁰ used both the MATH and HIC methods for CSH determination of five *Coniothyrium minitans* isolates. They tracked CSH changes over culture

age (0-42 days). Depending on the isolate CSH remained constant, increased or decreased with age. Typical relative hydrophobicity (day 10) was 95 % (by HIC) or 80 % (by MATH). The majority of the isolates dropped to 50 % (HIC) or 30 % (MATH) by day 40. Two isolates had a constant (HIC) or increased CSH reaching 95 % (MATH) by day 42. These values are much higher than any of those seen here with non-induced and induced cultures/HCD fermentations of *P. pastoris*. The *C. minutans* data should be used as an indication that the assay methods are able to determine changes in CSH, but not for direct comparison to the *P. pastoris* data as the expression systems are so different.

The data presented here suggest that rapidly-growing *P. pastoris* cells, *i.e.* those in mid-log phase, have higher cell surface hydrophobicities than “resting” cells in lag or stationary phases. This appears to be true for both non-induced and induced cultures. The assay technique (MATH) employed here provided quite variable results from the same sample pool, which could be attributed to inconsistent mixing of the aqueous and hydrocarbon phases. If the MATH assay is used in future work, the agitation rates and timings should have increased stringency. Other techniques are possible, such as flow cytometry¹³⁴ which may provide more consistent results in a timely fashion.

5.1.5 Implications for high cell density fermentations

The effects of variable cell surface hydrophobicity are known to influence other yeast species: flocculation in cultures of *S. cerevisiae*¹³² and surface adhesion in *Candida albicans*.^{136 137} It is not inconceivable therefore that cell aggregation in *P. pastoris* could also be influenced by CSH. This would have an effect on the dynamics of a HCD *P. pastoris* culture, especially when high biomasses are reached of >500 OD₆₀₀, equivalent to 130 g/L dry cell weight.¹ During culture, wet cells could represent 10-20 % of the total culture volume. If these formed aggregates (due to increased CSH) and thus increased the viscosity of the culture,¹³⁸ this would have an influence on the mass transfer coefficient (k_{La}) of the system¹³⁹ since increasing viscosity leads to a decrease in k_{La} .¹²⁸ Reducing the k_{La} will lead to decreases in cellular uptake, *e.g.* oxygen and media components, and so reduce cell growth. This could give scope for further optimisation of a HCD culture: altering process factors *e.g.* agitation rate to compensate for the deviation from designed or expected k_{La} during periods of rapid growth such as mid-log phase cultures. Variable agitation rate during mid-log phase growth could be included as an additional factor in a Design of Experiments (DoE) optimisation matrix, as used in Chapter 3. This approach may not be suited to large scale manufacture as

bioreactor vessels are likely to be already set-up (run conditions *e.g.* agitation and gas flow rates) to provide maximal k_La for the system. However this may be useful for bench-scale experiments where the system has not already been fully optimised.

5.2 Rationalisation for the use of antifoam agents

During the aerobic culture of micro-organisms, such as *E. coli* and *P. pastoris*, the cells require a source of oxygen to maintain efficient rapid growth and/or recombinant protein expression. Oxygen is required for energy generation (ATP) by electron transfer during the final step in catabolism (biodegradation of biomolecules such as glucose, amino acids and fatty acids), namely oxidative phosphorylation. Oxidative phosphorylation is the process of electron transfer from NADH and FADH₂ to molecular oxygen forming H₂O. This is “formally equivalent to burning of hydrogen in air to form water”¹³³ and produces a major portion of a cell’s ATP and heat energy.

The demand for oxygen by micro-organisms and cultured cells can be met by aerating the culture medium. This aeration of the medium is normally achieved by agitation and sparging sterile air or oxygen-enriched air into it. An unfortunate effect of both sparging gas through culture medium at high rates and agitation is the formation of foam. This is a particular problem when surface-active species, such as proteins, are present at high concentrations. Foaming reduces the efficiency of gas exchange at the surface of the culture, as a barrier is formed between the culture and the gases in the headspace of the vessel. Foaming can also be detrimental to the cells in the culture. For example, when bubbles burst they exert sheer forces which may damage cells or secreted proteins. Additionally cells and culture medium are lost to the foam phase which can lead to a decrease in process productivity. In extreme cases of foaming, ‘foam out’ can lead to loss of process sterility.⁷⁴

In order to minimise the deleterious effects of foaming, an antifoam agent can be included in the culture medium. Antifoams prevent foam forming by reducing the surface tension of the culture so that any bubbles that reach the surface cannot be supported and therefore disperse without bursting.⁵⁸ Control of foam formation in high cell density (HCD) fermentations is an important consideration with regard to maximizing product yield.⁷⁴

The use of antifoam agents in the research environment is often based on anecdotal evidence, *i.e.* “Will often says- I know three drops of this antifoam works because that is what we’ve

always done". Here, a systematic approach was undertaken to evaluate the use of a number of commercially-available antifoam agents. For this investigation two secreted target proteins were used: a recombinant Fc fusion protein and green fluorescent protein (GFP). The Fc fusion protein was obtained from a client of Alpha Biologics and was expressed in *P. pastoris* and *S. cerevisiae* cell lines. The Fc fusion protein was industrially relevant and was available for use in addition to the GFP_{uv} expressing *P. pastoris* cell-line as used in Chapter 3 for process condition optimisation.

Four antifoam agents were evaluated:-

1. Schill & Schelinger's Struktol SB2121 (polyalkylenglycol);
2. Schill & Schelinger's Struktol J673A (an alkoxyated fatty acid ester on a vegetable base);
3. Sigma Antifoam C (an aqueous emulsion of Sigma Antifoam A; siloxane polymer).
4. Fluka P2000 (polypropylene glycol). Fluka P2000 is a polypropylene glycol (PPG) with an average molecular weight of 2000 Da.

The structure of PPG is shown in Fig 5.10. The exact structures of the three other antifoams were not available.

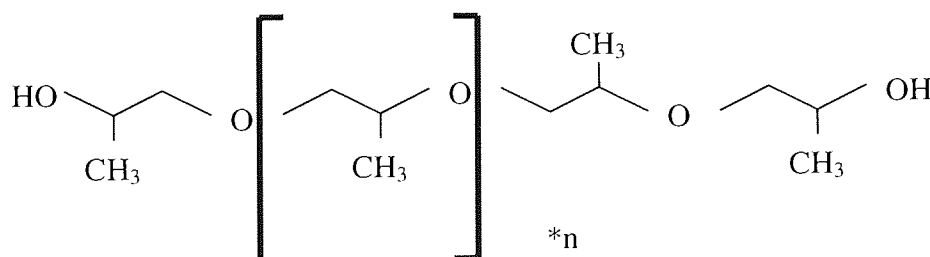


Figure 5.10. Structure of polypropylene glycol (Struktol P2000 antifoam)

5.2.1 Antifoam evaluation with an Fc fusion protein as an indication of yield

S. cerevisiae Alcofree™ transformed with the pYX212-Fc vector and *P. pastoris* GS115 with an integrated copy of Fc (from the pPICZα-C-Fc vector) were used. In *S. cerevisiae* Alcofree (a uracil auxotrophic strain discussed in Chapter 4) protein expression from the episomal plasmid was under the control of the *TPII* constitutive promoter, and in *P. pastoris* it was under inducible control of the *AOXI* promoter.

Each antifoam type was added at 0.5, 1.0, 2.0, 4.0 and 8.0 % (v/v) with 0 % (v/v) antifoam as a negative control. The culture medium was also varied to demonstrate any differences in antifoam effects between non-induced (YPD) and induced (SD^{-ura} or BMMY) cultures.

Triplicate shake flask cultures were used to evaluate each of the organism and antifoam agent combinations. Samples were taken from each shake flask to monitor cell growth (OD₅₉₅) and supernatant was assayed for Fc fusion protein expression by SDS-PAGE and silver stain visualisation. The analysis of Fc fusion protein yield was problematic as the protein was expressed at levels towards the lower limit of SDS-PAGE and silver stain detection. Therefore, small differences in yield were difficult to identify by product band densitometry from SDS-PAGE images.

It was found that the different types of antifoam affect *S. cerevisiae* and *P. pastoris* growth in different ways depending on the concentration and medium type being used. The Fc fusion protein data indicated that antifoam agents can be used at concentrations up to 1 % total volume. Higher concentrations can lead to an apparent decrease in protein yield. Additionally some antifoam agents (e.g. Struktol J673A) become difficult to work with at higher concentrations, producing precipitates which interfere with sampling and analysis. Table 5.1 highlights the main conclusions for each individual antifoam and application.

Antifoam Type	<i>S. cerevisiae</i> Alcofree™		<i>P. pastoris</i> GS115	
	Growth	Protein Production	Growth	Protein Production
Struktol SB2121	Increased by >0% <4%	Decreased by >1%	No effect 8%	No effect 8%
Struktol J673A	Increased	Decreased by >1%	Increased	Decreased by >1%
Sigma Antifoam C	No effect 8%	Decreased by >2%	No effect 8%	No effect 8%
Fluka P2000	Increased by >4%	Decreased by >1%	No effect 8%	Decreased by >4%

Table 5.1. Summary of antifoam influence on yeast (*S. cerevisiae* Alcofree and *P. pastoris*) growth and recombinant protein production (Fc fusion protein) in shake flask cultures. Beneficial interactions are highlighted

5.2.2 Antifoam evaluation with green fluorescent protein (GFP_{uv}) as an indication of yield

The GFP_{uv}-expressing *P. pastoris* strain X33GFP_{uv} described in Chapter 3 was seen as a convenient tool for the rapid evaluation of antifoam type and concentration. Culture supernatant could be reliably assayed using GFP fluorescence with a 96 well plate and a UV plate reader for yield determination, rather than SDS-PAGE and silver stain visualisation as used with the Fc fusion protein, so moving towards a high(er)-throughput system.

The previous study using the Fc fusion protein highlighted unwanted properties in two of the four antifoam agents used. J673A formed a white precipitate that adhered to plastic and glass making sampling and analysis difficult. Antifoam C suffered from phase separation following dilution and sterilisation by autoclave; making measurement of accurate volumes difficult (the siloxane polymer formed “droplets” that were not evenly dispersed). It was for these reasons Struktol J673A and Sigma Antifoam C were not used in the second evaluation.

GFP fluorescence is sensitive to pH,²⁴ therefore all samples involving GFP_{uv} were buffered to pH >7.0. To achieve this, the addition of 50 % sample volume 1 M potassium phosphate buffer pH 8.0 was identified as sufficient for all the culture pH values.

A calibration curve was needed to convert from the plate reader output of relative fluorescence units (RFU) to protein concentration (*e.g.* mg/mL). To achieve this, GFP_{uv}-containing culture supernatant was serially diluted and assayed by SDS-PAGE and fluorescence. The results were plotted as a XY scatter plot (see Figure 5.11) and a conversion factor calculated ($3.0 \times 10^{-4} \times \text{RFU} = \mu\text{g/mL GFP}_{\text{uv}}$).

Shake flask cultures of *P. pastoris* X33GFP_{uv} were used to evaluate how two antifoam agents (Fluka P2000 and Struktol SB2121) affected growth and protein expression at a range of concentrations (0, 0.2, 0.4, 0.6, 0.8, 1.0 % v/v), more typical of those used for bioreactor cultures.

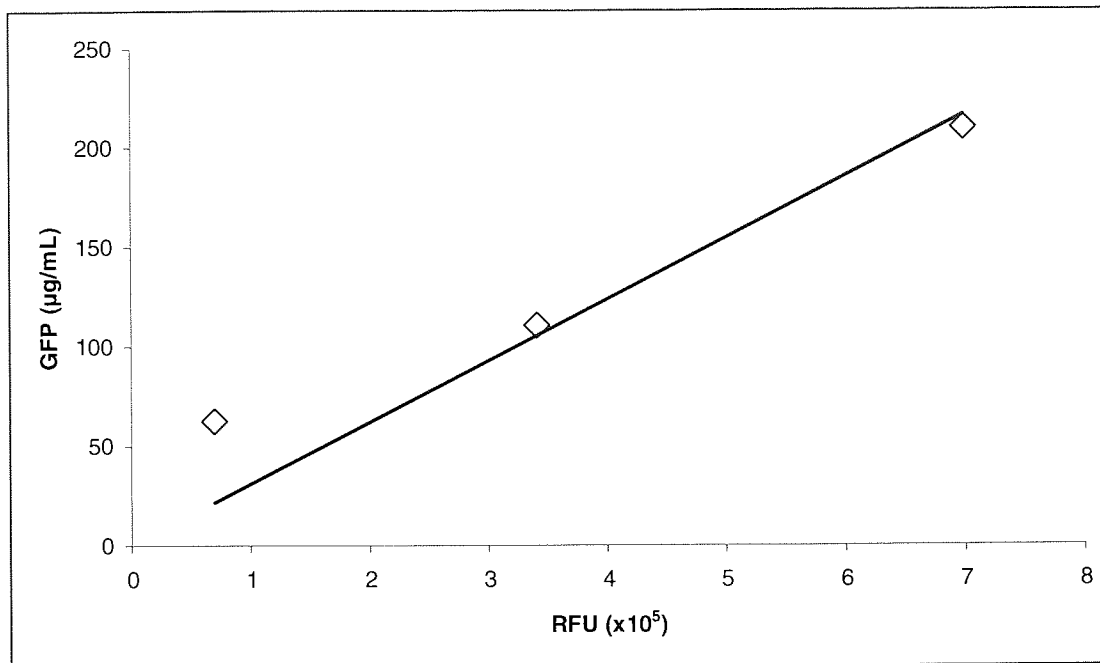


Figure 5.11. XY scatter plot of GFP_{uv} protein concentration ($\mu\text{g/mL}$) and associated fluorescence (RFU). Line of fit has the equation $y=3.0 \times 10^{-7}x$ giving the conversion of $3.0 \times 10^{-4} \times \text{RFU} = \mu\text{g/mL GFP}_{\text{uv}}$ so allowing GFP_{uv} yield to be calculated from fluorescence.

A single colony of *P. pastoris* X33GFP_{uv} from a freshly-streaked agar plate was used to inoculate an overnight BMGY seed culture. This was used to inoculate 400 mL BMMY to an OD₅₉₅ of 1.0, and then 20 mL of seeded BMMY was aliquoted into twenty 125 mL shake flasks, to ensure comparable results. The BMMY shake flasks contained 0, 0.2, 0.4, 0.6, 0.8 or 1.0 % v/v antifoam (SB2121 or P2000) in triplicate. These were incubated at 30 °C with 250 rpm agitation for 48 h. Sterile methanol was added (1 % v/v) to the shake flasks at 24 h to maintain induction. Samples were taken at 24 and 48 h post induction to determine cell growth by OD₅₉₅. At harvest (48 h post induction) a supernatant sample was taken from each culture and assayed for GFP yield. Yields were calculated as normalised to cell density (OD₅₉₅) so that direct effects on protein expression could be seen and not masked by any differences in cell growth (OD₅₉₅). Figure 5.12 shows the normalised GFP_{uv} yields for the cultures with P2000 antifoam. There is no significant difference in normalised GFP_{uv} yield when using P2000 antifoam at concentrations up to 1 % v/v in shake flask cultures.

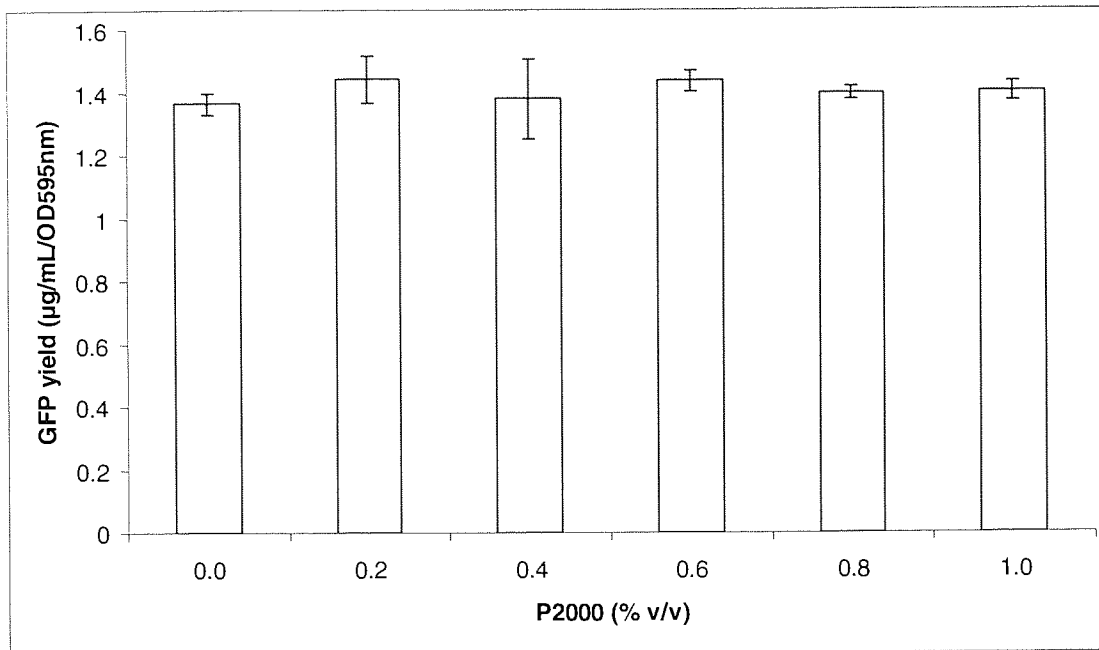


Figure 5.12. Normalised GFP_{uv} yield at 48 h post induction from triplicate shake flask cultures of *P. pastoris* X33GFP_{uv} containing a range of Fluka P2000 antifoam concentrations. Error bars represent standard deviation.

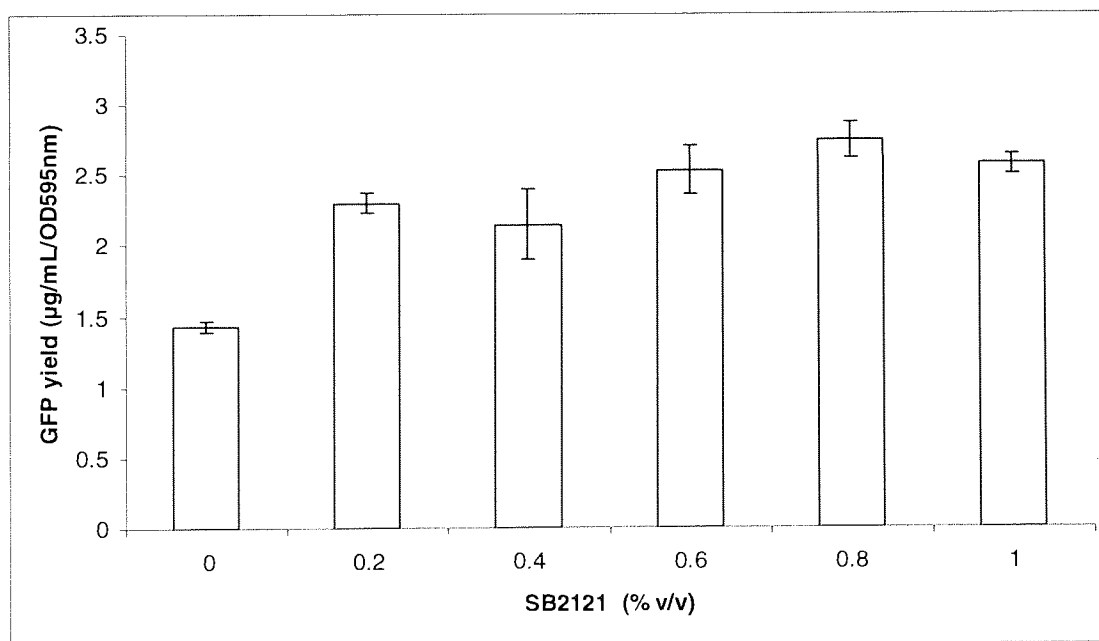


Figure 5.13. Normalised GFP_{uv} yield at 48 h post induction from triplicate shake flask cultures of *P. pastoris* X33GFP_{uv} containing a range of Struktol SB2121 antifoam concentrations. Error bars represent standard deviation.

Figure 5.13 shows the normalised GFP_{uv} yields for the cultures with SB2121 antifoam. There is an obvious difference in normalised GFP_{uv} yield when using SB2121 antifoam at all tested concentrations above 0.2 % v/v in shake flask cultures. There is no obvious difference in GFP_{uv} yield between 0.2 % v/v and the higher concentrations used here.

To confirm that these effects were not due to an auto-fluorescent property of the SB2121 antifoam, a blanking experiment was run. SB2121 antifoam additions were made to BMMY media and then assayed for fluorescence. To further determine whether SB2121 was interacting with GFP_{uv} to enhance fluorescence, SB2121 antifoam additions were made to culture supernatant already containing GFP_{uv}. The results from both blanking and GFP_{uv} interaction experiments are shown in Figure 5.14.

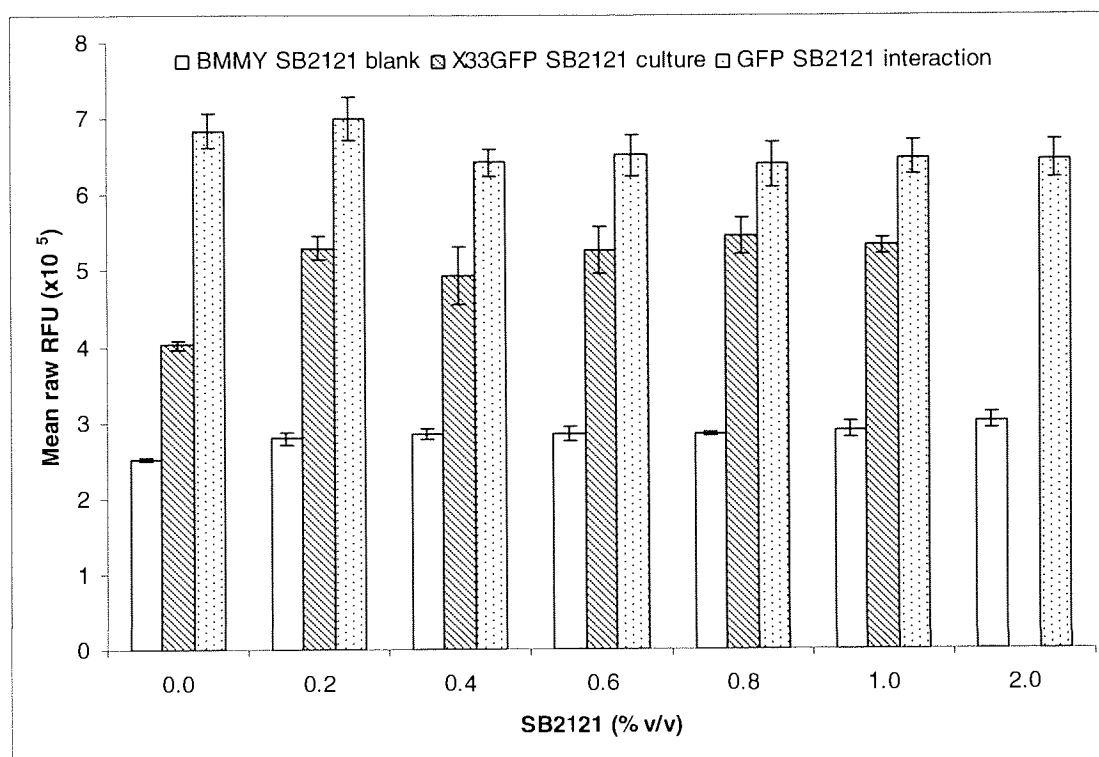


Figure 5.14. Fluorescence blanking experiments with Fluka SB2121 antifoam additions (white bars); BMMY medium, culture supernatant with existing GFP_{uv} (spotted bars). Raw RFU values of X33GFP_{uv} shake flask cultures are shown for comparison as grey hashed bars. Error bars represent standard deviation.

When SB2121 is added to culture supernatant with existing GFP_{uv} present a slight decrease in fluorescence can be seen, demonstrating the SB2121 does not interact with GFP_{uv} to increase the RFU. When SB2121 is added to BMMY, SB2121 possesses some fluorescence above the BMMY background fluorescence, but this does not account for the increases seen during the X33GFP_{uv} shake flask culture, where the raw RFU values were 40292.1 in the absence of SB2121 and 53044.6 in the presence of 0.2 %.

This suggests that the presence of SB2121 enhances recombinant protein expression of GFP_{uv}. This could be associated with an enhanced mass transfer coefficient (k_{La}) for the shake flask system. The mode of action for antifoam agents is to reduce the surface tension of the liquid, so that bubble formation cannot be supported. It is feasible that the reduction in

surface tension, in the presence of SB2121, results in more efficient oxygen transfer from gas to liquid phases and/or liquid phase to cells. To confirm this, further experiments in a controlled bioreactor system would be required to ascertain any change in dissolved oxygen (DO) level when antifoam is added to the medium.

The background fluorescence of both culture medium (BMMY) and antifoam agents highlight some of the difficulties that may be encountered when using on-line fluorescence for fermentation process characterisation/monitoring, as suggested in the literature.^{140, 27}

5.3 Many factors can affect a protein production process

The work presented in this chapter on cell physiology and rationalising antifoam use has shown that these factors are capable of influencing the characteristics of a protein production experiment. The cell surface hydrophobicity (CSH) and antifoam evaluations are not well documented in the literature, which may be an omission, as it has been demonstrated that these factors could provide scope for further optimisation of the process.

Ideally as many factors as possible should be considered and preferably included in a multi-factorial design of experiments (DoE) optimisation process such as that described in Chapter 3. Factors identified here are agitation rate during mid-log phase growth and antifoam type and concentration.

The viscosity of a culture may increase with CSH, which would decrease the mass transfer coefficient ($k_L a$) of the system, so reducing oxygen uptake of the cells in culture. Varying the agitation or gas (air/oxygen) flow rate used during periods of rapid growth, such as mid-log phase cells could allow improvements in cellular growth. This could decrease the overall process time, as cells would reach the target achievable biomass more quickly.

The addition of antifoam agents may also influence the $k_L a$ of the system. Antifoam agents reduce the surface tension of a culture medium, which prevents stable/standing foam formation. An additional result of reducing surface tension could be improved diffusion and therefore increasing $k_L a$.

“If you do not train hard whilst you are young,
You will have nothing when you are old”

(Wing Chung proverb)

CHAPTER 6

6 IMPROVING BIOMASS AND VOLUMETRIC PRODUCT YIELD IN WILD TYPE AND A NOVEL ALCOFREE STRAIN OF *SACCHAROMYCES CEREVISIAE*

The methylotropic yeast *P. pastoris* was initially developed as a potential source of Single Cell Protein (SCP) for animal feeds, a prerequisite for this being high achievable biomass in culture. During the 1980's a wealth of information for achieving continuous high cell density *P. pastoris* cultures was amassed, with the yeast being developed for the production of recombinant proteins.¹ The achievement of high density cultures with high yields (e.g. 15 g/L murine collagen⁶) of recombinant proteins has subsequently resulted in the *Pichia pastoris* system being the most frequently used host for heterologous protein expression in general.^{1, 6, 11}

The *S. cerevisiae* Alcofree strain was designed to combine the industrial benefits of *S. cerevisiae* (well understood genetics and physiology) and the high achievable biomass of *P. pastoris*. The unique respiratory nature of the *S. cerevisiae* Alcofree is a result of the modified glucose uptake system which was achieved by introducing chimeric hexose transporters so limiting the uptake of glucose into the cell. Therefore, the Alcofree strain does not produce ethanol under the standard fermentation conditions of aerobic cultivation at high glucose concentrations.¹³ This use of metabolic engineering allows high cell biomass to be achieved without the need for complex control mechanisms to avoid ethanol formation.¹⁴¹

In Chapter 4 it was highlighted that low *S. cerevisiae* Alcofree biomass hindered the detection of expressed recombinant proteins. This low volumetric yield proved problematic during the positive transformant selection as product (if present) was not readily detected by SDS-PAGE or Western blotting. A hundred-fold concentration procedure was required for target protein detection from positively-expressing cultures. If these transformants were to be used for further production of the target proteins, the low volumetric yield of the cultures would not be favourable for efficient manufacture. One of the proposed avenues of investigation was to improve the composition of the growth medium, by medium design, so allowing higher cell densities (biomass) to be met. To this end, alternative carbon sources to glucose and various additives were screened to try to improve biomass yield. The novel *S. cerevisiae* Alcofree strain and wild type strains W303-1A and VWK70-945 were used for this medium development.

Chapter 3 demonstrated the usefulness of small-scale bioreactors, *e.g.* the Micro 24 (Applikon Biotechnology), for performing parallel multi-factorial experiments. The controlled nature of these small-scale cultures, whilst being high-throughput unlike traditional stirred tank reactors, made the Micro 24 an appropriate option for performing this medium development study.

6.1 Constitutive expression and growth; linking yield to biomass

The plasmid vector (pYX212) used with the Alcofree strain utilises the constitutive triosephosphate isomerase (*TPI1*) promoter for expression of genes inserted into the multiple cloning site (MCS). Thus as cells harbouring the plasmid grow they express the target protein. This contrasts with the inducible system used with *P. pastoris* clones where the cells grow but do not express the inserted gene until they come into contact with the methanol induction substrate.¹⁰⁴ This inducible system allows high cell densities to be met prior to induction of protein expression.

The enzyme Tpi1 forms part of the glycolytic pathway and is responsible for the inter-conversion of dihydroxyacetone phosphate and glyceraldehyde 3-phosphate. These are isomers but dihydroxyacetone phosphate is not on the direct pathway of glycolysis whereas glyceraldehyde 3-phosphate is. Both molecules are formed when fructose 1, 6-bisphosphate is enzymatically split by aldolase.¹⁴² As Tpi1 is required for efficient continuation of glycolysis, using the *TPI1* promoter sequence will result in expression of the inserted recombinant protein whenever the cell up-regulates glycolysis *i.e.* is growing on glucose. The *TPI1* promoter is both constitutive and powerful, however, it may lead to down-regulation of copy number when used with high copy number plasmid systems.⁴²

The link between cell growth and productivity can be manipulated to achieve higher product yields. This can be done primarily by increasing the achievable biomass (number of cells) in a culture. The amount of biomass in a high cell density culture will be limited by one or a number of factors, for example: carbon source, nitrogen source, trace elements, vitamins and essential amino acids.^{143 2, 144} These are available at fixed levels in a batch culture, but by increasing the amount accessible to the cells, it may be possible to remove that limiting factor and increase the biomass. The goal is to provide enough of everything the cell requires for growth, whilst avoiding excess, which can lead to nutrient feedback repression or toxic inhibition. Catabolite repression is a form of nutrient feedback repression seen when cells are

grown in the presence of excess glucose. The expression of enzymes required for the metabolism of non-glucose carbon sources is repressed. This is also seen with nitrogen source utilisation; the cell preferentially uses high quality nitrogen sources (if present) in preference to poor ones and the enzymes required for the processing of the poor sources are repressed.¹⁴⁵

In summary the objective is to provide sufficient nutrients but not an excess, so avoiding the overload of one particular enzyme in a metabolic pathway. The nutrients must also be in a readily utilisable form.¹⁴⁴

6.2 Medium design

As the composition of a culture medium is critical to obtaining good growth and/or productivity from the cells, it has been the focus of many optimisation studies.^{72, 73 70 71, 90 146}

⁸⁸ In essence this should be carried out with each new cell-line as nutrient requirements may vary with the organisms and recombinant protein being expressed. Some of these medium optimisation studies use statistical design of experiments (DoE) to help reduce the number of experimental cultures required to identify the optimal composition. Many of the studies do not use DoE and rely on empirical determination which may overlook important interactions between the factors (medium components).⁹⁰ The approach used here was empirical and built on previous observations of the Alcofree strain.

The use of non-fermentable carbon sources has been shown to increase heterologous protein expression in *S. cerevisiae*. Teun van der Laar *et al*¹⁴⁷ used ethanol as the sole carbon source in fed-batch *S. cerevisiae* fermentations expressing an antibody fragment. They found an increase in specific productivity (1.2 to 4.8 fold) when they tested the ethanol limiting protocol against their standard glucose protocol; they tested this with a panel of seven target proteins.

It is not possible to supply all the required ethanol to the culture at the point of inoculation, as this would result in toxic levels of ethanol being present. Instead a fed-batch protocol must be implemented with either “slug” additions at time points during the culture or using a feed pump. The use of a feed pump is preferred as this allows control over the feeding rate and matching of carbon source addition to biomass accumulation.

6.2.1 Non-fermentable carbon sources

It should be possible to achieve biomasses comparable to *P. pastoris* with *S. cerevisiae* Alcofree cultures using ethanol as a sole carbon source as this should negate the ‘Crabtree

effect' observed when glucose is used as the carbon source. This would result in a much increased volumetric productivity with the Alcofree strain, possibly leading to product yields comparable to those seen in induced *P. pastoris* cultures.

The use of ethanol as a non-fermentable carbon source was investigated for growth of Alcofree and wild type *S. cerevisiae* strains. Ethanol-containing media; 2xCBSE (defined) and YPE (complex) were required for this work, based on standard 2xCBS and YPD media. In the first instance Alcofree and *wt* (VWK70-945) cells were streaked out on 1 % ethanol 2xCBSE and YPE agar. Both grew well on the new media.

The ethanol tolerance of the Alcofree cells was then evaluated in shake flasks, to determine working culture concentrations. Cultures of Alcofree were set-up using 2xCBSE with ethanol concentrations; 1, 2, 5, 10 and 15 % (v/v) and glucose (1 %) as a control. In the first instance the Alcofree cells used were transformed with an "empty" pYX212 plasmid. Figure 6.1 depicts the growth profiles (OD_{600}) obtained for these cultures.

Seed cultures were grown in standard 2xCBS and used to inoculate the experimental shake flask cultures to a target density of OD_{600} 1.0. Ethanol (or glucose), trace elements and vitamins (as per 2xCBS) were added at 31 hours post inoculation to see if further growth was possible and incubated for a further 90 h. It should be noted that both 1 and 2 % ethanol (v/v) have higher achievable biomass (OD_{600}) than the 1 % glucose (w/v) control. The cultures with higher (5, 10 and 15 % v/v) ethanol concentrations showed reduced growth. Therefore the ethanol tolerance of the Alcofree strain can be inferred as 2 % (v/v) in shake flask culture, which should not be exceeded in future work.

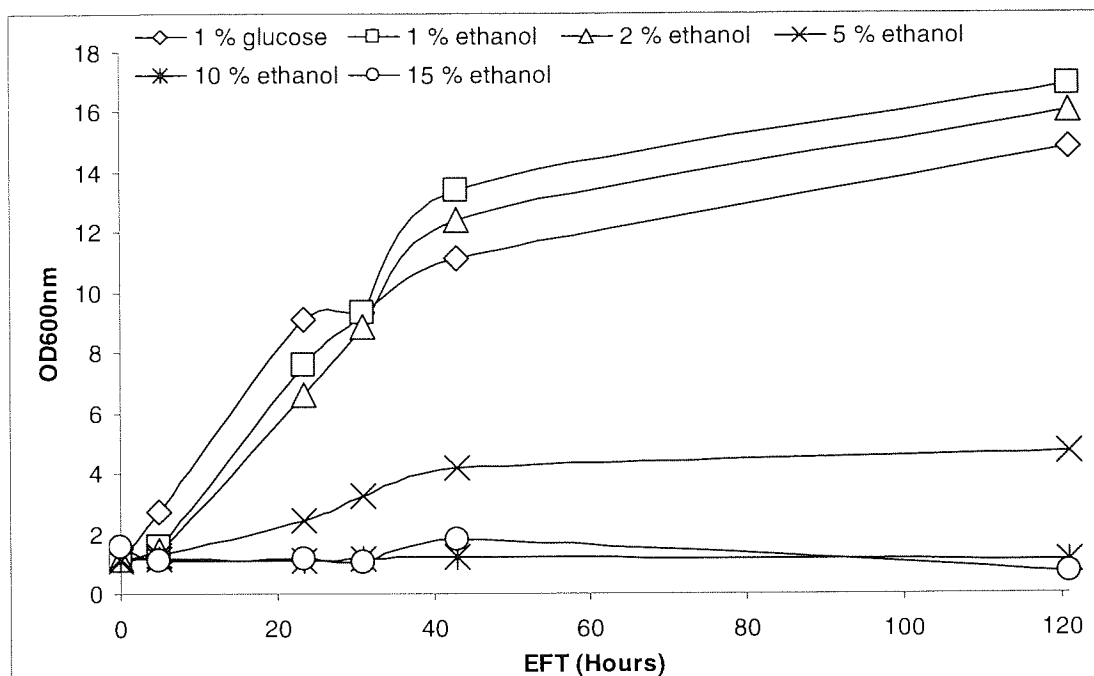


Figure 6.1. Growth profiles of *S. cerevisiae* Alcofree 2xCBSE shake flask cultures. The concentration of ethanol was varied to evaluate the ethanol tolerance of the novel strain. A glucose control culture was included for comparison of achievable biomass. Culture time is displayed as effective fermentation time (EFT) in hours.

In order to determine an ethanol feeding strategy for “slug” additions to shake flask cultures, ethanol depletion was monitored during YPE batch shake flask cultures. *S. cerevisiae* Alcofree and *wt* (W303-1A and VWK70-945) strains were used to further test the use of ethanol as a sole carbon source. Ethanol concentration was determined using a kit from Roche (Cat # 10 176 290 035). All three *S. cerevisiae* strains were grown in YPE medium with 1 and 2 % ethanol (1 and 2 % YPE). Figures 6.2A and 6.2B show the growth curves (OD₆₀₀) and corresponding ethanol consumption curves. It can be seen that all cultures reach a plateau in their growth profiles at ~40 hours post inoculation. This correlates with the ethanol depletion curves dropping to 3 g/L ethanol at ~40 hours post inoculation. The cultures with 2 % YPE reach a higher final cell density than those with 1 % YPE and take longer to exhaust the ethanol supply. The VWK70-945 strain exhibits the fastest growth on both 1 and 2 % YPE, depleting the ethanol sooner than the other strains in the same medium. VWK70-945 also achieves higher biomass than other strains in the same ethanol concentration. The cultures with the Alcofree strain show slower growth and ethanol utilisation than the VWK70-945 strain but faster than the W303-1A strain. Using Figure 6.2B as a guide, fed-batch 1 % YPE shake flask cultures of VWK70-945 and Alcofree should have additional carbon source (ethanol) added at ~48 hours post inoculation. At this point the ethanol content of the medium is much depleted, as indicated by the growth curve reaching plateau (stationary phase) and

additional ethanol can be supplied to the culture without reaching toxic concentrations (this was earlier determined as >2 % v/v).

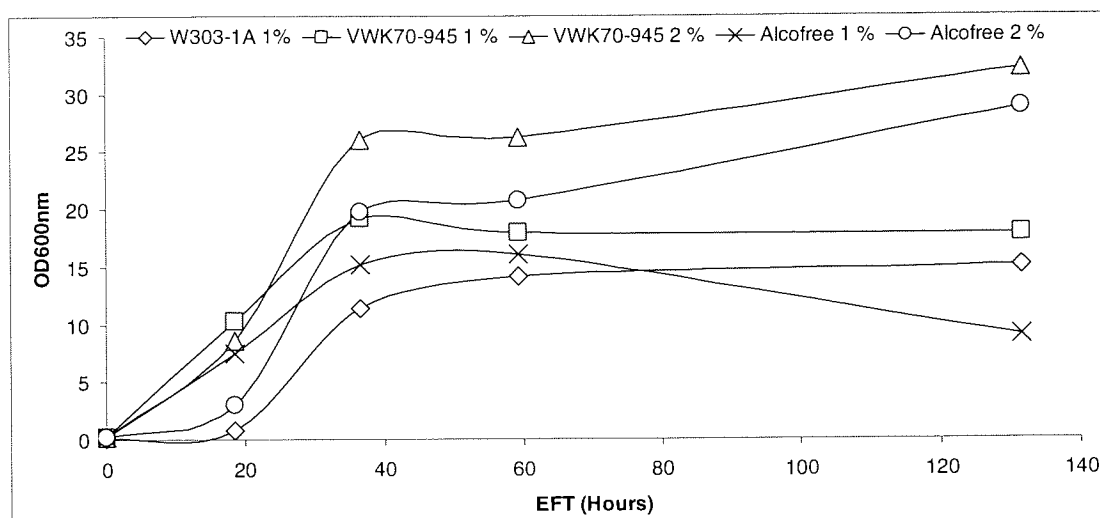


Figure 6.2A. Growth curves (OD₆₀₀) for *S. cerevisiae* 250 mL baffled shake flask cultures with 50 mL of 1 / 2 % YPE medium. Strains used are Alcofree and *wt* (W303-1A and VWK70-945)

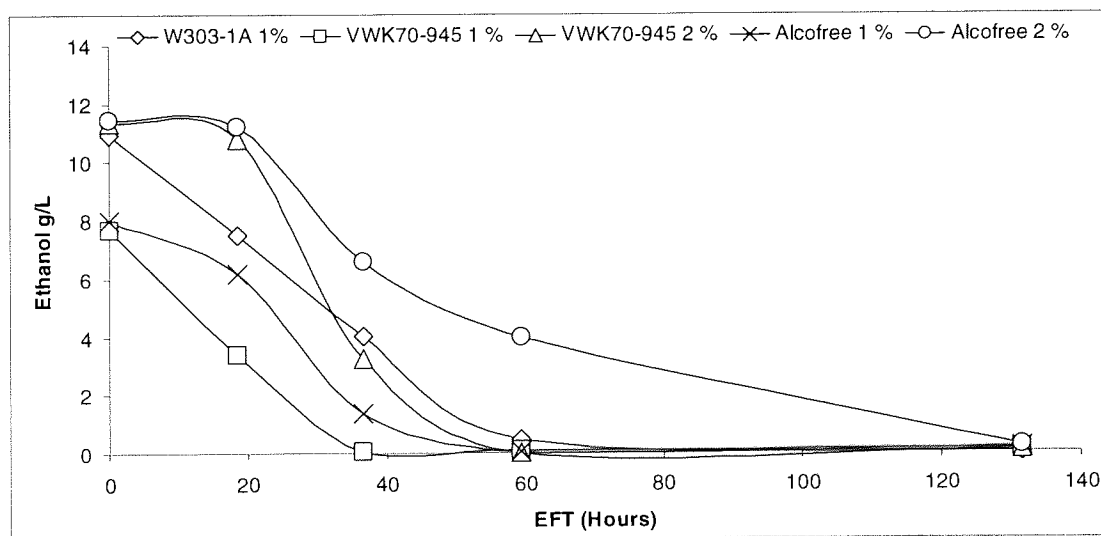


Figure 6.2B. Ethanol depletion curves for *S. cerevisiae* shake flask cultures shown in Figure 6.2A.

6.2.2 Optimisation of CSM media

The concept of using alternatives to glucose as non-fermentable carbon sources such as ethanol, glycerol, lactate and acetate, were further developed. Other means of improving the achievable biomass in CSM medium were also identified; other researchers in our laboratory had observed that the addition of inositol, biotin, copper zinc and cobalt were beneficial. During this period of work the Micro 24 bioreactor was available for use which was seen as an opportunity to gather data from a controlled culture (T, pH and DO). This was not possible

in the initial shake flask studies with ethanol, where growth conditions were assumed to be more heterologous and relied on medium component buffers for pH “control”. The Micro 24 allowed 24 cultures (6 mL working volume) to be run in parallel with growth conditions maintained at set-point so removing some of the variance in the growth data seen with the shake flask cultures.

6.2.2.1 Carbon source: Micro 24 experiments

The Alcofree and *wt* W303-1A strains were used to provide a comparison, as the Alcofree strain has an altered (chimeric) hexose transporter which confers its solely respiratory phenotype.

Both 2xCBS and CSM media were supplemented with a complete amino acid solution and used to evaluate ethanol, glycerol, lactate and acetate as replacement carbon sources for glucose. The Micro 24 parallel bioreactor was used to reduce the experimental time, compared to using a traditional stirred tank reactor (STR) such as the 7 L glass vessel used in the scale up/down experiments in Chapter 3. Each well on the incubation cassette has a total volume of 10 mL and a working volume of 4-7 mL; 6 mL was selected as the starting working volume as a number of samples would be withdrawn to determine biomass accumulation.

The total available carbon was maintained at 2 % w/v with the ratio of the individual carbon sources (ethanol, glycerol, lactate and acetate) being varied. A plate layout was designed and used for the experimental run as shown in Table 6.1.

Plate layout	1	2	3	4	5	6								
A														
B														
C														
D														

Well	Strain	Medium	% Carbon source					xDO	OD _{600nm}				
			Glucose	Ethanol	Glycerol	Lactate	Acetate		0	18	25	42	48
A1	Alcofree	2xCBS	100					1	0.2	8.0	18.4	22.5	26.0
A2	Alcofree	2xCBS	100					1	0.2	14.7	25.5	25.8	28.8
A3	Alcofree	2xCBS	100					1	0.2	16.9	24.3	23.4	29.4
A4	Alcofree	CSM	100					1	0.2	4.3	8.8	10.8	12.1
A5	Alcofree	CSM	100					1	0.2	4.1	8.3	10.6	10.2
A6	Alcofree	CSM	100					1	0.2	3.9	6.8	9.4	8.3
B1	Alcofree	CSM			100			1	0.2	2.6	3.3	4.4	4.7
B2	Alcofree	CSM			100			1	0.2	2.5	3.4	5.9	5.7
B3	Alcofree	CSM			100			4	0.2	2.4	2.6	3.2	3.8
B4	Alcofree	CSM			100			4	0.2	2.5	2.8	3.6	3.9
B5	Alcofree	CSM			100			4	0.2	2.2	2.6	2.9	3.5
B6	Alcofree	CSM			50		50	4	0.2	1.5	2.0	3.2	4.2
C1	W303-1A	2xCBS	100					1	0.2	1.6	2.2	4.0	4.5
C2	W303-1A	2xCBS	100					1	0.2	1.6	2.0	2.4	3.4
C3	W303-1A	CSM	100					1	0.2	2.4	3.3	4.0	4.4
C4	W303-1A	CSM	100					1	0.2	2.4	3.3	4.0	4.6
C5	W303-1A	2xCBS		100				1	0.2	1.4	2.2	3.0	3.7
C6	W303-1A	2xCBS			100			1	0.2	1.7	2.2	3.0	3.4
D1	W303-1A	2xCBS				100		1	0.2	1.4	1.9	2.6	2.7
D2	W303-1A	2xCBS					100	1	0.2	1.0	1.5	1.8	2.0
D3	W303-1A	2xCBS		33	33		33	1	0.2	1.6	2.2	3.0	3.8
D4	W303-1A	2xCBS		25	25	25	25	1	0.2	1.8	2.2	3.3	4.1
D5	W303-1A	2xCBS		50	30	10	10	1	0.2	1.7	2.3	2.8	3.6
D6	W303-1A	2xCBS		60	30	5	5	1	0.2	1.4	2.1	3.1	3.5

NOTE: A1 is at 28 °C not 30 °C

Table 6.1. Micro 24 well information for alternative carbon source experiment. *S. cerevisiae* strain, carbon source composition and OD₆₀₀ measurements are shown.

The cultures were set-up with 5.7 mL 2xCBS (w/o carbon source) in each well, 300 µL (total) carbon source was then added with 4.5 µL P2000 antifoam. The plate was equilibrated to the run condition set-points (30 °C, pH 6.0, 20 % DO) prior to inoculation to a target of OD₆₀₀ 0.2. The culture was incubated for 48 h with samples taken for biomass estimation by OD₆₀₀ at time points 18, 25, 42 and 48 h post inoculation. Figure 6.3 shows the growth curves for all cultures described in Table 6.1.

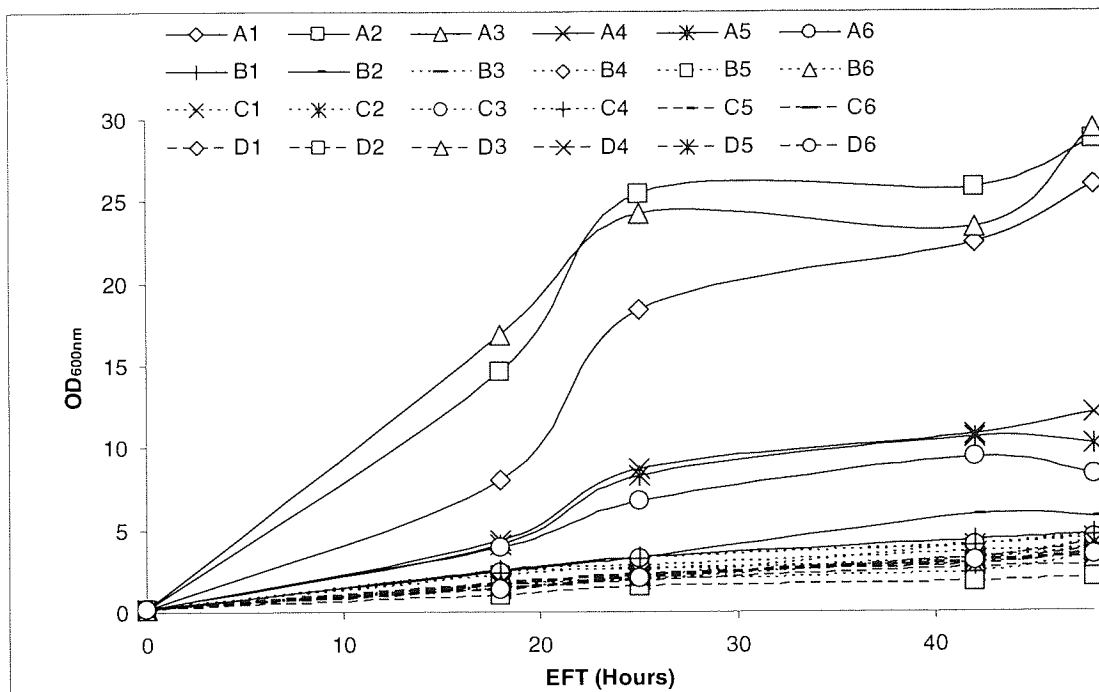


Figure 6.3. Micro 24 culture growth profiles for *S. cerevisiae* Alcofree and W303-1A strains grown on CSM and CBS media with variable amino acid and carbon source compositions (as indicated by Table 6.1)

The W303-1A cultures achieved higher biomasses than in the previous run; these are shown in isolation by Figure 6.4. The W303-1A strain had the best growth on CSM and 2xCBS media with glucose as the sole carbon source (OD_{600} 4.6 and OD_{600} 4.5, respectively) followed by 2xCBS with 25 % ethanol, 25 % glycerol, 25 % acetate, 25 % lactate (OD_{600} 4.1). The lowest biomass achieved was observed with the 2xCBS with acetate as the sole carbon source (OD_{600} 2.0).

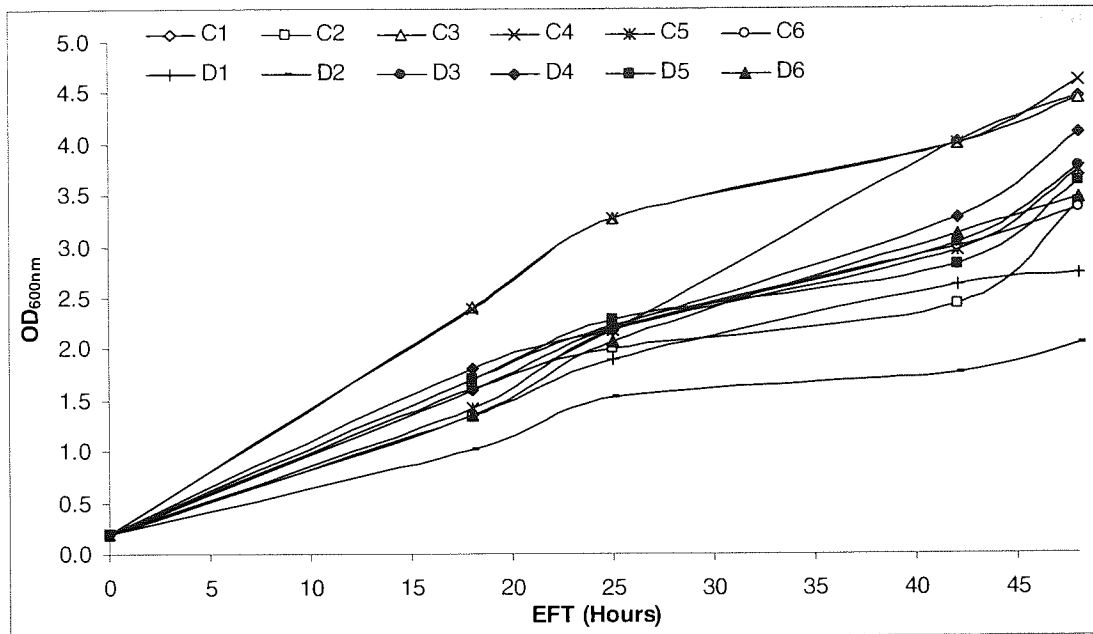


Figure 6.4. Micro 24 culture growth profiles for *S. cerevisiae* W303-1A grown on CSM and CBS media with variable carbon source compositions (well contents as indicated by Table 6.1) Cultures C1 and C2 are replicates as are C3 and C4.

The media compositions were ranked by achievable biomass at harvest (48 h post inoculation) and represented in Figure 6.5 The use of 2 % glycerol as the sole carbon source is favourable in both 2xCBS and (to a lesser extent) CSM media.

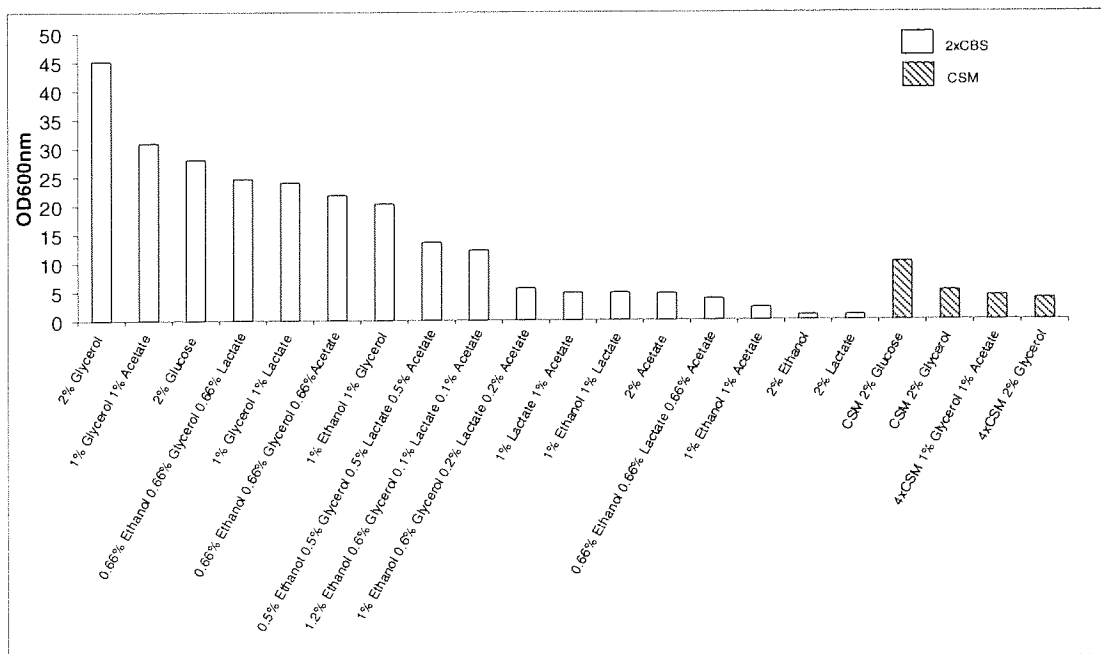


Figure 6.5. Ranked 2xCBS and CSM media carbon source composition, based on *S. cerevisiae* Alcofree Micro 24 culture biomass at 48 h post inoculation.

Further validation of these results was not carried out due to the constraints of equipment availability. The Micro 24 system was on loan from Applikon Biotechnology for a fixed

period. However, the use of glycerol with 2xCBS can be seen to provide increased achievable biomass, compared to standard 2 % glucose 2xCBS and this could be the subject of future work.

6.2.3 Nutrient addition: Micro24 experiments

The use of simple nutrient additions to boost achievable maximum Alcofree biomass had previously been shown to be effective by other researchers in our laboratory (data not shown here). To follow this up the addition of inositol, biotin, copper, zinc and cobalt to CSM medium was examined. It was hypothesised that the composition of standard CBS and CSM media may not be optimal for biomass accumulation. Therefore, the amino acid and trace element/vitamin concentration of the CBS and CSM were also investigated. This was achieved by developing the novel media 2xCSM, 3xCSM and 4xCSM. These contained 2, 3 and 4 fold more amino acid or vitamin/trace element solutions. Table 6.2 shows the initial experimental plan for the Micro 24 nutrient addition cultures.

Micro 24 incubation cassette well layout:

	1	2	3	4	5	6
A						
B						
C						
D						

CSM (2 % glucose)

	A1	B1	C1	D1	A2	B2	C2	D2	A3	B3	C3	D3	A4	B4	C4
inositol	+	+	+	+	+	+			+			+	+	+	+
copper	+	+				+	+		+		+	+	+		
zinc	+	+	+				+		+	+	+			+	
biotin	+	+		+				+		+	+	+		+	+
cobalt	+				+			+		+			+		+

Well	Medium
A5	2xCBS
B5	3xCBS
C5	4xCBS
D4, D5, A6	CSM
D6	2xCSM
B6	3xCSM
C6	4xCSM

Table 6.2. Initial Micro 24 incubation cassette well layout for the nutrient addition and novel media cultures. The CSM media additions required for each well are indicated. Additions were made to the level found in CBS medium with the exception of inositol (10 µg/mL). The amino acid and trace element/vitamin were the only components altered in the novel CBS and CSM media.

The cultures were set-up with 6 mL CSM in each well, 20 µL of each nutrient was then added with 4.5 µL P2000 antifoam. The plate was equilibrated to the run condition set-points (30 °C, pH 6.0, 50 % DO) prior to inoculation to a target of OD₆₀₀ 0.2. The culture was incubated for 48 h with samples taken for biomass estimation by OD₆₀₀ at time points 9, 24, 33 and 48 h

post inoculation. Table 6.3 shows the media, nutrient additions and OD₆₀₀ values for all cultures. Figure 6.6 shows the growth curves for all cultures described in Table 6.3.

Well	Media	Additions					OD600nm				
		Inisitol	Copper	Zinc	Biotin	Cobalt	0	9	24	33	48
A1	CSM	20	20	20	20	20	0.2	0.968	2.98	3.08	2.44
A2	CSM	20				20	0.2	1.188	3.18	3.4	2.74
A3	CSM	20	20	20			0.2	1.472	3.48	3.28	3
A4	CSM	20	20			20	0.2	1.16	3.04	3.34	2.7
A5	2xCBS						0.2	0.204	0.205	0.32	0.3
A6	CSM						0.2	0.928	2.9	2.7	2.32
B1	CSM	20	20	20	20		0.2	1.5	2.92	2.82	2.3
B2	CSM	20	20				0.2	1.332	3.84	3.7	3.1
B3	CSM			20	20	20	0.2	1.464	2.8	2.8	2.04
B4	CSM	20		20	20		0.2	1.64	3	2.76	2.42
B5	3xCBS						0.2	1.92	0.14	0.24	0.32
B6	3xCSM						0.2	1.248	6.22	6.48	6.88
C1	CSM	20		20			0.2	1.492	3.32	3.44	3.04
C2	CSM		20	20			0.2	1.492	3.32	3.32	2.92
C3	CSM		20	20	20		0.2	1.536	2.66	2.76	2.36
C4	CSM	20			20	20	0.2	1.52	2.68	2.6	1.96
C5	4xCBS						0.2	0.216	0.2	0.22	0.4
C6	4xCSM						0.2	1.308	6.88	7.84	8.42
D1	CSM	20			20		0.2	1.372	3.36	3.28	2.42
D2	CSM				20	20	0.2	1.208	2.66	2.88	2.58
D3	CSM	20	20		20		0.2	1.216	2.72	3.08	2.24
D4	CSM						0.2	1.4	3.46	3.74	3.18
D5	CSM						0.2	1.46	3.44	3.72	2.7
D6	2xCSM						0.2	1.532	5.5	5.94	5.74

Table 6.3. Micro 24 well information for nutrient addition and novel medium experiment. *S. cerevisiae* WT W303-1A strain was used in all wells. Medium composition and OD₆₀₀ measurements are shown.

The most noticeable improvements in biomass were observed with the CSM media cultures with an increased total amino acid concentration (See Figure 6.6). The biomass yields at 48 h being 4xCSM (OD₆₀₀ 8.4), 3xCSM (OD₆₀₀ 6.9), 2xCSM (OD₆₀₀ 5.7) and CSM (OD₆₀₀ 2.7). This improvement was not dramatic in CBS medium when the vitamin/trace element concentration was increased a similar way: 4xCBS (OD₆₀₀ 0.4), 3xCBS (OD₆₀₀ 0.32) and 2xCBS (OD₆₀₀ 0.3).

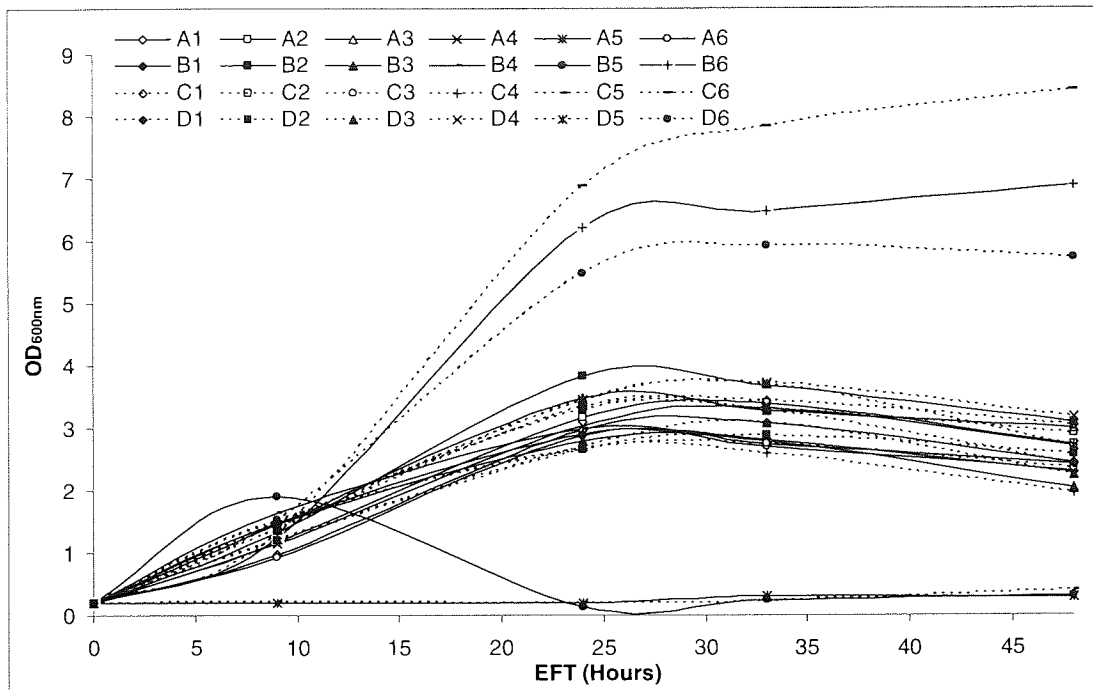


Figure 6.6. Micro 24 culture growth profiles for *S. cerevisiae* W303-1A grown in media as indicated by Table 6.3. The highest biomass was achieved 4xCSM followed by 3xCSM and 2xCSM. The simple nutrient additions lead to only small improvements to biomass in comparison to increasing the total amino acid composition of CSM the medium.

The simple addition of inositol, biotin, copper, zinc and cobalt to the CSM medium did not lead to any noticeable improvement in biomass accumulation during the 48 h culture. This was unexpected as these had been shown previously to improve biomass; therefore, a repeat of this experiment was made. Table 6.4 details the media, nutrient additions and OD₆₀₀ values for the repeat experiment.

Micro24 well layout						
	1	2	3	4	5	6
A						
B						
C						
D						

Well	Strain	Medium	Additions						OD _{600nm}						
			Inisitol	Copper	Zinc	Biotin	Cobalt	xDO	0	17.5	21	28	42	48	
A1	W303-1A	CSM	20	20	20	20	20	1	0.20	1.22	1.44	1.90	2.73	2.59	
A2	W303-1A	CSM	20				20	1	0.20	1.08	1.20	1.36	1.76	1.71	
A3	W303-1A	CSM	20	20	20			1	0.20	1.37	1.73	2.12	2.44	2.46	
A4	W303-1A	CSM	20	20				20	1	0.20	1.05	1.46	1.82	2.25	2.32
A5	W303-1A	2xCBS						1	0.20	1.60	2.08	2.52	3.06	3.09	
A6	W303-1A	CSM						1	0.20	1.42	1.93	2.26	2.37	2.20	
B1	W303-1A	CSM	20	20	20	20		1	0.20	1.37	1.80	2.25	2.47	2.43	
B2	W303-1A	CSM	20	20				1	0.20	1.64	2.20	2.84	3.25	3.25	
B3	W303-1A	CSM			20	20	20	1	0.20	1.51	1.81	2.16	2.42	2.46	
B4	W303-1A	CSM	20		20	20		1	0.20	1.49	1.93	2.16	2.41	2.29	
B5	W303-1A	3xCBS						1	0.20	1.72	2.15	2.68	3.40	3.23	
B6	W303-1A	CSM						3	0.20	3.50	4.62	6.62	6.34	6.36	
C1	W303-1A	CSM	20		20			1	0.20	1.45	1.78	2.05	2.34	2.40	
C2	W303-1A	CSM		20	20			1	0.20	1.47	1.84	2.24	2.46	2.43	
C3	W303-1A	CSM		20	20	20		1	0.20	1.47	1.71	2.15	2.45	2.42	
C4	W303-1A	CSM	20			20	20	1	0.20	1.19	1.39	1.58	1.66	1.69	
C5	W303-1A	4xCBS						1	0.20	1.68	2.08	2.41	2.81	2.81	
C6	W303-1A	CSM						4	0.20	3.97	5.16	6.62	7.70	7.34	
D1	W303-1A	CSM	20			20		1	0.20	1.50	2.04	2.65	3.24	3.42	
D2	W303-1A	CSM				20	20	1	0.20	1.10	1.22	1.48	1.56	1.50	
D3	W303-1A	CSM	20	20		20		1	0.20	1.13	1.23	1.42	1.52	1.45	
D4	W303-1A	CSM						1	0.20	1.68	1.88	2.67	3.08	3.27	
D5	W303-1A	CSM						1	0.20	1.56	1.76	2.51	3.06	3.17	
D6	W303-1A	CSM						2	0.20	2.60	3.52	4.96	5.48	5.47	

Table 6.4. Micro 24 well information for repeat nutrient addition and novel medium experiment. *S. cerevisiae* WT W303-1A strain was used in all wells, Medium composition and OD₆₀₀ measurements are shown.

The cultures were set-up with 6 mL CSM in each well, 20 μ L of each nutrient was then added with 4.5 μ L P2000 antifoam. The plate was equilibrated to the run condition set-points (30 $^{\circ}$ C, pH 6.0, 50 % DO) prior to inoculation to a target of OD₆₀₀ 0.2. The culture was incubated for 48 h with samples taken for biomass estimation by OD₆₀₀ at time points 0, 17.5, 21, 28, 42 and 48 h post inoculation. Table 6.4 shows the media, nutrient additions and OD₆₀₀ values for all cultures. Figure 6.7 shows the growth curves for all cultures described in Table 6.4.

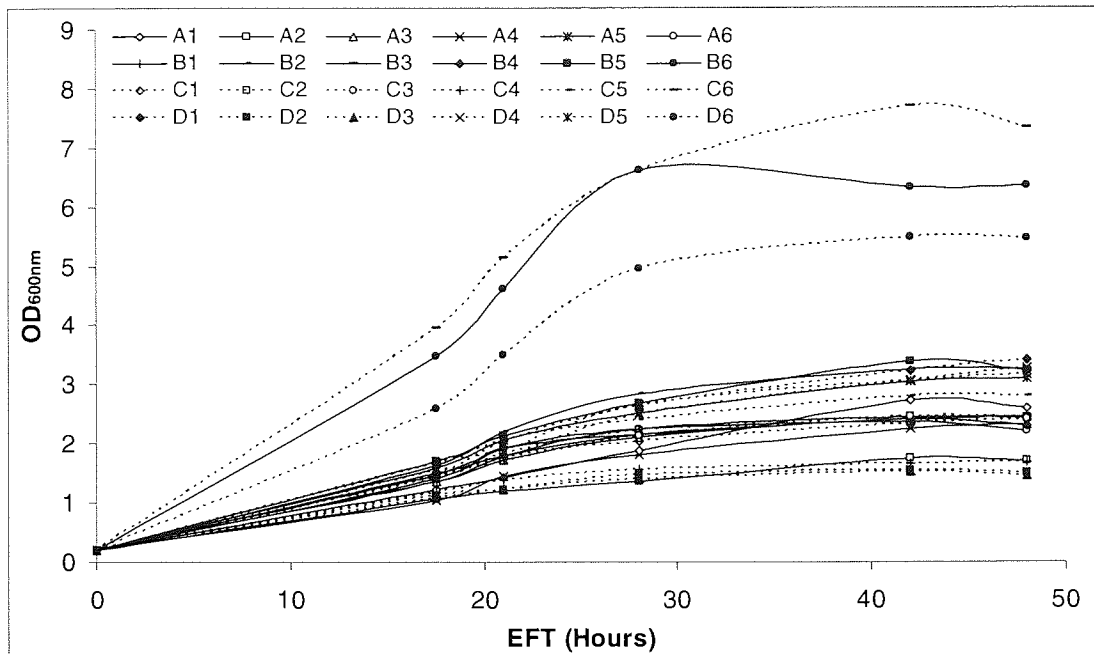


Figure 6.7. Repeat Micro 24 culture growth profiles for *S. cerevisiae* W303-1A grown in media as indicated by Table 6.4.

Again the most noticeable improvements in biomass were seen in the CSM media cultures with an increased total amino acid concentration (see Figure 6.7). The biomass yields at 48 h being 4xCSM (OD_{600} 7.3), 3xCSM (OD_{600} 6.4), 2xCSM (OD_{600} 5.5) and CSM (OD_{600} 3.4). This improvement was not dramatic in CBS medium when the vitamin/trace element concentration was increased a similar way: 4xCBS (OD_{600} 2.8); 3xCBS (OD_{600} 3.2) and 2xCBS (OD_{600} 3.1).

Figure 6.8 shows the combined data for *S. cerevisiae* W303-1A achieved biomass with simple additions of inositol, biotin, copper, zinc and cobalt to CSM media. With the exception of the inositol, biotin, cobalt combination all of the addition combinations showed, on average, an improvement over standard CSM. The increase in biomass is greatest with CSM + inositol and copper. CSM + inositol and biotin also showed an increase but with a larger error. These mean values are based on just two replicates, therefore additional repeats should be run to confirm these differences are not solely due to run to run variation.

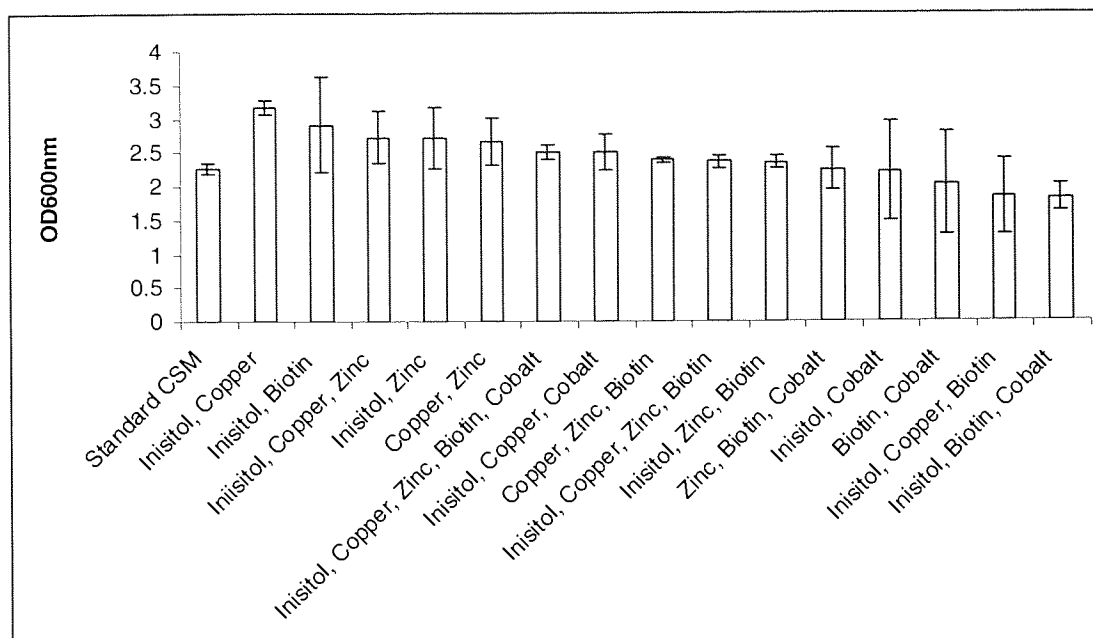


Figure 6.8. Combined biomass data from *S. cerevisiae* W303-1A Micro 24 cultures at 48 hours post inoculation. CSM medium was used with simple additions as indicated. Error bars represent the standard deviation.

6.3 Conclusions on improving *S. cerevisiae* biomass

The standard media (CSM and 2xCBS) used for *S. cerevisiae* cultures can be improved upon for biomass generation with the Alcofree and W303-1A strains. The work presented here has indicated that using non-fermentable carbon sources, such as ethanol and glycerol can permit a higher biomass to be attained. The amino acid and trace element components of widely-used media can be developed to support such increased cell densities. The estimation of biomass could have been calculated more accurately by using dry cell weights, although this would reduce the throughput achieved using the OD₆₀₀ measurements. The improvements in achievable biomass observed here, at the bench scale (< 7 L), may not be practical or possibly economically viable at larger scale industrial use, due to the cost implications of the additional amino acid and trace elements. One of the main “selling points” for using yeast as an expression system is the low cost of goods, removing this may make it less desirable. Nonetheless if volumetric yield is improved, this could counteract the increase in media cost, especially on smaller scales.

“Been dazed and confused for so long it's not true”

(Dazed and confused, Led Zeppelin)

This thesis is concerned with the production of therapeutically-applicable recombinant proteins and particularly in understanding how to design more effective production systems for optimising development parameters and shortening the development time frame. The work described focused on the use of a design-of-experiments approach in a multi-well mini-bioreactor to enable the rapid establishment of high yielding production conditions in yeast and mammalian cell lines.

This project aimed to generate a scalable predictive modelling methodology suitable for the optimisation of range of industrially relevant recombinant proteins and host systems. The yeast *P. pastoris* (expressing GFP) was used for developing the methods. The main attraction of using GFP as a target protein for developing our optimisation methods was the rapid assay techniques that could be applied to determine secreted product yield in the culture supernatant *i.e.* fluorescence analysis of culture supernatant. The approach was then evaluated using a mammalian cell line (hybridoma) expressing a monoclonal antibody to confirm the transferability of the modelling process. Additional factors (medium composition, antifoam agent use, induction strategy and cell surface hydrophobicity) which may also impact target protein yield were identified and investigated for further optimisation of the process *i.e.* maximising product yield. Table 7.1 lists the objectives and current status of the project

Objective	Status
Identify appropriate target proteins	GFP, anti-tumour factor antibody, insulin, hPGH, novel antibody based vaccine (Chapter 1)
Create strains/transformants for each host system to express the target proteins	X33-GFP, HD1, Alcofree MPI, <i>P. pastoris</i> and Alcofree expressing an antibody-based vaccine (Chapters 3 and 4)
Identify suitable factors/variables for process optimisation (applicable to small and large-scale)	T, pH, DO, glucose concentration (Chapters 1, 3 and 4)
Identify suitable DoE/modelling processes	Box-Behnken, response surface methods (Chapters 1, 3 and 4)
Implement proposed modelling process	X33-GFP, Micro 24 (DoE1 and DoE2) (Chapter 3)
Confirm the predictive capacity of the model at the small-scale	Micro 24 parallel mini-bioreactor (6 mL) (Chapter 3)
Confirm the predictive capacity of the model at the large(r)-scale	7 L stirred tank reactor (Chapter 3)
Confirm the scalable modelling process with an alternative target protein and expression system	Anti-tumour factor antibody, HD1 (DoE3) (Chapter 4)
Identify and investigate further factors applicable to process optimisation	Medium composition, cell surface hydrophobicity, antifoam use (Chapters 5 and 6)

Table 7.1. Thesis objectives and current status.

Chapter 3 describes how a DoE-derived model was generated using a parallel mini-bioreactor that was predictive not only of protein yield normalised for culture density in the same system, but is also directly scalable to cultures grown in a 7 L bioreactor. By further optimising both the accumulation of cell density in batch and improving the fed-batch induction regime, we then showed that additional yield improvement could be achieved. By essentially investigating the cell growth and productivity phases separately, our approach has the benefit of reducing process development time by allowing the direct scale-up of the best production conditions from bench to bioreactor.

The limits of the design space used to construct the DoE1 model were shown to not capture a truly optimised set of culture conditions (T, pH and DO). The DoE1 model building and validation process had shown, however, that the methods devised would work provided the correct question was asked of the DoE. A second round of model building (DoE2) was therefore devised based on the Box-Behnken design with the design space moved towards the summit of the DoE1 yield 'mountain': towards lower temperature and higher pH and DO values. The DoE2-predicted optimal conditions (20.2 °C, pH 7.76 and 90 % DO) were run at the small-scale and gave the highest experimental yield of those conditions tested.

The DoE2 model was subjected to a scale-up process aimed at proving that a model built using data from small-scale (6 mL) cultures could be directly used to optimise culture conditions with five hundred fold increase in volume (3 L). With the exception of one result from f004 "best small-scale experimental fit" conditions (27.5 °C, pH 6.5, 60 % DO) the predicted yields give a good approximation of both small and large scale experimental normalised yield.

It was noted that there was scope for further optimisation of the large-scale process, which was not possible at the small-scale due to limitations of the equipment *e.g.* no provision for on-line liquid additions to the cultures. Two improvements were suggested by the initial DoE2 scale-up experiments; change the induction protocol (*i.e.* methanol feed) and increase the pre-induction biomass. These alterations can be used in addition to the optimised process conditions (temperature, pH and DO) predicted by the small-scale modelling *e.g.* DoE2 optimised conditions (21.5 °C, pH 7.6, 90 % DO). Increasing pre-induction biomass by increasing the carbon source (glycerol) in the medium (BMGY) from 10 g/L to 40 g/L, during growth phase was found to improve the volumetric yield for DoE2-predicted optimal conditions. Additionally switching to a mixed feed (60 % w/v sorbitol, 40 % v/v methanol) induction strategy was seen to eliminate methanol accumulation and associated deleterious effects, leading to further improvements in total and normalised yield.

Figure 7.1 illustrates the optimisation process developed. As recombinant protein expression is complex involving many factors, a truly optimal process may be difficult to achieve. However, the cyclical nature of optimisation methods can provide substantial improvements to the process. The key consideration in using this work scheme is posing an appropriate question of the DoE which can be fulfilled by both the small and larger-scale equipment.

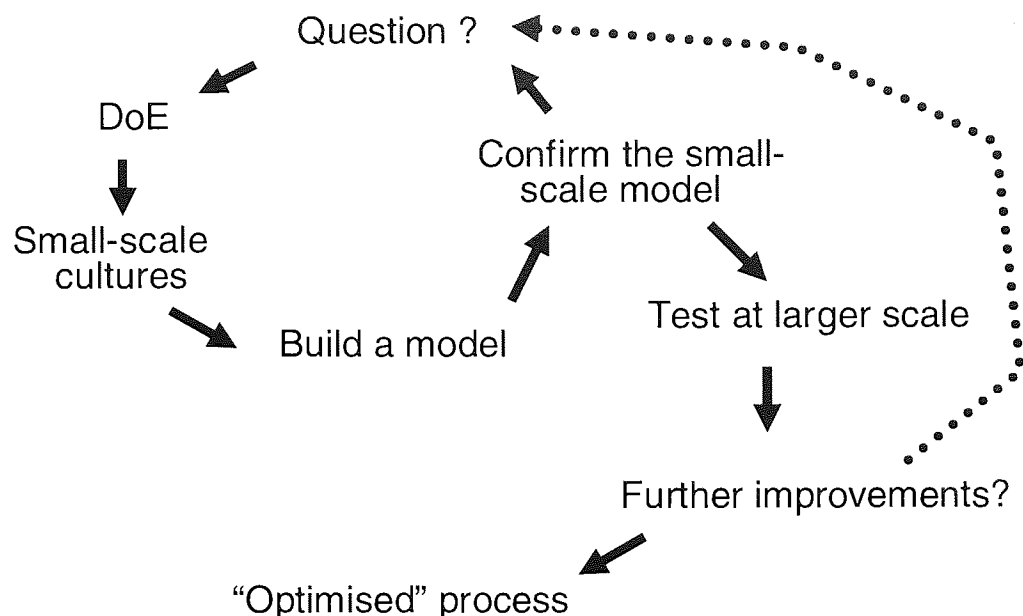


Figure 7.1. Overview of the optimisation process used for maximising GFP expression from *P. pastoris* X33GFP bioreactor cultures as well as other target proteins and expression systems.

As part of the partnership between Aston University and Alpha Biologics a hybridoma cell line (HD1) was made available for use. The cell line expressed a monoclonal antibody raised against a tumour-inducing factor; possible uses of such an antibody include attaching chemotherapy agents for direct targeting of tumour cells.³³ Applying our optimisation strategy to antibody expression in HD1 challenged the robustness of the optimisation methodologies.

The question asked of DoE3 in Chapter 4 was whether altering the process conditions (culture temperature, pH, DO and glucose concentration) would increase the product (antibody) yield from the expression system (HD1, hybridoma cell line). The Micro 24 was used to generate small-scale yield data for building the DoE3 predictive model. DoE3 suggested that optimal expression conditions were 38.0 °C, pH 6.6 and 10 mg/mL glucose, with DO not being significant. The *p*-values given by the DoE3 model for each term are much lower than those from the DoE2 (*P. pastoris* GFP expression) model indicating a higher quality model, statistically. However, the predictive capability for experimental values was poorer than expected, with the model overestimating the yield. This was most likely due to variability in the experimental model building data and no repeat data from the model confirmation runs. Despite this the predicted conditions did provide the highest yields.

Having shown the DoE3 model was able to predict optimal conditions at the small-scale in the Micro 24 system the model was tested at a larger scale as with the DoE2 model. The

DoE3 scale-up experiments used a bi-phase process with the check point conditions run for the first 4 days to provide comparative data for model scalability. The process conditions were then switched to the DoE3 predicted optimal conditions for the second phase of the experiment. This bi-phase process is more representative of a production experiment where cell biomass is generated using the optimal growth conditions (standard conditions) and then optimal expression conditions used to gain maximal product from the high biomass. BR01 demonstrated that yield can be increased by using a bi-phase process, the 2 day second phase using DoE3 optimised conditions gave a higher yield (1.717 µg/mL) than either standard (BR01 phase 1) or best small-scale experimental fit (BR03) conditions over 4 days. The DoE3 scale-up process highlighted some deficiencies in the process, most notably seeding density and cell survival at optimal expression conditions. The bi-phase conditions used in BR01 could be further developed to give a higher yielding process.

Further target proteins were selected: insulin (MPI), human placental growth hormone (hPGH), and a monoclonal antibody based influenza vaccine to extend this approach. Since a positive *P. pastoris* MPI clone was not produced it was not possible to carry out a comparison of MPI expression between the Alcofree strain and *P. pastoris*. The Alcofree MPI work indicated that increased biomass would be needed to easily detect a secreted recombinant protein when constitutive expression is utilised with the pYX2121 plasmid system. The failure to generate a positively-expressing MPI *Pichia* strain, indicated that strain creation is not as simple as suggested by the literature, especially Invitrogen's *Pichia* expression handbook. Further attempts to create hPGH expressing *Pichia sp.* and Alcofree clones/transformants were not pursued. It was noted that the cell densities achieved in shake flask cultures of Alcofree were approximately half those obtained with *Pichia sp.* This is not ideal as the pYX212 vector used with the Alcofree strain confers constitutive recombinant protein expression; therefore, product titer is proportional to cell density. In an attempt to address this optimisation of medium composition was identified as a route to increasing Alcofree biomass and thus product yield, as described in Chapter 6.

Many optimisation experiments described in the literature investigate obvious factors such as culture temperature¹²⁷, pH⁶¹, aeration^{128, 129} and medium composition⁹⁰. However, the optimisation of many other factors offers the possibility for further improvements to the process. In Chapter 5 we examined how cell surface hydrophobicity and antifoam agent use affect the characteristics of recombinant protein expression in yeast systems. If cell physiology changes (*e.g.* cell surface hydrophobicity), the protein production process could,

perhaps, be manipulated to exploit this. In a similar way if simply using alternative antifoam agents or concentrations of the selected antifoam can improve cell growth or product yield, this could prove a convenient way for improving the target protein yield.

Our data suggests that rapidly-growing *P. pastoris* cells, *i.e.* those in mid-log phase, have higher cell surface hydrophobicities (CSH) than cells in lag or stationary phases. This appears to be true for both non-induced and induced cultures. The effects of variable CSH are known to influence other yeast species: flocculation in cultures of *S. cerevisiae*¹³² and surface adhesion in *Candida albicans*.^{136 137} It is not inconceivable therefore that cell aggregation in *P. pastoris* could also be influenced by CSH. This would have an effect on the dynamics of a HCD *P. pastoris* culture, especially when high biomasses are reached of $>500 \text{ OD}_{600\text{nm}}$, equivalent to 130 g/L dry cell weight.¹ During culture, wet cells could represent 10-20 % of the total culture volume. If these formed aggregates (due to increased CSH) and thus increased the viscosity of the culture,¹³⁸ this would have an influence on the mass transfer coefficient (k_{LA}) of the system¹³⁹ since increasing viscosity leads to a decrease in k_{LA} .¹²⁸ Reducing the k_{LA} will lead to decreases in cellular uptake, *e.g.* oxygen and media components, and so reduce cell growth. This could give scope for further optimisation of a HCD culture: altering process factors *e.g.* agitation rate to compensate for the deviation from designed or expected k_{LA} during periods of rapid growth such as mid-log phase cultures. Variable agitation rate during mid-log phase growth could be included as an additional factor in a DoE optimisation matrix. This approach may not be suited to large scale manufacture as bioreactor vessels are likely to be already set-up (run conditions *e.g.* agitation and gas flow rates) to provide maximal k_{LA} for the system. However this may be useful for bench-scale experiments where the system has not already been fully optimised.

We also evaluated the use of four antifoam types as part of Chapter 5: Schill & Schelinger's Struktol SB2121 (Polyalkylenglycol); Schill & Schelinger's Struktol J673A (an alkoxyated fatty acid ester on a vegetable base); Sigma Antifoam C (an aqueous emulsion of Sigma Antifoam A; Siloxane polymer) and Fluka P2000 (Polypropylene glycol). It was found that the different types of antifoam affect *S. cerevisiae* and *P. pastoris* growth in different ways depending on the concentration and medium type being used. The data suggested that the presence of SB2121, although not the other antifoam types, enhanced recombinant protein expression of GFP in *P. pastoris*. This could be associated with an enhanced mass transfer coefficient (k_{LA}) for the shake flask system. The mode of action for antifoam agents is to reduce the surface tension of the liquid, so that bubble formation cannot be supported. It is

feasible that the reduction in surface tension, in the presence of SB2121, results in more efficient oxygen transfer from gas to liquid phases and/or liquid phase to cells.

In Chapter 4 it was highlighted that low *S. cerevisiae* Alcofree biomass hindered the detection of expressed recombinant proteins. This low volumetric yield proved problematic during the positive transformant selection as product was not readily detected. A hundred fold concentration procedure was required for target protein detection from positively expressing cultures. One of the proposed avenues of investigation was to improve the composition of the growth medium, so allowing higher cell densities (biomass) to be met. To this end, alternative carbon sources to glucose and various additives were screened for to try and improve biomass yield. The novel *S. cerevisiae* Alcofree strain and wild type strains W303-1A and VWK70-945 were used for the medium development, described in Chapter 6.

The use of ethanol as a non-fermentable carbon source was investigated for growth of Alcofree and wild type *S. cerevisiae* strains. The concept of using alternatives to glucose as non-fermentable carbon sources such as ethanol, glycerol, lactate and acetate, were further developed. Other means of improving the achievable biomass in CSM medium were also identified such as addition of inositol, biotin, copper zinc and cobalt. The use of 2 % glycerol as the sole carbon source was found to be favourable in both 2xCBS and CSM media. With the exception of the inositol, biotin, cobalt combination, all of the nutrient addition combinations showed, on average, an improvement over standard CSM. The increase in biomass was greatest with CSM + inositol and copper, but CSM + inositol and biotin also showed an increase. This demonstrated that standard media (CSM and 2xCBS) used for *S. cerevisiae* cultures are not fully optimised for biomass generation for the Alcofree and W303-1A strains. The work presented here has indicated that using non-fermentable carbon sources, such as ethanol and glycerol can achieve a higher biomass to be attained. The amino acid and trace element components of widely-used media can be developed to support increased cell densities. The improvements in achievable biomass observed here, may not be practical or economically-viable for larger scale industrial use due to the cost implications of the additional amino acid and trace elements. One of the main selling points for using yeast as an expression system is the low cost of goods. Removing this may make it less desirable. Nonetheless if volumetric yield is improved, this could counteract the increase in media cost.

The optimisation process described here can allow the rapid establishment of robust process conditions for improved recombinant protein yield in a range of expression systems. The

increase in throughput achieved by using the small-scale parallel DoE approach can provide sufficient time for investigating non-traditional process factors, which may otherwise be overlooked or excluded as trivialities. As optimisation of recombinant protein expression is a complex activity, a balance must be made between gaining a full understanding of the process, through inclusion of as many factors as possible and devising a process capable of meeting its targets within budget. The work described in this thesis has contributed to achieving this balance.

“When you can take the pebble from my hand
It will be time for you to leave”

(Master Kan, Kung Fu)

8 REFERENCES

1. Cereghino, J.L. & Cregg, J.M. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.* **24**, 45-66 (2000).
2. Sue Macauley-Patrick, M.L.F., Brian McNeil and Linda M. Harvey Heterologous protein production using the *Pichia pastoris* expression system. *Yeast* **22**, 249-270 (2005).
3. Fredric M. Steinberg, J.R. Biotech Pharmaceuticals and Biotherapy: An Overview. *J Pharm Pharmaceut Sci* **1**, 48-59 (1998).
4. Walsh, G. Biopharmaceutical benchmarks. *Nat. Biotechnol.* **18**, 831-833 (2000).
5. Brunt, J.V. in Signals biotechnology industry analysis 2001).
6. Schmidt, F.R. Recombinant expression systems in the pharmaceutical industry. *Appl. Microbiol. Biotechnol.* **65**, 363-372 (2004).
7. Schiff, L.J. Review: production, characterisation and testing of banked mammalian cell substrates used to produce biological products. *In Vitro Cell Dev Biol* **41**, 65-70 (2005).
8. Jeewon Lee et al. Novel Secretion System of Recombinant *Saccharomyces cerevisiae* Using an N-terminus Residue of Human IL-1 α as Secretion Enhancer. *Biotechnol. Prog.* **15**, 884-890 (1999).
9. Sherman, F. Getting started with yeast. *Methods Enzymol.* **350**, 3-41 (2002).
10. Koti Sreerishna et al. Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. *Gene* **190**, 55-62 (1997).
11. Mewes Boettner, Bianka Prinz, Caterina Holz, Ulf Stahl & Lang, C. High-throughput screening for expression of heterologous proteins in the yeast *Pichia pastoris*. *Biotechnol. J.* **99**, 51-62 (2002).
12. Roslyn M. Bill et al. Analysis of the Pore of the Unusual Major Intrinsic Protein Channel, Yeast Fps1p*. *Biol. Chem.* **276**, 36543-36549 (2001).
13. Karin Otterstedt et al. Switching the mode of metabolism in the yeast *Saccharomyces cerevisiae*. *EMBO reports* **5**, 532-537 (2004).
14. Ehrenberg, M. in Scientific Background on the Nobel Prize in Chemistry 2008(2008).
15. Ormo M et al. Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392-1395 (1996).
16. Tsien, R.Y. The green fluorescent protein. *Annu. Rev. Biochem* **67**, 509-544 (1998).
17. Te-Tuan Yang et al. Improved Fluorescence and Dual Color Detection with Enhanced Blue and Green Variants of the Green Fluorescent Protein. *Biol. Chem.* **273**, 8212-8216 (1998).
18. Heim, R., Prasher, D.C. & Tsien, R.Y. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. USA* **91**, 12501-12504 (1994).
19. Wan-ling Chiu et al. Engineered GFP as a vital reporter in plants. *Curr. Biol.* **6**, 325-330 (1996).
20. Cormack, B.P. (ed. U.S. Patent)USA; 1998).
21. Kirby R. Siemering, Ralph Golbik, Richard Sever & Haseloff, J. Mutations that suppress the thermosensitivity of green fluorescent protein. *Curr. Biol.* **6**, 1653-1663 (1996).
22. Alexander Schulte, Inken Lorenzen, Böttcher, M. & Plieth, C. A novel fluorescent pH probe for expression in plants. *Plant Methods* **2** (2006).
23. Surribas, A., Montesinos, J.L. & Valero, F.F. Biomass estimation using fluorescence measurements in *Pichia pastoris* bioprocess. *J. Chem. Technol. Biotechnol.* **81**, 23-28 (2006).
24. Malea Kneen, Javier Farinas, Yuxin Li & Verkman, A.S. Green Fluorescent Protein as a Noninvasive Intracellular pH Indicator. *Biophys. J.* **74**, 1591-1599 (1998).

25. Jincal Li et al. Green Fluorescent Protein in *Saccharomyces cerevisiae*: Real-Time Studies of the GAL1 Promoter. *Biotechnol. Bioeng.* **70**, 187-196.
26. Deo, S.K. & Daunert, S. Luminescent proteins from *Aequorea victoria*: applications in drug discovery and in high throughput analysis. *Fresen J Anal Chem* **369**, 258-266 (2001).
27. Anna Surribas, David Resina, Ferrer, P. & Valero, F. Limitations using GFP as a protein expression reporter in *Pichia pastoris*. *Microbial Cell Factories* **5**, 56 (2006).
28. George H. Patterson, S.M.K., Wallace D. Sharif, Steven R. Kain, and David W. Piston Use of the Green Fluorescent Protein and Its Mutants in Quantitative Fluorescence Microscopy. *Biophysical journal* **73**, 2782-2790 (1997).
29. Guoping Feng et al. Imaging Neuronal Subsets in Transgenic Mice Expressing Multiple Spectral Variants of GFP *Neuron* **28**, 41-51 (2000).
30. Abbas Belkhiri et al. A Noninvasive Cell-Based Assay for Monitoring Proteolytic Activity within a Specific Subcellular Compartment. *Anal. Biochem.* **306**, 237-246 (2002).
31. Day, R.N. & Periasamy, A. in *Multiphoton Microscopy in the Biomedical Sciences*, Vol. 4262 San Jose, CA, USA; 2001).
32. Takeharu Nagai et al. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **20**, 87-90 (2002).
33. Ghose, T. et al. Immunochemotherapy of Cancer with Chlorambucil-carrying Antibody. *British Medical Journal* **3**, 495-499 (1972).
34. Clark, P.M. Assays for insulin, proinsulin(s) and C-peptide. *Ann Clin Biochem* **36**, 541-564 (1999).
35. Caplan, S., Green, R., Rocco, J. & Kurjan, J. Glycosylation and Structure of the Yeast MFtd α -Factor Precursor Is Important for Efficient Transport through the Secretory Pathway. *J. Bacteriol* **173**, 627-635 (1991).
36. D C Williams, R M Van Frank, Muth, W.L. & Burnett, J.P. Cytoplasmic inclusion bodies in *Escherichia coli* producing biosynthetic human insulin proteins *Science* **215**, 687-689 (1982).
37. J. Markussen et al. in *Peptides*. (ed. D. Theodoropoulos) 189-194 (de Gruyter, Berlin; 1987).
38. Yan Wang et al. Human Insulin from a Precursor Overexpressed in the Methylophilic Yeast *Pichia pastoris* and a Simple Procedure for Purifying the Expression Product. *Biotechnol. Bioeng.* **73**, 74-79 (2001).
39. Thomas Kjeldsen, Annette Frost Pettersson & Hach, M. The role of leaders in intracellular transport and secretion of the insulin precursor in the yeast *Saccharomyces cerevisiae*. *Biotechnol. J.* **75**, 195-208 (1999).
40. Chang, S.-G., Kim, D.-Y., Choi, K.-D., Shin, J.-M. & Shin, H.-C. Human insulin production from a novel mini-proinsulin which has high receptor-binding activity. *Biochem. J.* **329**, 631-635 (1998).
41. Thomas Kjeldsen, Pettersson, A.F. & Hach, M. Secretory expression and characterization of insulin in *Pichia pastoris*. *Biotechnol. Appl. Biochem.* **29**, 79-86 (1999).
42. Kjeldsen, T. Yeast secretory expression of insulin precursors. *Appl. Microbiol. Biotechnol.* **54**, 277-286 (2000).
43. Igout, A. et al. Purification and biochemical characterization of recombinant human placental growth hormone produced in *Escherichia coli*. *Biochem J.* **295**, 719-724 (1993).
44. Fuglsang, J., Lauszus, F., Flyvbjerg, A. & Ovesen, P. Human Placental Growth Hormone, Insulin-Like Growth Factor I and -II, and Insulin Requirements during Pregnancy in Type 1 Diabetes. *J. Clin. Endocrinol. Metab.* **88**, 4355-4361 (2003).

45. Nickel BE, K.E., Cattini PA. The human placental growth hormone variant is mitogenic for rat lymphoma Nb2 cells. *Endocrinology* **126**, 971-976 (1990).
46. E. Alsat, J. Guibourdenche, A. Couturier & Evain-Brion, D. Physiological role of human placental growth hormone. *Mol. Cell. Endocrinol.* **40**, 121-127 (1998).
47. G. Baviera, C. Carbone, F. Corrado & Mastrantonio, P. Placental growth hormone in Down's syndrome screening *J. Matern. Fetal. Neonatal. Med.* **16**, 241-243 (2004).
48. Gibney, J. et al. The Effects of 10 Years of Recombinant Human Growth Hormone (GH) in Adult GH-Deficient Patients. *J. Clin. End. & Metab.* **84**, 2596-2602 (1999).
49. Bengtsson, B.-A. et al. Treatment of Adults with Growth Hormone (GH) Deficiency with Recombinant Human GH*. *J. Clin. End. & Metab.* **76**, 309-317 (1993).
50. M Saugy et al. Human growth hormone doping in sport *Br. J. Sports Med.* **40**, 35-39 (2006).
51. Perls, T.T. Anti-Aging Medicine: The Legal Issues. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* **59**, 682-691 (2004).
52. Health, N.I.o. in National Hormone and Pituitary Program 2004).
53. Sittiruk Roytrakul, Lily Eurwilaichitr, Suprasongsin, C. & Panyim, S. A Rapid and Simple Method for Construction and Expression of a Synthetic Human Growth Hormone Gene in *Escherichia coli*. *J. Biochem. Molec. Biol.* **34**, 502-508 (2001).
54. Ikehara, M. et al. Synthesis of a gene for human growth hormone and its expression in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **81**, 5956-5960 (1984).
55. Tomoko Tokunaga et al. Expression of a synthetic human growth hormone gene in yeast. *Gene* **39**, 117-120 (1985).
56. Hiramatsu, R., Yamashit, T., Aikawa, J.-I., Horinouchi, S. & Beppu, T. The Prepro-peptide of Mucor Rennin Directs the Secretion of Human Growth Hormone by *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **56**, 2125-2132 (1990).
57. Kyung-Soon Park et al. Identification and Use of Zinc Finger Transcription Factors That Increase Production of Recombinant Proteins in Yeast and Mammalian Cells. *Biotechnol. Prog.* **21**, 664-670 (2005).
58. Corning, D. in Europe product information 2005).
59. Nicklas Bonander et al. Design of improved membrane protein production experiments: Quantitation of the host response *Protein Sci.* **14**, 1729-1740 (2005).
60. Antje Eiden-Plach et al. Viral Prepro-toxin Signal Sequence Allows Efficient Secretion of Green Fluorescent Protein by *Candida glabrata*, *Pichia pastoris*, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe*. *Appl Environ Microbiol* **70**, 961-966 (2004).
61. Thiry, M. & Cingolani, D. Optimising scale-up fermentation processes. *Trends Biotechnol.* **20**, 103-105 (2002).
62. Micheletti, M. & Lye, G.J. Microscale bioprocess optimisation. *Curr. Opin. Biotechnol* **17**, 611-618 (2006).
63. Chinyuan Cheng & Yang, S.-T. Dynamics and Modeling of Temperature- Regulated Gene Product Expression in Recombinant Yeast Fermentation. *Biotechnol. Bioeng.* **50**, 663-674 (1996).
64. Oriol Cos, Ramón Ramón, José Luis Montesinos & Valero, F. Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: A review. *Microbial Cell Factories* **5** (2006).
65. E. C, elik, P.C.a., S.M. Halloran, S.G. Oliver Production of recombinant human erythropoietin from *Pichia pastoris* and its structural analysis. *Journal of Applied Microbiology* **103**, 2084-2094 (2007).
66. Ru-Shi Liu et al. High-yield expression of recombinant SARS coronavirus nucleocapsid protein in methylotrophic yeast *Pichia pastoris*. *World J Gastroenterol* **10**, 3602-3607 (2004).

67. P. A. Gibbs & Seviour, R.J. Does the agitation rate and/or oxygen saturation influence exopolysaccharide production by *Aureobasidium pullulans* in batch culture? *Appl. Microbiol. Biotechnol* **46**, 503-510 (1996).
68. Lu Han & Al-Dahhan, M.H. Gas-liquid mass transfer in a high pressure bubble column reactor with different sparger designs. *Chem. Eng. Sci.* **62**, 131-139 (2007).
69. Kaster, J.A., Michelsen, D.L. & Velandar, W.H. Increased Oxygen Transfer in a Yeast Fermentation Using a Microbubble Dispersion. *Appl. Biochem. Biotechnol.* **24/25**, 469-484 (1990).
70. P. Ambati & Ayyanna, C. Optimising medium constituents and fermentation conditions for citric acid production from palmyra jaggery using response surface method. *World J. Microbiol. Biotechnol.* **17**, 331-335 (2001).
71. Weuster-Botz, D. Experimental Design for Fermentation Media Development: Statistical Design or Global Random Search? *J. Biosci. Bioeng.* **90**, 473-483 (2000).
72. Quang D. Nguyen, Judit M. Rezessy-Szabó & Hoschke, Á. Optimisation of Composition of Media for the Production of Amylolytic Enzymes by *Thermomyces lanuginosus* ATCC 34626. *Food technol. biotechnol.* **38**, 229-234 (2000).
73. Olson, B.H. & Johnson, M.J. Factors producing high yeast yields in synthetic media. *J. Bacteriol.* **57**, 235-246 (1948).
74. Varley, J., Brown, A.K., Boyd, J.W.R., Dodd, P.W. & Gallagher, S. Dynamic multi-point measurement of foam behaviour for a continuous fermentation over a range of key process variables. *Biochem. Eng. J.* **20**, 61-72 (2004).
75. Marc C d'Anjou, A.J.D. mixed-feed exponential feeding for fed-batch culture of recombinant methylotrophic yeast. *biotechnology letters* **22**, 341-346 (2000).
76. Thorpe, E.D., d'Anjou, M.C. & Daugulis, A.J. Sorbitol as a non-repressing carbon source for fed-batch fermentation of recombinant *Pichia pastoris*. *Biotechnol. Lett.* **21**, 669-672 (1999).
77. Kupcsulik, B. & Sevelia, B. Effects of methanol concentration on the recombinant *Pichia pastoris* MUTS fermentation. *Periodica Polytechnica Ser. Chem. Eng.* **48**, 73-87 (2004).
78. Stefan Minning et al. Optimization of the high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris*. *biotechnol. J.* **86**, 59-70 (2001).
79. Wenhui Zhang, Mark A. Bevins, Bradley A. Plantz, Leonard A. Smith & Meagher, M.M. Modeling *Pichia pastoris* Growth on Methanol and Optimizing the Production of a Recombinant Protein, the Heavy-Chain Fragment C of Botulinum Neurotoxin, Serotype A. *Biotechnol. Bioeng.* **70**, 1-8 (2000).
80. Duetz, W.A. Microtiter plates as mini-bioreactors: miniaturization of fermentation methods. *Trends Microbiol.* **15**, 469-475 (2007).
81. Betts, J.I. & Baganz, F. Miniature bioreactors: current practices and future opportunities. *Microbial Cell Factories* **5** (2006).
82. Aaron Chen, Rajesh Chitta, David Chang & Amanullah, A. Twenty-Four Well Plate Miniature Bioreactor System as a Scale-Down Model for Cell Culture Process Development. *Biotechnol. Bioeng.* **102**, 148-160 (2009).
83. Burke, F. in Practical fermentation technology. (ed. B.M.a.L.M. Harvey) 231-269 (John Wiley and sons, Chichester; 2008).
84. Anurag Rathore et al. in Pharma. Tech. Europe2006).
85. Donatas Levišauskas & Tekorius, T. Model-based optimisation of fed-batch fermentation processes using predetermined type feed-rate time profiles. A comparative study. *Information Technology and Control* **34**, 231-236 (2005).
86. R. Sen & Swaminathan, T. Application of response-surface methodology to evaluate the optimum environmental conditions for the enhanced production of surfactin. *Appl. Microbiol. Biotechnol* **47**, 358-363 (1997).

87. Henry Lin, Tina Kim, Fei Xiong & Yang, X. Enhancing the Production of Fc Fusion Protein in Fed-Batch Fermentation of *Pichia pastoris* by Design of Experiments. *Biotechnol. Prog.* **23**, 621-625 (2007).
88. Mehmet Inan et al. Optimization of temperature-glycerol -pH conditions for fed-batch fermentation process for recombinant hookworm (*Ancylostoma caninum*) anticoagulant peptide (AcAP-5) production by *Pichia pastoris*. *Enzyme Microb. Technol.* (1999).
89. Ralph Berkholz, Dirk Rohlig & Guthke, R. Data and knowledge based experimental design for fermentation process optimization. *Enzyme Microb. Technol.* **27**, 784-788 (2000).
90. Emrah Nikerela, Ebru Toksoy, Betül Kırdar & Ramazan Yıldırım Optimizing medium composition for TaqI endonuclease production by recombinant *Escherichia coli* cells using response surface methodology. *Process Biochem.* **40**, 1633-1639 (2005).
91. Francis, F. et al. Use of response surface methodology for optimizing process parameters for the production of α -amylase by *Aspergillus oryzae*. *Biochem. Eng. J.* **15**, 107-115 (2003).
92. Dhanya Gangadharan, Swetha Sivaramakrishnan, K. Madhavan Nampoothiri, Rajeev K. Sukumaran & Pandey, A. Response surface methodology for the optimization of alpha amylase production by *Bacillus amyloliquefaciens*. *Bioresour. Technol.* **99**, 4597-4602 (2008).
93. Narang, S., Sahai, V. & Bisaria, V.S. Optimization of Xylanase Production by *Melanocarpus albomyces* IIS68 in Solid State Fermentation Using Response Surface Methodology. *J. Biosci. Bioeng.* **91**, 425-427 (2001).
94. Kiran M. Desai, Bhalchandra K. Vaidya, Rekha S. Singhal & Bhagwat, S.S. Use of an artificial neural network in modeling yeast biomass and yield of β -glucan. *Process Biochem.* **40**, 1617-1626 (2005).
95. NIST e-Handbook of Statistical Methods. *SEMATECH* (2006).
96. Box, G.E.P. Use and abuse of regression. *Technometrics* **8**, 625-629 (1966).
97. F Rodríguez-Quiñones & Benedí, V.J. *Escherichia coli* strain DH5a is a suitable for the study of phoA insertions. *Focus* **8**, 110-112 (1986).
98. Hanahan, D. DNA Cloning: A Practical Approach Vol. 1. (IRL Press, Virginia; 1985).
99. Grant, S.G., Jessee, J., Bloom, F.R. & Hanahan, D. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc Natl Acad Sci U S A* **87**, 4645-4649 (1990).
100. Meselson, M. & Yuan, R. DNA restriction enzyme from *E. coli*. *Nature* **217**, 1110-1114 (1968).
101. Studier FW, M.B. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113-130 (1986).
102. Studier FW, R.A., Dunn JJ, Dubendorff JW Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in enzymology* **185**, 60-89 (1990).
103. Casadaban, M.J. & Cohen, S.N. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J Mol Biol* **138**, 179-207 (1980).
104. Invitrogen *Pichia* expression kit handbook. **011102**.
105. Voth, W.P., Olsen, A.E., Sbia, M., Freedman, K.H. & Stillman, D.J. ACE2, CBK1, and BUD4 in Budding and Cell Separation. *Eukaryotic Cell* **4**, 1018-1028 (2005).
106. Fan, H.Y., Cheng, K.K. & Klein, H.L. Mutations in the RNA Polymerase II Transcription Machinery Suppress the Hyperrecombination Mutant hpr1 {Delta} of *Saccharomyces cerevisiae*. *Genetics* **142**, 749-759 (1996).
107. J. P. van Dijken et al. An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains *Enzyme Microb. Technol.* **26**, 706-714 (2000).

108. Elbing K et al. Role of hexose transport in control of glycolytic flux in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **70**, 5323-5330 (2004).
109. Lin-Cereghino, J. et al. Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast *Pichia pastoris*. *Biotechniques* **38**, 44-48 (2005).
110. S. N. Smith, R. Chohan, Armstrong, R.A. & Whipps, J.M. Hydrophobicity and surface electrostatic charge of conidia of the mycoparasite *Coniothyrium minitans*. *Mycol. Res* **102**, 243-249 (1998).
111. Islam, R.S., Tisi, D., Levy, M.S. & Lye, G.J. Framework for the rapid optimization of soluble protein expression in *Escherichia coli* combining microscale experiments and statistical experimental design. *Biotechnol. prog.* **23**, 785-793 (2007).
112. Islam, R.S., Tisi, D., Levy, M.S. & Lye, G.J. Scale-up of *Escherichia coli* growth and recombinant protein expression conditions from microwell to laboratory and pilot scale based on matched k(L)a. *Biotechnol. Bioeng.* **99**, 1128-1139 (2008).
113. Tessa N. Campbell & Choy, F.Y.M. Expression of two green fluorescent protein variants in citrate-buffered media in *Pichia pastoris*. *Anal. Biochem.* **311**, 193-195 (2002).
114. Guohong Zhang, Vanessa Gurtu & Kain, S.R. An Enhanced Green Fluorescent Protein Allows Sensitive Detection of Gene Transfer in Mammalian Cells. *Biochem. Biophys. Res. Commun.* **227**, 707-711 (1996).
115. H J Cha, Srivastava, R., Vakharia, V.N., Rao, G. & Bentley, W.E. Green Fluorescent Protein as a Noninvasive Stress Probe in Resting *Escherichia coli* Cells. *Appl Environ Microbiol* **65**, 409-414 (1999).
116. Andreas Cramer, Erik A. Whitehorn, Tate, E. & Stremmer, W.P.C. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol.* **14**, 315-319 (1996).
117. Ruppert, D. *Statistics and Finance: An Introduction*, Edn. 1. (Springer, 2004).
118. Vijay Chiruvolu, James M. Gregg & Meagher, M.M. Recombinant protein production in an alcohol oxidase-defective strain of *Pichia pastoris* in fedbatch fermentations. *Enzyme Microb. Technol* **21**, 277-283 (1997).
119. Technologies, I.L., Vol. Version B 1-11053002).
120. Wenhui Zhang, Jayanta Sinha & Meagher, M.M. Glycerophosphate as a phosphorus source in a defined medium for *Pichia pastoris* fermentation. *Appl Environ Microbiol* **72**, 139-144 (2006).
121. Emma D Thorpe, M.C.d.A., Andrew J Daugulis Sorbitol as a non-repressing carbon source for fed-batch fermentation of recombinant *Pichia pastoris*. *biotechnology letters* **21**, 669-672 (1999).
122. Thomas Kjeldsen et al. Engineering-enhanced Protein Secretory Expression in Yeast with Application to Insulin. *Biol. Chem* **277**, 18245-18248 (2002).
123. Alimjan Idiris, Kewei Bi, Hideki Tohda, Kumagai, H. & Giga-Hama, Y. Construction of a protease-deficient strain set for the fission yeast *Schizosaccharomyces pombe*, useful for effective production of protease-sensitive heterologous proteins. *Yeast* **23**, 83-99 (2006).
124. Hiramatsu, R., Horninouchi, S., Uchida, E., Hayakawa, T. & Beppu, T. The Secretion Leader of *Mucor pusillus* Rennin Which Possesses an Artificial Lys-Arg Sequence Directs the Secretion of Mature Human Growth Hormone by *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **57**, 2052-2056 (1991).
125. G.L. Gray & Heyneker, H.L. (ed. U.S. Patent)US; 1988).
126. Ecamilla-Treviño, L.L., Viader-Salvado, J.M., Barrera-Saldaña, H.A. & Guerrero-Olazarán, M. Biosynthesis and secretion of recombinant human growth hormone in *Pichia pastoris*. *Biotechnology letters* **22**, 109-114 (2000).

127. M. Sautour, P.M., N.E. Chihib, J.P. Hornez The effects of temperature, water activity and pH on the growth of *Aeromonas hydrophila* and on its subsequent survival in microcosm water. *Journal of Applied Microbiology* **95**, 807-813 (2003).
128. Meltem Demirtas, Kolhatkar, A. & Kilbane, J. Effect of aeration and agitation on growth rate of *Thermus thermophilus* in batch mode. *J. Biosci. Bioeng.* **95**, 113-117 (2003).
129. J. Hallborn et al. The influence of cosubstrate and aeration on xylitol formation by recombinant *Saccharomyces cerevisiae* expressing the XYLI gene. *Appl. Microbiol. Biotechnol* **42**, 326-333 (1994).
130. Sampermans, S., Mortier, J. & Soares, E.V. Flocculation onset in *Saccharomyces cerevisiae*: the role of nutrients. *J. App. Microbiol.* **98**, 525-531 (2005).
131. Frans M. Klis, Boorsma, A. & De-Groot, P.W.J. Cell wall construction in *Saccharomyces cerevisiae*. *Yeast* **23**, 185-202 (2006).
132. Smit, G., Straver, M.H., Lugtenberg, B.J.J. & Kijne, J.W. Flocculence of *Saccharomyces cerevisiae* Cells Is Induced by Nutrient Limitation, with Cell Surface Hydrophobicity as a Major Determinant. *Appl Environ Microbiol.* **58**, 3709-3714 (1992).
133. Bruce Alberts et al. *Molecular biology of the cell*, Edn. Third. (Garland publishing, New York; 1994).
134. Colling, L., Carter, R.N., Essmenn, M. & Larsen, B. Evaluation of relative yeast cell surface hydrophobicity measured by flow cytometry. *Infectious Diseases in Obstetrics and Gynecology* **13**, 43-48 (2005).
135. Sihua Qi et al. The Effects of Thiopental and Propofol on Cell Swelling Induced by Oxygen/Glucose Deprivation in the CA1 Pyramidal Cell Layer of Rat Hippocampal Slices *J. Neurosurg Anesthesiol* **94**, 655-660 (2002).
136. Glee, P.M., Sundstrom, P. & Hazen, K.C. Expression of Surface Hydrophobic Proteins by *Candida albicans* *In Vivo*. *Infection and Immunity* **63**, 1373-1379 (1995).
137. Hazen, B.W. & Hazen, K.C. Dynamic Expression of Cell Surface Hydrophobicity during Initial Yeast Cell Growth and before Germ Tube Formation of *Candida albicans*. *Infection and immunity* **56**, 2521-2525 (1988).
138. Lars Hamberg, Pernilla Walkenstroëm, Mats Stading & Hermansson, A.-M. Aggregation, viscosity measurements and direct observation of protein coated latex particles under shear. *Food Hydrocolloids* **15**, 139-151 (2001).
139. Neway, J.O. *Fermentation Process Development of Industrial Organisms*. (Marcel Dekker, 1989).
140. Aldo Roda, Patrizia Pasini, Mara Mirasoli, Michelini, E. & Guardigli, M. Biotechnological applications of bioluminescence and chemiluminescence. *Trends Biotechnol.* **22**, 295-303 (2004).
141. D. Riesenbergs & Guthke, R. High-cell-density cultivation of microorganisms. *Appl. Microbiol. Biotechnol.* **51**, 422-430 (1999).
142. Stryer, L. *Biochemistry*, Edn. 4th. (W. H. Freeman and company, New york; 1995).
143. Eggink, G. 185 (NRC Research Press, 1996).
144. Linda M. Harvey, B.M. in *Practical Fermentation Technology*. (ed. L.M.H. Brian McNeil) 97-1232008).
145. Demain, A.L. Regulation of secondary metabolism in fungi. *Pure & Appl. Chem.* **58**, 219-226 (1986).
146. Albers, E., Larsson, C., Liden, G., Niklasson, C. & Gustafsson, L. Influence of the Nitrogen Source on *Saccharomyces cerevisiae* Anaerobic Growth and Product Formation. *Appl. Environ. Microbiol.* **62**, 3187-3195 (1996).
147. Teun van de Laar et al. Increased Heterologous Protein Production by *Saccharomyces cerevisiae* Growing on Ethanol as Sole Carbon Source. *Biotechnol. Bioeng.* **96**, 483-494 (2007).

9 APPENDICES

9.1 Expression vector construction strategy

9.1.1 MPI

The full length human insulin cDNA contains sequences either side of the B-chain C-peptide A-chain construct. These were removed.

NNN Pre/Post coding sequence **NNN** B chain **NNN** C peptide **NNN** A chain

Full length human insulin cDNA sequence:

```
1  agccctccag gacaggctgc atcagaagag gccatcaagc agatcaactgt ccttctgcca 61
tgccctgtg gatgcgctc ctgccctgc tggcgctgct ggccctctgg ggacctgacc
121 cagccgcagc ctttgtgaac caacaectgt ggggctcaca cctgggtggaa gctctctacc
181 tagtgtgagg ggaacgaggc ttcttttaca cacccaagac ccgcccgggag gcagaggacc
241 tgcaggtggg gcaggtggag ctgggggggg gccctggtgc aggcagcctg cagcccttgg
301 ccttgagggg gtccctgcag aagcgtggca ttgtggaaca atgctgtacc agcatctgct
361 cctctacca gctgggaaac tactgcaadt agacgcagcc cgcaggcagc cccccaccg
421 cgcctcctg caccgagaga gatggaataa agcccttgaa ccaacaaaaa aaaaaaaaaa
481 aaaaaaaaaa aaaaa
```

Insulin(only) coding sequence:

```
1  ttttgaacc aacaectgtg cggetcacaac ctgggtggaag ctctctactt agtgtgggg 61
gaaagaggct tctttctaac acccaagacc cgcccgggagg cagaggacct gcaggtgggg 121
caggtggagc tggggggggg ccttgggtgca ggcagcctgc agcccttggc cctggagggg 181
tccttcgaga agcgtggcat tgtggaacaa tgctgtacca gcatctgctc cctctaccag 241
ctggagaact actgcaad
```

Translation of insulin(only) sequence:

```
FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSL
QKRGIVEQCCTSICSLYQLENYCN
```

To carry out this modification, PCR primers were designed. These bind to either end of the B-chain C-peptide A-chain coding sequence in the holding vector (pCMV6-XL5) and at the same time introduce *EcoRI* (5') and *XbaI* (3') sites.

Forward primer Round 1 *EcoRI*

```
5' ACGT GAA TTC TTT GTG AAC CAA CAC CTG
          EcoRI  F  V  N  Q  H  L
```

Reverse primer Round 1 *XbaI*

```
5' TGCA TCT AGA GTT GCA GTA GTT CTC CAG
          XbaI  Nr Cr  Yr Nr  Er Lr
```

The PCR product was inserted into cloning vector pUC19 (2686bp; Fermentas) using the newly introduced *EcoRI* and *XbaI* sites

Insulin (only) coding sequence with introduced *EcoRI* and *XbaI* sites:

```
1  ACGTGAATT C ttttgaacc caacaectgtc ggggctcaca cctgggtggaa gctctctacc 61
tagtgtgagg ggaacgaggc ttcttttaca cacccaagac ccgcccgggag gcagaggacc 121
tgcaggtggg gcaggtggag ctgggggggg gccctggtgc aggcagcctg cagcccttgg 181
```


ccctggaggg gtcacctgcag aagcgtggca ttgtggaaca atgctgtacc agcatctgct 241
 ccctctacca gctggagaac tactgcaacT CTAGATGCA

The C-peptide prevents efficient secretion of pro-insulin.
 The full length C-peptide (TRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKR) was replaced with a mini-C-peptide (AAK) referred to as C*.

NNN B chain NNN C peptide NNN A chain

Insulin(only) coding sequence:

1 tttgtgaacc aacacctgtg eggetcacaac ctggtggaag ctctctacct agtgtgaggg
 61 gaacgagget tctctacac acccaagacc cgccgggagg cagaggacct gcaggtgggg 121
 caggtggagc tgggcggggg ccctggtgca ggcagcctgc agcccttggc cctggagggg 181
 tcctgcaga agcgtggcat tgtggaacaa tgctgtacca gcactctgct cctctaccag 241
 ctggagaact actgcaac

Insulin (C*) coding sequence:

1 tttgtgaacc aacacctgtg eggetcacaac ctggtggaag ctctctacct agtgtgaggg
 61 gaacgagget tctctacac acccaaggct gctaaaagca ttgtggaaca atgctgtacc 121
 agcatctgct ccctctacca gctggagaac tactgcaac

1 GAATTCtttg tgaaccaaca cctgtgagggc tcacacctgg tggaaagctct ctacctagt
 61 tgcgggggaac gaggettctt ctacacacc aagctgcta aaggcattgt ggaacaatgc
 121 tgtaccagca ttgtctctct ctaccagctg gagaactact gcaacTCTAGA

Translation full length insulin:

FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSL
 QKRGIVEQCCTSICSLYQLENYCN

Translation insulin (C*)

FVNQHLCGSHLVEALYLVCGERGFFYTPKAAKGIVEQCCTSICSLYQLENYCN

This was achieved by PCR, adding the C* sequence (gctgctaaa: AAK) to the forward primer. The forward primer binds to the start of the A-chain (GIVEQ) and the reverse primer binds the end of the B-chain (K_r P_r T_r Y_r F_r) so resulting in exclusion of the C-peptide coding region and insertion of the C* sequence.

Forward primer Round 2 (C*-peptide)

5' GCTGCTAAAGGCATTGTGGAACAA
 A A K G I V E Q

Reverse primer Round 2

5' GAA GTA TGT GGG CTT
 F_r Y_r T_r P_r K_r

Attachment of the α -secretion factor, c-myc epitope, His₆ tag and stop codon to C* insulin:
 The α -mating factor, a peptide hormone which is the product of the *S. cerevisiae MFa1* gene, can be used to target recombinant proteins for secretion. The pPICZ α C plasmid contains a homolog of the α -factor (sequence comparison shown below). The pPICZ α C plasmid is available and has a high sequence similarity to the wild-type factor. There are other features of the pPICZ α C vector which are desirable, such as two affinity tags (c-myc and a poly His tag) and an in frame stop codon.

S. cerevisiae alpha factor:

AVNNTTTEDET AQIPAVIG YSDLEGDFDV AVLPPFSNSTN NGLLFINTTI ASIAAKEEGV SLKR

pPICZ α A alpha factor:

MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPA VIG YSDLEGDFDV AVL PFSNSTN NGLLFINTTI ASIAAKEEGV SLEKR*EAEA

* Kex2 signal cleavage site, to remove alpha factor from insulin peptide

The MPI sequence was introduced into the multi-cloning site of pPICZαC using the EcoRI and Xba1 sites (underlined)

XXX: α-factor XXX: Bchain XXX: C* peptide XXX: Achain XXX: C-myc XXX: His₆ TGA: Stop

ATGAGATTTCCTTCAATTTTACTGCTGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACA
ACAGAAGATGAAACGGCACAAATTCGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTCGATG
TTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAACTACTATTGCCAGCATTGCTG
CTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGCTGAAGCTGAATTCtttgtgaaccaaacctgtg
gtgctcacacctggtggaagctctctacctagtgtgctgggggaacgaggtctctctctacacaccccaaggtgctg
taaaagcattgtggaacaatgctgtaccagcatctgctccctctaccagctggagaactactgcaacTCTAGAAC
AAAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACATCATCATCATCATCATGA

PROBLEM: The sequence 3' of the A-chain is out of frame therefore both tags and the stop codon will not be expressed in the required manner.

SOLUTION: Another reverse PCR step to introduce two extra bases; GG immediately after the A-chain sequence. This generates an additional glycine residue. Also an alternative frame version of the pPICZα vector needs to be used pPICZαA

Insulin Fix PCR

Forward primer:

5' TAC TGC AAC GG TCT AGA GTC

Reverse primer:

5' GTT CTC CAG CTG GTA

Using pUC19-MPI as template DNA to introduce the additional bases, this "fixed" construct was denoted Mini-Pro-Insulin(Fixed) (MPI(F))

Round 3 PCR addition of *Hind*III and *Sac*I restriction sites to 5' and 3' ends of the MPI(F) expression cassette:

Insulin Round 3 PCR Forward primer

ACG TAA GCT TAT GAG ATT TCC TTC AAT TTT T
*Hind*III

Insulin Round 3 PCR Reverse primer

TGC AGA GCT CTC AAT GAT GAT GAT GAT G
*Sac*I

Template DNA pPICZαA-MPI (α-factor, MCS (including MPI),c-myc, His₆, STOP)

XXX: α-factor XXX: Bchain XXX: C* peptide XXX: Achain XXX: c-myc XXX: His₆ TGA: Stop

ATGAGATTTCCTTCAATTTTACTGCTGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACA
ACAGAAGATGAAACGGCACAAATTCGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTCGATG
TTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAACTACTATTGCCAGCATTGCTG
CTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGCTGAAGCTGAATTCtttgtgaaccaaacctgtgctg
gctcacacctggtggaagctctctacctagtgtgctgggggaacgaggtctctctctacacaccccaaggtgctgctaaa
gcattgtggaacaatgctgtaccagcatctgctccctctaccagctggagaactactgcaacGGTCTAGAACAAA
AACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACATCATCATCATCATCATGA

PCR Round 3 product (α -MPI(F)):

```

ACGTAAGCTATGAGATTTCCTCAATTTTACTGCTGTTTATTTCGCAGCATCCTCCGCATTAGCTGCTCCAGTC
AACACTACAACAGAAGATGAAACGGCACAAATTCGGCTGAAGCTGTCAICGGTACTCAGATTTAGAAGGGG
ATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAAATACTACTATTGCCA
GCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGGCTGAAGCTGAATTCtttTgtgaaccctaaag
acctgtgcccgtcacacctggttggaaagetotetacctagtgtgccccgggaaccgagccttcttctacacaccccaagg
ctgctaaaaggcaattgtggaacaatgctgtaccagcatctgctccctctaccagctggagaactactgcaacGGTC
TAGAACAAAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACATATCATATCATCATGAGAT
CTGCA
  
```

The α -MPI(F) construct can then be cloned into both *S. cerevisiae* expression vectors (pYX212 and pYES2) for transformation of competent cells and expression of a secreted and tagged mini-proinsulin. Both these plasmids contain the *Hind*III and *Sac*I restriction sites in the correct order, therefore the α -MPI(F) construct is in the correct orientation.

9.1.2 hPGH

The cDNA sequence for hPGH

```

1  aggatcccaa ggcccaactc cccgaaccac tcagggctct gtggacagct cacctagcgg
61  caatggctgc aggctcccgg acgtccctgc tectggcttt tggcctgctc tgccctgtct
121 ggcttcaaga gggcagtgcc ttcccaacca ttcccttacc caggcttttt gacaacgcta
181 tgctccgcgc ccgtcgctg taccagctgg catatgacac ctatcaggag tttgaagaag
241 cctatatcct gaaggagcag aagtattcat tcctgcagaa ccccagacc tccctctgct
301 tctcagagtc tattccaaca cctccaaca gggtgaaaac gcagcagaaa tctaacctag
361 agctgctccg catctccctg ctgctcatcc agtcatggct ggagcccgtg cagctcctca
421 ggagcgtctt cgccaacagc ctggtgtatg gcgctcggga cagcaacgct tatcgccacc
481 tgaaggacct agaggaaggc atccaaacgc tgatgtggag gctggaagat ggcagcccc
541 ggactgggca gatcttcaat cagtctaca gcaagtttga cacaaaatcg cacaaacgatg
601 acgcaactgct caagaactac gggctgctct actgcttcag gaaggacatg gacaaggtcg
661 agacattcct gcgcatcgtg cagtgccgct ctgtggaggg cagctgtggc ttctagctgc
721 ccgggtggca tcctctgtgac cctccccag tgctctctct ggtcgtggaa ggtgctactc
781 cagtgccccac cagccttgtc ctaataaaat taagttgcat c
  
```

Here we see an annotated version of hPGH cDNA

Key:

xxx: pre sequence (1-62) xxx: signal sequence(63-131) xxx: hGH2 coding sequence (141-173) tag: stop codon (174-176) xxx: post sequence (177-821)

```

AggatcccaaggcccaactccccgaaccactcagggctcctgtggacagctcacctagcggcaatggctgcaGgct
cccggacgtccctgctcctggcttttggcctgctctgctgtctctggcttcaagaaggcagtgcccttcccaacca
ttcccttaccaggctttttgacaacgctatgctccgcgcccgtcgctgtaccagctggcatatgacacctatc
aggagtttgaagaagcctatatcctgaaggagcagaagtattcattcctgcagaacccccagacctcctctgct
tctcagagtcctattccaacaccttccaacagggtgaaaacgcagcagaaatctaacctagagctgctccgcatct
cctgctgctcatccagtcctggctggagcccgtgcagctcctcaggagcgtcttcgccaacagcctgggtgatg
gcgctcggacagcaacgctctatcgccacctgaaggacctagaggaaggcatccaaacgctgatgtggagggctgg
aagatggcagccccggactgggcagatcttcaatcagtcctacagcaagtttgacacaaaatcgcacaaacgatg
acgcaactgctcaagaactacgggctgctctactgcttcaggaaggacatggacaaggtcgagacattcctggcga
tcgtgcagtgccgctctgtggagggcagctgtggctcttagctgccccgggtggcatcctctgacccctccccag
tgctctcctggctgtggaaggtgctactccagtgcccaccagccttgctcctataaaaat taagttgcatc
  
```

Sequence with amino acid translation

```

aggatcccaaggcccaactccccgaaccactcagggctcctgtggacagctcacctagcggcaatggctgca
MetAlaAla
ggctcgggacgtccctgctcctggcttttggcctgctctgctgtctctggcttcaagaaggcagtgccctt
GlySerArgThrSerLeuLeuLeuAlaPheGlyLeuLeuCysLeuSerTrpLeuGlnGluGlySerAlaPh
ccccaccattccttaccaggctttttgacaacgctatgctccgcgcccgtcgctgtaccagctggcat
  
```


eProThrIleProLeuSerArgLeuPheAspAsnAlaMetLeuArgAlaArgArgLeuTyrGlnLeuAlaT
 Atgacacctatcaggagtttgaagaagcctatatcctgaaggagcagaagtattcattcctgcagaacccc
 yrAspThrTyrGlnGluPheGluGluAlaTyrIleLeuLysGluGlnLysTyrSerPheLeuGlnAsnPro
 cagacctccctctgcttctcagagctctattccaacacctccaacagggtgaaaacgcagcagaaatctaa
 GlnThrSerLeuCysPheSerGluSerIleProThrProSerAsnArgValLysThrGlnGlnLysSerAs
 cctagagctgctccgcatctccctgctgctcatccagctcatggctggagcccgctgcagctcctcaggagcg
 nLeuGluLeuLeuArgIleSerLeuLeuLeuIleGlnSerTrpLeuGluProValGlnLeuLeuArgSerV
 tcttcgccaacagcctggtgtatggcgccctcggacagcaacgtctatcgccacctgaaggacctagagga
 alPheAlaAsnSerLeuValTyrGlyAlaSerAspSerAsnValTyrArgHisLeuLysAspLeuGluGlu
 ggcattccaaaacgctgagtgaggctggaagatggcagcccccggactgggcagatcttcaatcagtccta
 GlyIleGlnThrLeuMetTrpArgLeuGluAspGlySerProArgThrGlyGlnIlePheAsnGlnSerTy
 cagcaagtttgacacaaaatcgacacacgatgacgcactgctcaagaactacgggctgctctactgcttca
 rSerLysPheAspThrLysSerHisAsnAspAlaLeuLeuLysAsnTyrGlyLeuLeuTyrCysPheA
 ggaaggacatggacaaggtcgagacattcctgcgcatcgtgcagtgccgctctgtggagggcagctgtggc
 rgLysAspMetAspMetAspMetAspLysValGluThrPheLeuArgIleValGlnCysArgSerValGlu
 ttcctagctgcccgggtggcatccctgtgacctcccagctgctctcctggtcgtggaaggtgctactcc
 agtggccaccagccttgtcctaataaaaattaagttgcate

The HGH2 cDNA sequence does not contain the following restriction recognition sites: *EcoRI*, *Xba1*, *HinIII* or *Sac1* so these can be used for the manipulations required in constructing the expression vectors.

Removal of pre and post HGH2 coding sequences and addition of a 3' Glycine residue (218) and restriction sites 5' and 3' of HGH2 sequence

This was achieved by Polymerase Chain Reaction (PCR). Primers were designed to amplify the HGH2 coding sequence and simultaneously introduce *EcoRI* (5') and *Xba1* (3') restriction sites. The introduction of GG 5' of the *Xba1* site brings the tag sequences in pPICZalphaA into frame with the HGH2 sequence.

The restriction sites allow the HGH2 sequence to be ligated directly into pPICZalphaA expression vector.

Round 1 PCR: Addition of *EcoRI* and *Xba1* sites 5' and 3' of HGH2 coding sequence

Template DNA: HGH2 sequence

ttccaaccattcccttatccaggcttttgacaacgctatgctccgcgcccgtcgccctgtaccagctggcatat
 gacacctatcaggagtttgaagaagcctatatcctgaaggagcagaagtattcattcctgcagaacccccagacc
 tccctctgcttctcagagctctattccaacacctccaacagggtgaaaacgcagcagaaatctaacctagagctg
 ctccgcatctccctgctgctcatccagctcatggctggagcccgctgcagctcctcaggagcgtcttcgccaacagc
 ctggtgtatggcgccctcggacagcaacgtctatcgccacctgaaggacctagaggaaggcatccaaacgctgatg
 tggaggctggaagatggcagcccccggactgggcagatcttcaatcagtcctacagcaagtttgacacaaaatcg
 cacaaacgatgacgcactgctcaagaactacgggctgctctactgcttcaggaaggacatggacaaggtcgagaca
 ttctgcgcatcgtgcagtgccgctctgtggagggcagctgtggcttc

Forward primer Round 1 PCR:

5' ACGT GAA TTC TTC CCA ACC ATT CCC TTA ATC
 EcoRI F P T I P L S

Reverse Primer Round 1 PCR:

5' TGCA TCT AGA **GG** GAA GCC ACA GCT GCC CTC CAC
 Xba1

Round 1 PCR Product:

5' ACGTGAATTC ttccaaccattcccttatccaggcttttgacaacgctatgctccgcgcccgtcgccctgtac
cagctggcatatgacacctatcaggagtttgaagaagcctatatcctgaaggagcagaagtattcattcctgcag
aacccccagacctccctctgcttctcagagctctattccaacacctccaacagggtgaaaacgcagcagaaatct
aacctagagctgctccgcatctccctgctgctcatccagctcatggctggagcccgctgcagctcctcaggagcgtc
ttcgccaacagcctggtgtatggcgccctcggacagcaacgtctatcgccacctgaaggacctagaggaaggcatc
caaacgctgatgtggaggctggaagatggcagcccccggactgggcagatcttcaatcagtcctacagcaagttt
gacacaaaatcgacacacgatgacgcactgctcaagaactacgggctgctctactgcttcaggaaggacatggac
aaggtcgagacattcctgcgcatcgtgcagtgccgctctgtggagggcagctgtggcttc **GGTCTAGAT** TGCA

Round 1 PCR product and pPICZ α A digested with EcoR1 and Xba1 (sequential digests) can be ligated to for the complete expression vector for *Pichia pastoris*: pPICZ α A-HGH2

The pPICZ α vector contains several useful features: α -mating factor (secretion factor), C-myc epitope, His₆ tag and stop codon.

The pPICZ α A-HGH2 expression cassette is composed of:
 α -factor, HGH2, C-myc, His₆, STOP

Construction of HGH2 tagged secretion cassette

Using the pPICZ α A-HGH2 construct as template DNA a second round of PCR can be performed using primers designed to introduce: HindIII restriction site 5' of the α -factor and a SacI restriction site 3' of the Stop codon. (these primers are already available from the insulin work).

Template DNA pPICZ α A-HGH2 (α -factor, MCS (including HGH2), C-myc, His₆, STOP)

XXX: α -factor XXX: corrected HGH2 XXX: C-myc XXX: His₆ TGA: Stop

```
ATGAGATTTCCTTCAATTTTTACTGCTGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCTCCAGTCAACACTACA
ACAGAAGATGAAACGGCACAATAATCCGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGGGATTTTCGATG
TTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTG
CTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGCTGAAGCTGAATTCttcccaaccattcccttatcca
ggotttttgacaacgctatgctccgcgcccgtcgctgtaccagctggcatatgacacctatcaggagtttgaag
aagcctatatcctgaaggagcagaagtattcattcctgcagaacccccagacctccctctgcttctcagagtcta
ttccaacaccttccaacagggtgaaaacgcagcagaaatctaacctagagctgctccgcattctccctgctgctca
tccagctcatggctggagcccgtgcagctcctcaggagcgtcttcgccaacagcctgggtgatggcgcccteggaca
gcaacgctctatcgccacctgaaggacctagaggaaggcatccaaacgctgatgtggaggctggaagatggcagcc
cccggactgggcagatcttcaatcagtcctacagcaagtttgacacaaaatcgcaaacgatgacgcactgctca
agaactacgggctgctctactgcttcaggaaggacatggacaaggctcgagacattcctgcgcactcgtgcagtgcc
gctctgtggaggggcagctgtggtctcGGTCTAGAACAATACTCATCTCAGAAGAGGATCTGAATAGCGCCGTC
GACATCATCATCATCATCATTTGA
```

HGH2 (same as Insulin) Round 2 PCR Forward primer

ACG TAA GCT TAT GAG ATT TCC TTC AAT TTT T
HindIII

HGH2 (same as Insulin) Round 2 PCR Reverse primer

TGC AGA GCT CTC AAT GAT GAT GAT GAT GAT G
SacI

Round 2 PCR Product:

```
ACGTAAGCTTATGAGATTTCCTTCAATTTTTACTGCTGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCTCCAGT
CAACACTACAACAGAAGATGAAACGGCACAATAATCCGGCTGAAGCTGTCATCGGTTACTCAGATTTAGAAGGG
GATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAATAACGGGTTATTGTTTATAAATACTACTATTGCC
AGCATTGCTGCTAAAGAAGAAGGGGTATCTCTCGAGAAAAGAGAGGCTGAAGCTGAATTCttcccaaccatt
cccttatccaggctttttgacaacgctatgctccgcgcccgtcgctgtaccagctggcatatgacacctatcag
gagtttgaagaagcctatcctgaaggagcagaagtattcattcctgcagaacccccagacctccctctgcttc
tcagagtctattccaacaccttccaacagggtgaaaacgcagcagaaatctaacctagagctgctccgcattctcc
ctgctgctcatccagctcatggctggagcccgtgcagctcctcaggagcgtcttcgccaacagcctgggtgatggc
gctcggacagcaacgctctatcgccacctgaaggacctagaggaaggcatccaaacgctgatgtggaggctggaa
gatggcagcccccggactgggcagatcttcaatcagtcctacagcaagtttgacacaaaatcgcaaacgatgac
gcactgctcaagaactacgggctgctctactgcttcaggaaggacatggacaaggctcgagacattcctgcgcactc
gtgcagtgccgctctgtggaggggcagctgtggtctcGGTCTAGAACAATACTCATCTCAGAAGAGGATCTGAA
TAGCGCCGTCGACATCATCATCATCATCATTTGAGATCTGCA
```

The round 2 PCR product was digested with *Hind*III and *Sac*I then ligated into pYES2 and pYX212 *S. cerevisiae* expression vectors

PROEOMICS

Yeast as a Production Host

Rationalising Eukaryotic Membrane Protein Production

Understanding membrane proteins



Membrane proteins are essential for the function of most cells and are the target of many drugs.



Astron University

Content has been removed for copyright reasons

what is happening in the host cell itself, and use that rationally to improve product yields. In a step towards a rational

The notable exception to this is the case of rhodopsin, which can be isolated from bovine retina in very large (>100 mg)



Aston University

Content has been removed for copyright reasons

obligatory first step in their exploitation.

www.aon.com

EC Directive 90/269 - 3

9.3 Poster presentations

9.3.1 Recombinant protein production (RPP) conference 2006, Barcelona, Spain.



Evaluation of antifoams in the expression of a recombinant Fc fusion protein in shake flask cultures of *Saccharomyces cerevisiae* & *Pichia pastoris*

William Holmes¹, Rodney Smith¹, Roslyn Bill²

¹Aston Academy of Life Sciences, Aston University, Aston Triangle, Birmingham, United Kingdom, B4 7ET.

²CTM Biotech Ltd, Babraham Research Campus, Babraham,

Cambridge, United Kingdom, CB2 4AT.



Aim

The aim of this study was to determine the effect on growth rate and protein production yield for four antifoam types at varying concentrations. This was achieved by monitoring the growth and target protein production in *S. cerevisiae* and *P. pastoris*.

The effectiveness of 4 antifoam agents was examined:

- Schill & Scheininger's Struktol SB2121 (Polyalkylenglycol)
- Schill & Scheininger's Struktol J673A (an alkoxylated fatty acid ester on a vegetable base)
- Sigma Antifoam C (Siloxane polymer)
- Fluka P2000 (Polypropylene glycol)

A secreted Fc fusion protein expressed by both *P. pastoris* and *S. cerevisiae* was used as the target protein. The protein itself has a molecular weight of approximately 48kDa.

Background

Optimisation of culture conditions for the expression and production of important therapeutic biologics is a key element in the rapid and cost effective manufacture of these important molecules [1]. The demand for oxygen by a microorganism can be met by aerating the culture medium, which is often done by sparging sterile air through the medium.

An unfortunate effect of both sparging through culture medium at high rates and intense agitation is the formation of foam. Foam formation reduces the efficiency of gas exchange at the surface of the culture, as a barrier is formed between the culture and the gases in the headspace of the vessel. This is a particular problem when surface active species such as proteins are present at high concentrations. Foaming can also be detrimental to the cells; when bubbles burst they exert shear forces, which can damage cells and/or any secreted proteins. Additionally cells and culture medium are lost to the foam phase which can lead to a decrease in process productivity. In extreme cases a 'foam out' situation can lead to loss of process sterility [2].

In order to minimise the deleterious effects of foaming, antifoam agents are used which prevent foam forming by reducing the surface tension of the culture [1]. There is a wide range of antifoam agents available from various suppliers.

Figure 1: Growth curve for *P. pastoris* in YPD medium with Struktol J673A antifoam at 30 °C with elliptical agitation at 200 rpm.

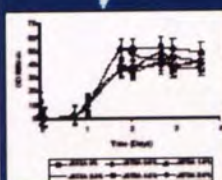


Figure 2: Growth curve for *S. cerevisiae* in SD-URA medium with Struktol SB2121 antifoam at 30 °C with elliptical agitation at 200 rpm.

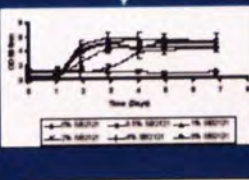
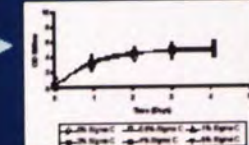


Figure 3: Growth curve for *S. cerevisiae* in SD-URA medium with Sigma antifoam C at 30 °C with elliptical agitation at 200 rpm.



Materials and Methods

Duplicate cultures of SD-URA (*S. cerevisiae*), YPD and BMMY (*P. pastoris*), were incubated at 30 °C with elliptical agitation at 200 rpm. 125 mL vented shake flasks were used with a working volume of 50 mL. Antifoam was added to the media prior to inoculation, at set concentrations 0%, 0.5%, 1%, 2%, 4% and 8% total culture volume. Cultures were analysed by optical density at 695nm (OD_{695nm}) for monitoring growth. Product in culture samples were precipitated using Ethanol/Acetone (2:1:1). The re-suspended precipitates were analysed by SDS-PAGE and visualised using Safestain (Invitrogen) and Silverstain (Invitrogen).

Results

The different types of antifoam affect *S. cerevisiae* and *P. pastoris* growth in different ways depending on the concentration and medium type being used. When Struktol J673A is used with YPD medium for *P. pastoris* growth, increasing antifoam concentration increases OD_{695nm} (see Figure 1). Conversely when Struktol SB2121 is used with SD-URA medium for *S. cerevisiae* (strain: ALCOFREE™ Yeast 01 [3]) protein production, increasing antifoam concentration reduces OD_{695nm} measurements of the cultures (see Figure 2). When Antifoam C is used with YPD medium, *S. cerevisiae* growth is not affected by concentrations up to 8% (see Figure 3). The effect on protein production is less variable, with the trend being that concentrations over 1% total volume decrease the yield of recombinant protein in the cultures (See Figure 4).

Figure 4: Silver stain image of protein samples taken 120 hours post-inoculation, from *P. pastoris* culture. Struktol J673A antifoam concentration increases from left to right (0, 0.25, 0.5, 1, 2, 4 and 8% total volume). Target protein is indicated by an arrow.



Conclusions

The data indicate that antifoam agents can be used effectively at concentrations up to 1% total volume. Higher concentrations can lead to a decrease in protein yield. Additionally some of the antifoam agents become difficult to work with at higher concentrations, producing precipitates which interfere with sampling and analysis. The table below highlights the main conclusions for each antifoam and application.

Antifoam Type	<i>S. cerevisiae</i> ALCOFREE™		<i>P. pastoris</i> GS115	
	Growth	Protein Production	Growth	Protein Production
Struktol SB2121	Increased by [SB2121] >0% <4%	Decreased by [SB2121] >1%	No effect [SB2121] <=8%	No effect [SB2121] <=8%
Struktol J673A	Increased by [J673A]	Decreased by [SB2121] >1%	Increased by [J673A]	Decreased by [J673A] >1%
Sigma Antifoam C	No effect [Antifoam C] <=8%	Decreased by [Antifoam C] >2%	No effect [Antifoam C] <=8%	No effect [Antifoam C] <=8%
Fluka P2000	Increased by [P2000] <4%	Decreased by [P2000] >1%	No effect [P2000] <=8%	Decreased by [P2000] >4%

Further work

The present evaluation highlights differences in organism performance and yield. However, the range of antifoam concentrations are not process concentrations. Further work will be done to determine the effect of lower antifoam concentrations (e.g. 0-1% in 0.2% increments) and use in high cell density fermentations.

Acknowledgements

The work is part of a PhD project by William Holmes, Aston University and in conjunction with CTM BioTech Ltd. We would like to thank the Engineering and Physical Sciences Research Council (EPSRC) for supporting this work.



References

- [1] Dow Corning. [http://20057408-PaperCard02.pdf] Dow Corning Antifoam product information, created 15/02/2005
- [2] Vetter, J., Brown, A. G., Sneyd, J. W. R., Good, P. W., Gossage, S. Dynamic multipoint measurement of foam behaviour for a continuous fermentor over a range of low process variables. *Biotechnol Bioengineering Journal* 2004, 20, 61-72
- [3] ALCOFREE™ Yeast 01 derived from the CTM PK yeast family. www.gottliebsteint.com/Gottlieb_Yeast_Solutions_AB_Tamagastan-7_411-33_Gothenburg_Sweden

9.3.2 Poster presentations at Aston University, School of life and health sciences post-graduate research days

9.3.2.1 2007



Improved production of industrially relevant recombinant proteins in yeast systems

William Holmes^{a*}, Roslyn Bill^a, Rodney Smith^a

^aAston Academy of Life Sciences, Aston University, Aston Triangle, Birmingham, United Kingdom, B4 7ET.

^{*}Alpha Biologics Ltd, Braham Research Campus, Braham, Cambridge, United Kingdom, CB2 4AT.



Purpose:

The production of recombinant human protein therapeutics is well suited to yeast systems, which provide the benefits of eukaryotic hosts but are as easy to handle as prokaryotes. Here we examine a unique respiratory *Saccharomyces cerevisiae* (ALCOFREE™; Otterstedt, 2004) and compare it against an industry standard, *Pichia pastoris*.

An evaluation of secreted protein yield is being carried out using important recombinant proteins; mini-pro-insulin (MPI) (Figure 1), human growth hormone – placental specific (hGH2) (Figure 2) and Green fluorescent protein - Cycle 3 mutant (GFPuv) (Figure 3). The optimisation of process conditions (temperature, pH and aeration) is a key element in the expedient, reproducible and cost-effective production of these important molecules. The aim is to identify critical parameters to aid optimal process design in a systematic manner.



Figure 1: Model structure of MPI

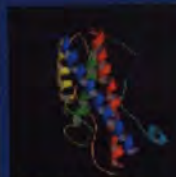


Figure 2: Model structure of hGH2



Figure 3: Model structure of GFP

Methods:

Initial experiments were carried out to examine the robustness of standard production conditions, and to learn more about the chosen test systems. These were carried out using shake flask cultures with 100 mL culture volume.

Optimisation studies on a small scale (~5 mL cultures) are planned using Box-Behnken experimental design (DoE) in an Applikon µ24 bioreactor. These will provide growth profile and protein yield data to construct models for each combination of system and target. These models can then be tested at both small scale (~5 mL) and larger scale (5 L) using a fermentor.

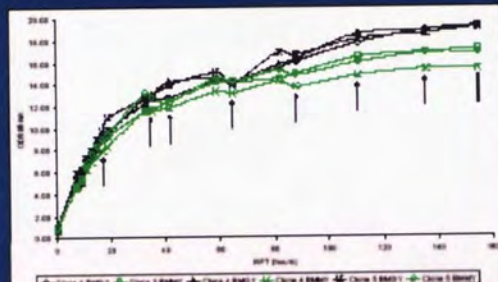


Figure 4: *P. pastoris* X33-GFPuv shake flask cultures. Inoculated at 30 °C with agitation at 250rpm.

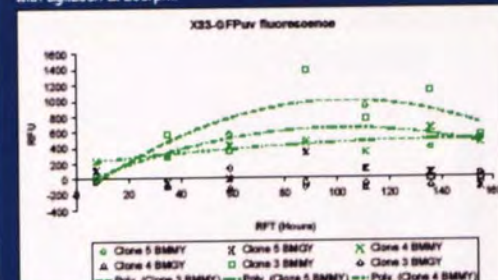


Figure 7: *P. pastoris* X33-GFPuv culture supernatant fluorescence measurement of samples taken during cultures shown in Figure 4

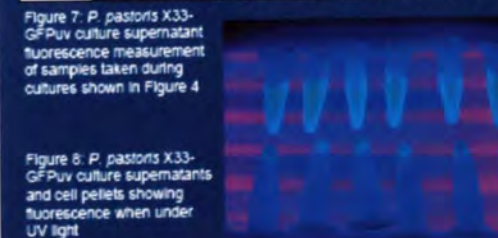


Figure 6: *P. pastoris* X33-GFPuv culture supernatants and cell pellets showing fluorescence when under UV light

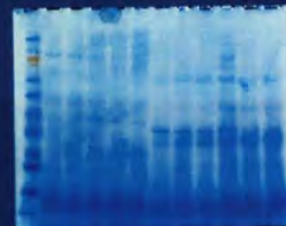


Figure 5: SDS-PAGE analysis of culture supernatant from non-induced and induced *P. pastoris* X33-GFPuv Clone 5 culture



Figure 6: Western blot analysis of culture supernatant from non-induced and induced *P. pastoris* X33-GFPuv Clone 5 culture

Results:

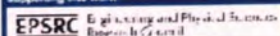
Expression vectors and strain selection has been carried out for GFPuv and hGH2 expressing ALCOFREE™ and *P. pastoris* X33 clones. Work on insulin vectors/strains is ongoing. We have analysed the growth and harvesting of our cells to ensure reproducibility under standard conditions. Detailed growth curves have been plotted. Figure 4 shows the growth curves for induced and non-induced *P. pastoris* X33-GFPuv shake flask cultures. Figures 5 and 6 show SDS-PAGE and Western Blot analysis of supernatant samples from *P. pastoris* X33-GFPuv Clone 5 taken during the growth curve in Figure 4. Figure 7 shows the fluorescence analysis of *P. pastoris* X33-GFPuv Clones supernatant samples taken during the growth cultures shown in Figure 4. Figure 8 shows *P. pastoris* X33-GFPuv culture supernatant samples exposed to UV light.

Conclusions:

DoE will greatly reduce experimental time compared to a full factorial method. We expect that conditions for optimal yield may not support optimal growth, as protein yield varies with time post inoculation/induction. If a "generic" model for optimising protein yield can be proven this would be invaluable for process development.

Acknowledgements

The work is part of a PhD project by William Holmes, Aston University and in conjunction with Alpha Biologics. We would like to thank the Engineering and Physical Sciences Research Council (EPSRC) for supporting this work.



References

- Bill R M. Yeast - a panacea for the structure-function analysis of membrane proteins? *Curr Genet* 40: 157-171 2001
- Boettner M. High-throughput screening for expression of heterologous proteins in the yeast *Pichia pastoris*. *Journal of Biotechnology* 99 51-62 2002
- Cerghino J L. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiology Reviews* 24 45-66 2000
- Otterstedt K et al. Switching mode of metabolism in the yeast *Saccharomyces cerevisiae* EMBO Rep May 2004 5: 532-537 2004
- Schmidt P R. Recombinant expression systems in the pharmaceutical industry (mini-review). *Appl Microbiol Biotechnol* 65 383-372 2004

9.3.2.2 2008 (Additional handout information for poster presentation)



A Predictive Method for the Rapid Optimisation and Scale-up to Fermentation for Protein Production in *Pichia pastoris*

William Holmes*, Rodney Smith*, Roslyn Bill*

*School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham, United Kingdom, B4 7ET.
*Alpha Biologics Ltd, Babraham Research Campus, Babraham, Cambridge, United Kingdom, CB2 4AT.



Here we demonstrate how the use of statistical design of experiments (DoE) and a parallel multi-well mini-bioreactor (Applikon micro24) permits the rapid establishment of high yield conditions that can be directly transferred into a fermentor.

We show that a predictive model can be constructed for secreted protein yield, in this case GFP expressed in *Pichia pastoris*. The model was derived from 13 sets of conditions in which 3 factors (temperature, pH and dissolved oxygen) were varied each at 3 levels. MiniTab statistical software was used for model building and refining processes. Fig 1 shows a graphical representation of the model, demonstrating how yield changes with process conditions.

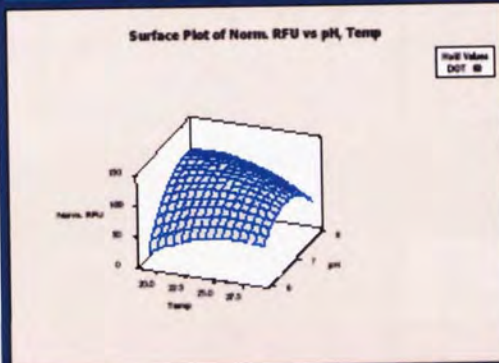


Figure 1. Surface plot of predictive model

In the micro24, the model predicted the experimental yield. Having established the robustness of the model in small volumes (5 mL), scalability was examined using fermentors with a 3 L working volume. Figure 2 shows the experimental yields from the micro24 and fermentors have a good fit to model predictions.

The methanol induction strategy used in the micro24 was not optimised, due to equipment limitations, however, it was predictive of yields in sub-optimal fermentations. Figure 3 shows how improving the fermentation induction strategy leads to a further 3 fold increase in yield.

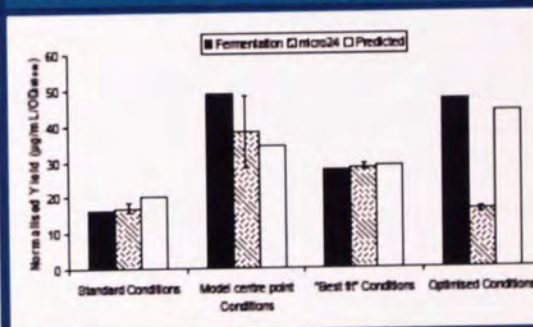


Figure 2. Comparison of model prediction to small (micro24) and large scale (Fermentation) experimental yield

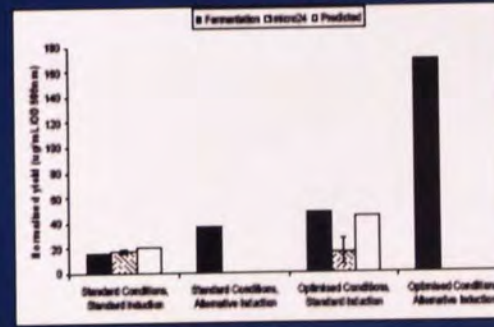


Figure 3. Alternative induction strategy (60 % sorbitol 40 % methanol) improved GFP yield over standard induction (100 % methanol)

This approach is particularly relevant for bio-processing applications in yeast, as high cell densities are a prerequisite for high productivity, which in turn depend on carefully-controlled growth regimes. In particular, we show that our approach separates conditions for growth and production, allowing maximal cell density and per cell yield to be reached whilst reducing overall process development time.

9.3.3 BioProcessUK meeting Cardiff, United Kingdom, 2007. (Travel bursary awarded)



Improved production of industrially-relevant recombinant proteins in yeast systems

William Holmes^{1*}, Rodney Smith², Roslyn Bill³

¹School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham, United Kingdom, B4 7ET.

²Alpha Biologics Ltd, Babraham Research Campus, Babraham, Cambridge, United Kingdom, CB2 4AT.

Purpose: The production of recombinant human protein therapeutics is well suited to yeast systems, which provide the benefits of eukaryotic hosts but are as easy to handle as prokaryotes. Here we examine a unique respiratory *Saccharomyces cerevisiae* (ALCOFREE™; Otterstedt, 2004) and compare it against an industry standard, *Pichia pastoris*.

The optimisation of process conditions (temperature, pH and aeration) is a key element in the expedient, reproducible and cost-effective production of these important molecules. The aim is to identify critical parameters and to build yield models to aid optimal process design in a systematic manner. An evaluation of secreted protein yield was therefore carried out on green fluorescent protein - Cycle 3 mutant (GFPuv) (Figure 1), mini-pro-insulin (MPI) and human growth hormone - placental specific (hGH2).



Figure 1. Model structure of GFP (<http://www.rcsb.org/pdb/explore.do?structureId=2G6E>)

Methods: Initial experiments were carried out to examine the robustness of standard production conditions and to learn more about the chosen test systems. These were carried out using shake flask cultures (100 mL culture volume). GFPuv expression was monitored using a plate reader (λ_{exc} 397 nm, λ_{em} 506 nm) from 100 μ L of culture supernatant. Optimisation studies were carried out on a small scale (~5 mL culture) using a Box-Behnken DoE matrix and an Applikon Micro 24 Bioreactor. These provided growth profiles and protein yield data to construct models for each combination of system and target. These models will be tested at both small- and larger-scale (5 L) using a fermenter.

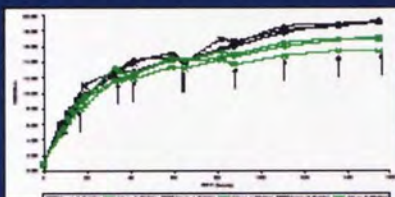


Figure 2. *P. pastoris* X33-GFPuv shake flask cultures. Incubated at 30°C with agitation at 250rpm

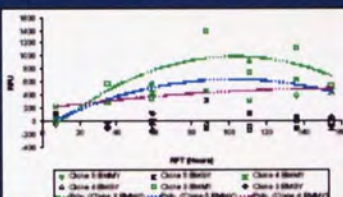


Figure 3. *P. pastoris* X33-GFPuv culture supernatant fluorescence measurement of samples taken during cultures shown in Figure 2



Figure 4. Applikon Micro 24 Bioreactor in use; online monitoring/control and data logging provided by laptop computer



Figure 5. Applikon Micro 24 Bioreactor in use; showing 24 well incubation cassette with *S. cerevisiae* cultures

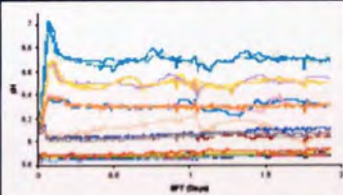


Figure 6. On-line pH traces for individual wells of a *P. pastoris* X33-GFPuv process optimisation run using the Micro 24 Bioreactor

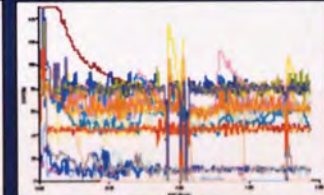


Figure 7. On-line DOT traces for individual wells of a *P. pastoris* X33-GFPuv process optimisation run using the Micro 24 Bioreactor

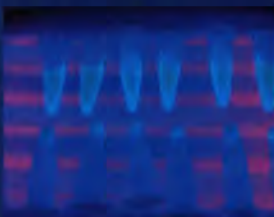


Figure 8. *P. pastoris* X33-GFPuv culture supernatants and cell pellets showing fluorescence when under UV light

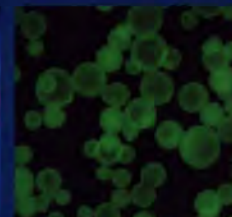


Figure 9. Z-stack confocal image of secreted GFPuv bound to Ni-Agarose chromatography resin

Results: Expression vector generation and strain selection have been carried out for GFPuv and hGH2 expressing ALCOFREE™ and *P. pastoris* X33 clones. Work on insulin vectors/strains is ongoing. We have analysed the growth and harvesting of our cells to ensure reproducibility under standard conditions. Detailed growth curves have been plotted. Figure 2 shows the growth curves for induced and non-induced *P. pastoris* X33-GFPuv shake flask cultures.

The Micro 24 Bioreactor (Figures 4 and 5) has been used for growing cultures at the various conditions required for the model. Analysis and model building is ongoing. Figures 6 and 7 show the online pH and DOT control over a 48 hour run with the Micro 24. Figure 3 shows the fluorescence analysis of *P. pastoris* X33-GFPuv supernatant samples taken during the growth cultures shown in Figure 2. Figure 8 shows *P. pastoris* X33-GFPuv culture supernatant samples exposed to UV light. Figure 9 shows secreted GFPuv bound to activated Ni-Agarose following incubation (2 hours) with X33-GFPuv culture supernatant as seen by confocal microscopy (Z-stack image).

Conclusions: DoE has greatly reduced experimental time compared to a full factorial method. It appears that conditions for optimal yield may not support optimal growth, as protein yield varies with time post inoculation/induction. If a "generic" model for optimising the yield of a given protein can be proven, this would be invaluable for process development.

Acknowledgements
The work is part of a PhD project by William Holmes, Aston University and in conjunction with Alpha Biologics. We would like to thank the Engineering and Physical Sciences Research Council (EPSRC) for supporting this work.

EPSRC Engineering and Physical Sciences Research Council

References
Bill R M. Yeast - A paradise for the structure-function analysis of membrane proteins? *Curr Opin* 40: 157-171 2001
Bostlier M. High-throughput screening for expression of heterologous proteins in the yeast *Pichia pastoris*. *Journal of Biotechnology* 99 51-62 2002
Cereghino J L. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiology Reviews* 24 45-66 2000
Otterstedt K et al. Switching mode of metabolism in the yeast *Saccharomyces cerevisiae* *EMBO Rep* May 2004 5: 532-537 2004
Schmidt F R. Recombinant expression systems in the pharmaceutical industry (mini-review). *Appl Microbiol Biotechnol* 65 363-372 2004

- 9.4 Applikon Biotechnology; Fermentation and cell culture symposium 2008 (Liverpool). Use of the Micro 24 system in predictive scale-down modelling of recombinant protein expression in *P. pastoris* and a hybridoma cell-line. Slides for presentation given as an invited speaker



Use of the Micro 24 system in predictive scale-down modelling of recombinant protein expression in *P. pastoris* and a hybridoma cell line

William Holmes

Areas covered

- Scale-down modelling for process optimisation
- Expression systems (*Pichia* and hybridoma)
- Target proteins (secreted GFP and antibody)
- Design of experiments (DoE)
- Small scale culture using a parallel multi-well micro bioreactor
- Initial method planning and experimental trial (X33GFP)
- Direct process scale-up
- Confirmation of the methodology (Hybridoma antibody)

Scale-down modeling for process optimisation

- Process optimisation
 - Culture conditions
 - Media
 - Induction regime
- Optimisation at production scale (10 -10000 L)
 - Time, Cost, Equipment availability
- Small scale equivalent
 - Key: Parallel cultures

Scale-down modelling for process optimisation

- Appropriate question
- Design of experiments (DoE)
- Data from small (6 mL) cultures
- Statistical modelling and analysis
- Experimental confirmation
- Transfer to larger scale (7L)
- Refine
- Transfer to larger scale (>100L)
- Manufacture

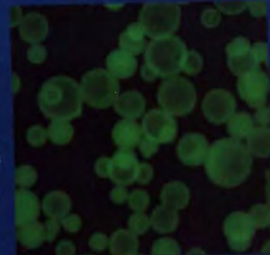


Expression systems

- *Pichia pastoris* (X33 strain)
 - Industry standard
 - Tightly regulated and highly inducible expression
 - High cell density fermentation
 - Low cost of goods
- Hybridoma cell line
 - Immortal fusion of B-cell and myeloma
 - Monoclonal antibody expression
 - Used as part of an Alpha Biologics project

Target proteins

- Green fluorescent protein (GFP)
 - Tool for method development
 - Simple and rapid assay techniques
- Antibody raised against a tumour inducing factor
 - Industrially relevant
 - Part of an ongoing project



GFP bound to Hi agarose
(Z stack image)

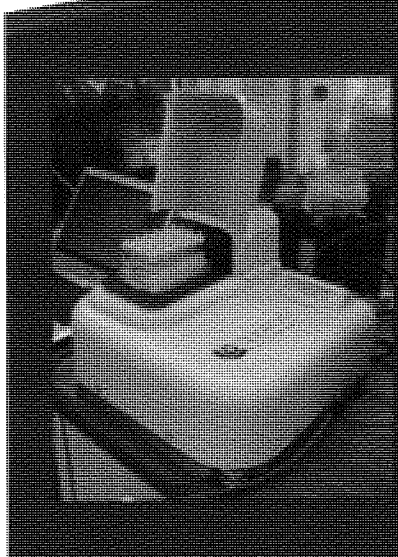
Statistical design of experiments (DoE)

- DoE can maximise "useful" information available from an experimental period
- "Black Box" approach



- Box-Behnken Design
- Response Surface Methodologies
 - Estimate factor interactions

Small scale culture using a parallel multi-well micro bioreactor



- Applikon Biotechnology Micro 24 Bioreactor
- Full experimental matrix and model confirmation completed in under one month
- Used with
 - *S. cerevisiae*, Media design
 - *P. pastoris*, process condition optimisation
 - Hybridoma cell line, process condition optimisation

Small scale culture using a parallel multi-well micro bioreactor

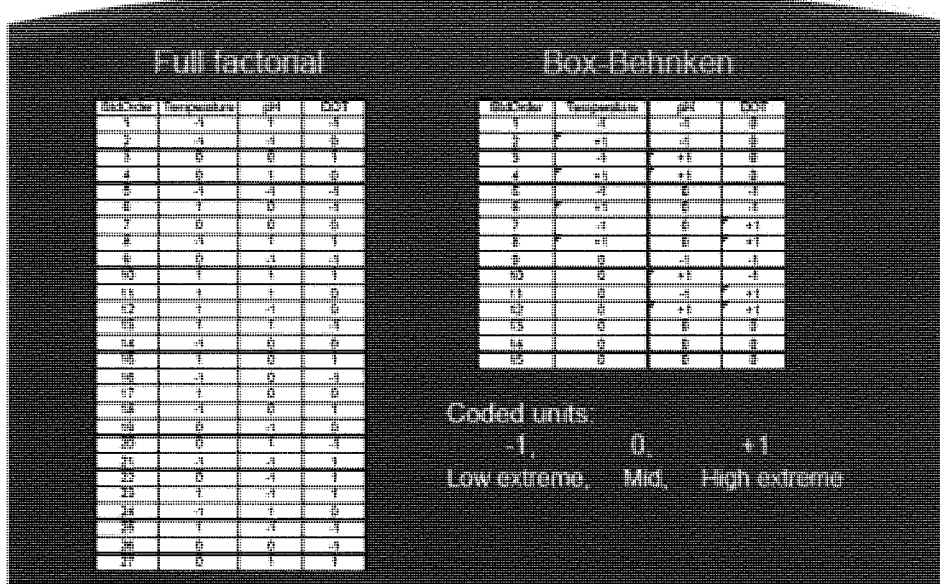


- Micro 24 allows control of culture conditions (Temperature, pH and DO)
- Not possible in shake flask cultures
- Higher through put than traditional stirred tank vessels
 - "Non-traditional" miniaturised stirred tank vessels
- Fed-batch cultures not currently possible

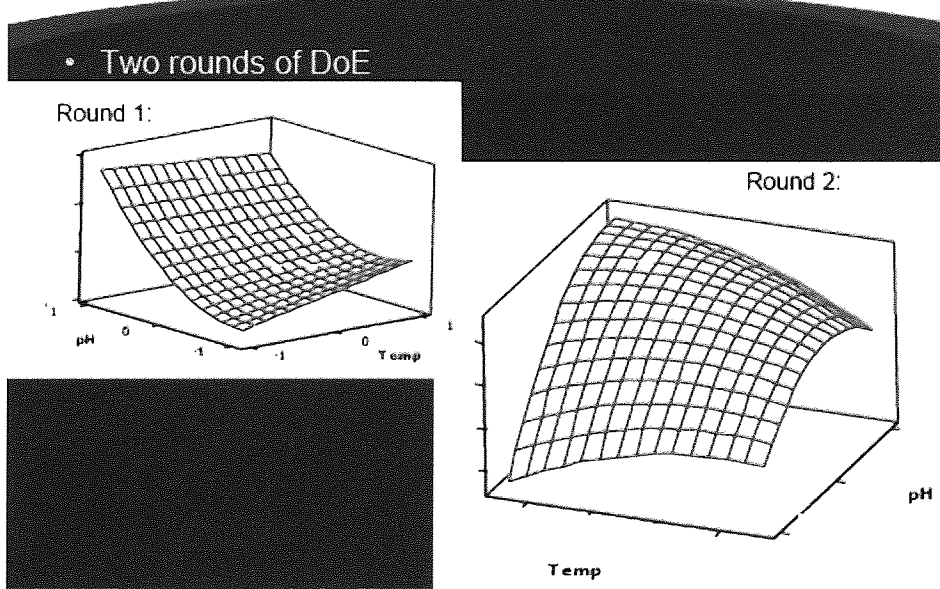
Initial method planning and experimental trial (X33GFP)

- Question: Can alternate process conditions improve yield?
- Optimising GFP expression in *P. pastoris* cultures
- Target protein: GFP (Cycle 3 mutant)
- Strain creation: X33GFP
- Factors: Temperature, pH and DO
- Small scale culture using Micro 24 bioreactor
- Induction by "slug" methanol addition

Initial method planning and experimental trial (X33GFP)

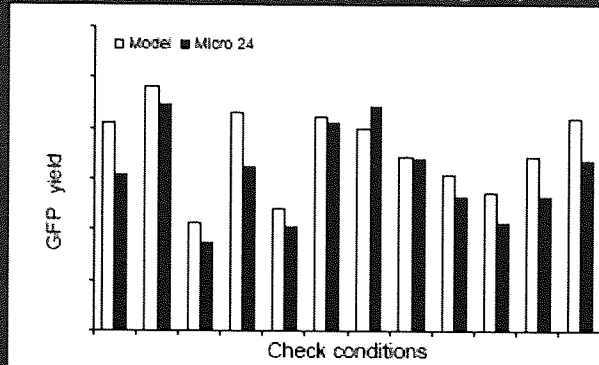


Initial method planning and experimental trial (X33GFP)



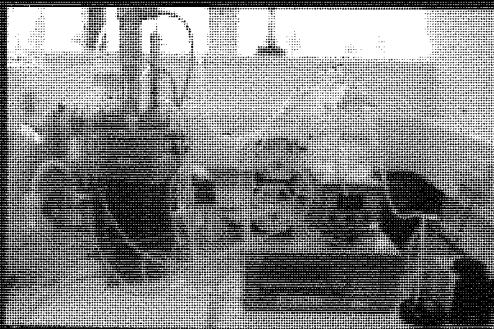
Initial method planning and experimental trial (X33GFP)

- Models assessed experimentally
 - Predicted optimal conditions confirmed experimentally
 - “Check point” conditions
 - Conditions out side of model design space



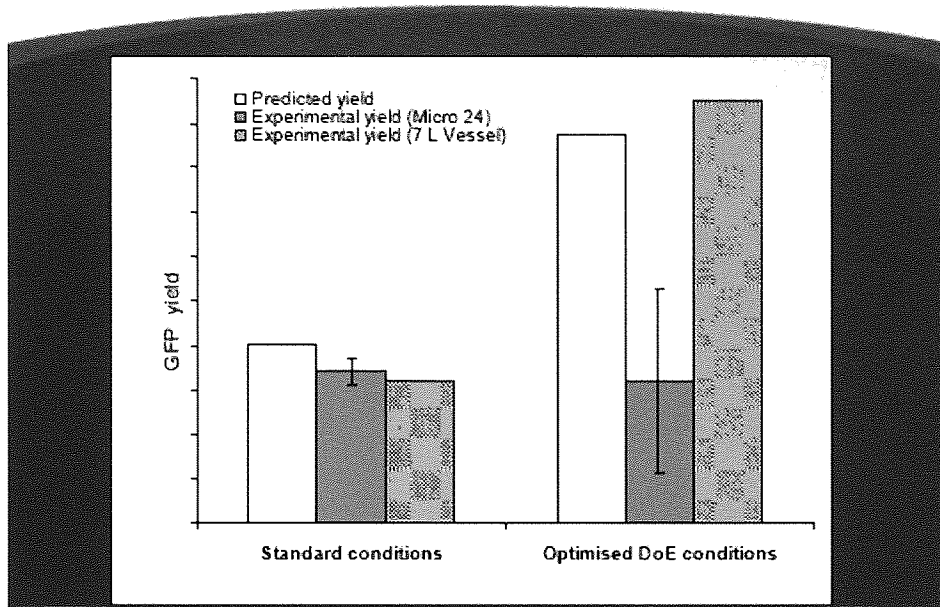
Direct process scale-up

- Process directly transferred to 7 L vessel



- “Standard” conditions
- DoE predicted optimal
- Best predicted to experimental fit
- Further optimisation possible
 - Increase pre-induction biomass (Total yield)
 - Alternative induction protocols (Methanol accumulation)

Direct process scale-up



Testing the methodology with an alternate expression system and target protein

- Mammalian cell line, Hybridoma
- Constitutive antibody expression
- Included medium composition
- 4 factors, 3 levels
- Yield determination via Protein A purification

Coded units:
 -1, 0, +1
 Low extreme, Mid, High extreme

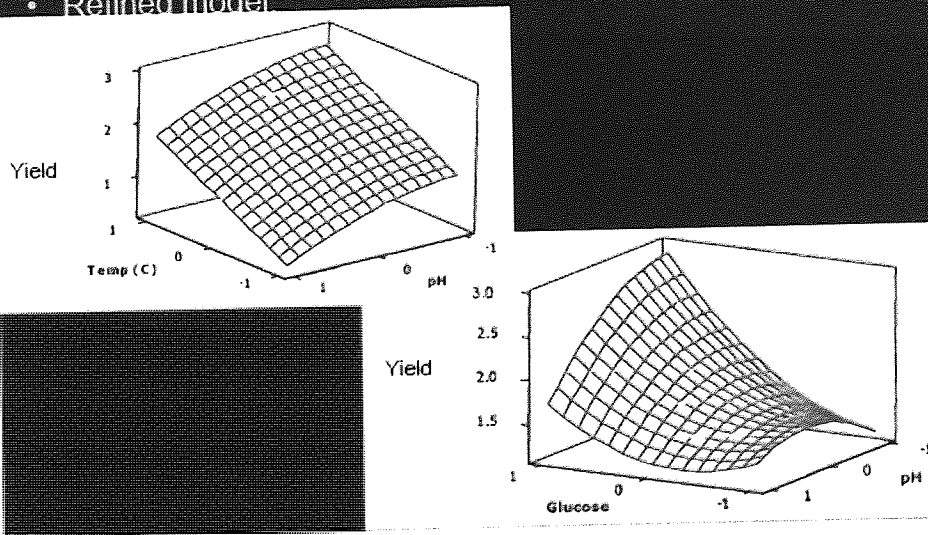
Run Order	Temperature (°C)	pH	DO	Glucose
1	-1	0	0	1
2	1	-1	0	0
3	1	-1	0	1
4	0	0	0	1
5	0	0	0	-1
6	0	0	-1	-1
7	0	0	0	0
8	0	0	0	0
9	0	-1	-1	0
10	0	0	0	0
11	0	1	0	1
12	-1	0	0	-1
13	-1	-1	-1	0
14	0	1	0	-1
15	1	1	0	-1
16	1	1	0	0
17	-1	-1	0	0
18	0	-1	0	1
19	-1	1	0	0
20	0	1	1	0
21	1	0	-1	0
22	1	0	1	0
23	1	0	0	-1
24	0	0	-1	1
25	0	1	-1	0
26	0	-1	1	0
27	0	-1	0	-1

Testing the methodology with an alternate expression system and target protein

- Predictive model generated and confirmed at small scale
- Monitored off-line: Cell number, Viability, Glucose conc.
- 2×10^5 seeding density
- 96 hour duration
- "Slug" glucose addition
- Complete data set per run
- 3 replicates: 3 "raw" models

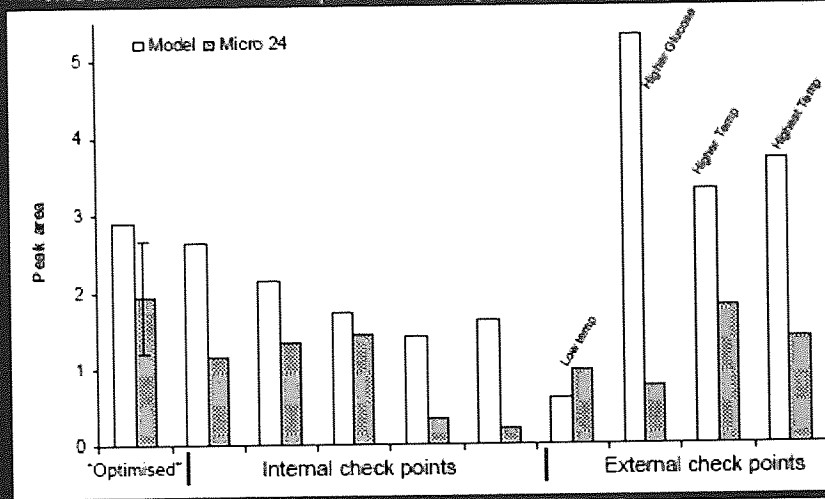
Testing the methodology with an alternate expression system and target protein

- Refined model:



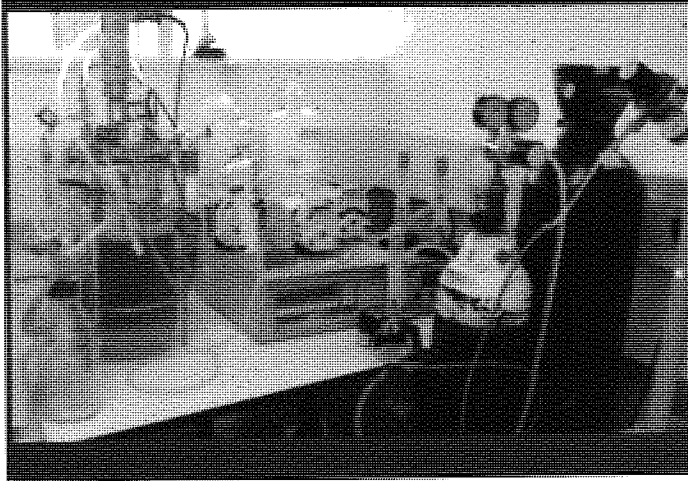
Testing the methodology with an alternate expression system and target protein

- Model assessed experimentally



Testing the methodology with an alternate expression system and target protein

- Process directly transferred to 7 L vessel
- Confirmation runs on-going



- Further optimisation possible
 - Bi-phase conditions
 - Glucose Feed

In conclusion

