

DOCTOR OF PHILOSOPHY

Improving recombinant human adenosine
A2A receptor production in yeast

Zharain Bawa

2014

Aston University

Some pages of this thesis may have been removed for copyright restrictions.

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

**IMPROVING RECOMBINANT HUMAN ADENOSINE A_{2A} RECEPTOR
PRODUCTION IN YEAST**

Zharain Bawa

Doctor of Philosophy

ASTON UNIVERSITY

May, 2014

©Zharain Bawa, 2014

Zharain Bawa asserts her moral right to be identified as the author of this thesis

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

IMPROVING RECOMBINANT HUMAN ADENOSINE A_{2A} RECEPTOR
PRODUCTION IN YEAST

Zharain Bawa

Ph.D

2014

Thesis summary

Over 50% of clinically-marketed drugs target membrane proteins; in particular G protein-coupled receptors (GPCRs). GPCRs are vital to living cells, performing an active role in many processes, making them integral to drug development. In nature, GPCRs are not sufficiently abundant for research and their structural integrity is often lost during extraction from cell membranes.

The objectives of this thesis were to increase recombinant yield of the GPCR, human adenosine A_{2A} receptor (hA_{2A}R) by investigating bioprocess conditions in large-scale *Pichia pastoris* and small-scale *Saccharomyces cerevisiae* cultivations. Extraction of hA_{2A}R from membranes using novel polymers was also investigated.

An increased yield of hA_{2A}R from *P. pastoris* was achieved by investigating the methanol feeding regime. Slow, exponential feed during induction (μ_{low}) was compared to a faster, exponential feed (μ_{high}) in 35 L pilot-scale bioreactors. Overall hA_{2A}R yields were increased for the μ_{low} cultivation (536.4 pmol g⁻¹) compared to the μ_{high} 148.1 pmol g⁻¹. hA_{2A}R levels were maintained in cytotoxic methanol conditions and unexpectedly, pre-induction levels of hA_{2A}R were detected. Small-scale bioreactor work showed that Design of Experiments (DoE) could be applied to screen for bioprocess conditions to give optimal hA_{2A}R yields. Optimal conditions were retrieved for *S. cerevisiae* using a d-optimal screen and response surface methodology. The conditions were 22°C, pH 6.0, 30% DO without dimethyl sulphoxide. A polynomial equation was generated to predict hA_{2A}R yields if conditions varied.

Regarding the extraction, poly (maleic anhydride-styrene) or PMAS was successful in solubilising hA_{2A}R from *P. pastoris* membranes compared with dodecyl- β -D-maltoside (DDM) detergent. Variants of PMAS worked well as solubilising agents with either 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) or cholesteryl hemisuccinate (CHS). Moreover, esterification of PMAS improved solubilisation, suggesting that increased hydrophobicity stabilises hA_{2A}R during extraction.

Overall, hA_{2A}R yields were improved in both, *P. pastoris* and *S. cerevisiae* and the use of novel polymers for efficient extraction was achieved.

Keywords: Yeast; human adenosine A_{2A} receptor, methanol-induction, poly (maleic anhydride-styrene), Design of Experiments

I would like to dedicate this thesis

To my loving parents, Rabia and Hassan

To my loving sister, Alia, and Adam and Aden

And to my precious husband, Mathew who gave me his undying love and support throughout
the whole of this PhD

And for Sofia, my beautiful daughter, you are my world.

Acknowledgements

Firstly, I would like to give special thanks to my supervisor Professor Roslyn Bill for all her time, guidance and support throughout the course of this Ph.D. and the writing up of this thesis. My thanks also go to Dr. Markus Ganzlin for his industrial supervision at AstraZeneca Ltd. I would also like to thank Professor David Poyner for his guidance with the radio-ligand binding work. Thanks also go to Dr. Sarah Routledge, Dr. Richard Darby, Dr. Matthew Conner and Dr. James Barwell for helping and teaching me many laboratory techniques and for many fruitful discussions. My thanks also go to my lab colleagues, Michelle Clare, Stephanie Cartwright, Marvin Dilworth, Debasmita Sarkar, Charlie Bland, Bitu Tajarodkhah and Gabriel Kuteyi for making this Ph.D an enjoyable experience. I also would like to thank the BBSRC and AstraZeneca Ltd for funding this project.

Finally I would like to thank everyone involved in collaborations concerning this thesis:

- The PMAS polymers research was done in collaboration with Professor Brian Tighe and Dr. Anisa Mahomed, Chemical Engineering and Applied Chemistry, Aston University, Birmingham, United Kingdom (Chapter 5).
- The pilot-scale cultivations were carried out at AstraZeneca Ltd., Alderley Park, United Kingdom and directed by Dr. Markus Ganzlin, Industrial Supervisor (current address: Lonza BioPharma Ltd CH-3930 Visp, Switzerland) as part of the BBSRC Industrial CASE studentship award (Chapter 4).
- Helpful advice and discussions regarding the DoE work were provided by Dr. Martin D.B. Wilks, Smallpeice Enterprises, Leamington Spa, United Kingdom, (Chapter 3).
- Advice and guidance on the SMA co-polymer work was given by Dr. Yu-pin Lin and Dr. Mohammed Jamshad, University of Birmingham, United Kingdom, (Chapter 5).
- hA_{2A}R construct design work was done in collaboration with Dr. Arjan Snijder, AstraZeneca, Mölndal, Sweden and Dr. Markus Ganzlin (Chapter 5).

CONTENTS

	Page
Abbreviations	17
List of figures	21
List of tables	25
Chapter 1: Introduction	28
1.1. Overview	28
1.2. Recombinant protein production in a commercial environment	28
1.2.1. Challenges in recombinant protein production for commercial use	29
1.2.2. Yeast as a host system for recombinant protein production – <i>Saccharomyces cerevisiae</i> and <i>Pichia pastoris</i>	30
1.2.2.1. The microbiology of <i>S. cerevisiae</i> and <i>P. pastoris</i>	32
1.2.2.2. Growth characteristics of <i>S. cerevisiae</i> and <i>P. pastoris</i>	35
1.2.2.3. Genomic characterisation and strain selection of <i>S. cerevisiae</i> and <i>P. pastoris</i>	37
1.2.2.4. <i>S. cerevisiae</i> and <i>P. pastoris</i> as cell factories for recombinant protein production	39
1.2.3. Molecular biological considerations	41
1.2.3.1. Vector choices and promoters for <i>S. cerevisiae</i>	41
1.2.3.2. Vector choices and promoters for <i>P. pastoris</i>	42
1.3. Recombinant protein production in yeast: from clone to culture	43
1.3.1. Screening for high yielding clones	44
1.3.2. Culture vessels	45
1.3.3. Bioprocessing parameters	46
1.3.3.1. Temperature	47
1.3.3.2. pH	47
1.3.3.3. Dissolved oxygen concentration (DO)	47
1.3.3.4. Additives	47
1.3.3.5. Antifoams	48

1.3.4. <i>P. pastoris</i> bioreactor cultivations	48
1.3.5. The application of Design of Experiments (DoE) for recombinant protein production	51
1.3.5.1. Factors, levels and responses for a DoE set-up	51
1.3.5.2. DoE process screening	52
1.3.5.3. DoE process characterisation and optimisation	53
1.4. The challenges of recombinant membrane protein production in yeast	55
1.4.1. The target protein: the GPCR, human adenosine A _{2A} receptor (hA _{2A} R)	56
1.4.1.1. Ligand-binding assay for hA _{2A} R	61
1.4.2. The yeast plasma membrane	63
1.4.3. Extraction of recombinant membrane proteins from the yeast plasma membranes	66
1.4.3.1. Detergent use in membrane protein solubilisation	66
1.4.3.2. Alternative solubilising agents	69
1.4.3.2.1. Bicelles	69
1.4.3.2.2. Amphipols	69
1.4.3.2.3. Nanodiscs	70
1.4.3.2.4. Maltose-neopentyl glycol (MNG) amphiphiles	70
1.4.3.2.5. Responsive hydrophobically associating polymers	70
1.5. Project aims	75
Chapter 2: Material and methods	76
2.1. Materials	76
2.1.1. Reagents and buffers	76
2.1.1.1. Stock solutions and buffers	76
2.1.1.1.1. 10× YNB (13.4% yeast nitrogen base with ammonium sulphate without amino acids)	76
2.1.1.1.2. 500× biotin (0.02%)	76

2.1.1.1.3. 10× glycerol (10%)	76
2.1.1.1.4. 10× methanol (5%)	76
2.1.1.1.5. 40× glucose (40%)	76
2.1.1.1.6. 1 M potassium phosphate buffer, pH 6.0	76
2.1.1.1.7. 1× T.A.E. buffer	77
2.1.1.1.8. 1 M DTT (dithiothreitol)	77
2.1.1.1.9. 1 M HEPES (2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid) buffer pH 8.0	77
2.1.1.2. <i>E. coli</i> culture medium	77
2.1.1.2.1. LB (Luria-Bertani)	77
2.1.1.3. <i>S. cerevisiae</i> culture media	77
2.1.1.3.1. Yeast peptone dextrose (YPD)	77
2.1.1.3.2. 2× CBS (Centralbureau voor Schimmelcultures) medium	77
2.1.1.3.3. CSM (complete synthetic medium)	77
2.1.1.3.4 10× Drop out solution (DO solution) (minus uracil)	78
2.1.1.3.5 10× Drop out solution (DO solution) (minus histidine)	78
2.1.1.3.6 Trace elements solution	78
2.1.1.3.7. Vitamin solution	78
2.1.1.4. <i>P. pastoris</i> culture media	79
2.1.1.4.1. Buffered complex glycerol medium (BMGY)	79
2.1.1.4.2. Buffered complex methanol medium (BMMY)	79
2.1.1.4.3. Basal salts medium (BSM)	79
2.1.1.4.4. PTM ₁ trace salts	79
2.1.1.4.5. Fermentation medium 22 (FM22)	79

2.1.1.4.6. PTM ₄ trace salts	79
2.1.1.4.7. 1 M Sorbitol	80
2.1.1.4.8. YPDS (yeast extract, peptone, dextrose, sorbitol) plus 100 µg mL ⁻¹ zeocin culture plates	80
2.1.1.4.9. YPDS (yeast extract, peptone, dextrose, sorbitol) plus 250 µg mL ⁻¹ zeocin culture plates	80
2.1.1.4.10. YPDS (yeast extract, peptone, dextrose, sorbitol) plus 500 µg mL ⁻¹ zeocin culture plates	80
2.1.1.4.11. YPDS (yeast extract, peptone, dextrose, sorbitol) plus 1000 µg mL ⁻¹ zeocin culture plates	80
2.1.1.4.12. Antifoams	81
2.1.1.4.13. Glycerol analysis	81
2.1.1.4.14. Methanol analysis	81
2.1.1.5. Antibiotics	81
2.1.1.5.1. Zeocin	81
2.1.1.5.2. Ampicillin	81
2.1.1.5.3. Doxycycline	81
2.1.1.6. Molecular biology reagents	82
2.1.1.6.1. pPICZB vector	82
2.1.1.6.2. Restriction enzyme <i>PmeI</i>	82
2.1.1.6.3. Mini-preparation kit	82
2.1.1.6.4. Purification of <i>PmeI</i> digestion	82
2.1.1.6.5. XL-10 Gold <i>E.coli</i> cells	83
2.1.1.7. Yeast strains	83
2.1.1.7.1. <i>P. pastoris</i> X33	83
2.1.1.7.2. <i>P. pastoris</i> SMD1163	83

2.1.1.7.3. <i>S. cerevisiae</i> BY4741	83
2.1.1.7.4. <i>S. cerevisiae</i> KOY-TM6*	84
2.1.1.8. Membrane preparation reagents and materials	84
2.1.1.8.1. Breaking buffer pH 7.4	84
2.1.1.8.2. Buffer A pH 7.0	84
2.1.1.8.3. Glass beads	84
2.1.1.9. Protein quantification reagents	84
2.1.1.9.1. Bovine serum albumin (BSA) standard	84
2.1.1.9.2. Copper (II) sulphate solution	84
2.1.1.9.3. Bicinchoninic acid (BCA)	84
2.1.1.10. Immunoblot reagents	85
2.1.1.10.1. 5 × Laemmli sample buffer	85
2.1.1.10.2. SDS Tris buffer	85
2.1.1.10.3. Immunoblot Tris buffer	85
2.1.1.10.4. Phosphate buffered saline (PBS)	85
2.1.1.10.5. PBS-Tween buffer	85
2.1.1.10.6. PBS-5% milk	85
2.1.1.11. Solubilisation reagents	85
2.1.1.11.1. n-dodecyl-β-d-maltopyranoside (DDM)	85
2.1.1.11.2. Cholesteryl hemi-succinate (CHS)	85
2.1.1.11.3. 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC)	85
2.1.1.11.4. Poly (maleic anhydride-styrene) (PMAS)	86
2.1.1.11.5. Esterification of poly (maleic anhydride-styrene) (PMAS)	86
2.1.1.11.6. Styrene maleic acid (SMA)	86

2.1.1.12. Radio-ligand binding reagents	86
2.1.1.12.1. Binding buffer	86
2.1.1.12.2. Gel filtration columns	86
2.1.1.12.3. Tritiated ZM241385 ($[^3\text{H}]$ ZM241385)	87
2.1.1.12.4. Unlabelled ZM241385 (cold ZM241385)	87
2.1.1.12.5. Unlabelled theophylline (cold theophylline)	87
2.1.1.12.6. Unlabelled NECA (cold NECA)	87
2.1.1.12.7. Unlabelled XAC (cold XAC)	87
2.1.1.12.8. Scintillant	87
2.1.1.12.9. Soluene	87
2.1.1.13. Protein purification reagents	87
2.1.1.13.1. Purification columns	87
2.1.1.13.2. Lysis buffer	88
2.1.1.13.3. Wash buffer	88
2.1.1.13.4. Elution buffer	88
2.1.1.14. Green fluorescent protein (GFP) assay	88
2.1.1.14.1. Recombinant GFP	88
2.1.2. Equipment	89
2.1.2.1. Pilot-scale 35 L bioreactors AstraZeneca Ltd. Alderley Park, United Kingdom	89
2.1.2.2. Bench-top 2 L bioreactors	92
2.1.2.3. Micro-24 microreactor (small-scale bioreactors)	94
2.1.2.4. Centrifuges	94
2.1.2.5. Gas chromatograph (GC)	94
2.1.2.6. Cell lyser for small scale membrane preparations	95

2.1.2.7. Cell lyser for large scale membrane preparations	95
2.1.2.8. Scintillation counter for measurement of radioactivity	95
2.1.3. Software	95
2.1.3.1. Graphpad Prism software	95
2.1.3.2. Minitab statistical software	95
2.1.3.3. Clone Manager	95
2.1.3.4. Origin software	95
2.1.3.5. Microsoft Excel	95
2.2. Methods	95
2.2.1. Molecular biology techniques	95
2.2.1.1. Construct design and virtual cloning	95
2.2.1.2. Mini-preparations (Minipreps) of vector DNA	96
2.2.1.3. DNA quantification	96
2.2.1.4. Agarose gel electrophoresis	96
2.2.1.5. Linearisation of pPICZB vector-hA _{2A} R construct DNA	97
2.2.1.6. Purification of linearised pPICZB vector-hA _{2A} R construct DNA	97
2.2.1.7. Preparation of <i>P. pastoris</i> electrocompetent cells	97
2.2.1.8. Electroporation and recombinant clone selection	97
2.2.1.9. Transfection of hA _{2A} R vector into human embryonic kidney cells (HEK)	98
2.2.2. Yeast cultivations	100
2.2.2.1. Shake flask cultures	100
2.2.2.2. Pilot-scale 35 L bioreactor cultivations	100
2.2.2.2.1. Preparation of 35 L bioreactor	100
2.2.2.2.1.1. Vessel pressure hold testing	100

2.2.2.2.1.2. pH and DO probe calibration	101
2.2.2.2.1.3. Bioreactor sterilisation	101
2.2.2.2.2. Performing a 35 L bioreactor cultivation for <i>P. pastoris</i> producing hA _{2A} R	102
2.2.2.2.3. Finishing the 35 L bioreactor run	103
2.2.2.3. Bench-top 2 L bioreactor cultivations	103
2.2.2.3.1. Preparation of the 2 L bioreactor	103
2.2.2.3.2. Connecting the bioreactor to the control unit	104
2.2.2.3.3. Performing a 2 L bioreactor cultivation for <i>P. pastoris</i> producing hA _{2A} R	104
2.2.2.3.4. Glycerol batch and fed batch phase in 2 L bench-top bioreactor	105
2.2.2.3.5. Induction phase in 2 L bench-top bioreactor	105
2.2.2.4. Micro-24 microreactor cultures	105
2.2.2.4.1. Preparation of Micro-24 microreactor cultivations	105
2.2.2.4.2. Antifoam and medium assessment in Micro-24 well plates	107
2.2.2.4.3. Performing the Micro-24 microreactor cultivations	107
2.2.3. Glycerol assay	108
2.2.4. Methanol assay	109
2.2.5. Green Fluorescent Protein (GFP) assay	110
2.2.6. Small scale membrane preparations	110
2.2.7. Large scale membrane preparations	111
2.2.8. Protein quantification	111
2.2.9. Immunoblot analysis	112
2.2.9.1. Sodium dodecyl sulphate polyacrylamide electrophoresis (SDS PAGE) gels	112
2.2.9.2. Immunoblots	113
2.2.10. Yeast membrane solubilisation with DDM	114

2.2.11. Yeast membrane solubilisation with styrene maleic acid (SMA) co-polymer	114
2.2.12. Yeast membrane solubilisation with poly (maleic anhydride-styrene) (PMAS)	115
2.2.13. Radio-ligand binding for hA _{2A} R	115
2.2.13.1. Single-point binding assays for hA _{2A} R	116
2.2.13.2. Saturation binding assays for hA _{2A} R	116
2.2.13.3. Competition binding assays for hA _{2A} R	117
2.2.13.4. Competition binding curve on membrane bound hA _{2A} R from HEK cells	120
Chapter 3: The application of ‘Design of Experiments’ (DoE) to membrane protein production	121
3.1. Defining of DoE model for optimising hA_{2A}R yields in yeast	127
3.1.1. Experimental set-up	128
3.1.2. Input parameters	128
3.1.3. <i>Saccharomyces cerevisiae</i> strains	131
3.2. Execution of the d-optimal DoE design	132
3.3. Using response surface methodology (RSM) to improve the statistical outcome of the d-optimal data	144
3.3.1. Examination of residuals	147
3.4. The importance of a robust experimental set-up to minimise error	148
3.4.1. Instrument challenges	141
3.4.2. Biological sample error	152
3.5. Discussion	152
Chapter 4: An investigation into the induction phase of <i>P. pastoris</i> cultures: production of hA_{2A}R as a case study	156
4.1. Analysis of growth rates during methanol induction in pilot-scale bioreactors	156
4.1.1. Expression strain	156

4.1.2. Methanol feed regimes	159
4.1.3. Pilot-scale cultivations	161
4.2. A low methanol induction feed profile increases biomass yield leading to improved overall volumetric dG-hA_{2A}R yields	163
4.2.1. The μ_{low} <i>P. pastoris</i> cultivation yields increased biomass when compared to the μ_{high} <i>P. pastoris</i> cultivation	163
4.2.2. Increased volumetric dG-hA _{2A} R yields are achieved in the μ_{low} cultivation	165
4.3. Increased dG-hA_{2A}R yields correlate with low residual methanol in the cultivation but cytotoxic levels of residual methanol maintain baseline levels of active dG-hA_{2A}R	166
4.4. Yield co-efficients differ between μ_{low} and μ_{high} cultivations during induction	168
4.5. Calculated specific growth rates versus set growth rates for μ_{low} and μ_{high} cultivations	169
4.6. Pre-induction activity in <i>P. pastoris</i> cultivations	169
4.6.1. dG - hA _{2A} R binding activity is present in the pre-induction glycerol phases of the μ_{low} and μ_{high} cultivations	169
4.6.2. dG - hA _{2A} R binding activity is also present in the pre-induction glycerol phases of a 2L bench-top bioreactor cultivation	170
4.6.3. Soluble GFP is produced in all phases of <i>P. pastoris</i> bioreactor cultivations	172
4.7. Discussion	176
Chapter 5: The use of novel polymers to extract recombinant hA_{2A}R from <i>P. pastoris</i> membranes	179
5.1. Design of a novel hA_{2A}R construct	179
5.1.1. Construct design	179
5.1.2. Validation of construct design	181
5.1.3. Generation of a high yielding <i>P. pastoris</i> strain for production of recombinant MT-hA _{2A} R	183

5.1.4. Analysis of functional recombinant MT-hA _{2A} R in <i>P. pastoris</i>	186
5.2. Solubilisation of MT-hA_{2A}R from SMD1163 <i>P. pastoris</i> membranes with the detergent, n-dodecyl-β-d-maltopyranoside (DDM)	188
5.3. Solubilisation of MT-hA_{2A}R from SMD1163 <i>P. pastoris</i> membranes with responsive hydrophobically associating polymers	189
5.3.1. Preliminary optimisation of PMAS solubilisation conditions	193
5.3.2. Screening PMAS solutions for solubilisation of MT-hA _{2A} R from SMD1163 <i>P. pastoris</i> membranes	195
5.3.3. Comparative pharmacology of PMAS solubilised MT-hA _{2A} R	200
5.3.4. Investigation of the lipid composition in PMAS solubilisation mixtures	203
5.3.5. Purification of MT-hA _{2A} R after 2000P pH 11.0 solubilisation	205
5.4. Discussion	206
Chapter 6: General discussion	209
6.1. Controlling methanol feeds during membrane protein induction in <i>P. pastoris</i> to increase yields	211
6.1.1. Cytotoxic levels of methanol in the cultivation are not completely detrimental to hA _{2A} R activity	211
6.1.2. Pre-induction activity is present in pilot-scale and bench-top bioreactors	213
6.2. Using novel polymers to extract hA_{2A}R from <i>P. pastoris</i> membranes	215
6.3. Using DoE for hA_{2A}R production improvement in yeast	219
7. References	223
8. Appendices	241
A1. Classification of GPCRs	241
A2. MT-hA _{2A} R DNA and amino acid sequence alignment	243
A3. MT-hA _{2A} R sequence alignment	246
A4. Linear regression analysis for d-optimal DoE (Minitab [®] output data)	250

A5. Response surface regression analysis for RSM DoE (Minitab [®] output data)	253
A6. Patent application for PMAS work showing claims	256
A7. Oral presentation at Aston University post-graduate research day	262
A8. Poster presentations	271
A9. Publications arising from this thesis	272

Abbreviations

A _{2A} R	Adenosine A _{2A} receptor
APS	Ammonium persulphate
BCA	Bicinchoninic acid
B _{max}	Maximum binding capacity
BMGY	Buffered complex glycerol medium
BMMY	Buffered complex methanol medium
BSA	Bovine serum albumin
BSM	Basal salts medium
FM22	Fermentation Medium 22
PTM	<i>Pichia</i> trace metals
CCD	Charged coupled device
CDA	Clean dry air
CHS	Cholesteryl hemi-succinate
CBS	Centralbureau vorr Schimmelcultures
CSM	Complete synthetic medium
DDM	n-dodecyl-β-d-maltopyranoside
DMEM	Dulbecco's Modified Eagle's medium
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DO	Dissolved oxygen concentration
DO	Drop out solution
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EC ₅₀	Half the effective concentration
ECL	Extracellular loop
GC	Gas chromatography
GFP	Green fluorescent protein
GPCR	Guanine nucleotide-binding protein coupled receptor
³ H	Tritium
hA _{2A} R	Human adenosine A _{2A} receptor
HEK	Human embryonic kidney cells
HEPES	(2-[4-(2-hydroxyethyl) piperazin-1-yl]ethanesulfonic acid)
ICL	Intracellular loop
K _d	Concentration of ligand required to bind half of the receptors

K _i	Concentration of competing ligand required to bind half of the receptors
LB	Luria-Bertani
MEK	Methyl ethyl ketone
MFC	Mass flow controller
NECA	5'- <i>N</i> -Ethylcarboxamidoadenosine
Ni-NTA	Nickel nitrilotriacetic acid
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline Tween 20
PEI	Poly(ethylenimine)
pH	Negative logarithm of the hydrogen ion concentration
PID	Proportional – integral - derivative
pK _d	- log ₁₀ of K _d
pK _i	- log ₁₀ of K _i
PMAS	Poly (maleic anhydride-styrene)
<i>P. pastoris</i>	<i>Pichia pastoris</i>
PTFE	Polytetrafluoroethylene
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SCADA	Supervisory control and data acquisition software
SDS	Sodium dodecyl sulphate
SMA	Styrene maleic anhydride co-polymer
SMALP	Styrene maleic acid lipid particle
T	Time
TAE	Tris acetate EDTA
TEMED	N,N,N',N'-tetramethyl- ethane-1,2-diamine
TEV	Tobacco etch virus
TM	Transmembrane domain
Tris	Tris(hydroxymethyl)aminoethane
Tween 20	Polyoxyethylene sorbitan monolaurate
UPS	Uninterruptable power supply
v/v	Volume/volume
WT	Wild type
w/v	Weight/volume
XAC	Xanthine amine congener
YPD	Yeast peptone dextrose
YPDS	Yeast peptone dextrose sorbitol

YNB	Yeast nitrogen base
ZM241385	4-(2-(7-amino-2-(furan-2-yl)-[1,2,4]triazolo[1,5-a][1,3,5]triazin-5-ylamino)ethyl)phenol
μ	Specific growth rate
λ_{exe}	Excitation wavelength
λ_{em}	Emission wavelength

Units

°C	Celsius
fmol	Femtomole
g	Gram
g	Gravitational force
Hz	Hertz
h	Hour
kDa	Kilo Dalton
Kg	Kilogram
L	Litre
mg	Milligram
min	Minute
mL	Millilitre
mmol	Millimole
mol	Mole
ms	Millisecond
nM	Nanomolar
nm	Nanometre
Nm	Newton metre
nmol	Nanomole
pmol	Picomole
psi	Pounds per square inch
RCF	Reactive centrifugal force
RFU	Relative fluorescence units
rpm	Revolutions per minute
s	Second
U	Units
V	Volts
vvm	Volume per volume per minute
μL	Microlitre

μg Microgram
μM Micromolar

Amino acids

Table 1 Standard amino acid abbreviations. Standard amino acid abbreviations used interchangeably throughout text.

Amino acid	3-letter	1-letter	Amino acid	3-letter	1-letter
Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamic acid	Glu	E	Serine	Ser	S
Glutamine	Gln	Q	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

List of figures

	Page
Chapter 1	
Figure 1.1 Confocal microscopy image of <i>S. cerevisiae</i> cells	31
Figure 1.2 Confocal microscopy image of <i>P. pastoris</i> cells	31
Figure 1.3 Yeast cell wall	33
Figure 1.4 Methanol metabolism in <i>P. pastoris</i>	34
Figure 1.5 Theoretical growth curve of microbial cells	36
Figure 1.6 pPICZ α expression vector series map from Life Technologies Corporation	43
Figure 1.7 Micro-24 microreactor at Aston University	44
Figure 1.8 Micro-24 microreactor software screen shot	45
Figure 1.9 Bench-top 2 L bioreactor set-up at Aston University	46
Figure 1.10 Pilot-scale 35 L bioreactor set-up at AstraZeneca Ltd.	46
Figure 1.11 Typical <i>P. pastoris</i> bioreactor cultivation for recombinant protein production	49
Figure 1.12 One Factor At a Time (OFAT) versus full factorial design	52
Figure 1.13 Fractional factorial design	53
Figure 1.14 CCF design and CCC design	54
Figure 1.15 Summary of G protein cycle	57
Figure 1.16 Diagrammatic representation of Family A GPCR	58
Figure 1.17 Crystal structure of hA _{2A} R with ZM241385 antagonist bound	61
Figure 1.18 General structure of a phospholipid	64
Figure 1.19 Chemical structure of cholesterol and ergosterol	65
Figure 1.20 Yeast plasma membrane	65
Figure 1.21 Scheme to show stages of lipid bilayer solubilisation by detergent	67
Figure 1.22 Images of proposed structures formed with alternative solubilising agent, lipid and membrane protein	74
Chapter 2	
Figure 2.1 pPICZB expression vector developed by Life Technologies Corporation	82
Figure 2.2 Pilot-scale fermentation facility set-up	90

Figure 2.3 Sartorius 35 L pilot-scale bioreactor	91
Figure 2.4 2 L Bench-top bioreactor set-up at Aston University	93
Figure 2.5 Micro-24 microreactor (Pall Corporation) at Aston University	94
Figure 2.6 pcDNA3.1 expression vector developed by Life Technologies Corporation	99
Figure 2.7 Biostat C valve settings	101
Figure 2.8 Micro-24 microreactor set-up at Aston University	106
Figure 2.9 Micro-24 microreactor 24 well plate	106
Figure 2.10 Antifoam and media test set up of the Micro-24 microreactor 24 well plate	107
Figure 2.11 A representative calibration curve for determining the concentration of methanol	110
Figure 2.12 Emulisflex-C3 (Avestin) pressure homogeniser	111
Figure 2.13 BCA representative standard curve	112
Chapter 3	
Figure 3.1 Carrying out a DoE - steps to follow in a bioprocess setting	124
Figure 3.2 Illustration of residuals	126
Figure 3.3 Micro-24 microreactor set up at Aston University	128
Figure 3.4 hA _{2A} R construct design schematic for <i>S. cerevisiae</i> transformations	131
Figure 3.5 Micro-24 microreactor plate set-up	134
Figure 3.6 Main effects plots of OD600 values for each hA _{2A} R strain, for each input factor at the end of the Micro-24 microreactor run	135
Figure 3.7 Main effects, interaction and ANOVA analysis for WT-hA _{2A} R	138
Figure 3.8 Main effects, interaction and ANOVA analysis for BMS1-hA _{2A} R	139
Figure 3.9 Main effects, interaction and ANOVA analysis for TM6-hA _{2A} R	140
Figure 3.10 Residual plots for WT-hA _{2A} R, BMS1-hA _{2A} R and TM6-hA _{2A} R after linear regression analysis from d-optimal design	141
Figure 3.11 Actual set-points for DO and pH for WT-hA _{2A} R replicate 1	142
Figure 3.12 Actual set-points for DO and pH for BMS1-hA _{2A} R and TM6-hA _{2A} R replicate 1	143

Figure 3.13 Response optimisation plot for WT-hA _{2A} R strain after RSM analysis	146
Figure 3.14 Contour and surface wire-frame plot for WT-hA _{2A} R strain after RSM analysis	147
Figure 3.15 Residual plots for WT-hA _{2A} R strain after RSM analysis showing actual DO and pH set-points for the whole experiment	148
Figure 3.16 Actual DO and pH data for WT-hA _{2A} R experiment	150
Figure 3.17 DO calibration test	151
Figure 3.18 Membrane preparation process for Micro-24 yeast cultures	152
Chapter 4	
Figure 4.1 Schematic diagram for dG-hA _{2A} R construct	156
Figure 4.2 Immuno-blot from X33-dG-hA _{2A} R colony screen	157
Figure 4.3 Bioreactor cultivation of <i>P. pastoris</i> X33 dG-hA _{2A} R	158
Figure 4.4 Immuno-blot from X33-dG-hA _{2A} R 2 L bioreactor cultivation (clone 26)	159
Figure 4.5 Bioprocess parameters for the μ_{low} and μ_{high} cultivations	164
Figure 4.6 Biomass yield of the μ_{low} and μ_{high} cultivation	165
Figure 4.7 Cumulative addition of methanol to μ_{low} and μ_{high} cultivations	168
Figure 4.8 dG - hA _{2A} R binding activity is present in all phases of a subsequent 2 L cultivation	171
Figure 4.9 dG - hA _{2A} R binding activity is present in all phases in a methanol-free cultivation	172
Figure 4.10 Recombinant GFP is produced in the pre-induction phases of two independent 2 L cultivations	174
Figure 4.11 Recombinant GFP is produced in methanol-free cultivations of <i>P. pastoris</i> grown on either glycerol or glucose as the carbon source	175
Chapter 5	
Figure 5.1 MT-hA _{2A} R construct design and sequence	180
Figure 5.2 pPICZB-MT-hA _{2A} R vector map	181
Figure 5.3 Competition binding curves for untagged, His-tagged and MT-hA _{2A} R from HEK cell membranes	182

Figure 5.4 Agarose gel (1%) showing purified pPICZB-MT-hA _{2A} R before and after linearisation	183
Figure 5.5 Scheme for screening MT-hA _{2A} R expression	184
Figure 5.6 Growth curves for several colonies of SMD1163-MT-hA _{2A} R in BMGY cultured in shake flasks	185
Figure 5.7 Immuno-blot from SMD1163-MT-hA _{2A} R colony screen	186
Figure 5.8 Specific binding activities from SMD1163 MT-hA _{2A} R colony expression screen	187
Figure 5.9 Saturation binding curve and summary of wet cell weight, specific growth rates (μ), B _{max} and pK _d data for X33-dG-hA _{2A} R and SMD1163-MThA _{2A} R in FM22 minimal media	188
Figure 5.10 Saturation binding curve for DDM solubilised MThA _{2A} R from SMD1163 <i>P. pastoris</i> cells	189
Figure 5.11 Chemical structure of poly (maleic anhydride-styrene) and poly (maleic acid-styrene)	190
Figure 5.12 Clarity of solutions in solubilisation experiments	195
Figure 5.13 Cracked ultracentrifugation tubes with 1000F pH 11.0 solubilisation mixture after 1 h spin at 100000 \times g	196
Figure 5.14 Graphical representation of clarity of solubilisation solution versus MT-hA _{2A} R activity	198
Figure 5.15 Percent recovery of MT-hA _{2A} R from PMAS solubilisations	199
Figure 5.16 Competition binding curves for MT-hA _{2A} R (membrane bound, DDM solubilised, 2000P pH 7.0 solubilised, 2000P pH 11.0 solubilised, 1600ME pH 7.0 solubilised and 1600ME pH 11.0 solubilised) with ZM241385, XAC, NECA and theophylline agonist and antagonists	202
Figure 5.17 Competition binding curve after solubilisation in (i) 2000P pH 11.0 with CHS and DMPC and (ii) DDM with CHS	204
Chapter 6	
Figure 6.1 Schematic for proposed automation on of membrane protein production using DoE and Micro-24 microreactor	221
Figure 6.2 Sketch of Micro-24 well plate with modifications	222

List of tables

	Page
Abbreviations	
Table 1. Standard amino acid abbreviations	20
Chapter 1	
Table 1.1 <i>P. pastoris</i> strains commonly used in recombinant protein production and their phenotypes	38
Table 1.2 Comparison of some main characteristics of <i>S. cerevisiae</i> and <i>P. pastoris</i>	39
Table 1.3 Examples of clinically relevant recombinant proteins produced from <i>S. cerevisiae</i> or <i>P. pastoris</i> host cells	40
Table 1.4 Examples of DoE in bioprocess development	55
Table 1.5 Human adenosine A _{2A} crystal structures resolved	60
Table 1.6 Four main groups of detergents used for solubilisation of membrane proteins	68
Table 1.7 Alternative solubilising agents	72
Chapter 2	
Table 2.1 Summary of bioprocess events during two simultaneous <i>P. pastoris</i> cultivations producing recombinant dG - hA _{2A} R with different exponential methanol feed rates (μ_{set})	102
Table 2.2 10% separating gel	113
Table 2.3 4% stacking gel	113
Table 2.4 Single-point binding assay reaction set up	116
Table 2.5 Saturation binding assay reaction set-up for membrane bound hA _{2A} R	117
Table 2.6 Saturation binding assay reaction set-up for solubilised hA _{2A} R	117
Table 2.7 Competition binding assay reaction set-up for membrane bound or solubilised hA _{2A} R with ZM241385	118
Table 2.8 Competition binding assay reaction set-up for membrane bound or solubilised hA _{2A} R with theophylline	119
Table 2.9 Competition binding assay reaction set-up for membrane bound or solubilised hA _{2A} R with NECA	119
Table 2.10 Competition binding assay reaction set-up for membrane bound or solubilised hA _{2A} R with XAC	120

Chapter 3

Table 3.1 Glossary of relevant DoE terminology	122
Table 3.2 Types of DoE design and their characteristics	127
Table 3.3 Input factors used in the DoE screen for hA _{2A} R production	130
Table 3.4 <i>S. cerevisiae</i> strains producing hA _{2A} R used for DoE research	131
Table 3.5 d-optimal design of 4 input factors, 3 numeric factors and 1 attribute factor	133
Table 3.6 Total protein yields of membrane preparations from each input combination and <i>S. cerevisiae</i> strain	136
Table 3.7 Specific binding activity from each input combination and <i>S. cerevisiae</i> strain	137
Table 3.8 ANOVA analysis output for WT-hA _{2A} R strain after RSM analysis	145

Chapter 4

Table 4.1 Bioprocess events during two simultaneous <i>P. pastoris</i> cultivations producing recombinant dG - hA _{2A} R with different exponential methanol feed rates (μ_{set})	162
Table 4.2 The yield of dG-hA _{2A} R is higher in the μ_{low} cultivation than the μ_{high} cultivation	166
Table 4.3 dG-hA _{2A} R activity, residual methanol and cumulative methanol values for μ_{low} and μ_{high} cultivations	167
Table 4.4 Calculated yield co-efficients for the μ_{low} and μ_{high} cultivations	168
Table 4.5 Calculated specific growth rates for the μ_{low} and μ_{high} cultivations	169
Table 4.6 Residual glycerol in the culture medium does not repress dG-hA _{2A} R expression	170

Chapter 5

Table 5.1 Four PMAS solutions used for testing solubilisation of hA _{2A} R in <i>P. pastoris</i> membranes	192
Table 5.2 Summary of solubilisation condition tests using PMAS 2000P pH 7.0	194
Table 5.3 B _{max} and pK _d values for DDM and PMAS solubilisations	197
Table 5.4 Summary of pK _i values as an indication of the pharmacological properties of MT-hA _{2A} R after PMAS and DDM solubilisations	201
Table 5.5 Summary of lipid (DMPC and CHS) composition in PMAS and DDM solubilisation experiments	203
Table 5.6 MT-hA _{2A} R activity of native membrane, solubilised with either DDM or 2000P pH 11.0 and at each purification step	205

Chapter 6

Table 6.1 Thesis objectives, optimal conditions and major findings

210

Chapter 1: Introduction

1.1. Overview

Over 50% of clinically marketed drugs target membrane proteins; in particular G-protein coupled receptors (GPCRs) (Gudermann et al., 1995). GPCRs are vital to living cells due to their active role in many cellular processes, including signal transmission. Many diseases centre on the action of these GPCRs, making an understanding of their structures and functions integral to drug design (Flower, 1999). In nature, GPCRs are not sufficiently abundant for such studies and their structural integrity is often lost during extraction from the cell membrane. Therefore, a major challenge pharmaceutical industries face is to reproducibly produce high yields of structurally stable recombinant GPCRs (Baker, 2010, Flower, 1999).

This thesis is concerned with a specific GPCR - recombinant human adenosine A_{2A} receptor (hA_{2A}R) - and aims to improve the tractability of its production in yeast. The research focussed on two main areas: the bioprocess of yeast cultures and the extraction of hA_{2A}R from yeast cell membranes. For the bioprocess aspect, the work was further divided into two areas: an examination of feeding regimes and their impact on hA_{2A}R production in the yeast, *Pichia pastoris*, and an investigation into optimising hA_{2A}R production conditions in the yeast, *Saccharomyces cerevisiae* via a statistical Design of Experiments (DoE) approach. For the extraction aspect, novel polymers called poly (maleic anhydride-styrene) (PMAS) were used to solubilise recombinant hA_{2A}R from *P. pastoris* membranes to investigate an alternative to detergent solubilisations.

1.2. Recombinant protein production in a commercial environment

Sir Alexander Fleming's discovery of penicillin from a common mold in 1928 and its subsequent development for use in World War II for treatment of injured soldiers prompted the need to produce the organism in large quantities in tanks of broth, thereby signifying the beginnings of pharmaceutical biotechnology (Rhodes, 1985). Several decades later, in the 1970s, gene cloning and recombinant DNA technology were developed and define modern day pharmaceutical biotechnology (Doran, 1995). Recombinant DNA technology provides a solution to producing proteins in non-native host cells. In general, recombinant protein production is performed if the native protein of interest is present low quantities, if the protein of interest is innately unstable or if mutant forms of the protein are required for *in vitro* studies (Schmidt, 2004, Doran, 1995). In terms of clinical and pharmaceutical use, recombinant proteins are essential for the development of drugs (either as drugs or drug targets), the engineering of antibodies and enzymes as well as other therapeutic protein production e.g. insulin, human growth hormone. In fact, in the U. S. market alone, the volume of recombinant

protein products on sale recently is more than \$ 50-60 billion (Schmidt, 2004). Also, it is not only the large pharmaceutical companies that develop, produce and market recombinant protein products, but many small and medium enterprises (SMEs) follow the recombinant protein production business model (Repasi, 2013). SMEs tend to focus on a single disease area and develop therapies and drugs. Their popularity is increasing since the world financial crisis started in the late 2000s and big pharmaceutical companies began outsourcing and downsizing (Repasi, 2013). Notable recombinant protein products include erythropoietin (developed by Janssen), insulin (developed by Novo-Nordisk) and Hepatitis B vaccine (developed by Glaxo Smith Kline) (Doran, 1995, Schmidt, 2004, Demain and Vaishnav, 2009).

Typically, recombinant protein production involves amplifying a DNA sequence which encodes a protein of interest. It is inserted into an appropriate vector and then transformed or transfected into the chosen host cell. The combination of the DNA encoding the gene of interest, usually from another organism, and the vector DNA results in a recombinant DNA product and hence the protein that is produced by the cells is termed a recombinant protein (Doran, 1995).

1.2.1. Challenges in recombinant protein production for commercial use

The demand for therapeutic and prophylactic recombinant protein treatment continues to increase (Schmidt, 2004). Yet, new challenges present themselves with regards to recombinant protein production using current technologies. On a commercial stance, many applications would benefit from higher production efficiencies with a better quality of product whilst lowering the costs of the final product (Doran, 1995).

Common issues encountered with recombinant protein production include loss of expression caused by structural changes or disappearance of the recombinant gene; issues in post-translational processing including protein mis-folding, protein aggregation, proteolytic processing and glycosylation, depending on the host system used (Mattanovich, 2012). Strategies are usually in place to minimise these unwanted effects of recombinant protein production in a commercial setting in order to maximise quality product at a low cost whilst increasing yield (Schmidt, 2005). These include optimising bioreactor operations such as controlling the temperature, pH and aeration delivered to the host organism which is producing the recombinant protein (section 1.3.3.); (Schmidt, 2004, Schmidt, 2005, Doran, 1995). In this thesis, Chapter 3 investigates this approach by applying it to recombinant hA_{2A}R production in yeast. Another strategy routinely employed to overcome issues in recombinant protein production in a commercial setting are by applying induction strategies and controlling growth of the host organism (section 1.2.2.2. and 1.3.4.); (Potvin et al., 2012). This is addressed in Chapter 4 of this thesis where controlling the induction feeding regimes of yeast producing hA_{2A}R are examined.

20-30% of all open reading frames encode for membrane proteins (Wallin and Heijne, 1998). The largest family of membrane proteins, GPCRs, transduce extracellular stimuli into intracellular signals and are therefore vital drug targets when implicated in disease making them attractive proteins for pharmaceuticals to study (Gudermann et al., 1995, Flower, 1999). Since membrane proteins are not secreted out of the cell and are embedded in the cell membrane, extraction of the recombinant protein must be performed in order for further study to be carried out; this comes with its own challenges (section 1.4.); (Seddon et al., 2004). Chapter 5 studies the use of a novel compound to improve extraction of recombinant hA_{2A}R from yeast membranes.

1.2.2. Yeast as a host system for recombinant protein production - *Saccharomyces cerevisiae* and *Pichia pastoris*

S. cerevisiae (Figure 1.1) is more commonly referred to as Bakers' or Brewers' yeast and its uses span beer, wine and bread making (Lodolo et al., 2008), being a model organism in molecular biology (Botstein, 1997), being a biological tool for studying aging (Longo et al., 2012) and as a host system for recombinant protein production (Mattanovich, 2012). It has been classified as 'Generally Regarded As Safe' or GRAS by the FDA in the U.S.A. and is regularly used in food stuff production (Mattanovich, 2012).

P. pastoris (Figure 1.2) was first reported along with other yeast species, *Candida*, *Hansenula* and *Torulopsis* by Koichi Ogata and colleagues in 1969 to utilise methanol as a sole carbon source (Ogata, 1969). They coined the term 'methylotroph' to describe these types of yeast. At the beginning of the 1970s, acquiring methanol from methane was a cheap process; therefore these species were an attractive source of single cell protein (SCP) for animal feed (Cregg, 1985). However, later in the 1970s a global oil crisis occurred and the cost of methane increased and the use of SCP was abandoned. Phillips Petroleum, USA subsequently developed the first recognised methods and protocols for culturing *P. pastoris* to high cell densities using methanol (Wegner, 1990) and it has since been utilised as a popular host system (available from Life Technologies Corporation) for recombinant protein production due to its regulated alcohol oxidase (AOX) promoter (Cereghino and Cregg, 2000).

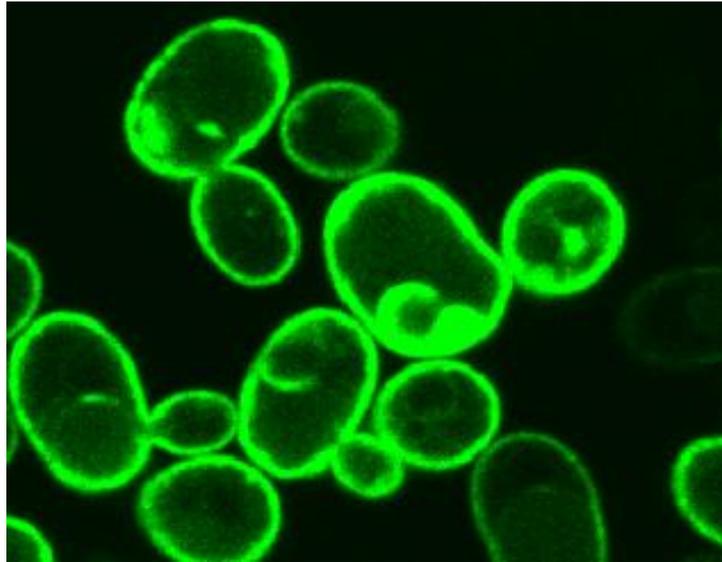


Figure 1.1 Confocal microscopy image of *S. cerevisiae* cells. The cells shown are producing a membrane protein tagged with recombinant green fluorescent protein and are actively budding. The cells are approximately 5-10 μm in size and were taken on a Leica TCS SP5 confocal microscope at 100 \times magnification (Image provided by Dr. Debasmita Sarkar, Aston University).

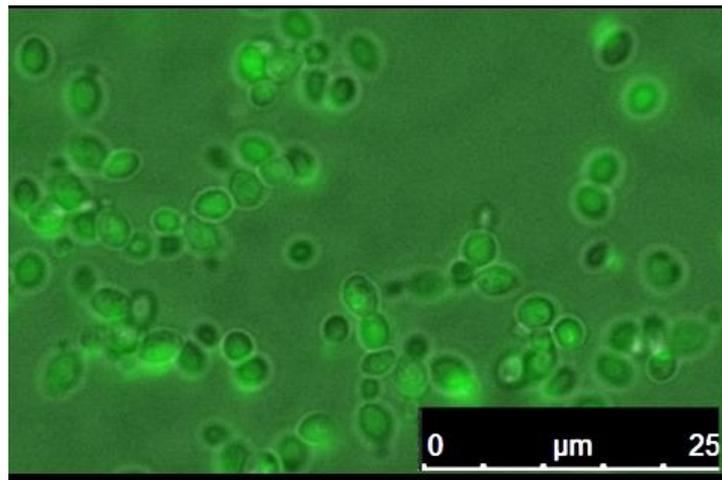


Figure 1.2 Confocal microscopy image of *P. pastoris* cells. Cells are secreting recombinant green fluorescent protein, are actively budding and are approximately 5-10 μm in size and were taken on a Leica TCS SP5 confocal microscope at 100 \times magnification (Image provided by Dr. Sarah Routledge, Aston University).

1.2.2.1. The microbiology of *S. cerevisiae* and *P. pastoris*

Saccharomyces cerevisiae and *Pichia pastoris* are eukaryotic organisms belonging to the Fungi kingdom, the *Ascomycota* phylum and the *Saccharomycetaceae* family. *S. cerevisiae* belongs to the *Saccharomyces* genus and *P. pastoris* belongs to the *Komagatella* genus (Dujon, 2010). This was originally the *Pichia* genus but was re-classified after phylogenetic sequencing was performed (Mikata and Yamada, 1995) but is still frequently referred to as *Pichia*.

S. cerevisiae and *P. pastoris* are between 5-10 μm in size and can reproduce sexually and asexually (Herskowitz, 1988). It is more common for the cells to reproduce asexually which is known as 'budding' (Herskowitz, 1988). There are three cell types for *S. cerevisiae* and *P. pastoris*, a type, α type and a/ α type and all of them are capable of undergoing mitotic cell division or budding. The haploid cells (a and α) can however mate giving rise to diploid cells and these diploid cells can undergo meiosis and yield haploid cells via the production of spores (Herskowitz, 1988). For both, it is known that if nitrogen is limited when the cells are under stress, sexual reproduction can occur (Shen et al., 1998).

Yeast has a three component cell envelope comprising the cell wall, the periplasmic space and the plasma membrane (Walker, 1998). The function of this envelope is to keep the yeast cell intact and to protect the cell. It also regulates and enables transport of material in and out of the cell. The yeast periplasmic space is very thin ($\sim 35\text{-}45 \text{ \AA}$) with the cell wall residing externally to it and the plasma membrane internally to it. The space contains secreted proteins called mannoproteins. These are unable to permeate the cell wall but perform functions such as hydrolysis of substrates (Walker, 1998).

The cell wall is present in yeast, which is a major difference between yeast and mammalian cells. The cell wall needs to be removed before any membrane protein extraction can take place. Methods include enzymatic removal with zymolase, lyticase or helicase. These resulting structures are termed spheroplasts or protoplasts and are useful for whole cell immunological studies. For example, Bonander and colleagues in 2013 used protoplasts of *P. pastoris* cells expressing the recombinant tetraspanin CD81 (Bonander et al., 2013). In some instances, the enzymatic removal of the cell wall is not appropriate as the enzyme can interfere with extracellular loops of GPCR structures (Salazar and Asenjo, 2007). Therefore the cell wall is removed by mechanical means by either glass bead agitation or high pressure homogenisation (Darby et al., 2012). The cell wall is 100-200 nm thick and consists of glucans, mannans or mannoproteins and less so, chitins. The β -(1,6) and β -(1,3)-glucans give strength to the cell wall by forming a microfibrillar network. The mannans include α -1,6 linked inner core with α -1,2 and α -1,3 side chains. The chitin, which is a polymer of N-acetylglucosamine is mainly located

in the bud scars and only found in 2-4% of the cell wall (Zinser and Daum, 1995). Figure 1.3 shows a representation of a yeast cell wall.

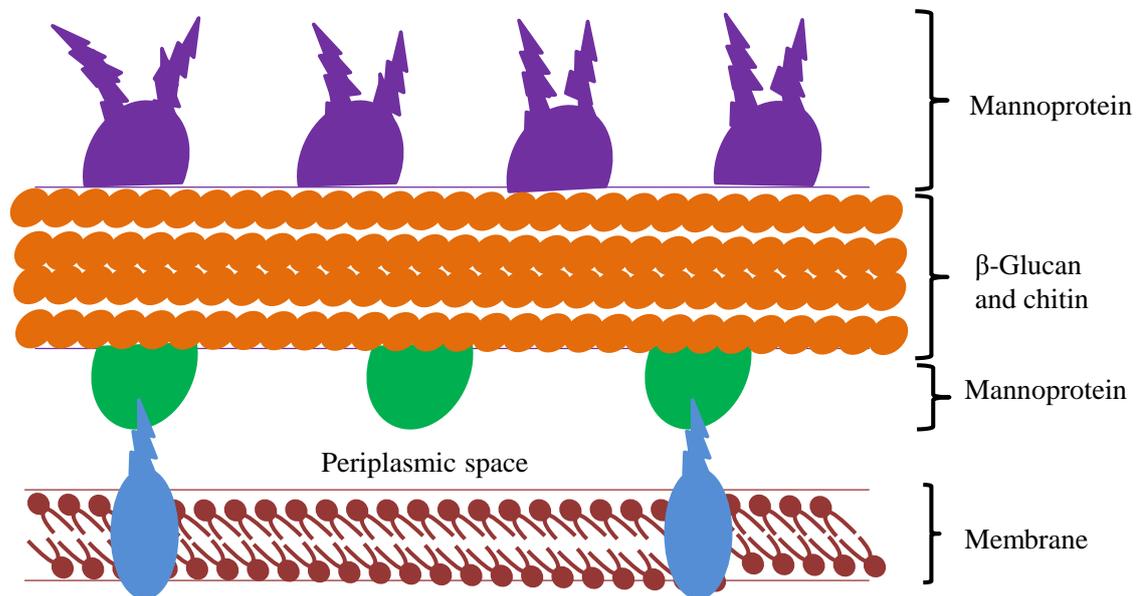


Figure 1.3 Yeast cell wall. Picture shows a representation of a yeast cell wall with the glucan, mannan or mannoproteins and chitin constituents. Image also shows the periplasmic space and plasma membrane. Adapted from Zinser & Daum, 1995.

S. cerevisiae has a respiro-fermentative metabolism if in the presence of a fermentable carbon source such as glucose (ethanol is the by-product). Once the glucose has been consumed, a diauxic shift occurs where the respiro-fermentative metabolism switches to respiration and the cells start to consume the ethanol (Al-mhanna, 2010, Werner-Washburne et al., 1996). *S. cerevisiae* is a Crabtree positive yeast (De Deken, 1966, Díaz-Ruiz et al., 2008) since it can ferment glucose in the presence of oxygen. *P. pastoris* is Crabtree negative and is strictly aerobic, utilising glucose or glycerol as a carbon source, which is metabolised via the glycolytic pathway; glycerol is phosphorylated by a cytosolic glycerol kinase to 3-phosphoglycerol. This is then oxidised to dihydroxyacetone by FAD-dependent glycerol phosphate ubiquitone oxidoreductase and is then used in pyruvate synthesis and gluconeogenesis. In recombinant protein production, glycerol is typically chosen for biomass generation since *P. pastoris* has a high glycerol uptake rate (Cereghino and Cregg, 2000).

In addition to glycerol and glucose, as mentioned earlier, *P. pastoris* has the capacity to metabolise methanol. Two *AOX* genes encode enzymes that metabolise methanol. *AOX1* accounts for more than 85% of the methanol metabolism activity. *AOX2* possesses the same specific activity as *AOX1* but has a much lower expression level and is therefore the weaker

promoter (Daly and Hearn, 2005). The methanol utilisation pathway for *P. pastoris* is complex and is represented in Figure 1.4.

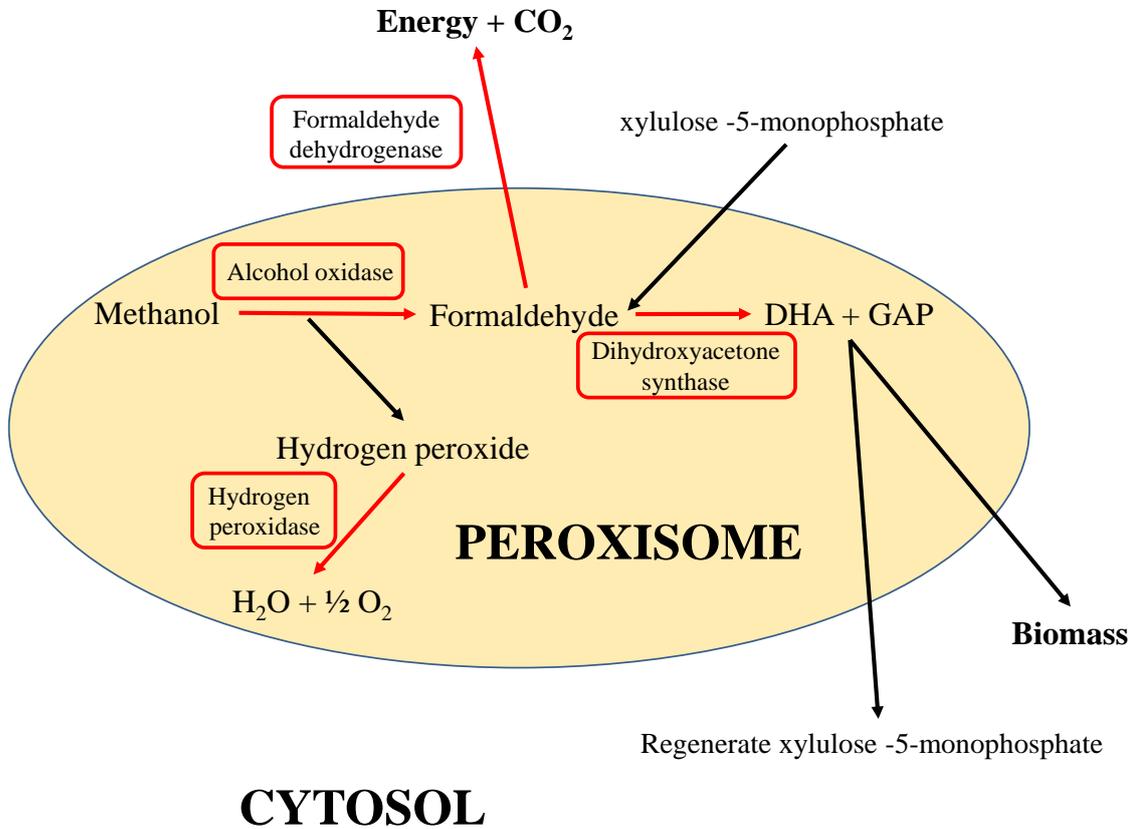


Figure 1.4 Methanol metabolism in *P. pastoris*. Diagrammatic representation of the *Pichia pastoris* peroxisome outlining the methanol metabolism pathway. Red boxes indicate name of enzyme involved in reaction. DHA is dihydroxyacetone and GAP is glyceraldehyde-3-phosphate.

P. pastoris first oxidises methanol to formaldehyde by the action of a flavoprotein, AOX, whose expression is strongly suppressed by carbon sources such as glucose, glycerol and ethanol but is induced in the presence of methanol. The reaction occurs within specialised compartments called peroxisomes and generates hydrogen peroxide. This metabolic by-product is toxic to the cell and it is subsequently catalysed by hydrogen peroxidase, into water and oxygen (Gellissen, 2000).

The formaldehyde generated from the breakdown of methanol enters a cytosolic dissimilatory pathway where it is catalysed to generate energy for the cells, as well as simultaneously activating an assimilatory pathway which results in an increase in cellular biomass. During this assimilatory pathway, residual formaldehyde reacts with xylulose-5-monophosphate and is catalysed to dihydroxyacetone and glyceraldehyde-3-phosphate by dihydroxyacetone synthase

(DAS). These products leave the peroxisome and are used to regenerate xylulose 5-monophosphate and glyceraldehyde-3-phosphate (Gellissen, 2000, Cereghino and Cregg, 2000).

1.2.2.2. Growth characteristics of *S. cerevisiae* and *P. pastoris*

A typical growth curve for *S. cerevisiae* and *P. pastoris* consists of a lag, log or exponential and a stationary phase. The lag phase is where the cells adapt to grow on the carbon source. The exponential phase is where the cells grow in a doubling manner. The stationary phase is when the cells stop increasing in number due to limiting nutrient availability and build-up of toxins and waste products (Werner-Washburne et al., 1996). From the exponential phase, the specific growth rate can be calculated. This is defined as the increase in cell mass per unit time (h^{-1}) and can be described by the equation:

$$\frac{dx}{dt} = \mu x$$

where x = cell biomass, t = time and μ = specific growth rate (h^{-1})

Upon integration, the equation becomes:

$$x_t = x_0 e^{\mu t}$$

where x_0 is the original cell biomass, x_t is the cell biomass after time (t) and e is the base of the natural logarithm. If the natural logarithm is taken for all the cell biomass values, a straight line plot should be achieved and hence the specific growth rate, μ , can be resolved with the following equation:

$$y = mx + c$$

where c = the y intercept when x is 0 and m is the slope or gradient of the line. In this instance, $m = \mu$ or the specific growth rate (Stanbury, 1988).

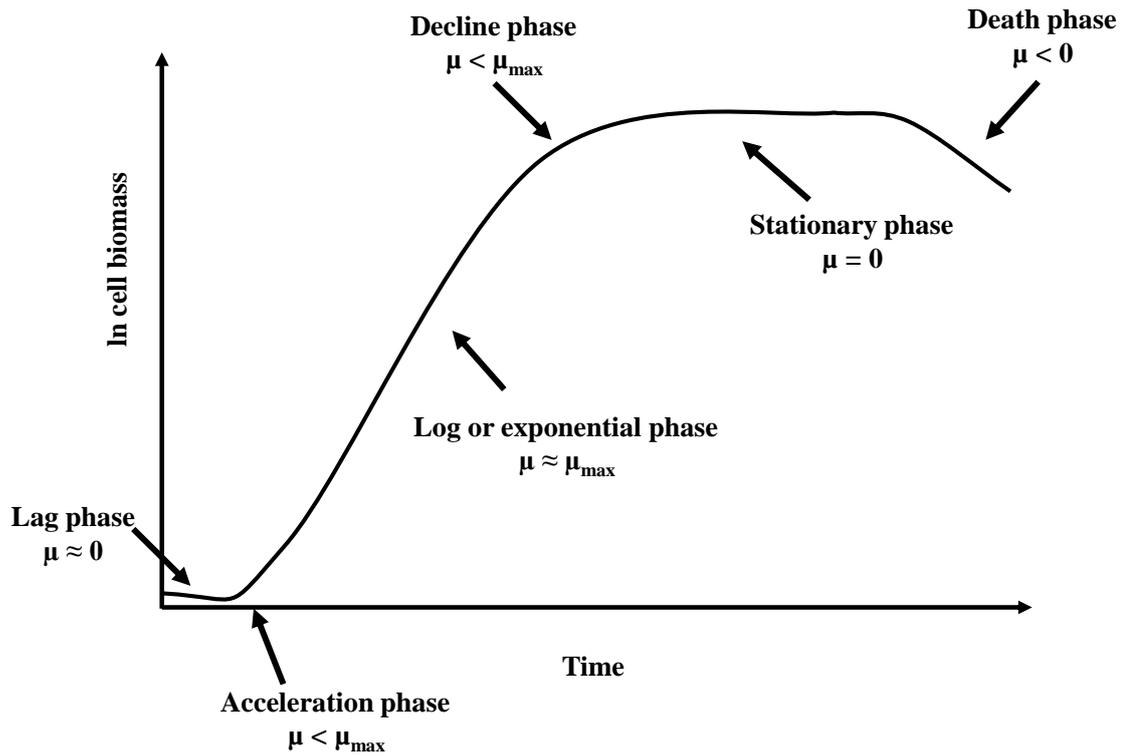


Figure 1.5 Theoretical growth curve of microbial cells. The curve shows the lag, acceleration, log, decline, stationary and death phases. For each phase the specific growth rate (μ) is predicted to be either equal or approximate to 0 or less than or approximate to the maximum growth rate (μ_{max}). (Adapted from Doran, 1995).

Figure 1.5 shows a theoretical representation of a yeast growth curve. The cells are initially in the lag phase and this is where they are adapting to the new medium environment and so there is little or no growth, giving rise to a specific growth rate which is approximately zero ($\mu \approx 0$). Next is the acceleration phase and this is where the cells start growing and therefore the specific growth rate is less than the maximum growth rate ($\mu < \mu_{max}$). The third phase is the log or the exponential phase where the cells achieve their maximum growth rate and so the specific growth rate is approximately equal to the maximum growth rate ($\mu \approx \mu_{max}$). This phase is important as it forms the linear part of the curve and hence specific growth rate data can be derived from this section. The decline phase shows that the cells start to slow down in growth due to nutrient limitation and build-up of toxins and therefore the specific growth rate is less than the maximum specific growth rate ($\mu < \mu_{max}$). In the stationary phase, growth ceases and therefore the specific growth rate is equal to zero ($\mu = 0$). The final phase is the death phase where the cells start to lyse and die and so the specific growth rate is in theory less than zero ($\mu < 0$) (Doran, 1995).

1.2.2.3. Genomic characterisation and strain selection of *S. cerevisiae* and *P. pastoris*

S. cerevisiae was the first eukaryotic organism to be sequenced in 1996 (Goffeau et al., 1996). The genome's annotations are regularly updated and can be viewed on the World Wide Web (www.yeastgenome.org). The sequence includes ~12000 kb and from those, ~5900 protein encoding genes. It was found that approximately, 140 genes were ribosomal RNA; 40 genes were small nuclear RNA molecules and 275 genes were transfer RNA genes. The sequencing project also organised the 16 chromosomes (Goffeau et al., 1996).

The *P. pastoris* genome has been fully sequenced more recently and it has been estimated that it is 9.7 Mbp in total and organised into four chromosomes and one mitochondrial genome (De Schutter et al., 2009, Mattanovich et al., 2009). Some of the main findings which came from this project were that although all the genes coding for enzymes and their promoters involved in methanol metabolism were identified, there were no common sequence motifs or promoter organisation. Many of the endogenous signal sequences for the secretory pathways of *P. pastoris* were revealed and many sequences of vacuolar and secreted proteases were reported, thereby aiding the development of protease deficient strains (De Schutter et al., 2009, Mattanovich et al., 2009).

Both *S. cerevisiae* and *P. pastoris* genomes have sequences for 2 homologs of low affinity sugar transporters genes but *P. pastoris* has 4 H⁺/glycerol transporters present and *S. cerevisiae* do not (De Schutter et al., 2009). It was also found that where the 5S rRNA gene was localised to the rDNA locus in the *S. cerevisiae* genome, 21 copies of the same gene were located across the entire length of all the chromosomes for *P. pastoris* (De Schutter et al., 2009). The genomic sequencing of *P. pastoris* revealed a multitude of endogenous signal sequences. This mitigates the use of the alpha-mating factor signal sequence from *S. cerevisiae* which is frequently employed to induce Sec61p-mediated translocation of protein in the endoplasmic reticulum of *P. pastoris* during recombinant protein production (Prabha et al., 2009). The sequencing also confirmed that many highly immunogenic terminal α -1 3-mannosyl glycotypes were present for *S. cerevisiae* but were not detected at all in the *P. pastoris* genome (Bretthauer and Castellino, 1999, De Schutter et al., 2009).

Consideration should be given to the *S. cerevisiae* and *P. pastoris* strains selected for recombinant protein production. *P. pastoris* has three major strain backgrounds based on their methanol metabolising phenotypes and subsequent strain types with additional phenotypes are based on these three. They are the wild-type methanol utilisation plus phenotype (Mut⁺); a strain with deleted a *AOX1* gene, so only the *AOX2* gene is active, which is termed a methanol utilisation slow phenotype (Mut^S); and a third strain with a deletion of both the *AOX1* and *AOX2* genes resulting in a methanol utilisation minus phenotype (Mut⁻) (Daly and Hearn, 2005). Table

1.1 lists the common *P. pastoris* strains used in recombinant protein production and their phenotypes.

Table 1.1 *P. pastoris* strains commonly used in recombinant protein production and their phenotypes. In this thesis, the X33 and SMD1163 strain were used. Adapted from (Bora, 2012).

Strain	Phenotype
X-33	Mut ⁺ His ⁺
GS115	Mut ⁺ His ⁻
KM71	Mut ^s His ⁻ Arg ⁺
SMD1163	Mut ⁺ His ⁻ , proteinases A, B and carboxypeptidase Y deficient
SMD1165	Mut ⁺ His ⁻ , proteinase B deficient
SMD1168	Mut ⁺ His ⁻ , proteinase A and carboxypeptidase Y deficient; partial reduction in proteinase B activity
MC100-3	Mut ⁻ His ⁻

For *S. cerevisiae* a plethora of strains are available for use in recombinant protein production. A set of single, non-essential gene deletion strains are available from EUROSCARF as well as a collection of tetracycline-regulated essential genes (Open Biosystems). *S. cerevisiae* yTHCBMS1 is a strain from the tetracycline-regulated gene strain and is used in Chapter 3 since it is therefore considered a high-yielding strain. Bonander and colleagues (Bonander et al., 2005) showed that 39 host cell genes' expression levels were significantly altered when Fps1, a glycerol facilitator, was produced under high yielding conditions compared to standard growth conditions. This included elevated levels of *BMS1* (involved in ribosome biogenesis). Further studies showed that tuning the amount of *BMS1* transcript levels by varying doxycycline amounts had an effect on the yields of other functional proteins such as hA_{2A}R and also soluble GFP (Bonander et al., 2009). Another strain that was considered high-yielding and is used in Chapter 3 is the *S. cerevisiae* TM6* strain developed by Otterstedt and colleagues (Otterstedt et al., 2004). This strain was produced by integrating a gene encoding for a chimeric hexose transporter which mediates decreased sugar uptake into the genome of hexose transporter null yeast strain and is called *TM6** (Otterstedt et al., 2004). It is fully respiratory and does not switch to fermentative metabolism even at high glucose concentrations. This leads to a higher biomass yield thereby increasing production of recombinant proteins. Commercially, this strain is available as AlcoFree™ (Cereduce, Sweden) and is used in heterologous protein production,

fine chemical productions and alcohol-free production of beers and wines (Henricsson et al., 2005).

1.2.2.4. *S. cerevisiae* and *P. pastoris* as cell factories for recombinant protein production

Both *S. cerevisiae* and *P. pastoris* are excellent host systems for recombinant protein production as they are cheap to cultivate, can grow to high biomass yields and can perform post-translational modifications (Darby et al., 2012). Such modifications include disulphide bond formation, glycosylation (both O-linked and N-linked), phosphorylation, acetylation and acylation of the recombinant protein and would not be possible in prokaryotic systems such as *E. coli* (Bonander and Bill, 2012). However, it is important to consider that the type of glycosylation carried out by *S. cerevisiae* differs from that in mammals. For example, O-linked oligosaccharides carry only mannose moieties rather than sialylated O-linked chains (Bretthauer and Castellino, 1999). *S. cerevisiae* is known to hyperglycosylate N-linked sites which potentially modify binding sites leading to altered immunogenic responses in therapeutic applications (Gerngross, 2004). However there is an added advantage when glycosylation is carried out in the *P. pastoris* host, where the oligosaccharides are shorter in length (Bretthauer and Castellino, 1999). Furthermore, a strain has been developed that can produce terminally sialylated humanised proteins (Li et al., 2006).

Table 1.2 summarises the differences between *S. cerevisiae* and *P. pastoris* as host systems for recombinant protein production and Table 1.3 shows some examples of recombinant protein production for clinically relevant proteins that target specific diseases in both *S. cerevisiae* and *P. pastoris*.

Table 1.2 Comparison of some main characteristics of *S. cerevisiae* and *P. pastoris*.

Characteristic	<i>S. cerevisiae</i>	<i>P. pastoris</i>
Industrial application	Yes	Yes
Recombinant protein products	Yes	Yes
Food grade (GRAS)	Yes	No
Annotated genome	Yes	Yes
Metabolism	Respiro-fermentative	Aerobic
Secretion efficiency	Low	High
Hyperglycosylation	Yes	No

Table 1.3 Examples of clinically relevant recombinant proteins produced from *S. cerevisiae* or *P. pastoris* host cells. Table section highlighted in grey colour show recombinant proteins produced as biopharmaceuticals in either *S. cerevisiae* or *P. pastoris*. Table section in white show recombinant proteins as drug targets produced in either *S. cerevisiae* or *P. pastoris*.

<i>S. cerevisiae</i>	Disease targeted	<i>P. pastoris</i>	Disease targeted	
Insulin (Lindholm et al., 2002)	Diabetes	Angiostatin (Chen et al., 2010)	Cancer treatments	Biopharmaceuticals
Interferon- α -2a, hepatitis B surface antigen (Ryff, 1993)	Hepatitis B	Anti-HBs Fab fragment (Ning et al., 2003)	Liver diseases	
Human papilloma virus vaccine (Siddiqui and Perry, 2006)	Papilloma virus	Granulocyte-macrophage colony-stimulating factor (hGM-CSF) (Pal et al., 2006)	Non-Hodgkin's lymphoma, HIV, Crohn's disease	
Human α (2)-adrenergic receptor subtype 2C (Blaxall et al., 1991)	Hypertension, vasoconstriction	Tetanus toxin fragment C (Clare et al., 1991)	Tetanus	
M5 muscarinic acetylcholine receptor (Huang et al., 1993)	Central and peripheral nervous system diseases e.g. Parkinson's	Human μ -opioid receptor (Sarramegna et al., 2002)	Receptor for analgesia	
Rat adenosine 2_A receptor (Price et al., 1996)	Heart disease, inflammation, cancer, epilepsy	Human dopamine D2S receptor (de Jong et al., 2004)	Depression, psychosis, Parkinson's	
Human adenosine 2_A receptor (hA 2_A R) (Ferndahl et al., 2010)	Heart disease, inflammation, cancer, epilepsy	Human muscarinic acetylcholine receptor M2 sub-type (CHRM2) (Asada et al., 2011)	Heart disease	
SERCA1a (Jidenko et al., 2005)	Malaria	Human histamine H $_1$ receptor complex with doxepin (Shimamura et al., 2011)	Inflammation, allergies	

1.2.3. Molecular biological considerations

When a target protein is chosen for recombinant production, its DNA sequence is usually amplified via PCR from genomic DNA or cDNA or the gene can be synthesised. It is then cloned into a suitable expression vector within an expression cassette which contains a yeast promoter and a termination sequence (Darby, 2010). The vector is a vehicle by which the target DNA can be introduced into the yeast strain of choice by transformation (Mattanovich, 2012). Transformation can be carried out by spheroplast preparation (Burgers and Percival, 1987); lithium acetate preparation (Gietz and Woods, 2002) or electroporation (Sanchez et al., 1993).

The target DNA sequence is usually designed with sequences to encode for purification tags and signal sequences and the complete sequence is termed as a DNA construct. When designing constructs for target proteins, it is valuable to design sequences for several tags which aid in the purification and detection of the protein. The most common tags used for these types of proteins are poly-histidine, FLAG, haem-agglutinin, Biotin and c-myc tags (Terpe, 2003). Larger fusion proteins are increasingly used for recombinant membrane protein production, and these include green fluorescent protein (GFP) (Drew et al., 2005), maltose binding protein (MBP)(Duplay et al., 1984, di Guana et al., 1988), thioredoxin (TRX)(LaVallie et al., 1993) or glutathione-S-transferase (GST)(Taylor et al., 1994, Smith and Johnson, 1988). They are used to improve stability as certain proteases such as the *E.coli* FtsH complex can degrade proteins from their free amino (N-) or carboxy (C-) ends (Wagner et al., 2006, Mancina et al., 2004). Tucker and Grisshammer tested various combinations of fusion proteins and tags in an attempt to improve recombinant MBP-neurotensin (NTR) receptor production and they found that a fusion with TRX significantly improved the yields of the receptor (Tucker and Grisshammer, 1996). Other studies found that a combination of an N- terminal MBP and a poly-histidine tagged C- terminal TRX fusion protein was successfully used to produce the cannabinoid 2 receptor and the adenosine receptor in *E. coli* (Yeliseev et al., 2005, Weiß and Grisshammer, 2002).

GFP is 24kD, naturally occurring secreted protein that was first isolated from Pacific jellyfish where it interacts with another protein called aequorin to produce green light. Since its discovery, it has been extensively used for recombinant protein studies to assist in localisation of target proteins. There are many variations of the recombinant version of GFP, but the wild type version has a major excitation peak at 395 nm and an emission peak at 509 nm (Prasher, 1995).

1.2.3.1. Vector choices and promoters for *S. cerevisiae*

There are a number of expression vectors available and for *S. cerevisiae* that may be episomal or integrative (Darby, 2010). Promoter sequences and selection markers are usually included in

the vector. The promoters can be categorised as inducible or constitutive. Inducible promoters prompt target protein expression when an inducer is added to the culture medium, thereby controlling the recombinant protein production. Constitutive promoters allow continuous expression of the target protein and therefore offer less control. There are three inducible promoters for *S. cerevisiae*; *GAL1*, *ADH2* and *CUP1* (Schneider and Guarente, 1991). The main constitutive promoter available for *S. cerevisiae* is the triose-phosphate isomerase (*TPI*) promoter (Zhang et al., 1996).

1.2.3.2. Vector choices and promoters for *P. pastoris*

P. pastoris vectors tend to be integrative. Integrative vectors allow the target DNA to homologously recombine into the host genome and can be quite stable. However, most transformants often exhibit heterogeneous expression levels and therefore extensive colony screening is required (Darby et al., 2012). The types of promoters available for *P. pastoris* vectors like *S. cerevisiae* vectors can be either inducible or constitutive. *AOX1* is a tightly regulated, inducible promoter used in *P. pastoris* host systems and is induced in the presence of methanol. Although this is the most common promoter to be exploited, it does come with disadvantages, such as the need to use toxic and flammable methanol. Therefore alternative promoters have been investigated (Cos et al., 2006a). These include constitutive expression via the glyceraldehyde 3-phosphate dehydrogenase (*GAP*) promoter (Waterham et al., 1997, Kim, 2009) whilst strong induction via the formaldehyde dehydrogenase (*FLDI*) promoter has also been reported in the presence of methylamine as well as methanol (Resina et al., 2005) Figure 1.6 shows a typical *P. pastoris* expression vector system from Life Technologies Corporation. It shows an example of the pPICZ series of vectors for *P. pastoris* systems.



Figure 1.6 pPICZ α expression vector series map from Life Technologies Corporation. Showing AOX1 promoter, terminator and a choice of multiple cloning sites.

This vector includes an α -factor signal sequence. It is thought that the presence of this sequence may improve the chances of a membrane protein being inserted into the yeast membrane via the yeast secretory pathway. This signal peptide is derived from the *S. cerevisiae* mating factor $\alpha 1$ (MF $\alpha 1$) (Weiss et al., 1998). This peptide contains a 'pre' and 'pro' protein signal which are 22 and 61 residues, respectively. In the endoplasmic reticulum, the pre-protein signal is cleaved by peptidases and then the pro-protein signal is cleaved by the cells' Kex proteases in the Golgi apparatus. It is important to note that if these signal sequences are not cleaved, the activity of the membrane protein may be compromised (Zhang et al., 2007).

1.3. Recombinant protein production in yeast: from clone to culture

Once the molecular biology of the target protein and successful transformation has occurred into the yeast cell, a change in discipline is required where microbiology and bioengineering are given thought in order to produce the highest yield and quality of the target protein from yeast cultures.

1.3.1. Screening for high yielding clones

Traditionally, colony screening and screening for the best culture environmental conditions for recombinant protein production are carried out in shake flasks or deep well plates. This is a cheap and quick method to ascertain optimal growth and production conditions for the target protein in question. However, when it is time to scale-up the growth to larger bioreactors, the conditions do not necessarily translate. This is due to the lack of control and monitoring of environmental conditions available when grown in shake flasks (Schapper et al., 2009).

Therefore the development of culture vessels with the capability of control and monitoring at a small scale and high throughput level is an attractive option. The Micro-24 microreactor (Pall Corporation) is a technology which encompasses this by enabling temperature, pH and dissolved oxygen (DO) control and monitoring for 24 5-10 mL bioreactors in a plate well format. Figure 1.7 shows the Micro-24 microreactor set-up at Aston University.



Figure 1.7 Micro-24 microreactor at Aston University. Photograph on the left shows the 24 well plate of individually controlled mini bioreactors. Photograph on the right shows the plate sitting inside the instrument.

The Micro-24 wells can be controlled individually for temperature, pH and DO and also be run in parallel. In the bottom of each well, DO, pH optical sensors, a sparging port and thermal conduction pads are present to provide control and monitoring. The pH and DO of the culture are controlled by a set point by sparging gas through the medium, in this instance, O₂ for DO and CO₂ or ammonium hydroxide gas for pH. Figure 1.8 shows the online monitoring capability of all the individual wells for temperature, pH and DO.

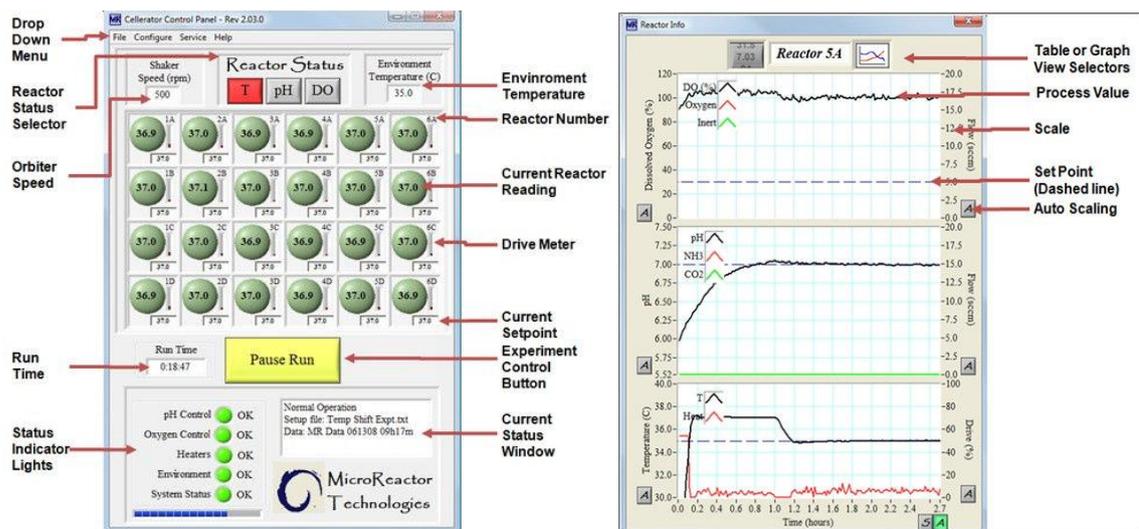


Figure 1.8 Micro-24 microreactor software screen shot. Screen-shots show the real-time collection of the data for DO in this instance (left side of screen). On the right side, real-time plots are constructed for a specific well for temperature, pH and DO.

With access to such technology, it is feasible to more systematically establish optimal recombinant protein production conditions or parameters such as temperature, pH and DO for a host cell. Chapter 3 uses this technology to improve hA_{2A}R yields in *S. cerevisiae* strains.

1.3.2. Culture vessels

Once a high-yielding clone has been identified, recombinant proteins produced in yeast can be generated in shake flasks or bioreactors (sometimes referred to as fermenters). Shake flasks are a laboratory standard consumable and small scale screening of recombinant protein production tends to begin using these vessels. However, they do not allow control and monitoring of culture parameters such as a temperature, pH, and DO of the culture medium but bioreactors do. Bioreactors are culture vessels of varying size and allow control of these parameters via control of input gases and air supplies, acid, base and temperature which are monitored via probes and sensors. Set-points of a desired input condition are entered and the bioreactor software and controllers maintain the set-points. The difference between bench-top and pilot-scale bioreactors is minimal and the main difference is the vessel capacity (Baumann et al., 2010, Schmidt, 2005, Abad et al., 2010). Bench-top bioreactors typically have a culture capacity between 1-10 L and pilot-scale bioreactors have a culture capacity between 10-100 L. An example of a bench-top bioreactor set-up at Aston University is shown in Figure 1.9 and an example of a pilot-scale bioreactor at AstraZeneca Ltd, is shown in Figure 1.10.

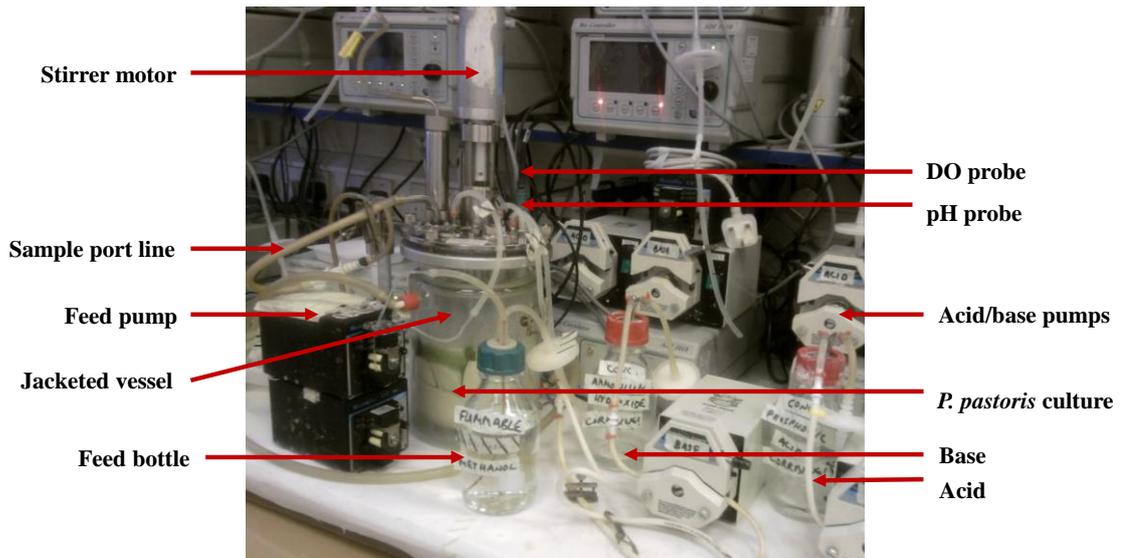


Figure 1.9 Bench-top 2 L bioreactor set-up at Aston University. Photograph shows 2 L vessel with growing *P. pastoris* culture and acid, base and methanol feed pumps.

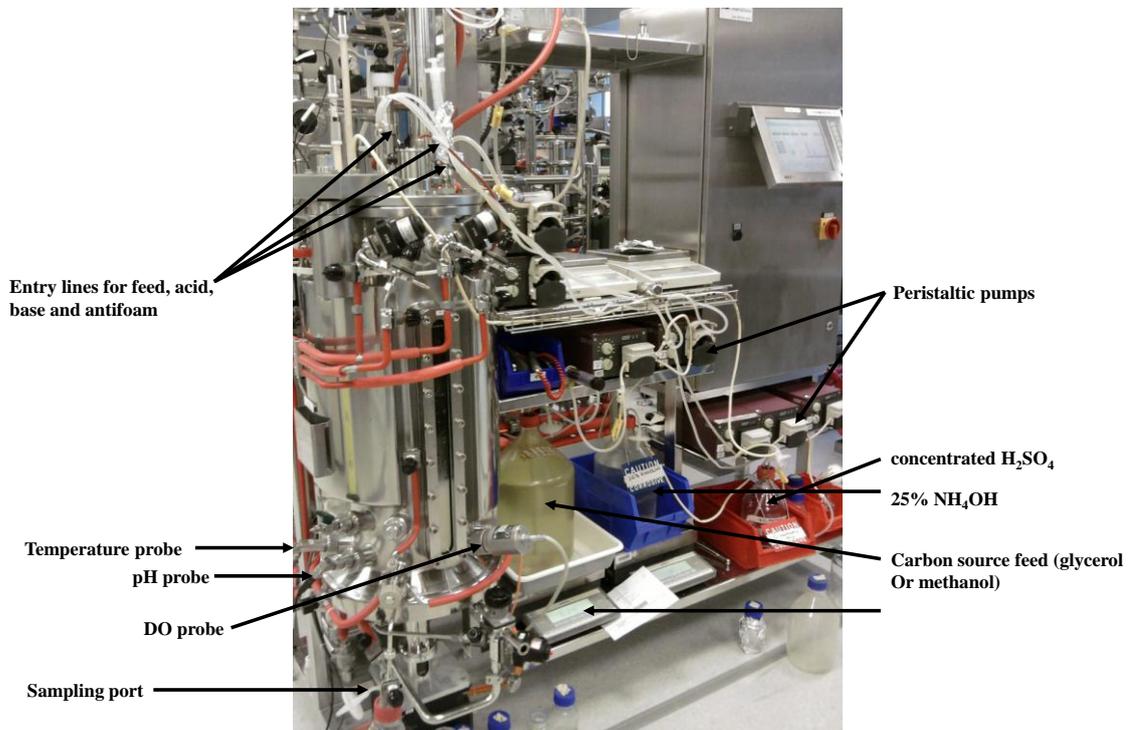


Figure 1.10 Pilot-scale 35 L bioreactor set-up at AstraZeneca Ltd. Photograph shows 35 L vessel with growing *P. pastoris* culture. Several pumps are employed for acid, base, feed and anti-foam addition via the use of controller software.

1.3.3. Bioprocessing parameters

The use of bioreactors is vital when developing the production of a recombinant protein. They allow the precise control of temperature, pH, aeration and addition of carbon source. This is important to achieve in order to maximise recombinant protein yields but also to maintain the

integrity of the protein. Such requirements are essential in industrial, pharmaceutical and commercial settings (Macauley-Patrick et al., 2005).

1.3.3.1. Temperature

It is important to maintain optimal temperature for growth of any micro-organism. For yeast, this is 30°C. However, the optimal temperature for growth is not necessarily the same as that for recombinant protein production (Cos et al., 2006a). For example for the recombinant production of hA_{2A}R, the temperature is usually lowered to 22°C at the induction phase of *P. pastoris* cultivation (Singh et al., 2008, Singh et al., 2010). Contrastingly, the optimal recombinant production of the human tetraspanin, CD81 in *P. pastoris* was found to be at growth temperature of 30°C (Bonander et al., 2013).

1.3.3.2. pH

pH is an important factor that must be considered when growing yeast as an optimal pH will aid in high biomass and recombinant protein production (Routledge, 2012). A desired set-point is often stated, so that the pH is maintained. Often the pH will change as the yeast grows (usually metabolites are released creating an acidic environment during yeast growth); therefore it is important for a system to maintain it. Optimal pH will give high yields for growth and recombinant protein expression (Çalık et al., 2010). For secreted proteins, the pH will also have an effect on their stability via their iso-electric point (pI) (Schmidt, 2005).

1.3.3.3. Dissolved oxygen concentration (DO)

Yeast require oxygen as they grow and this requirement can be met in bioreactors by supplying external air and oxygen directly into the culture medium. The amount of dissolved oxygen (DO) is a measure of the oxygen available to the growing culture (Visser et al., 1990). The air flow rate is regulated via a mass flow controller and air bubbles enter the vessel through a sparger at the vessel base. Impellers then disperse the bubbles and the oxygen dissolves in the medium as the bubbles travel up the surface of the culture medium. DO in the medium can be increased in three ways; by increasing the stirrer speed, by increasing the airflow and by increasing the oxygen-enriched airflow. A 'cascade' mechanism can be employed where each method can be activated in turn until the DO set-point is reached and maintained (Visser et al., 1990, Schmidt, 2005).

1.3.3.4. Additives

Chemical additives can be added to a defined culture medium to further assist in maximising recombinant protein yields. Murata and colleagues (Murata et al., 2003) showed that dimethyl sulphoxide (DMSO) can change cell membranes physically and therefore have a downstream

effect on intracellular biochemical pathways. This was further investigated to see if the presence of DMSO in yeast culture medium had an effect on GPCR yields. It was reported that the binding activity of more than half of the GPCRs in an expression screen were increased in the presence of DMSO (Andre et al., 2006, Lundstrom et al., 2006).

1.3.3.5. Antifoams

Antifoams are almost always required when using bioreactors since continuous stirring will cause the formation of bubbles and foam in the culture and this can have a detrimental effect on recombinant protein production. Routledge and colleagues carried out an extensive study on the effects of antifoams for bioreactor and shake flask cultures and showed that specific antifoam agents could increase GFP yields in *P. pastoris* cultures (Routledge and Bill, 2012, Routledge et al., 2011).

1.3.4. *P. pastoris* bioreactor cultivations

P. pastoris is often cultured in media with non-limiting amounts of a repressing carbon source such as glycerol to generate biomass, followed by an induction period with limiting amounts of methanol. The distinct feeding phases of *P. pastoris* tend to follow the specific growth rate trends for the typical growth described. During the first or batch phase, cells grow at their maximum growth rate (μ_{\max}) (section 1.2.2.2) until the initial carbon source, typically glycerol, has been depleted. In a subsequent fed-batch phase, the same carbon source is fed continuously with the objective of yielding high pre-induction biomass; during this phase, growth is nutrient limited and a constant specific growth rate, lower than μ_{\max} , is achieved. A transition phase, when the glycerol feed is stopped and the cells are monitored for glycerol depletion, allows the cells to adapt to low concentrations of inducer (typically methanol); in some cases temperature changes are also applied to facilitate induction at a temperature optimised for a given target protein. Finally, in the induction phase, methanol is added in a controlled manner to induce *AOX1*-driven recombinant protein production (Cereghino and Cregg, 2000) (Minning et al., 2001). Figure 1.11 illustrates this cultivation process for *P. pastoris*.

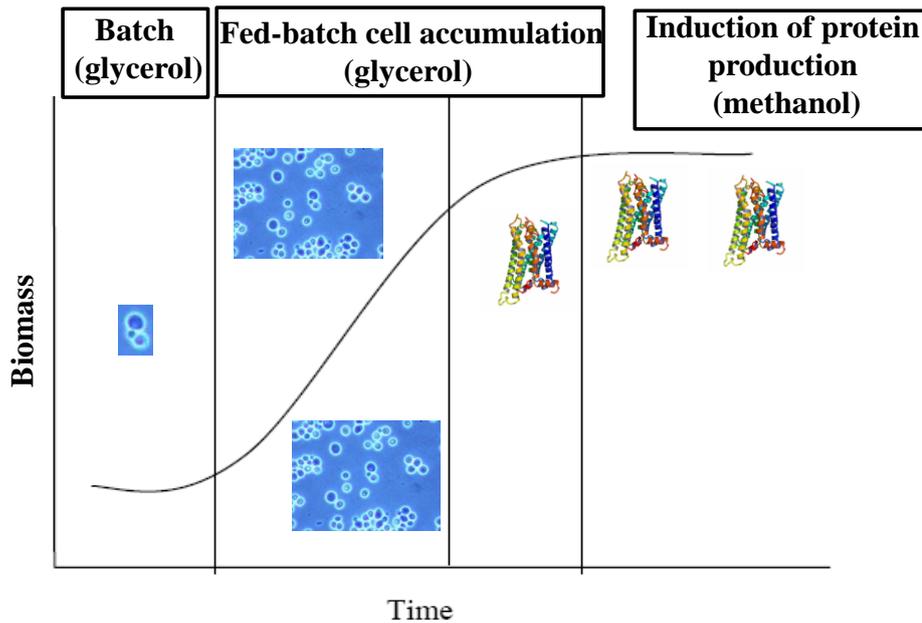


Figure 1.11 Typical *P. pastoris* bioreactor cultivation for recombinant protein production. Theoretical growth curve shows the batch phase where the cells adapt to the glycerol carbon source (lag phase). In the fed-batch phase, glycerol is continuously fed in order to increase the cell biomass. The induction phase is carried out with controlled addition of methanol (once all the glycerol has been consumed) to induce target protein production.

Although the simple addition of methanol to a *P. pastoris* culture induces protein production, careful consideration must be given to its addition rate and duration as methanol can be toxic to the cells if accumulation occurs within the culture (Guarna et al., 1997). Conversely insufficient methanol will result in sub-optimal protein yields; therefore it is imperative to strike the optimal balance (Thorpe et al., 1999). The general “rule-of-thumb” is that the concentration of methanol within the culture should be maintained below 5 g L^{-1} to avoid cyto-toxicity (Guarna et al., 1997). Although it is common practice to induce expression in *P. pastoris* cultures solely using methanol, some reports have demonstrated increased protein yields by inducing with a mixed feed of glycerol and methanol (d’Anjou and Daugulis, 2000, Jungo et al., 2007) and sorbitol and methanol (Cos et al., 2006b) (Holmes et al., 2009). It is possible that both these strategies decrease the potential toxicity to the yeast cells caused by methanol overload as well as permitting enhanced biomass generation in the induction phase compared with a solely methanol fed system.

Since the specific growth rate of *P. pastoris* during the induction phase can influence both recombinant protein and cell biomass yields, a theoretical constant specific growth rate (μ_{set}) can be applied to a *P. pastoris* cultivation, in a “feed forward strategy”, to control methanol uptake (Potvin et al., 2012).

The equation upon which these calculations can be made is:

$$F(t) = F_0 e^{\mu_{\text{set}} t}$$

where $F(t)$ is the feed rate (g h^{-1}) at time, t (h); F_0 is the initial feed rate and μ_{set} is the desired theoretical constant specific growth rate (h^{-1}).

For example, (Celik et al., 2009) reported that when producing recombinant human erythropoietin (hEPO) in *P. pastoris*, higher protein yields were achieved at $\mu_{\text{set}} = 0.03 \text{ h}^{-1}$ than $\mu_{\text{set}} = 0.02 \text{ h}^{-1}$, while the cell biomass yield was lower. This strategy, which employed a mixed feed of methanol and sorbitol, gave similar results to another study by the same team on the recombinant production of human growth hormone (Celik et al., 2010). Notably, hEPO appears to have been produced prior to the onset of the methanol feed in these studies (Celik et al., 2009). Three other studies also demonstrated the effect of specific induction phase growth rates on recombinant soluble protein yields (Kobayashi et al., 2000, Zhang et al., 2000). While this suggests that the yield of a given soluble target protein may increase as a function of μ_{set} , the reasons for this are not fully understood. It could be, as previously reported by Holmes and colleagues, that there is a requirement to match the methanol feeding strategy with the metabolic capacity of the cells, as exemplified by the recombinant production of soluble green fluorescent protein (GFP); (Holmes et al., 2009).

Few, if any, detailed studies have been published on the influence of the methanol feed profile during the induction phase for membrane proteins. This contrasts with the numerous studies on the influence of parameters such as the temperature and pH of the culture, the amount of dissolved oxygen (DO) in the culture medium and, specifically, the addition of chemical additives and ligands for optimal GPCR production in the pre-induction and induction phases of *P. pastoris* cultivations (Andre et al., 2006, Cos et al., 2006b, Schmidt, 2005, Lundstrom et al., 2006, Singh et al., 2010). Notably, Singh and colleagues observed active hA_{2A}R in bioreactor cultures prior to the methanol feed (Singh et al., 2008), while this was not apparent in shake flasks (Singh et al., 2012). These findings are especially noteworthy as glycerol is a known repressor of the *AOX1* promoter: glycerol, glucose, ethanol and acetate have all been shown to support growth of *P. pastoris* cells without inducing the *AOX1* promoter (Inan and Meagher, 2001). An example of this, is work by Hellwig and colleagues in 2001 who demonstrated that glycerol in the culture medium inhibited production of a recombinant single-chain antibody in mixed feed bioreactor cultures. They also noted that ethanol and acetate accumulated (Hellwig et al., 2001).

In Chapter 4, the impact of set growth rates (μ_{set}) on hA_{2A}R production in *P. pastoris* cultivations are investigated during the induction phase.

1.3.5. The application of Design of Experiments (DoE) for recombinant protein production

Statistical Design of Experiments (DoE) is a concept that is frequently applied in process optimisation in the biotechnology and other industries. It provides an effective way of investigating the impact of multiple conditions whilst reducing the overall number of experiments and not compromising the quality of the data (Mandenius and Brundin, 2008). DoE was first proposed by Sir Ronald A. Fisher in 1935 as an alternative to the changing of ‘one factor at a time’ (OFAT) approach, which is time consuming and costly. He based his approach on the statistical method ‘analysis of variance’ (ANOVA). This concept was further developed in the 1950s by Genichi Taguchi to improve the quality of manufactured goods and now the Taguchi DoE method is used (along with other developed DoE methods) in biotechnology. In particular, DoEs are viewed as integral components of industrial bioprocess development and are recognised as valid methods by the US Food and Drug Administration (Bora et al., 2012). Information on the relationship between the parameter temperature and the response e.g. specific binding activity of a receptor, is derived in the form of an equation. This means that it is not necessary to carry out all the possible experimental combinations of the parameters since the equation will predict outcomes of a response if the statistics of the DoE model are robust. For example, Holmes and colleagues investigated increasing GFP yields in *P. pastoris* by exploring temperature, pH and DO as input parameters and found that with the use of a DoE only 13 experimental combinations were required out of a possible 27 and furthermore since the model was statistically robust, the equation generated was able to predict accurate responses (Holmes et al., 2009).

1.3.5.1. Factors, levels and responses for a DoE set-up

A typical DoE set-up includes the input factors (also referred to as conditions or parameters interchangeably) to be tested usually at a number of levels with a number of replicates specified in a design matrix. Factors are usually variables that have defined set-points, e.g. temperature, pH, DO or the components of a growth medium. Input factors may also be an ‘attribute’ e.g. the presence or absence of a medium component at a specific level. Other factors referred to as ‘noise factors’ should be considered in the DoE as they distort the data. Their effects can be minimised by applying ‘blocking’ or ‘randomisation’ in a DoE (Bora et al., 2012, Isar et al., 2007).

The simplest DoE designs are ones where the factors are studied at 2 levels only, a high and low level. They are known as 2^k designs and the levels can be coded as +1 for the high level and -1 for the low level in a design matrix. The levels can vary in number and in general more than 2 levels will enable detection of non-linear relationships (Burdick et al., 2005). However, the

levels should be considered carefully as certain levels may not be biologically practical to be performed such as very low pH levels or very high temperatures.

A response in a recombinant protein production experiment is typically protein yield, protein activity, culture density or biomass and tends to be a continuous variable. Once the response has been generated by running the specific experimental combinations of input factors as defined by the matrix, statistical analysis is used to fit this response to a model which is either linear or non-linear. The effect of each input factor on this response is determined quantitatively and the amount of error in the model is calculated to see if there is a significant lack of fit (Bora et al., 2012).

1.3.5.2. DoE process screening

The DoE steps are usually screening, characterisation and optimisation for recombinant protein production (Bora et al., 2012). Screening designs are used to reduce the initial number of input factors that are to be tested, for example if there are between 4-12 or more. Alternatively, a screen could test 3-5 input factors for a more detailed study. Usually, full factorial designs are run at the screening stage (Montgomery, 2006). They take into account all the input factors in the experiments and all the possible combinations associated with them. Because of this they tend to require many experimental runs; however the results retrieved are valuable as they can give information regarding any main effects or interactions between factors (Bora et al., 2012, Mandenius and Brundin, 2008). Figure 1.12 illustrates the concept of a full factorial design when compared to an OFAT approach. More of the response surface is covered by carrying out a full factorial design compared to the OFAT approach.



Figure 1.12 One Factor At a Time (OFAT) versus full factorial design. Graph on left shows X_1 variable versus X_2 variable. If either X_1 or X_2 are varied and the other is kept constant, 5 separate experiments would need to be performed. This is called achieving the 'quasi-optimum' and the correct optimum is never reached. The graph on the right however, shows that if

simultaneous variations were carried out for both X_1 and X_2 the true optimum may be found by software analysis. Taken from Mandenius & Brundin, 2008.

Fractional factorials take a 'fraction' of the full factorial experiments in order to reduce the number of experiments to be run, however this impacts the design power of the model as certain experimental runs will be missing from the design matrix (Figure 1.13); (Mandenius and Brundin, 2008); (Chen et al., 1993).



Figure 1.13 Fractional factorial design. 3 variables X_1 , X_2 and X_3 investigated at only 4 points in the design space, a high and low level for each variable. Taken from Mandenius & Brundin, 2008.

Other screening designs are available such as Plackett-Burman designs (Vindevoegel and Sandra, 1991) (Montgomery, 2006) where only 2 levels are investigated for each factor (Plackett-Burman) and the number of experimental runs are reduced via a computer based method (d-optimal) (de Aguiar et al., 1995).

1.3.5.3. DoE process characterisation and optimisation

The process characterisation goal is to identify and quantify the influence of the key factors (established from the screening designs) in order to improve the bioprocess by predicting an optimal response under a range of operating conditions via an equation. Process optimisation involves zooming in on a particular portion of the design space or by exploring any non-linear behaviour (e.g. quadratic behaviour) that was observed in the previous stages. This is achieved by carrying out a response surface methodology (RSM) (Montgomery and Myers, 1995). Box-Behnken, Composite Face Centred (CCF) or a Central Composite Circumscribed (CCC) are different examples of RSM designs (Mandenius and Brundin, 2008). Figure 1.14 illustrates a CCF and CCC design where a 3 factor experiment is shown.



Figure 1.14 CCF design (left) and CCC design (right). For 3 factors including triplicate points in the centre of the design space (red spheres), corner points on the cube are black spheres and the face points are green spheres. Taken from Mandenius & Brundin, 2008.

The design points for the CCC design appear to spread beyond the confines of the design space. Because of this, CCC may be a better design than CCF, as it covers more volume. Additionally, the CCC design covers 5 levels for each factor and hence will enable the investigation of the cubic response behaviour even more than the CCF design. Once the CCF or CCC experiments are run, a contour plot or a response surface plot may be generated where the optimum can be clearly visualised (Mandenius and Brundin, 2008).

RSM allow a close examination of each factor and its interactions and what relationship exists between them. Data from one round of results in a model can provide information for an improved design in subsequent rounds. Table 1.4 gives examples of DoE used in bioprocess applications including recombinant protein production experiments.

Table 1.4 Examples of DoE in bioprocess development.

Protein	DoE goal	DoE design used
Recombinant GFP from <i>P. pastoris</i>	Maximise GFP yield as a function of temperature, pH and DO in the culture medium	RSM (Box-Behnken) (Holmes et al., 2009)
Polyglutamic acid from <i>Bacillus subtilis</i>	Maximise polyglutamic acid yield as a function of the composition of the growth medium	Fractional factorial and RSM (Shi et al., 2006)
Recombinant Fab' fragment from <i>E. coli</i>	Maximise Fab' fragment yield as a function of agitation rate and DO in the culture medium	Full factorial (2 ²) (García-Arrazola et al., 2005)
Clavulanic acid from <i>Streptomyces clavuligerus</i>	Maximise clavulanic acid yield by optimising the composition of the growth medium	Fractional factorial and RSM (Wang et al., 2005)
Recombinant cystatin C mutant from <i>P. pastoris</i>	Maximising yield and protein glycosylation as a function of three nitrogen sources	Full factorial (2 ³) (Pritchett and Baldwin, 2004)
Neomycin from <i>Streptomyces marinensis</i>	Maximising neomycin yield by optimising the composition of the growth medium	Full factorial and RSM (Adinarayana et al., 2003)

1.4. The challenges of recombinant membrane protein production in yeast

Producing recombinant membrane proteins is more challenging than producing recombinant soluble proteins. This is because recombinant membrane proteins must be inserted into the cell membrane *in vivo*, and then removed from the cell membrane for further downstream processing such as X-ray crystallography for structural determination. This is a difficult process as removing the membrane protein from the lipid bilayer generally causes the protein to lose its integrity (Kalipatnapu and Chattopadhyay, 2005). For G Protein-Coupled receptors (GPCRs),

the largest family of membrane proteins, various strategies have been used to overcome this difficulty and other issues associated with X-ray structural determination (Moraes et al., 2014). These include: recombinant membrane protein overexpression (Wagner et al., 2008, Drew et al., 2005, Bonander et al., 2013, Fraser, 2006, Tate, 2001); novel solubilisation approaches with new detergents or chemicals (Prive, 2007, Chae et al., 2010, Serebryany et al., 2012, Knowles et al., 2009); improvement of protein stability via mutations, engineering of fusion partners and monoclonal antibodies (Tate and Schertler, 2009, Serrano-Vega et al., 2008, Serrano-Vega and Tate, 2009, Chun et al., 2012); automation and high-throughput screening of initial crystallisation conditions (Stevens et al., 2001) and synchrotron radiation and beamline developments (Duke and Johnson, 2010).

1.4.1. The target protein: the GPCR, human adenosine A_{2A} receptor (hA_{2A}R)

GPCRs moderate many physiological processes and are therefore targeted by many clinical drugs (Foord et al., 2005). The signalling pathways for these physiological processes are controlled by heterotrimeric guanine-nucleotide-binding proteins or G proteins constituting α , β and γ subunits. These proteins act as molecular switches by coupling the activation of GPCRs at the cell surface to intracellular signalling pathways. In the resting state, G proteins are inactive and $G\alpha$ binds guanine diphosphate (GDP) and $G\beta\gamma$. When extracellular stimuli such as hormones, neurotransmitters, chemokines, light or odourants activate the GPCR, a conformational change occurs in the receptor which allows G protein binding and GDP is released from the $G\alpha$ sub-unit. This results in a more stable, high affinity complex between the activated receptor and the G protein. Guanine triphosphate (GTP) then binds to $G\alpha$ and makes the complex unstable which leads to a dissociation of the α and $\beta\gamma$ sub-units from the receptor. The $G\alpha$ (GTP) and $G\beta\gamma$ sub-units interact with downstream effector proteins such as enzymes or channels that promote intracellular changes leading to a biological response. The $G\alpha$ subunit then hydrolyses GTP to GDP and re-associates with $G\beta\gamma$ which completes the G protein cycle and ends the cellular response (Oldham and Hamm, 2006). Figure 1.15 summarises G protein cycling.

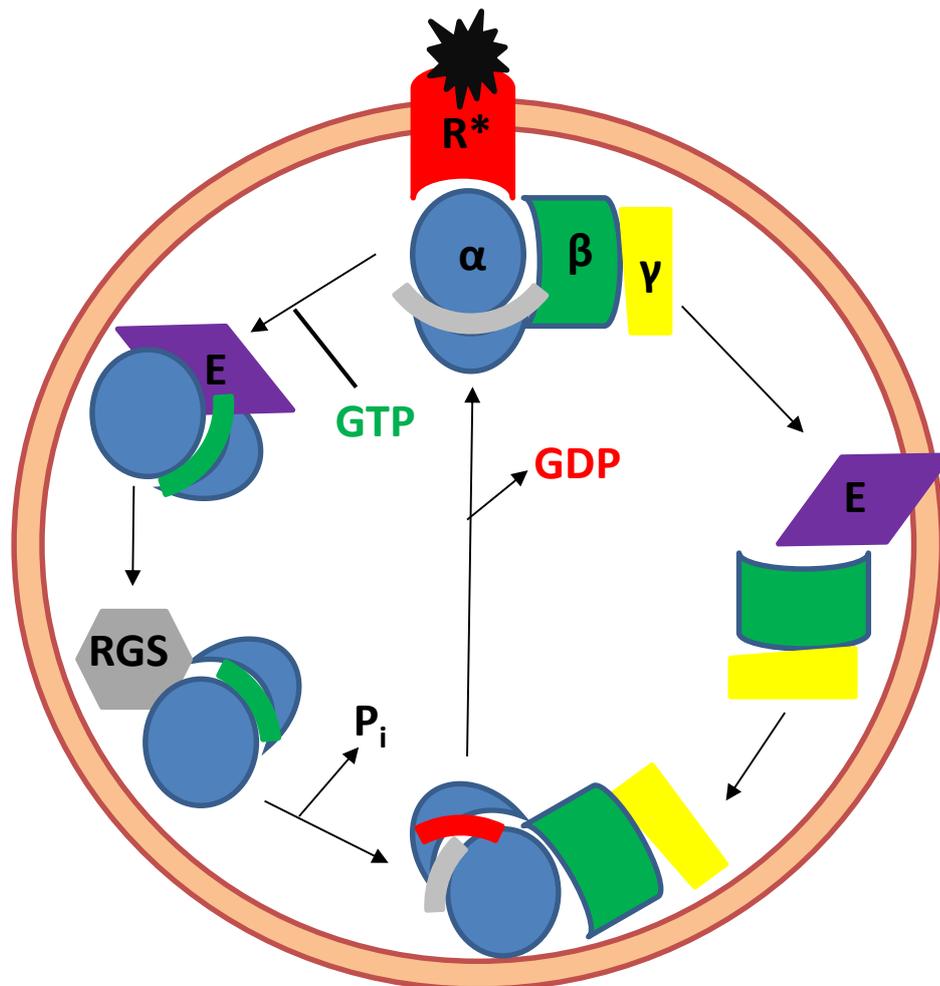


Figure 1.15. Summary of G protein cycle. G proteins are heterotrimeric of GDP-bound alpha (blue) beta (green) and gamma (yellow) subunits in the resting state. As an agonist (black) binds to the receptor (red) in the cell membrane, a conformational change results leading to G protein binding and subsequent GDP release. This is stable until GTP binding causes dissociation of R^* , $G\alpha$ (GTP) and $G\beta\gamma$. The subunits then activate a variety of effector proteins (E, purple). The signal ends when GTP is hydrolysed to GDP by $G\alpha$ and then maybe catalysed by RGS proteins (grey).

Structurally, GPCRs possess seven transmembrane α -helices. These α -helices are connected by three intracellular loops (ICL) and three extracellular loops (ECL); an intracellular carboxy or C-terminus and an extracellular amino or N-terminus are also present (Ahuja and Smith, 2009, Congreve and Marshall, 2010). Within different GPCR families, there are many amino acid sequence similarities and conserved residues which has led to the identification of motifs such as the DRY motif at the cytoplasmic end of the third transmembrane domain (Foord et al., 2005). The extracellular portions of a GPCR may be subjected to biochemical modifications such as glycosylation and disulphide bond formation (Karchin et al., 2002, Jacoby et al., 2006). GPCRs can also interact with other proteins such as GPCR kinases (GRKs), arrestin molecules and receptor-activity modifying proteins (RAMPs) which lead to specific actions such as trafficking (Brady and Limbird, 2002). The first mammalian GPCR structure to be resolved via

crystallisation was bovine rhodopsin by Palczewski and colleagues (Palczewski, 2000). The next milestone for GPCR structure resolution was the human $\beta 2$ adrenergic receptor (Rasmussen et al., 2007, Cherezov et al., 2007).

Figure 1.16 shows a representation of a Family A GPCR with the transmembrane portion (the 7 α -helices) embedded in the lipid bilayer of the plasma membrane and also ICLs and ECLs. Appendix A1 describes in more detail the other GPCR families and their characteristics (Bockaert and Philippe Pin, 1999).

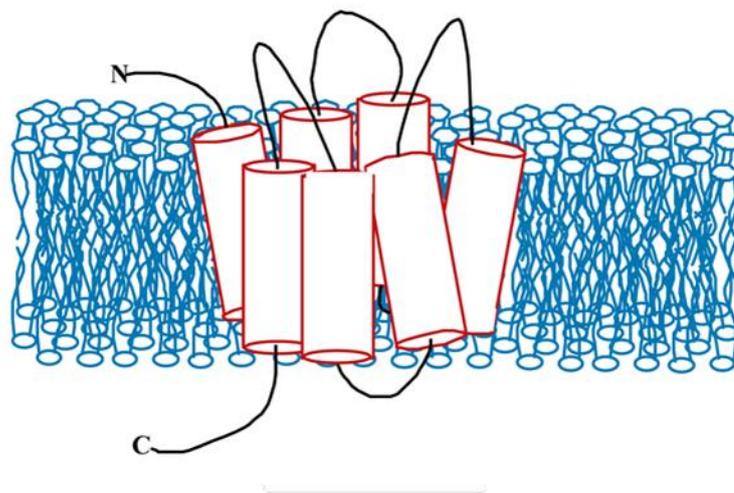


Figure 1.16 Diagrammatic representation of Family A GPCR. 7 cylindrical shapes in red outline represent 7 α -helices embedded in lipid bilayer (blue). Black lines represent ICLs (in the cell cytosol) or ECLs (outside the cell membrane). The ECL has the N-terminus and the ECL has the C-terminus (adapted from <http://structbio.vanderbilt.edu/>)

The human adenosine A_{2A} receptor ($hA_{2A}R$) is a Family A or rhodopsin-like GPCR (Palczewski, 2000) and is 47 kDa in size. The $hA_{2A}R$ structure was first determined at a 2.6 Å resolution with the high-affinity antagonist ZM241385 bound. The receptor was recombinantly produced from *S. frugiperda* insect cells (Figure 1.17) (Jaakola et al., 2008). This was an example of a GPCR fusion protein where T4 lysozyme replaced the conformationally dynamic ICL3 between the TM helices V and VI. Table 1.5 summarises the $hA_{2A}R$ crystal structures that have been resolved since 2008. It can be seen that all the $hA_{2A}R$ receptors were recombinantly produced as fusion proteins and/or contained mutations and were additionally co-crystallised with ligands. This was necessary as the receptor is highly dynamic and therefore these modifications increased its stability. Xu and colleagues (Xu et al., 2011) resolved the first $hA_{2A}R$ structure with an agonist bound. Previously it had been shown that agonist-bound receptors formed poorly diffracting crystals due to highly increased dynamics of the receptor.

However, Xu and colleagues (Xu et al., 2011) showed that careful selection of the agonist, UK-432097, which is a conformationally selective ligand, provided a less dynamic receptor for crystallisation trials. Furthermore, they showed that agonist binding at the extracellular domain triggers only small changes at the binding pocket but generally led to large scale seven transmembrane rearrangement that is required for G protein binding. Lebon and colleagues (Lebon et al., 2011) crystallised two versions of the hA_{2A}R with the bound agonists, adenosine and NECA. These agonist bound forms were able to be crystallised due to a thermostabilised construct (A_{2A}R-GL3) which contained 4 point mutations. In general, the structures revealed that when these agonists are bound to the receptor, the ligand binding pocket is narrowed or contracted due to the helices III, V and VII moving inwards. Dore et al., 2011 also used a different thermostable hA_{2A}R construct (A_{2A}-StaR2) for crystallising three structures with bound ZM241385 or the xanthines, XAC or caffeine. These structures showed inactive conformational states which were characterised by an ionic lock (a salt bridge) and the visibility of the third intracellular loop which is responsible for G protein coupling. Hino and colleagues (Hino et al., 2011) reported the first crystal structure of hA_{2A}R that was recombinantly produced in *P. pastoris*, whereas all other structures were recombinantly produced in insect cell lines. The receptor was in complex with an antibody Fab fragment (Fab2838) which gave a stable, inactive conformation when bound to the antagonist, ZM241385 demonstrating that ICL3 did not need to be replaced with the T4 lysozyme fusion protein. The general features of the structure were that the Fab2838 recognised the intracellular surface of the hA_{2A}R and that one of the complementarity determining regions (CDR) of the Fab2838 locks the hA_{2A}R in an inactive conformation. The most recent structure of hA_{2A}R to be resolved was by Liu and colleagues (Liu et al, 2012) to an increased resolution of 1.80 Å, which was not previously achieved. Due to this higher resolution, more information regarding the water and lipid molecules was revealed. It was found that 57 ordered water molecules were present inside the receptor and formed three clusters. The middle cluster contained a sodium ion which was bound to the highly conserved aspartate residue. Furthermore it was found that two cholesterol molecules stabilised the helix VI conformation. One lipid from the 23 ordered lipids was present inside the binding pocket. Functional studies revealed that sodium ion binding and antagonist binding in hA_{2A}R was non-competitive whereas agonist binding and sodium ion binding require different conformational states of the receptor for binding to occur. This suggests that the concentration of sodium ions affect functionally-relevant conformational states and therefore new ligands could be designed to exploit the sodium ion binding pocket (Gutiérrez-de-Terán et al., 2013).

So far these structures have revealed novel information about the differences between active and inactive conformations as well as some atomic detail of water and lipid positions in the receptor. However, the need for higher resolution structures is increasingly desired as these can provide

more atomic level information which is highly valuable as demonstrated by Liu et al., 2011 and Gutiérrez-de-Terán et al., 2013. Furthermore, developing methods for stable forms of the receptor for crystallisation without native loops or sections being replaced with fusion proteins is sought after to investigate the receptor in as native-like form as possible.

Table 1.5 Human adenosine A_{2A} crystal structures resolved. The table summarises the resolution of the structure, the recombinant host from which the receptor was produced for crystallisation studies and the main features and findings of the structures.

Reference	Resolution (Å)	Recombinant host	Features
(Xu et al., 2011)	2.71	<i>S. frugiperda</i>	Agonist, UK-432097; T4 lysozyme inserted between helices V and VI; structural changes in helices III, V and VI compared to (Jaakola et al., 2008) structure
(Lebon et al., 2011)	3.00	<i>Trichoplusia ni</i>	Agonist NECA and adenosine; thermostable receptor
(Dore et al., 2011)	3.60	<i>S. frugiperda</i>	Caffeine bound; thermostable receptor
	3.30	<i>S. frugiperda</i>	ZM241385 bound; thermostable versions of receptor
	3.31	<i>S. frugiperda</i>	XAC bound; thermostable receptor
(Hino et al., 2012)	2.70	<i>P. pastoris</i>	ZM241385 bound; in complex with inverse-agonist antibody (mouse Fab2838)
(Liu et al., 2012)	1.80	<i>S. frugiperda</i>	ZM241385 bound; apocytochrome b562 replaces ICL3; 23 lipids and 57 water molecules



Figure 1.17 Crystal structure of hA_{2A}R with ZM241385 antagonist bound. A. The transmembrane domain is coloured brown, ZM241385 is coloured light blue and the four lipid molecules bound to the receptor are coloured red. The four disulphide bonds are yellow. The extracellular loops (ECL1-3) are coloured green and the intracellular loops are coloured blue. The T4L is coloured as cyan. B. The molecule rotated 180° around the x-axis (image taken from Jaakola et al, 2008)

The hA_{2A}R has been implicated in diseases such as Parkinson's disease (Bara-Jimenez et al., 2003), Huntington's disease (Cha, 2000), asthma (Luijk et al., 2008), epilepsy (Boison, 2005) and numerous other neurological disorders (Boison, 2008). In mammalian cells, its natural agonist is adenosine. There are three other adenosine receptors including : the adenosine A₁, A_{2b} and A₃ receptors (Fredholm et al., 2001). Adenosine is known to have role in reducing inflammation (Sitkovsky et al., 2004) and therefore the hA_{2A}R is a target for therapeutic drugs treating diseases such as heart disease (Yang et al., 2006) and cancer (Stagg and Smyth, 2010). The main antagonists include caffeine, theophylline (a caffeine derivative) (Kulisevsky et al., 2002), ZM241385 (Poucher et al., 1995), 5'-N-Ethylcarboxamidoadenosine (NECA) (Arslan et al., 1999) and xanthine amine congener (XAC) (Jacobson et al., 1992).

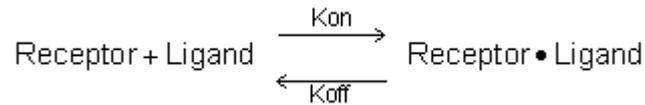
1.4.1.1. Ligand-binding assay for hA_{2A}R

The hA_{2A}R can be assayed accurately by ligand binding analysis (Singh et al., 2012). This is where known agonists or antagonists of the receptor can be radio-labelled and the amount of bound radioactivity can be measured (Hulme and Trevethick, 2010).

Agonists are synthetic or naturally occurring compounds that bind to receptors and activate them, therefore triggering a response. Antagonists are compounds that bind to the same

receptors without causing activation and can therefore block the action of an agonist (Leach, 2010).

Most radio-ligand binding experiments are based on the law of mass action:



where the ligand binds to the receptor to form a receptor - ligand complex. The rate at which the ligand binds to the receptor is defined by the association rate constant, k_{on} . As this equation is reversible, one can also define the dissociation rate constant, k_{off} . Equilibrium between association and dissociation is reached when the rate of formation of new receptor - ligand complexes equals the rate at which the receptor - ligand complexes dissociate (Motulsky, 1995). At equilibrium, the ratio of k_{on} and k_{off} values can provide information regarding the strength of the ligand-receptor interaction and this is termed the equilibrium dissociation constant, K_d :

$$K_d = \frac{k_{\text{off}}}{k_{\text{on}}}$$

K_d is equivalent to the concentration of ligand which binds 50% of the receptors. From this information, the Hill-Langmuir Binding Isotherm equation is derived:

$$[\text{Receptor} \cdot \text{Ligand}] = \frac{[\text{Receptor}_T] \times [\text{Ligand}]}{[\text{Ligand}] + K_d}$$

Where $[\text{Receptor}_T] = [\text{Receptor}] + [\text{Receptor} \cdot \text{Ligand}]$ and $K_d = k_{\text{off}} / k_{\text{on}}$

From the above fundamentals, three types of radio-ligand binding experiments are possible. In saturation binding experiments, the binding of an increasing concentration of radio-ligand, L, is measured at equilibrium to determine its binding constant (K_d) and the total number of the specific binding sites for the radio-ligand (B_{max}). In competition binding experiments, a fixed concentration of radio-ligand is measured at equilibrium in the presence of increasing concentrations of non-labelled ligand. The data derived from this can determine the binding constant (K_i) of a compound for the un-liganded receptor using the Cheng – Prusoff equation:

$$K_i = \frac{IC_{50}}{1 + [Radioligand] / K_d}$$

where K_i = dissociation constant; IC_{50} = the concentration of competing ligand that displaces 50% of the specific binding of the radioligand and $[Radio\text{-}ligand]$ = concentration of the radio-ligand. Finally, kinetic binding experiments are where the binding of one or more concentrations of radio-ligand is measured at increasing time points to determine the association (k_{on}) and dissociation (k_{off}) rate constants (Hulme and Trevethick, 2010). For the purposes of this research both saturation and competition binding experiments were employed (Motulsky, 1995).

1.4.2. The yeast plasma membrane

In order to understand the complexities of recombinant membrane protein production, it is important to understand the physical structure of the yeast plasma membrane as this is where the recombinant human membrane protein is inserted. It is also important to distinguish the differences between yeast and mammalian plasma membranes (Spira et al., 2012).

The yeast plasma membrane is about 7 nm thick with some cytosolic invaginations. It is a phospholipid bilayer with hydrophobic or non-polar (tails) and hydrophilic or polar (head) layers (Zinser and Daum, 1995, Spira et al., 2012). Figure 1.18 shows the chemical structure for the phospholipid in general (Figure 1.18A) where the hydrophobic tails are hydrocarbons of varying length that are attached to the hydrophilic head where a phosphate group is present. Figure 1.18B shows a simplified diagram of the phospholipid and Figure 1.18C shows how the phospholipid forms a bilayer in the plasma membrane.

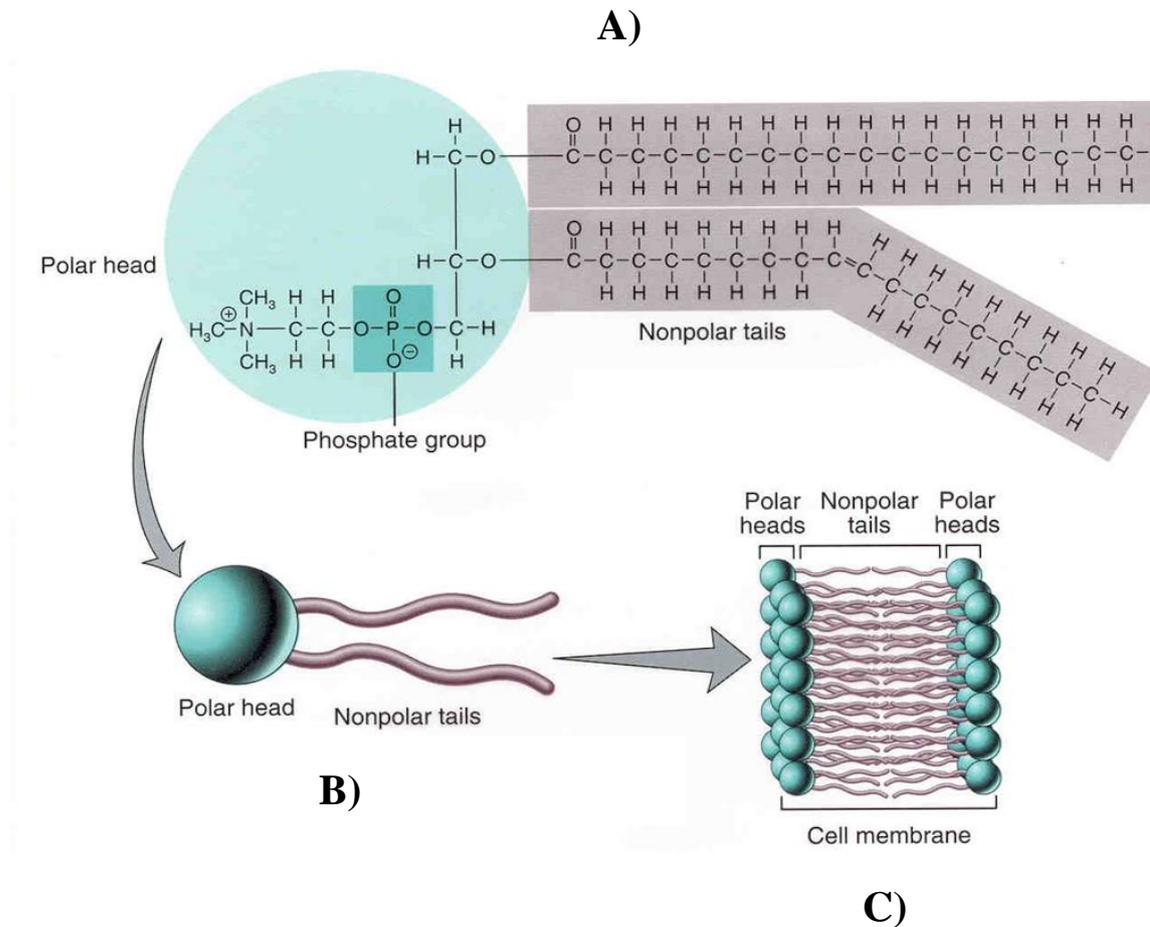


Figure 1.18 General structure of phospholipid. A) shows the chemical structure of the phospholipid with the hydrophobic (non-polar) tails which are hydrocarbon chains and hydrophilic (polar) heads with a phosphate group. B) shows a representation of the phospholipid, with one polar head and 2 non-polar tails. C) shows the phospholipids organised within a plasma membrane to form a lipid bilayer. Adapted from http://homepage.smc.edu/wissman_paul/anatomy2textbook/phospholipids.html.

The main phospholipids found in yeast plasma membranes are phosphatidylcholine such as 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC); 1,2-dilauroyl-*sn*-glycero-3-phosphocholine (DLPC); 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) and phosphatidylethanolamine such as 1,2-dimyristoyl-*sn*-glycero-3-phosphoethanolamine (DMPE) and 1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine (DPPE). Other phospholipids and lipids found in the plasma membrane but in small amounts include phosphatidylinositol, phosphatidylserine, phosphatidylglycerol, unsaturated fatty acids and sterols such as ergosterol, zymosterol, fecosterol and episterol (Zinser and Daum, 1995).

Native yeast membrane proteins are categorised into cytoskeleton anchors, cell wall synthesis enzymes, signal transduction proteins, solute transport proteins (permeases, transport channels and ATPases) and transport facilitators. It can be noted that the sterol, cholesterol is absent from

yeast plasma membrane but this is a major sterol in mammalian cells. Cholesterol and ergosterol are not biologically equivalent as ergosterol possesses two additional double bonds at positions C7 and C22 and a methyl group at C24 of the side chain (Figure 1.19) (Tierney et al., 2005).



Figure 1.19 Chemical structure of cholesterol and ergosterol. Ergosterol has two additional double bonds at positions C7 and C22 and a methyl group at C24 of the side chain when compared to cholesterol (Image from Tierney et al., 2005).

Furthermore, some human GPCRs are only active in the presence of cholesterol. Efforts are therefore being made in engineering yeast strains that make cholesterol-like sterols (Kitson et al., 2011). More commonly, cholesterol derivatives e.g. cholesteryl hemi-succinate are added when the extraction of the membrane protein is being performed (Section 1.4.3). Figure 1.20 shows a simple representation of a yeast plasma membrane.

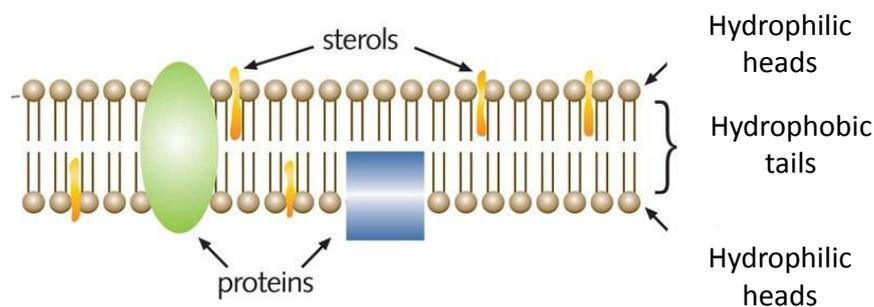


Figure 1.20 Yeast plasma membrane. Picture shows a representation of the phospholipid bilayer with the hydrophobic and hydrophilic regions. Membrane protein along with sterols are also shown. Image adapted from www.distillique.co.za

The way in which lipids are organised in the plasma membrane are also important when considering membrane proteins and their interactions with the lipids. This helps with efficient

co-ordination of functions if the membrane proteins are segregated into distinct domains in the lipid bilayer (Spira et al., 2012). There are several models for this theory and these include the lipid-raft theory where liquid-ordered domains (cholesterol and sphingolipids, the ‘rafts’) are separated from liquid-disordered domains (phospholipids) (Lingwood and Simons, 2010); protein-protein interactions (Douglass and Vale, 2005); cortical actin (Kusumi et al., 1993) and the extracellular matrix (Sackmann, 1996). These theories should be given some thought when producing recombinant membrane proteins, although separate and detailed research would need to be performed in order to fully characterise the target membrane protein in the yeast lipid bilayer.

1.4.3. Extraction of recombinant membrane proteins from the yeast plasma membranes

Helenius and Simons in 1975 first developed a method to extract membrane proteins from cell membranes referred to as surfactant solubilisation (Helenius and Simons, 1975). The term surfactant (an abbreviation of surface active agents) is used less commonly than detergent even though a detergent is defined as a formulation of a surfactant or cleaning product (Jamshad et al., 2011). However, for the purpose of this thesis, the term detergent will be used. Detergent use in membrane protein extraction involves maintaining the membrane protein of interest in a functional, correctly folded state without its native membrane present. This process is required for any purification methods of the membrane protein and for any further biophysical studies such as X-ray crystallography (Prive, 2007).

1.4.3.1. Detergent use in membrane protein solubilisation

Detergents are typically used to extract membrane proteins from their native lipid bilayer. This process, which is often referred to as membrane protein solubilisation, involves the replacement of lipids with detergent molecules as shown in Figure 1.21. Detergents possess a hydrophilic or polar head group and a hydrophobic tail group similar to the phospholipids in the membrane bilayer (Kalipatnapu and Chattopadhyay, 2005, Duquesne and Sturgis, 2010). Following a solubilisation experiment, the resulting aqueous solution contains membrane protein in complex either with detergent or with detergent and lipid (Figure 1.21) which are then suitable for purification and further analysis. The most common issues with detergent solubilisations are protein aggregation and protein denaturation which lead to loss of protein function and will prevent any further structural studies (Prive, 2007). Detergents are amphipathic in nature due to their hydrophilic head groups and hydrophobic alkyl tail groups. Detergent molecules exist as monomers in aqueous solutions when at low concentrations. When the concentration of the detergent increases to a certain level they start to self-assemble into small spherical and thermostable structures called micelles. At this minimum concentration when the micelles are formed it is termed the Critical Micelle Concentration (CMC) (le Maire et al., 2000). In order

for a detergent to solubilise a membrane protein, the detergent concentration must reach above its CMC (Prive, 2007). Moreover, the critical solubilisation concentration (CSC) is the minimum detergent concentration required to disperse the lipid bilayer into micelles (Prive, 2007)

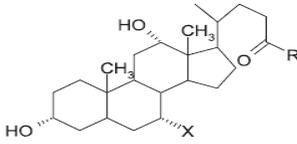
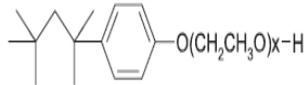
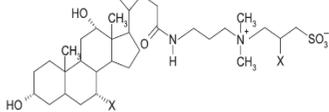
The head group of the detergent typically influences the interaction with proteins and the alkyl chain affects the CMC and the aggregation number. The aggregation number is the number of molecules in a micellar particle (Prive, 2007). Therefore in general, shorter chain detergents (C8) tend to be harsher and cause denaturation of a protein when compared to more gentle longer chained detergents (C12). When considering the head group, the smaller, highly charged head group will be a more harsh detergent than one with a larger more neutral head group. These properties must be considered when solubilising the lipid bilayer as the more harsh detergents may denature the protein and render them non-functional or conversely a more mild detergent may be poorly soluble (Prive, 2007).



Figure 1.21 Scheme to show stages of lipid bilayer solubilisation by detergent. (A) shows the lipid bilayer of a membrane (B) shows low concentrations of detergent (grey with single tails) and lipid bilayer (C) shows higher concentrations of detergent and the lipid bilayer disrupts (D) shows the mixed populations retrieved as result of high concentrations of detergents (taken from (Kalipatmapu and Chattopadhyay, 2005)).

There are many detergents available that can be classified into four groups (ionic, bile salts, non-ionic and zwitterionic detergents) according to their chemical structure. Table 1.6 summarises the structures and properties of given examples and their effectiveness.

Table 1.6 Four main groups of detergents used for solubilisation of membrane proteins.

	Ionic	Bile-acid	Non-ionic	Zwitterionic
Chemical structure	$\text{CH}_3(\text{CH}_2)_{10}\text{CH}_2\text{O}-\text{S}(=\text{O})_2\text{O}^-$			
Properties	Polar head group (anionic or cationic); hydrophobic tail (hydrocarbon chain or steroidal backbone)	Same as ionic detergents with rigid steroidal group backbone, results in polar and apolar face instead of well-defined group	Uncharged hydrophilic head groups of either polyoxyethylene or glycosidic groups	Combination of ionic and non-ionic detergents and more denaturing than non-ionic detergents
Examples	Sodium dodecyl sulphate (SDS)	Sodium deoxycholate, sodium cholate	Alkylglucosides: n-octyl-β-D-glucopyranoside, decyl-β-D-maltoside and dodcecyl-β-D-maltoside, all (DDM)	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPSO)
Effectiveness	Extremely effective but almost always denatures membrane proteins (Seddon et al., 2004)	Quite mild and therefore less denaturing (De Foresta et al., 1989)	Very mild detergents and so do not denature proteins readily. Very popular. (Kragh-Hansen et al., 1993)	More denaturing than non-ionic detergents. (Sardet et al., 1976)

1.4.3.2. Alternative solubilising agents

Although, detergents have been used to solubilise membrane proteins for more than 50 years, protein-detergent micelles are unstable and tend to aggregate leading to loss of membrane protein function. Furthermore, detergent molecules tend to be disordered and therefore the formation of crystal lattices may be compromised (Alguet et al., 2010). Biophysical studies have shown that detergent micelles can only give an approximation of the native environment provided by the lipid bilayer. Neutron-scattering studies have shown that the membrane bilayer is made up of many layers running perpendicular to the membrane normal (Lin and Guidotti, 2009) (Wiener and White, 1992). This is something that a detergent micelle will struggle to replicate along with alterations that can occur in the composition of the lipid bilayer such as changes to the phospholipid head or acyl chain, which can lead to varied interactions with the membrane protein (Charalambous et al., 2008). Also the presence of lipid rafts and other lipid organisational theories, as discussed in section 1.4.3., contribute to the challenges of using detergents for solubilisation. Furthermore, a study on two membrane protein transporters; the *E. coli* transporter, EmrE and the *Mycobacterium tuberculosis* transporter, TBsmr (both multi-drug resistance family members) showed that the phospholipid, phosphatidylethanolamine can alter the lateral pressure profile of the lipid bilayer. These lipid bilayer lateral pressures can affect membrane protein insertion, folding and activity and hence contribute to further difficulties in using detergents for solubilisation (Charalambous et al., 2008).

Research in this area has focused on identifying a more robust system to solubilise membrane proteins in a more stable manner using the tools of nanoscience and nano-self-assembly (Jamshad et al., 2011). Several types of novel systems some still in development, have been proposed for solubilising membrane proteins as outlined in Table 1.7.

1.4.3.2.1. Bicelles

Bicelles are composed of a central planar bilayer of long-chain phospholipids, such as DMPC, surrounded by a rim of short-chain phospholipids, such as DHPC, which shield the long-chain lipid tails from water (Sanders and Prestegard, 1990). As the long-chain phospholipid molecules in the bicelles are positioned in the planar core region, this region is thought to mimic natural membranes much better than micelles that are formed by detergents (Glover et al., 2002).

1.4.3.2.2. Amphipols

Amphipols are milder forms of detergents that allow a membrane protein to be surrounded by annular lipids (Picard et al., 2006), which are a shell of lipid molecules that surround the membrane penetrating surface of the membrane protein (Lee, 2011b). Amphipols are amphipathic polymers consisting of hydrophilic backbones and hydrophobic side chains. They

have the ability to hypercoil around the membrane protein region and therefore helps them to stay folded correctly (Popot, 2010).

1.4.3.2.3. Nanodiscs

Nanodiscs and nanodisc technology are also referred to as nanoscale apolipoprotein bound bilayers (NABB) or high-density lipoprotein (HDL) particles. The technology was developed by Bayburt and colleagues in 2002 (Bayburt and Sligar, 2003). Nanodiscs contain a central lipid bilayer (with the membrane protein within) and two molecules of membrane scaffold protein (MSP) which is a helical repeat protein with a hydrophobic and hydrophilic face. The MSP wraps itself around the hydrophobic perimeter of the lipid disc and stabilises it (Bayburt and Sligar, 2010).

1.4.3.2.4. Maltose-neopentyl glycol (MNG) amphiphiles

Maltose-neopentyl glycol (MNG) amphiphiles comprise a tetra-substituted carbon replaced with one hydrophilic and three hydrophobic substituents (Chae et al., 2010). Furthermore this quaternary carbon causes a subtle restraint on the conformational flexibility of the amphiphile. The quaternary carbon is derived from neopentyl glycol and the hydrophilic groups are derived from maltose, hence the name MNG amphiphiles (Chae et al., 2010).

1.4.3.2.5. Responsive hydrophobically associating polymers

Responsive hydrophobically associating polymers contain both hydrophilic and hydrophobic groups rendering them amphipathic. The polymers used in this thesis are poly (maleic acid-styrene) or PMAS. The maleic acid is a weakly negatively charged (hydrophilic) carboxylic acid and the styrene is an aromatic compound and highly hydrophobic. When the pH is high (above 7.0), the polymer is in chain form due to repulsive interactions between negatively charged hydroxyl groups and is soluble in water but as the pH decreases (lower than 6.0), the maleic acid becomes protonated and the hydrophobicity of the molecule increases. This causes the formation of a compact and insoluble molecule i.e. the polymer chain collapses and hyper-coiling occurs. It is thought that as the hyper-coiling takes place it encapsulates some of the lipid bilayer and also the membrane protein embedded within it (Tonge and Tighe, 2001). Variations of these polymers are achieved by altering the molecular weight and also the ratio of styrene to maleic acid. Modified side chains may also be attached.

Previously, Knowles and colleagues had worked with one version of the PMAS which they named SMA (styrene-*co*-maleic acid) and when mixed with the plasma membrane, the resulting structure was termed SMALP (poly(styrene-*co*-maleic acid) lipid particle (Knowles et al., 2009) (Jamshad et al., 2011). They were able to solubilise membrane proteins such as PagP and

bacteriorhodopsin with success leading to positive biophysical experiments (Knowles et al., 2009; Jamshad et al., 2011). It should be noted for this thesis, that the polymer named PMAS 2000P, has the same molecular structure as SMA (Chapter 5).

There are several advantages known when considering the use of PMAS for solubilisation of membrane proteins. These include the production of homogenous particles with a diameter of 9-11 nm and that these particles can be studied using a wide range of biophysical techniques such as CD, AUC, DSC (differential scanning calorimetry) and fluorescence. Furthermore, CD studies have shown that the PMAS solubilised material can be carried out at lower wavelengths (Jamshad et al., 2011). The polymer itself is very cheap to purchase and easy to prepare when compared to the detergents. The main disadvantage of these polymers is that there is limited published data for solubilisation of membrane proteins.

Table 1.7 summarises the chemical structure, the main properties of the solubilising agent, main advantages and disadvantages and an example of any membrane proteins that were solubilised with solubilising agent. Figure 1.22 shows a schematic diagram of what the specific solubilising agent, the membrane bilayer and the membrane protein may look like.

Table 1.7 Alternative solubilising agents

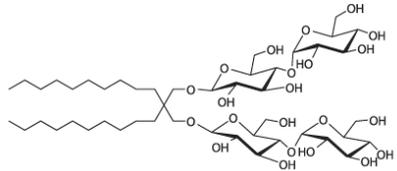
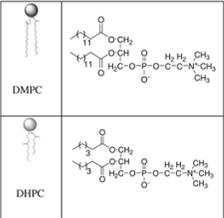
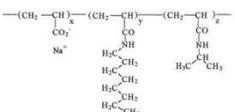
	Chemical structure	Properties	Advantage	Disadvantage	Example of membrane protein solubilisation
Maltose neopentyl glycol (MNG) amphiphiles		Tetra-substituted carbon with one hydrophilic and three hydrophobic substituents	Successful in crystallisation trials	Little information on size of micelles	Human β_2 adrenergic receptor –T4L, muscarinic M ₃ acetylcholine receptor (Chae et al, 2011)
Bicelles		Long chain and short chain phospholipids mimic natural membrane	Successful in NMR, where the presence of some surfactant not an issue	Intermediate detergent step required	Opsin protein to form the GPCR, rhodopsin (Reeves et al, 1999)
Amphipols		Form tight interactions with membrane protein, increased stability	Successful solubilisation, no interfering with membrane protein function	Certain pH lead to aggregated amphipol	Bacteriorhodopsin (Tribet et al, 1996)

Table 1.7 Alternative solubilising agents (continued)

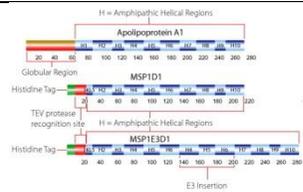
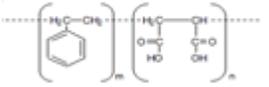
	Chemical structure	Properties	Advantage	Disadvantage	Example of membrane protein solubilisation
Nanodiscs	 <p>The diagram illustrates the structure of Nanodiscs. It shows Apolipoprotein A1 (ApoA1) with amphipathic helical regions (H) and a globular region. Below it, MSP1D1 is shown with a histidine tag and a TEV protease recognition site. At the bottom, MSP1E3D1 is shown with a histidine tag and an E3 insertion site. Residue numbers are indicated along the protein sequences.</p>	Membrane scaffold protein (MSP) wraps itself around the hydrophobic perimeter of the lipid disc and stabilises it	Similar constituents as lipid bilayer, more efficient integration	Scaffold protein remains a contaminant	Human β_2 adrenergic receptor (Leitz et al, 2006)
Responsive hydrophobically associating polymers (styrene maleic acid)	 <p>The chemical structure shows a copolymer of styrene and maleic acid. The styrene units are represented by a benzene ring attached to a $-CH_2-CH-$ group. The maleic acid units are represented by a $-CH-CH-$ group with two carboxylic acid groups ($-COOH$) attached to the same carbon atoms. The units are linked together in a chain.</p>	Responsive hypercoiling amphiphiles	Produces homogenous particles, able to solubilise GPCR (unpublished); little evidence for interfering with CD or AUC techniques	No published studies on GPCRs	Bacteriorhodopsin and PagP (Knowles et al, 2009); Human adenosine A_{2A} receptor (unpublished, Chapter 5 of this thesis)



Figure 1.22 Images of proposed structures formed with alternative solubilising agent, lipid and membrane protein A) Bicelle – membrane protein (blue) is embedded inside the bicelle made from long-chain lipids (red) and a mixture of short-chain lipids and detergent molecules (grey) (reproduced from Serebryany et al, 2012). B) Amphipol - the amphipol (red) is hypercoiled around the lipid bilayer (blue) (reproduced from Jamshad et al, 2011). C) Nanodisc – membrane protein (blue) embedded in a nanodisc composed of a lipid bilayer (red) and a membrane scaffold protein (MSP) (reproduced from Serebryany et al, 2012). D) Poly(maleic-anhydride styrene) or PMAS hyper-coiled (green) and wrapped around the lipid bilayer and (Image supplied by Dr. Paul Topham, Aston University).

1.5. Project aims

This project aimed to improve recombinant membrane protein production yields in yeast and study novel polymers for extracting a specific membrane protein from the plasma membrane.

The studies primarily utilised the methylotrophic yeast species, *Pichia pastoris* but also *Saccharomyces cerevisiae*. The membrane protein that was chiefly studied in this thesis was the GPCR, human adenosine _{2A} receptor, hA_{2A}R. This was selected as it is a well-characterised GPCR and therefore served as an appropriate model protein to study the novel extraction process and also the yield improvement processes which were the focus of the study.

More specifically, the optimisation strategies that were the primary focus of this research were the application of statistical Design of Experiments in a small scale, high –throughput system for hA_{2A}R production in *Saccharomyces cerevisiae* were investigated as a method of bioprocess optimisation prior to scale-up (Chapter 3). An investigation into the methanol feeding regimes, the pre-induction and cytotoxic culture environments for hA_{2A}R production in *Pichia pastoris* (Chapter 4). An in-depth study into the extraction of the hA_{2A}R from *Pichia pastoris* membranes using novel polymers (Chapter 5).

In summary this project sought to address three major challenges in recombinant protein production in yeast:

- Achieving optimised production with minimal trial and error through a DoE approach
- Maximising production through controlled feeding regimes
- Using novel polymers to improve the process of extracting proteins from cell membranes

Chapter 2: Material and methods

2.1. Materials

2.1.1. Reagents and buffers

2.1.1.1. Stock solutions and buffers

2.1.1.1.1. 10× YNB (13.4% yeast nitrogen base with ammonium sulphate without amino acids)

134.0 g yeast nitrogen base was dissolved in distilled water to a total volume of 1 L and filter sterilised. The medium was stored at 4°C.

2.1.1.1.2. 500× biotin (0.02%)

20 mg biotin was dissolved in distilled water to a total volume of 100 mL and filter sterilised. It was stored at 4°C.

2.1.1.1.3. 10× glycerol (10%)

100 mL glycerol was mixed with 900 mL distilled water. It was filter sterilised and stored at room temperature.

2.1.1.1.4. 10× methanol (5%)

5 mL methanol was mixed with 95 mL distilled water and filter sterilised. The medium was stored at 4°C.

2.1.1.1.5. 40× glucose (40%)

400 g glucose was dissolved in distilled water to a total volume of 1 L and filter sterilised. The solution was stored at 4°C.

2.1.1.1.6. 1 M potassium phosphate buffer, pH 6.0

1 M solution of K_2HPO_4 was made by dissolving 174.2 g in distilled water to a total volume of 1 L. 1 M solution of KH_2PO_4 was made by dissolving 136.1 g in distilled water to a total volume of 1 L. 132 mL 1 M K_2HPO_4 was mixed with 868 mL KH_2PO_4 and the pH adjusted to 6.0. The solution was autoclaved and stored at room temperature.

2.1.1.1.7. 1× T.A.E. buffer

40 mM Tris acetate and 1 mM EDTA were mixed to a final volume of 1 L and the pH adjusted to 8.2-8.4.

2.1.1.1.8. 1 M DTT (Dithiothreitol)

1.54 g DTT powder was dissolved in double distilled water to a final volume of 10 mL. The solution was filter sterilised and stored at 4°C.

2.1.1.1.9. 1 M HEPES (2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid) buffer pH 8.0

238.3 g HEPES powder was dissolved in double distilled water to a final volume of 1 L. The solution was adjusted to pH 8.0, filter sterilised and stored at 4°C.

2.1.1.2. *E. coli* culture medium

2.1.1.2.1. LB (Luria-Bertani)

20 g LB powder was dissolved in distilled water to a total volume of 1 L. For culture plates, 20 g agar was added and the solution autoclaved at 121°C for 20 min then cooled to room temperature and stored at 4°C.

2.1.1.3. *S. cerevisiae* culture media

2.1.1.3.1. Yeast peptone dextrose (YPD)

20 g peptone and 10 g yeast extract were dissolved in distilled water to a total volume of 900 mL. For the agar plates, 20 g agar was added. The solution was autoclaved and then cooled to room temperature before adding 100 mL 10× glucose and stored at 4°C.

2.1.1.3.2. 2× CBS (Centralbureau voor Schimmelcultures) medium

10 g ammonium sulphate, 6 g potassium dihydrogen phosphate and 1 g magnesium sulphate heptahydrate were dissolved in 646 mL distilled water. This was autoclaved and then 50 mL 40× glucose, 100 mL MES, pH 6.0, 200 mL 10× DO solution (-uracil for *S.cerevisiae* wild type and *S. cerevisiae* TM6* strains, and -histidine for *S.cerevisiae* BMS1 strains), 2 mL vitamin solution and 2 mL trace elements were added.

2.1.1.3.3. CSM (complete synthetic medium)

1.7 g yeast nitrogen broth without amino acids, 5 g ammonium sulphate, 20 g agar and 750 mL were dissolved and then autoclaved. 100 mL 10× DO solution (-uracil for *S.cerevisiae* wild type

and *S. cerevisiae* TM6* strains, and -histidine for *S. cerevisiae* BMS1 strains). 100 mL MES pH 6.0 and 50 mL 40× glucose added to the solution and plates were then poured, cooled and stored at 4°C.

2.1.1.3.4 10× Drop out solution (DO solution) (minus uracil)

200 mg L-adenine hemi-sulphate salt, 200 mg L-arginine HCl, 200 mg L-histidine HCl monohydrate, 300 mg L-isoleucine, 1000 mg L-leucine, 300 mg L-lysine HCl, 200 mg L-methionine, 500 mg L-phenylalanine, 2000 mg L-threonine, 200 mg L-tryptophan, 300 mg L-tyrosine and 1500 mg L-valine were added to 1 L distilled water then autoclaved and stored at 4°C.

2.1.1.3.5 10× Drop out solution (DO solution) (minus histidine)

200 mg L-adenine hemi-sulphate salt, 200 mg L-arginine HCl, 300 mg L-isoleucine, 1000 mg L-leucine, 300 mg L-lysine HCl, 200 mg L-methionine, 500 mg L-phenylalanine, 2000 mg L-threonine, 200 mg L-tryptophan, 300 mg L-tyrosine, 200 mg L-uracil and 1500 mg L-valine were added to 1 L distilled water then autoclaved and stored at 4°C.

2.1.1.3.6 Trace elements solution

3.75 g EDTA and 1.125 g zinc sulphate heptahydrate in 190 mL water were dissolved and the pH was adjusted to 6.0 with 1 M NaOH. Whilst maintaining the pH, the following were added: 0.25 g magnesium chloride tetra-hydrate, 0.075 g cobalt (II) chloride hexahydrate, 0.075 g copper (II) sulphate pentahydrate, 0.1 g sodium molybdenum dehydrate, 1.125 g calcium chloride dehydrate, 0.75 g iron sulphate heptahydrate, 0.25 g boric acid, 0.025 g potassium iodide. The pH was adjusted to 4.0 with 1 M HCl and topped up to 250 mL with water. The bottle was covered in foil to exclude light, autoclaved and stored at 4°C.

2.1.1.3.7. Vitamin solution

0.0125 g biotin was dissolved in 2.5 mL 0.1 M NaOH. 190 mL water was added and the pH adjusted to 6.5 with 1 M HCl. Whilst maintaining the pH, the following were added: 0.25 g calcium D-pantothenate, 0.25 g nicotinic acid, 6.25 g myo-inositol, 0.25 g thiamine HCl, 0.25 g pyridoxol HCl and 0.05 g d-amino benzoic acid. This was topped up to 250 mL, filter sterilised and stored at 4°C in the dark.

2.1.1.4. *P. pastoris* culture media

2.1.1.4.1. Buffered complex glycerol medium (BMGY)

10 g yeast extract and 20 g peptone were dissolved in distilled water to a total volume of 700 mL. The solution was autoclaved then cooled to room temperature. The following was then added: 100 mL 1 M potassium phosphate buffer pH 6.0, 100 mL 10× YNB, 2 mL 500× biotin and 100 mL 10× glycerol. The medium was stored at 4°C.

2.1.1.4.2. Buffered complex methanol medium (BMMY)

10 g yeast extract and 20 g peptone were dissolved in distilled water to a total volume of 700 mL. The solution was autoclaved then cooled to room temperature. The following was then added: 100 mL 1 M potassium phosphate buffer pH 6.0, 100 mL 10× YNB, 2 mL 500× biotin and 100 mL 10 × methanol. The medium was stored at 4°C.

2.1.1.4.3. Basal salts medium (BSM)

0.93 g calcium sulphate, 18.2 g potassium sulphate, 14.9 g magnesium sulfate heptahydrate, 4.13 g potassium hydroxide, 40 g glycerol and 26.7 ml 85% phosphoric acid were dissolved in distilled water to make 1 L. The medium was autoclaved and used immediately or stored at 4°C.

2.1.1.4.4. PTM₁ trace salts

6 g cupric sulphate pentahydrate, 0.08 g sodium iodide, 3 g manganese sulfate monohydrate, 0.2 g sodium molybdate dihydrate, 0.02 g boric acid, 0.5 g cobalt chloride, 20 g zinc chloride, 65 g ferrous sulfate heptahydrate, 0.2 g biotin, 5.0 mL sulphuric acid were dissolved in a final volume of 1 L distilled water. The solution was filter sterilised and stored at 4°C in the dark.

2.1.1.4.5. Fermentation medium 22 (FM22)

42.9 g potassium dihydrogen phosphate, 5 g ammonium sulphate, 1 g calcium sulphate dihydrate, 14.3 g potassium sulphate, 11.7 g magnesium sulphate heptahydrate and 10 g glycerol were dissolved in 1 L of distilled water. The medium was autoclaved and used immediately or stored at 4°C.

2.1.1.4.6. PTM₄ trace salts

2 g cupric sulphate pentahydrate, 0.08 g sodium iodide, 3 g manganese sulfate monohydrate, 0.2 g sodium molybdate dihydrate, 0.02 g boric acid, 0.5 g cobalt chloride, 7 g zinc chloride, 22 g ferrous sulfate heptahydrate, 0.2 g biotin and 1.0 mL sulphuric acid were dissolved in a

final volume of 1 L distilled water. The solution was filter sterilised and stored at 4°C in the dark.

2.1.1.4.7. 1 M Sorbitol

182.2 g sorbitol was dissolved in double distilled water to a final volume of 1 L. The solution was filter sterilised and stored at 4°C.

2.1.1.4.8. YPDS (yeast extract, peptone, dextrose, sorbitol) plus 100 µg mL⁻¹ zeocin culture plates

1% yeast extract (5 g), 2% peptone (10 g), 1 M sorbitol (91.1 g) and 10 g agar were added to distilled water to 487 mL and autoclaved. After cooling, 12.5 mL 40× glucose was added with 0.5 mL 100 mg mL⁻¹ zeocin. The plates were poured aseptically, stored at 4°C in the dark and had a shelf-life of 2 weeks.

2.1.1.4.9. YPDS (yeast extract, peptone, dextrose, sorbitol) plus 250 µg mL⁻¹ zeocin culture plates

1% yeast extract (5 g), 2% peptone (10 g), 1 M sorbitol (91.1 g) and 10 g agar were added to distilled water to 486.25 mL and autoclaved. After cooling, 12.5 mL 40× glucose was added with 1.25 mL 100 mg mL⁻¹ zeocin. The plates were poured aseptically, stored at 4°C in the dark and had a shelf-life of 2 weeks.

2.1.1.4.10. YPDS (yeast extract, peptone, dextrose, sorbitol) plus 500 µg mL⁻¹ zeocin culture plates

1% yeast extract (5 g), 2% peptone (10 g), 1 M sorbitol (91.1 g) and 10 g agar were added to distilled water to 485 mL and autoclaved. After cooling, 12.5 mL 40× glucose was added with 2.5 mL 100 mg mL⁻¹ zeocin. The plates were poured aseptically, stored at 4°C in the dark and had a shelf-life of 2 weeks.

2.1.1.4.11. YPDS (yeast extract, peptone, dextrose, sorbitol) plus 1000 µg mL⁻¹ zeocin culture plates

1% yeast extract (5 g), 2% peptone (10 g), 1 M sorbitol (91.1 g) and 10 g agar were added to distilled water to 482.5 mL and autoclaved. After cooling, 12.5 mL 40× glucose was added with 5 mL 100 mg mL⁻¹ zeocin. The plates were poured aseptically, stored at 4°C in the dark and had a shelf-life of 2 weeks.

2.1.1.4.12. Antifoams

Two antifoams were used in this research: Mazu DF 204 (Sigma) and P2000 polyethylene glycol (Fluka).

2.1.1.4.13. Glycerol analysis

A glycerol quantitation kit was used to evaluate residual glycerol concentration in culture media (r-biopharm, Roche) and was carried out according to manufacturer's instructions.

2.1.1.4.14. Methanol analysis

Methanol standards were prepared for gas chromatography analysis and were made with 100% methanol and double distilled water. Dilutions included: 0, 0.013, 0.065, 0.13, 0.25, 0.38, 0.5, 1, 2, 5 and 10%.

2.1.1.5. Antibiotics

2.1.1.5.1. Zeocin

250 mg zeocin was dissolved in 10 mL sterile water to give a working concentration of 25 mg mL⁻¹. It was stored at -20°C in the dark.

2.1.1.5.2. Ampicillin

500 mg ampicillin was dissolved in 10 mL sterile water to give 50 mg mL⁻¹. It was stored at -20°C.

2.1.1.5.3. Doxycycline

5 mg mL⁻¹ stock was made with sterile water and stored at -20°C.

2.1.1.6. Molecular biology reagents

2.1.1.6.1. pPICZB vector

The pPICZB vector was used as the backbone for the novel hA_{2A}R construct (Chapter 5). The key features are shown in Figure 2.1.



Figure 2.1 pPICZB expression vector developed by Life Technologies Corporation. Vector contains a multiple cloning site with a choice of restriction enzymes and includes the AOX1 promoter region (reproduced from www.lifetechnologies.com).

2.1.1.6.2. Restriction enzyme *PmeI*

The source is an *E.coli* strain that carries the *PmeI* gene from *Pseudomonas mendocina*. The reaction was carried out at 37°C (New England Biolabs)

2.1.1.6.3. Mini-preparation kit

The GeneJET plasmid DNA mini-prep kit (Fermentas) was used according to the manufacturer's protocol.

2.1.1.6.4. Purification of *PmeI* digestion

The QIAquick PCR purification kit (Qiagen) was used according to the manufacturer's protocol.

2.1.1.6.5. XL-10 Gold *E.coli* cells

E.coli competent cells were purchased from Agilent Technologies Inc. and were used to amplify the plasmid DNA for transformation into *P. pastoris* cells. XL-10 Gold yield high efficiency transformations of large plasmids. They are tetracycline and chloramphenicol resistant. The strain is endonuclease (*endA*) and recombination deficient (*recA*). The genotype is:

Tet^rΔ(*mcrA*)183Δ(*mcrCB-hsdSMR-mrr*)173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte*
[F' *proAB lacIUPSIDEDOWNbZΔM15 Tn10* (Tet^r)Amy Cam^r]

2.1.1.7. Yeast strains

2.1.1.7.1. *P. pastoris* X33

The X33 strain is a wild-type strain with a His⁺, Mut⁺ phenotype. This strain was obtained from laboratory stocks at Aston University but is also available from Life Technologies Inc. A de-glycosylated version of a hA_{2A}R construct (Fraser, 2006) (Chapters 4 and 5) developed by Dr. Niall Fraser, Glasgow University was transformed into the X33 strain by Dr. Richard A.J. Darby, Aston University. This strain was also the background to a green fluorescent protein (GFPuv) construct cloned into the pPICZα vector created by Dr. William J. Holmes, Aston University (Holmes et al., 2009). This X33-GFP strain was used in Chapter 4.

2.1.1.7.2. *P. pastoris* SMD1163

The genotype of this protease deficient strain is: *his4 pep4 prb1*. *PEP4* encodes for proteinase A, which is a vacuolar aspartyl protease required for the activation of other vacuolar proteases such as carboxypeptidase Y and proteinase B. The *PRB1* gene encodes for proteinase B (Cregg, 1985). This strain was used as the background for the novel hA_{2A}R construct (Chapter 5) and was a kind gift from Dr. Shweta Singh, Evotec (UK) Ltd., Abingdon, United Kingdom.

2.1.1.7.3. *S. cerevisiae* BY4741

Wild type strain BY4741 has the genotype : *MATα, ura3Δ0, leu2Δ0, met15Δ0, his3Δ1*. The vector pYX212 cloned with the hA_{2A}R construct was obtained from Dr. Renaud Wagner, Université de Strasbourg, France and was transformed into the BY4741 strain. Uracil selection was used to identify expressing colonies.

BY4741 is also used as the parental strain of an over-expression mutant developed by Dr. Nicklas Bonander, Chalmers University, Goteborg, Sweden (Bonander et al., 2005, Bonander et al., 2009). The endogenous promoter of the essential gene, *BMS1* was replaced by a tetracycline titratable promoter and hence levels of the *BMS1* gene could be regulated by the addition of doxycycline (a tetracycline derivative) (yTHCBMS1) (Bonander et al., 2009). The vector

pYX222 containing the hA_{2A}R construct was transformed into the strain. Histidine selection was used to identify expressing colonies.

2.1.1.7.4. *S. cerevisiae* KOY-TM6*

The strain used was KOY-TM6* as the parental strain which is a respiratory strain containing the gene that encodes a chimeric hexose transporter, TM6* (Otterstedt et al., 2004, Ferndahl et al., 2010). The pYX212 vector containing the hA_{2A}R construct was transformed into the strain and uracil selection was used to identify expressing colonies.

2.1.1.8. Membrane preparation reagents and materials

2.1.1.8.1. Breaking buffer pH 7.4

50 mM Na₂HPO₄, 50 mM NaH₂PO₄, 2 mM EDTA, pH 7.4, 100 mM NaCl and 5% glycerol were added to a final volume of 1 L with double distilled water, the pH adjusted and autoclaved.

2.1.1.8.2. Buffer A pH 7.0

20 mM HEPES, 50 mM NaCl and 10% glycerol were added to a final volume of 1 L with double distilled water, the pH adjusted to 7.0 and autoclaved.

2.1.1.8.3. Glass beads

Acid-washed glass beads 212 - 300 µm in size (Sigma-Aldrich) were used for small scale membrane preparations.

2.1.1.9. Protein quantification reagents

2.1.1.9.1. Bovine serum albumin (BSA) standard

BSA (Sigma) was diluted to a final amount in each well of 0-10 µg from a 1mg mL⁻¹ stock for all protein determinations.

2.1.1.9.2. Copper (II) sulphate solution

1:50 4% (w/v) copper (II) sulphate (Sigma) solution was used for the BCA assay in conjunction with BSA (Sigma).

2.1.1.9.3. Bicinchoninic acid (BCA)

BCA solution (Sigma) was used in the BCA assays.

2.1.1.10. Immunoblot reagents

2.1.1.10.1. 5 × Laemmli sample buffer

1.25 mL 0.5 M Tris-HCl pH 6.8, 1 mL 100% glycerol, 2 mL 10% SDS, 0.5 mL β -mercaptoethanol, 10 μ L bromophenol blue were mixed with double distilled water to a final volume of 8 mL.

2.1.1.10.2. SDS Tris buffer

100 mL 10× SDS Tris buffer (GeneFlow) was added to 900 mL double distilled water.

2.1.1.10.3. Immunoblot Tris buffer

100 mL 10× Tris buffer (GeneFlow) and 200 mL methanol was added to 700 mL double distilled water.

2.1.1.10.4. Phosphate buffered saline (PBS)

5 PBS tablets (Sigma) were dissolved in 1 L of double distilled water.

2.1.1.10.5. PBS-Tween buffer

2 mL Tween-20 (0.2%) was added to 1 L PBS.

2.1.1.10.6. PBS-5% milk

5 g powdered milk (generic brand) was dissolved in 100 mL PBS which is sufficient for 2 nitrocellulose membranes

2.1.1.11. Solubilisation reagents

2.1.1.11.1. n-dodecyl- β -d-maltopyranoside (DDM)

DDM (Anatrace) was diluted to 5% (w/v) with double distilled water and stored at 4°C.

2.1.1.11.2. Cholesteryl hemi-succinate (CHS)

0.5% CHS (w/v) (Sigma) solution was made with 50 mM Tris-HCl pH 8.0 and sonicated for 10 s, cooled in ice and repeated 3 times. The solution was stored at 4°C.

2.1.1.11.3. 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC)

10% DMPC (w/v) (Avanti Polar Lipids) solution was made with 50 mM Tris-HCl pH 8.0 and sonicated for 10 s cooled in ice and repeated 3 times. The solution was stored at 4°C.

2.1.1.11.4. Poly (maleic anhydride-styrene) (PMAS)

To a powder form of PMAS (Sigma), 1 M NaOH was added drop-wise, continuously stirred and heated at 80°C to dissolve the powder and also to achieve a pH of 11.0. The reaction was refluxed if the anhydride rings were difficult to open. The dissolved PMAS solutions were checked that the pH were at 11.0 prior to use and also a batch of the same PMAS were prepared at pH 7.0 by adding concentrated HCl drop-wise. The final concentration of the PMAS was 3% and was supplied by Dr. Anisa Mahomed, Aston University. Variations in molecular weight in the PMAS were supplied by the manufacturer.

2.1.1.11.5. Esterification of poly (maleic anhydride-styrene) (PMAS)

3 g of PMAS was dissolved in 6 mL methyl ethyl ketone (MEK) and heated to 60 - 70°C whilst stirring. 1.5 mL methanol was added and the solution was refluxed at 70 - 80°C with stirring for 14 h. An additional 3 mL MEK was added followed by a further 2.25 mL aliquot of methanol 15 min later. The solution was allowed to reflux for a further 6 h before being separated out using petroleum ether (60 - 80°C). The PMAS was precipitated out, filtered and dried in a vacuum oven. This was prepared and supplied by Dr. Anisa Mahomed, Aston University.

2.1.1.11.6. Styrene maleic acid (SMA)

10% SMA solution was dissolved in 1 M NaOH and stirred overnight at room temperature. The solution was refluxed for 2 h and incubated at 4°C for 48 h. The SMA was then dialysed overnight against 50 mM Tris-HCl, pH 8.0 using dialysis membranes (3500 MWCO, Thermo Scientific) to remove the NaOH. The final concentration was 2.5% at pH 8.0 and was prepared by Dr. Mohammed Jamshad, University of Birmingham. One molecular weight version of SMA was used in experiments.

2.1.1.12. Radio-ligand binding reagents

2.1.1.12.1. Binding buffer

This solution was prepared by mixing 11.3 mL 1M NaH₂PO₄, 38.7 mL 1M Na₂HPO₄ and 1 mL 0.5 M EDTA, adjusting to a final volume of 1 L (final concentration 50 mM sodium phosphate, 0.5 mM EDTA) with deionised water and adjusting to pH 7.4.

2.1.1.12.2. Gel filtration columns

P30 mini-spin gel filtration columns (Bio-Rad Laboratories Inc.) and Illustra G50 gel filtration columns (GE Healthcare) were used.

2.1.1.12.3. Tritiated ZM241385 ($[^3\text{H}]$ ZM241385)

Tritiated ZM241385 was purchased from American Radio Chemicals (ARC Inc). Dilutions were carried out using binding buffers and daughter products were stored at 4°C in a radiochemical laboratory.

2.1.1.12.4. Unlabelled ZM241385 (cold ZM241385)

Stock dilutions of cold ZM241385 (Tocris) were made with 100% DMSO and included: 0.1 mM, 10 μM , 1 μM , 0.1 μM , 100 nM, 10 nM and 0.1 nM where the final concentration of the cold ZM241385 was a 1:100 dilution of these stock solutions.

2.1.1.12.5. Unlabelled theophylline (cold theophylline)

Stock dilutions of cold theophylline (Sigma) were made with 100% DMSO and included: 1 M, 500 mM, 100 mM, 50 mM, 10 mM, 5 mM and 1 mM where the final concentration of the cold theophylline was a 1:100 dilution of these stock solutions.

2.1.1.12.6. Unlabelled NECA (cold NECA)

Stock dilutions of cold NECA (Tocris) were made with 100% DMSO and included: 1 mM, 500 μM , 100 μM , 50 μM , 10 μM , 5 μM and 1 μM where the final concentration of the cold NECA was a 1:100 dilution of these stock solutions.

2.1.1.12.7. Unlabelled XAC (cold XAC)

Stock dilutions of cold XAC (Sigma) were made with 100% DMSO and included: 10 mM, 5 mM, 1 mM, 500 μM , 100 μM , 10 μM , 5 μM and 1 μM where the final concentration of the cold XAC was a 1:100 dilution of these stock solutions.

2.1.1.12.8. Scintillant

ScintiSafe (Fisher Scientific) scintillant was used for the radio-labelled counting procedure.

2.1.1.12.9. Soluene

Soluene®-350 (Perkin-Elmer) was used as a tissue solvent to solubilise membrane fraction pellets.

2.1.1.13. Protein purification reagents

2.1.1.13.1. Purification columns

Ni-NTA agarose resin columns (Qiagen) were used according to the manufacturer's instructions.

2.1.1.13.2. Lysis buffer

This solution was prepared by mixing 11.3 mL 1M NaH₂PO₄, 38.7 mL 1M Na₂HPO₄, 1 mL 3 M NaCl and 10mL 1M imidazole (final concentration 50 mM NaH₂PO₄, 300 mM NaCl and 10 mM imidazole) adjusting to a final volume of 1 L with deionised water and adjusting to pH 8.0.

2.1.1.13.3. Wash buffer

This solution was prepared by mixing 11.3 mL 1M NaH₂PO₄, 38.7 mL 1M Na₂HPO₄, 1 mL 3 M NaCl and 20mL 1M imidazole (final concentration 50 mM NaH₂PO₄, 300 mM NaCl and 20 mM imidazole) adjusting to a final volume of 1 L with deionised water and adjusting to pH 8.0.

2.1.1.13.4. Elution buffer

This solution was prepared by mixing 11.3 mL 1M NaH₂PO₄, 38.7 mL 1M Na₂HPO₄, 1 mL 3 M NaCl and 300mL 1M imidazole (final concentration 50 mM NaH₂PO₄, 300 mM NaCl and 300 mM imidazole) adjusting to a final volume of 1 L with deionised water and adjusting to pH 8.0.

2.1.1.14. Green fluorescent protein (GFP) assay

2.1.1.14.1. Recombinant GFP

Recombinant GFP (Vector Labs) standards were made: 0, 1, 2 and 4 mg mL⁻¹.

2.1.2. Equipment

Equipment and instruments used in this research are listed.

2.1.2.1. Pilot-scale 35 L bioreactors AstraZeneca Ltd. Alderley Park, United Kingdom

35 L pilot-scale bioreactors, Biostat C (Sartorius) (Figure 2.2 and 2.3) were controlled by Supervisory Control And Data Acquisition software (SCADA) software. They were located in a dedicated fermentation facility and the parts include:

1. 35 L jacketed stainless steel culture vessel with impellers and baffles, Biostat C (Sartorius).
2. MatLab software for input of feeding rate conditions.
3. Mass spectrometer for off-gas analysis (Sartorius).
4. SCADA controller software and hardware (Sartorius).
5. Oxygen and sterile air gas supply (BOC).
6. 100 mL Schott-Duran sterile bottles for sampling.
7. Peristaltic pumps for acid, base, feed and antifoam addition (EasyLoad Masterflex).
8. Sterile filters (Sartorius, midistart 2000 0.2 μm PTFE) for tubing.
9. Dissolved oxygen (DO) probe (Broadley-James).
10. pH probe (Broadley-James).
11. pH 4.0 and 7.0 standard buffers (Sigma).
12. Clamps (Fisher).
13. Tubing for peristaltic pumps (PharMed).
14. 25% ammonium hydroxide (Sigma)
15. Concentrated sulphuric acid (Sigma)
16. P2000 antifoam (Fluka)
17. 50% methanol (Sigma)
18. 50% glycerol (Sigma)



Figure 2.2 Pilot-scale fermentation facility set-up. Image shows the Sartorius bioreactor system with several others in the fermentation dedicated facility. Main features from this image include a 35 L culture vessel, the SCADA control software and the off-gas lines that lead directly to a mass spectrometer for analysis. The pilot plant was located at AstraZeneca Ltd., Alderley Park, United Kingdom.

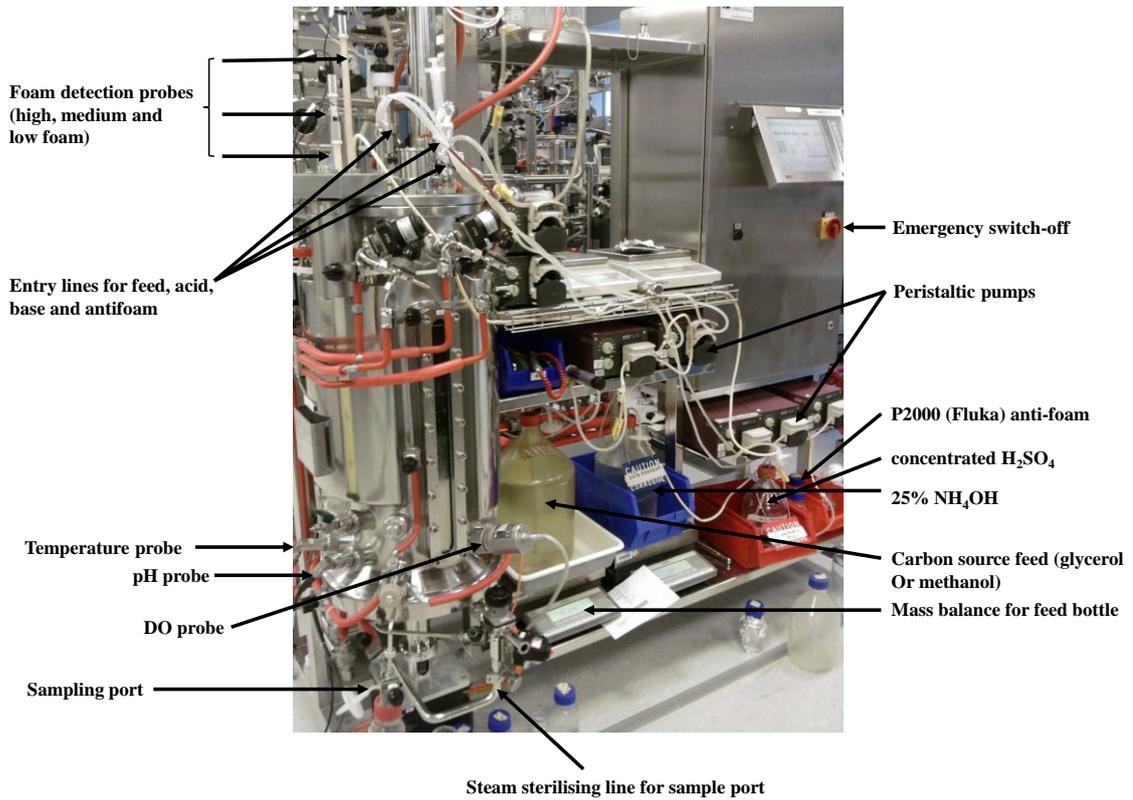


Figure 2.3 Sartorius 35 L pilot-scale bioreactor. Image shows the Sartorius 35 L bioreactor including: foam detection probes for different heights of foam production, tubing lines for feeds, acid, base and antifoam addition, temperature, pH and DO probes, a sampling port with a steam sterilising line, separate peristaltic pumps for the feed, acid, base and antifoam tubing lines. P2000 antifoam (Fluka), concentrated sulphuric acid and 25% ammonium hydroxide (acid and base) are also required. The carbon source, in this instance, the glycerol and methanol are pre-weighed before the cultivation begins and the decrease in mass as the feed is pumped into the culture is monitored and recorded by the mass balances that are connected to the output software.

2.1.2.2. Bench-top 2 L bioreactors

2 L bench-top bioreactors (Applikon) (Figure 2.4) controlled by BioXpert version 2 software and their parts include:

1. 2 L jacketed glass culture vessel (Applikon).
2. Heads plate for 2 L bioreactor with ports and impellers (Applikon).
3. BioXpert version 2 software (Applikon).
4. Thermo circulator AD 1018 (Applikon).
5. TanDem off gas analyser (Applikon).
6. Gas supply unit ADI 1026 (Applikon).
7. 60%:40% oxygen : nitrogen supply (BOC).
8. Recirculating chiller (Grant, LTL1).
9. Air compressor 75/150 (Bambi).
10. Peristaltic pumps for acid and base addition (EasyLoad Masterflex).
11. Peristaltic pumps for feed addition (Masterflex C/L).
12. Dissolved oxygen (DO) probe (Broadley-James).
13. pH probe (Applikon).
14. Temperature probe (Applikon).
15. Silicon tubing (Fisher).
16. Sterile filters (Sartorius, midistart 2000 0.2 µm PTFE) for tubing.
17. Tubing for acid/base addition (PharMed).
18. pH 4.0 and 7.0 standard buffers (Sigma).
19. Clamps (Fisher).
20. 25% ammonium hydroxide (Sigma)
21. Concentrated phosphoric acid (Sigma)
16. P2000 antifoam (Fluka)
17. 50% methanol (Sigma)

18. 50% glycerol (Sigma)
19. Optek controller (Applikon).
20. Optek probe (Applikon).
21. Tubing connectors and Y connectors (Fisher).
22. Clamps (Fisher).
23. Needles (Fisher).
24. Plastic syringes (Fisher).
25. Sample collection tubes (20 mL).
26. Tin foil (Fisher).

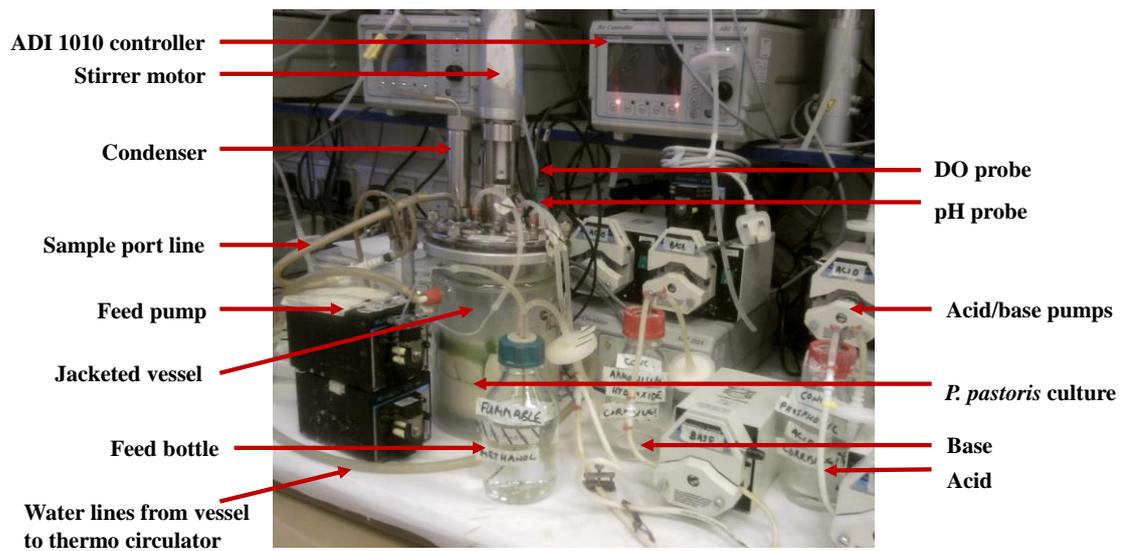


Figure 2.4 2 L Bench-top bioreactor set-up at Aston University. Photograph shows *P. pastoris* culture with methanol feeding in progress. Main pieces of associated equipment are labelled.

2.1.2.3. Micro-24 microreactor (small-scale bioreactors)

The parts required for the Micro-24 microreactor (Figure 2.5) small-scale bioreactor included:

1. Micro-24 microreactor (Pall Corporation).
2. 24 well culture plate or cassette with 10 mL culture capacity (Pall Corporation).
3. Gas supplies of oxygen, nitrogen and carbon dioxide (BOC).
4. Light duty nylon tubing for gas connections (RS components).
5. Pressure valve controlled ammonia bubbler vessel (Pall Corporation).
6. Air compressor (Jun Air).

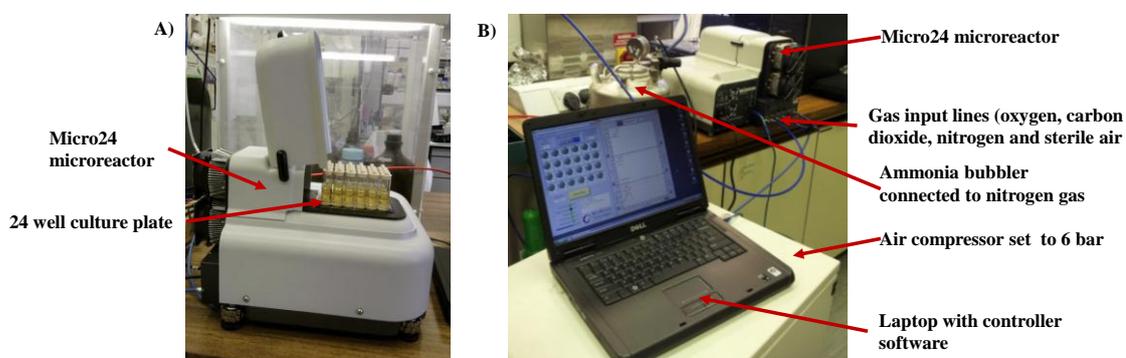


Figure 2.5 Micro-24 microreactor (Pall Corporation) bioreactor at Aston University. A) Micro-24 microreactor with a 24 well culture plate. B) The general set-up at Aston University including the controller software on the laptop, air compressor, ammonia bubbler and the gas input lines at the back of the instrument.

2.1.2.4. Centrifuges

The centrifuges used for this research were an Optima TLX bench-top ultra-centrifuge, Allegra 25R centrifuge, Avanti J-20 XP floor centrifuge and Optima XE floor ultra-centrifuge (all Beckman Coulter, Inc.)

2.1.2.5. Gas chromatograph (GC)

The GC used in this work was Thermo Scientific FOCUS Gas Chromatograph and utilised QuanLab® software for identification of peaks.

2.1.2.6. Cell lyser for small scale membrane preparations

The Tissue Lyser LT (Qiagen) was used for small scale membrane preparations.

2.1.2.7. Cell lyser for large scale membrane preparations

The Emuliflex-C3 pressure homogeniser (Avestin) was used for large scale membrane preparations.

2.1.2.8. Scintillation counter for measurement of radioactivity

A liquid scintillation counter (Packard 1600TR Liquid Scintillation Analyser) for counting tritium was used.

2.1.3. Software

Software used in this research is outlined below.

2.1.3.1. Graphpad Prism software

GraphPad Prism[®] 4 software was used primarily for radio-ligand binding analysis.

2.1.3.2. Minitab statistical software

Minitab[®] 15.1.30.0 was used primarily for Design of Experiments work.

2.1.3.3. Clone Manager

Clone Manager[®] v5.02 was used primarily for construct design work.

2.1.3.4. Origin software

OriginPro 8.5[®] software was used primarily for construction of multi-axes graphs.

2.1.3.5. Microsoft Excel

Microsoft Excel[®] 2010 was used primarily for general data analysis, table and graph construction.

2.2. Methods

2.2.1. Molecular biology techniques

2.2.1.1. Construct design and virtual cloning

The design of a novel hA_{2A}R construct was carried out with the aid of specific software programmes such as ExPASy Proteomics Server (<http://www.expasy.ch>) and a vector design

tool, Clone Manager v5.02. Once the design was confirmed to work virtually in the pPICZB vector (Life Technologies), the vector was sent to Life Technologies and the DNA sequence of the hA_{2A}R construct was synthesised by them and cloned into the vector. The vector with the cloned sequence of the construct was returned to Aston University where the molecular biology work continued. As a control, the pPICZB vector only was transformed into the SMD1163 *P. pastoris* cells using the same transformation protocol described in 2.2.1.5. to 2.2.1.8.

2.2.1.2. Mini-preparations (Minipreps) of vector DNA

Minipreps were carried out using Fermentas GeneJET plasmid miniprep kit. 5 mL over-night *E. coli* (LB zeocin (25 µg mL⁻¹) culture at 37°C were centrifuged at 3000 rpm for 5 min. The supernatant was discarded and the cells were re-suspended in 250 µL resuspension solution. 250 µL lysis solution was mixed by inverting the tube 6-8 times or until the solution became slightly viscous. 350 µL neutralisation solution was added and left for a few seconds. The solution was pipetted and a precipitation was seen. The solution was transferred to a clean tube and centrifuged at 13000 rpm for 5 min to pellet the cell debris. The supernatant was transferred to a GeneJET spin column whilst not disturbing the precipitation. The column was spun for 1 min at 13000 rpm and the flow through was discarded. 500 µL wash (containing ethanol) solution was added to the column and spun again at 13000 rpm for 1 min. A second spin was carried out to remove any traces of ethanol. 50 µL deionised water were added directly onto the column membrane and incubated at 37°C for 4 min. The column was then spun at 13000 rpm for 2 min and the eluted DNA was stored at -20°C.

2.2.1.3. DNA quantification

DNA quantification was carried out using a Nanodrop 1000 spectrophotometer (Thermo Scientific) with ND2000 software. 2 µL deionised water was loaded onto the Nanodrop platform and was measured as a calibration blank. 2 µL water was added again to the Nanodrop platform and was measured as a sample blank. Next 2 µL of the sample DNA was added to the platform and measured with the DNA-50 function (double-stranded DNA) at 260 nm.

2.2.1.4. Agarose gel electrophoresis

For one gel, 1 g high resolution agarose was dissolved in 100 mL 1× T.A.E. buffer and melted in a microwave. Once cooled, but not set, 5 µL ethidium bromide was added to the solution and mixed carefully. The gel was poured and an appropriate comb was added to form the wells. The DNA ladder was typically the 1 kb GeneRuler (Fermentas) ladder.

2.2.1.5. Linearisation of pPICZB vector-hA_{2A}R construct DNA

5 µg of vector DNA was linearised with 1.5 µL *PmeI* (25U, New England Biolabs), 20 µL 10× NEB buffer 4, 2 µL BSA (100 µg mL⁻¹) and deionised water to a final volume of 200 µL. The reaction mixture was heated to 37°C for 1 h and then the enzyme heat inactivated by incubating for a further 20 min at 65°C.

2.2.1.6. Purification of linearised pPICZB vector-hA_{2A}R construct DNA

A QIAquick purification kit was used to purify the linearised DNA. 1 mL Buffer PB (a proprietary buffer that contains a high concentration of guanidine hydrochloride and isopropanol) was added to the completed linearisation reaction and applied to a QIAquick column. The tube was centrifuged for 30-60 s at 17900 ×g in a table-top centrifuge. The flow-through was discarded and 0.75 mL Buffer PE (a proprietary buffer containing 96-100 % ethanol) was added to the same column and centrifuged for 1 min at the same centrifugal force. The column was transferred to a clean collection tube and 50 µL deionized water was added to the centre of the column and allowed to equilibrate for 1 min. The column and tube were centrifuged for a further 1 min to elute the purified linear vector-construct DNA.

2.2.1.7. Preparation of *P. pastoris* electrocompetent cells

A fresh SMD1163 *P. pastoris* colony was used to inoculate 100 mL YPD medium which was incubated overnight in a shaker incubator at 30°C. The optical density was measured at 600 nm and the culture was diluted to an OD₆₀₀ of 0.25 in a total of 400 mL YPD medium and was incubated at 30°C. The OD₆₀₀ was monitored and when it reached 1 after about 4 h incubation at 30°C, the cells were harvested by centrifugation in sterile tubes at 2000 × g for 5 min at 4°C. The cells were re-suspended in 100 mL YPD, 20 mL 1 M HEPES, pH 8.0 and 2.5 mL 1 M DTT with gentle mixing. These cells were incubated for a further 15 min at 30°C. The cells were then transferred onto ice and sterile, cold water was added to a final volume of 500 mL. The cells were harvested by centrifugation at 2000 × g for 5 min at 4°C. The cell pellet was washed with 250 mL of sterile, cold water and then centrifuged again at 2000 × g for 5 min at 4°C. The pellet was re-suspended in 20 mL cold 1 M sorbitol by gentle mixing and then the cells were harvested by centrifuging at 2000 × g for 5 min at 4°C. The pellet was re-suspended in 500 µL of cold 1 M sorbitol by gentle mixing and 40 µL aliquots were made and stored at -80°C for use within 6 months.

2.2.1.8. Electroporation and recombinant clone selection

An electroporation cuvette (Fisher) was placed on ice for at least 15 min prior to performing the transformation. 40 µL electrocompetent cells (section 2.2.1.7.) were mixed with 7.5 µL

linearised pPICZB-hA_{2A}R DNA (section 2.2.1.6.) in the electroporation cuvette by gently pipetting and then incubating on ice for 5 min. The electroporator (Eppendorf multiporator™) was set at 1800 V 15 ms pulse length, the cuvette was placed into the electroporator chamber and the electric pulse was applied once. The electroporated mixture was immediately re-suspended in 1 mL cold sorbitol and the cells were transferred to a sterile tube. The cells were allowed to recover for about 1 h at 30°C then they were harvested at 2000 × g for 10 min at 4°C. The supernatant was discarded and the pellet was re-suspended in 500 µL 1 M sorbitol and 0.5 mL YPD. The cells were transferred to 1.5 mL tubes and placed in a 30°C incubator at the lowest shaking speed for 1.5 h. The culture was plated onto YPDS plates with increasing concentrations of zeocin (100, 250, 500 and 1000 µg mL⁻¹). These plates were incubated at 30°C for up to 8 days.

2.2.1.9. Transfection of hA_{2A}R vector into human embryonic kidney cells (HEK)

HEK 293 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% v/v fetal calf serum and 1% penicillin/streptomycin in a tissue culture treated 75 cm² cell culture flask and incubated at 37°C with 5% carbon dioxide. Once 70-90% confluency was reached, the cells were seeded at 30000 cells in 200 µL per well in a 48 well plate. 24 h after the seeding, cells were transiently transfected with 1 µg vector DNA (pcDNA3.1 with MT-hA_{2A}R, Chapter 5) per well. Figure 2.6 shows the annotated pcDNA3.1 vector used. The transfection mix consisted of 4.5 µL 20 mM poly(ethylenimine) (PEI) per 1 µg vector DNA, 40 µL 5% glucose and DMEM to a total volume of 200 µL per well. DNA was first added to the appropriate volume of 5% glucose, mixed and incubated for 10 min at room temperature. The appropriate volume of PEI was added to 5% glucose, mixed and incubated for 10 min at room temperature. The PEI mixture was added to the DNA, mixed and incubated at room temperature for 20 – 30 min. DMEM was added to the required volume. The growth medium was removed and replaced with the transfection mixture. The plates were agitated and returned to the incubator. Binding assays were performed 48 h after transfection. (This protocol was performed by Dr. Sarah Routledge, Aston University).



Figure 2.6 pcDNA3.1 expression vector developed by Life Technologies Corporation. Vector contains a multiple cloning site with a choice of restriction enzymes and includes the T7 promoter region (reproduced from www.lifetechnologies.com).

2.2.2. Yeast cultivations

2.2.2.1. Shake flask cultures

Yeast cells from frozen glycerol stocks were streaked onto YPD plus zeocin (100 $\mu\text{g mL}^{-1}$) plates and incubated for 2 days at 30°C. 5 mL appropriate medium (YPD or BMGY) was then inoculated with a single yeast colony, grown to logarithmic phase and used to inoculate baffled 250 mL shake flasks containing 50 mL growth medium (YPD, BMGY or FM22) at a culture to flask volume ratio of no more than 1 in 5. Cultures were incubated at 30 °C, 220 rpm until the required OD₆₀₀ was obtained. Growth rates of yeast were determined using the equation defining a logarithmic curve:

$$y=ce^{\mu t}$$

Where c is a constant, e is the exponential, μ is the specific growth rate (h^{-1}) and t is the time (h)

The doubling time of the yeast was then calculated from the specific growth rate:

$$\text{Doubling time} = \ln(2)/\mu$$

If methanol induction was carried out in shake flasks, the required volume of log phase cells taken from a BMGY culture in order to achieve an OD₆₀₀ of 1 was calculated and the cells were centrifuged at 5000 \times g for 10 min and the pellet re-suspended in the required volume of BMMY.

2.2.2.2. Pilot-scale 35 L bioreactor cultivations

2.2.2.2.1. Preparation of 35 L bioreactor

Before the experimental runs began, several preparative steps were taken in accordance with the laboratory operating procedure from AstraZeneca Ltd. and also from the Biostat C operating manual.

2.2.2.2.1.1. Vessel pressure hold testing

To ensure the system could withstand high pressure levels safely, any pierced septa from the head-plate were replaced, the head-plate bolts were tightened with a torque wrench set to 50 Nm, the vessel side ports were tightened and all probes (DO, pH) were fitted and tightened. The test applied 1000 mBar pressure for 30 min via the controller software. The actual pressure values were noted after 5 min and the system was monitored for any leaks. At the end of the test, the actual pressure was compared to the pressure set-point after 30 min. Any deviation above 10 mBar was a system fail.

2.2.2.2.1.2. pH and DO probe calibration

The pH probe was calibrated with pH standards, 4.0 and 7.0. The DO probe was calibrated by unplugging the DO cable from the controller unit and then setting it to 0. The bioreactor was then aerated with 1 vvm of sterile air and the slope was calibrated when the DO signal was stable.

2.2.2.2.1.3. Bioreactor sterilisation

The valves on the bioreactor were positioned as shown in Figure 2.7 (closed if horizontal, open if vertical). High pressure steam was applied to the bioreactor at a temperature of 121° for 30 min. Care was taken during this process, adequate warning signs were in place and no lone working was permitted. Once the sterilisation was completed and the temperature of the system had cooled to 37°C or less, the over pressure of 200 mBar on the pressure controller was switched off and the cooling water valve condenser was opened. A final check before the bioreactor was inoculated with the *P. pastoris* culture was priming the pH control lines by inspecting the interior of the pump head for obstructions.

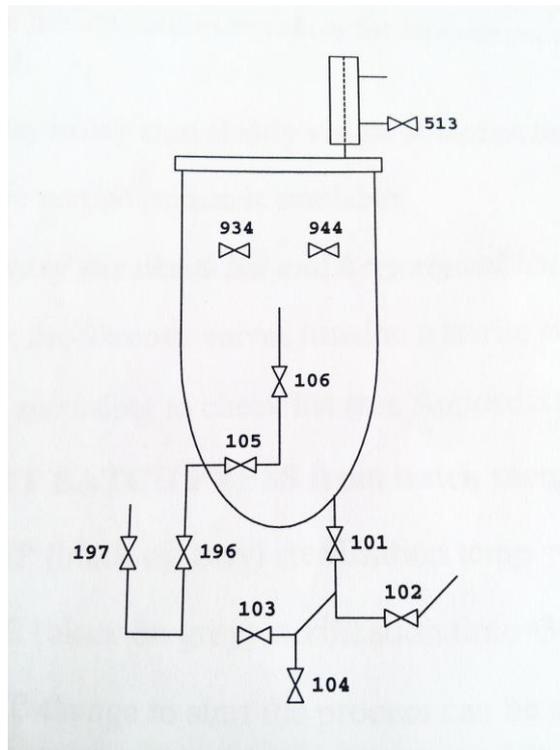


Figure 2.7 Biostat C valve settings. Valves numbered according to Biostat C manufacturing show horizontal (closed) or vertical (open) positions.

2.2.2.2.2. Performing a 35 L bioreactor cultivation for *P. pastoris* producing hA_{2A}R

10 L FM22 medium supplemented with 1 mL P2000 antifoam (Fluka), 40 mL PTM₄ salts and 8 mL 10 g L⁻¹ biotin were transferred into each of two 35 L (total volume) jacketed steel bioreactors (Biostat C, Sartorius Ltd.). Culture temperature was maintained at 30°C and pH at 5. The maximum agitation rate was 1500 rpm. X33 *P. pastoris* producing hA_{2A}R (dG-hA_{2A}R, Chapter 4) was inoculated in each bioreactor to a starting OD₆₀₀ of 1. For both cultivations, the end of the glycerol batch phase (phase I) was indicated by a spike in DO to 100%. A glycerol fed-batch phase (phase II) was then maintained for 40 h by employing an exponential feed rate of 50% aqueous glycerol (v/v) 4 g L⁻¹ h⁻¹ increasing exponentially first at a rate of 0.15 h⁻¹ for 10 h and subsequently at a rate of 0.03 h⁻¹ via the use of MatLab[®] analysis software. The first section of the transition phase (phase IIIA; starvation), during which no further carbon source was fed into the bioreactor, was maintained for 1 h. The second section of the transition phase (phase IIIB) comprised a further 1 h, where a constant methanol feedstock (50% (v/v) aqueous methanol) was applied at 8 g L⁻¹ h⁻¹. The culture temperature was lowered to 22°C and an exponential methanol (50% (v/v) aqueous methanol) feed profile was applied by exponentially increasing the feed rate at 0.01 h⁻¹ (μ_{low}) or 0.03 h⁻¹ (μ_{high}). The induction phase continued for 40 h. Table 2.1 summarises these cultivation events for both μ_{low} and μ_{high} set-ups.

Table 2.1 Summary of bioprocess events during two simultaneous *P. pastoris* cultivations producing recombinant dG - hA_{2A}R with different exponential methanol feed rates (μ_{set}). I denotes batch phase; IIA and B denotes fed-batch phase; IIIA and B denotes transition phase and IV denotes the induction phase. The μ_{high} cultivation has another induction phase, V where the μ_{set} is adjusted to 0.05 h⁻¹.

Age of Cultivation (h)	Feed Event		Phase		Temperature (°C)					
0.0	↑ Glycerol Batch (10 g) ↓		I		↑ 30 ↓					
14.4										
16.4										
18.4										
19.1										
20.4	Glycerol Fed-batch (Exponential feed at $\mu_{set} = 0.15 \text{ h}^{-1}$)		IIA		↓					
22.4										
23.6										
28.8										
38.4										
40.6	Glycerol Fed-batch (Exponential feed at $\mu_{set} = 0.03 \text{ h}^{-1}$)		IIIB		↓					
42.2										
42.9						Transition (No feed)		IIIA		↓
44.4										
44.4										
45.4										
46.4	μ_{low} ↑ Induction (Methanol exponential feed at $\mu_{set} = 0.01 \text{ h}^{-1}$) ↓		IV		↓					
47.4										
62.4										
64.4										
66.4										
68.4	μ_{high} ↑ Induction (Methanol exponential feed at $\mu_{set} = 0.03 \text{ h}^{-1}$) ↓		IV		↓					
70.4										
86.4										
86.4										
87.4										
88.9	Induction (Methanol exponential feed at $\mu_{set} = 0.05 \text{ h}^{-1}$)		V		↓					
90.4										
91.2										

Samples (1 mL) were taken for optical density measurements at 600 nm. For dry cell weight measurements, 1.5 mL culture was sampled in triplicate, placed in pre-weighed tubes and

centrifuged at $5000 \times g$ for 5 min. The supernatant was removed and stored at -20°C for residual glycerol and/or residual methanol analysis and the tubes were placed in a 100°C oven with the lids open and dried overnight. The dried tubes were then moved to a desiccator for 2-3 days and then the tubes weighed on a microbalance. The dry cell weights were calculated for g DCW per L of culture and were reported as g L^{-1}). Wet cells were stored at -80°C for membrane preparations and specific binding activity of the hA_{2A}R. Specific growth rates (μ) were calculated as described in section 2.2.2.1. Yield co-efficients were calculated according to the equation:

$$Y_{\frac{x}{s}}$$

where Y = yield co-efficient, x = biomass (g) and s = substrate (g).

2.2.2.2.3. Finishing the 35 L bioreactor run

Once the runs were completed, the bioreactor was inactivated (via steam, 30 min at 121°C) and cleaned once cooled and the pressure had been released. This was done by harvesting or draining the vessel of the culture. All the contaminated spears from the vessel head-plate and side ports were removed and prepared for autoclaving. The acid and base were reverse pumped until the lines were clear. The lines were then detached from the head-plate. All other tubing fixtures were removed. The spray ball water attachment was connected to the head-plate. 20 g of haemosol[®] was added into the vessel and 60°C water was applied. Once the vessel was $\frac{3}{4}$ full with water, the stirrer was set at 400 rpm for 1 h at 60°C . The water was drained and the vessel was checked for any residual debris.

2.2.2.3. Bench-top 2 L bioreactor cultivations

2.2.2.3.1. Preparation of the 2 L bioreactor

1 L BSM or FM22 medium was prepared and poured into the glass vessel. The head plate was attached and secured with bolts. Silicon tubes were attached to each of the ports on top of the head plate, clamped and the ends wrapped in foil. The pH probe was added after calibrating using pH 4.0 and pH 7.0 buffers. The DO and Optek (optical density measurement) probes were also connected. A $0.2 \mu\text{m}$ PTFE (polytetrafluoroethylene) gas filter was added to the inlet gas sparger and left unclamped to allow pressure equalisation and avoid vessel damage during autoclaving. A 250 mL glass sample bottle was attached to the sample port and a length of silicon tubing with a filter attached to the fork. The bioreactor was then autoclaved and sterilized at 121°C for 20 min with a slow cool cycle. A 0.5 L glass liquid addition bottle

containing 50% (v/v) phosphoric acid and a 0.5 L glass liquid addition bottle of 28 % ammonium hydroxide were prepared for pH control. PharMed[®] tubing was used and filters attached to the caps. Figure 2.4 shows the set-up at Aston University.

2.2.2.3.2. Connecting the bioreactor to the control unit

The bioreactor was removed from the autoclave and placed next to the control unit. The DO probe was connected to the ADI 1010 controller and allowed to polarize for a minimum of 6 h. The pH probe was also connected to the ADI 1010 controller and the Optek probe connected to the Optek controller. 5 mL 50% (w/v) glycerol was added to the port for the temperature probe (the thermo-well tube) and the temperature probe inserted. The chiller was turned on and connected by silicone tubing to the condenser. The silicone tubing attached to the condenser was connected to the off-gas analyser. The compressor and the 60%: 40% oxygen: nitrogen cylinder were connected to the gas supply ADI 1026 unit. The gas supply was attached to the sparger line on the bioreactor with a length of silicon tubing from the ADI 1026 unit with a filter also added. The stirrer motor was attached to the head plate. Acid and base bottles were connected to the controller via pumps. The foil was removed from the acid and base lines on the bioreactor and sprayed with 70% ethanol before connecting to the lines on the acid and base bottles. The water jacket lines were connected to the thermo circulator ADI 1018 unit. The feed bottles were also connected to the bioreactor by inserting the tubing from the bottles into a peristaltic pump and connecting the tubing to the feed line tubing on the bioreactor, spraying the line ends with 70% ethanol.

2.2.2.3.3. Performing a 2 L bioreactor cultivation for *P. pastoris* producing hA_{2A}R

Recombinant hA_{2A}R production in *P. pastoris* was carried out in 2 L bioreactors at 30 °C, pH 5.0, 30% DO and stirrer speed of 700 rpm. The temperature was reduced from 30°C to 22°C during the induction phase. The DO was initially maintained at 30% as the stirrer speed was in cascade mode of 700-1250 rpm. When the stirrer was no longer able to maintain the set-point, the mass flow controller (MFC) increased the proportion of air drawn from the 40%: 60% nitrogen: oxygen cylinder. The air flow into the bioreactor was set to 2 L min⁻¹ and the cylinder pressure was set to 2.5 Bar. The flow rate of exit gas to the off-gas analyser was set to approximately 0.4 L min⁻¹ by adjusting the clamp on the open end of the forked tubing. 5 mL PTM₁ or PTM₄ trace salts were added using a syringe and needle in a sterile manner through the septum and into the vessel. The pH was then adjusted to pH 5.0 by entering this set-point into the ADI1010 controller, causing the pumps to add the required volume of acid or base. The bioreactor was left running at these settings for approximately 1 h before inoculating.

To start the experiment, a pre-culture was prepared as described in section 2.2.2.1. and the required volume was inoculated into the 1 L medium in the vessel via the septum via aseptic technique to achieve an OD_{600} of 1. The BioXpert software was activated immediately prior to the inoculation.

2.2.2.3.4. Glycerol batch and fed batch phase in 2 L bench-top bioreactor

The cultures initially grew on the 40 g L⁻¹ glycerol present in the BSM or FM22 media. Once consumed, usually ~20 h and indicated by a DO spike, a fed batch phase was started. 50% w/v glycerol with 12 mL PTM₁ trace salts L⁻¹ was fed into the vessel at a flow rate of 12 mL h⁻¹ for 4 h for the BSM run and 14 mL h⁻¹ for 40 h for the FM22 run. The transition phase was 1-2 h after the feed was stopped and before methanol induction began.

2.2.2.3.5. Induction phase in 2 L bench-top bioreactor

Induction was begun (after starvation) using 20% v/v methanol with 12 mL PTM₁ trace salts L⁻¹. The pump was set at 2.04 mL h⁻¹ for 40 h and remained at that speed until the run ended at ~100 h. Optical density, dry cell weight and wet cell weights were monitored as in section 2.2.2.2.2.

2.2.2.4. Micro-24 microreactor cultures

2.2.2.4.1. Preparation of Micro-24 microreactor cultivations

The Micro-24 bioreactor from Pall Corporation comprises a control unit, a lap top and a 24 well plate. Figure 2.8 and 2.9 shows the set-up of the Micro-24 microreactor at Aston University and also a close up of the 24 well plate. The control unit both monitored and controlled the culture conditions of each well independently via the laptop. The 24 well plates held a working volume of 3-7 mL and in these experiments 5 mL was used. The 0.22 µm filter caps or closures for the wells used were Type D and were required for experiments that used more than 40% DO. Type A 0.22 µm filter caps or closures with airlock valves were required for experiments that used less than 40% DO. In the 24 well plates, two optical sensor spots and a sparging port were present in the bottom of each well. Culture pH and DO was monitored by the sensor spots and controlled by the sparging port with allowed oxygen, carbon dioxide and ammonium hydroxide gas (produced via a pressure vessel containing 15% ammonium hydroxide) all at constant gas cylinder pressures of 2 Bar. The culture temperature was controlled and monitored via the use of a thermo-cycler type heating element and coolant fans. It was important to have a temperature difference of no more than 2°C between adjacent wells. The vacuum was generated via the use of sterile clean dry air (CDA) compressor which was set at 6 Bar. The 24 well plate was set at 500 rpm agitation with a circular orbit of 5 mm. The Micro-24 microreactor

instrument was controlled using the Micro cellerator software installed on the laptop. These were both connected via an uninterruptable power supply (UPS).

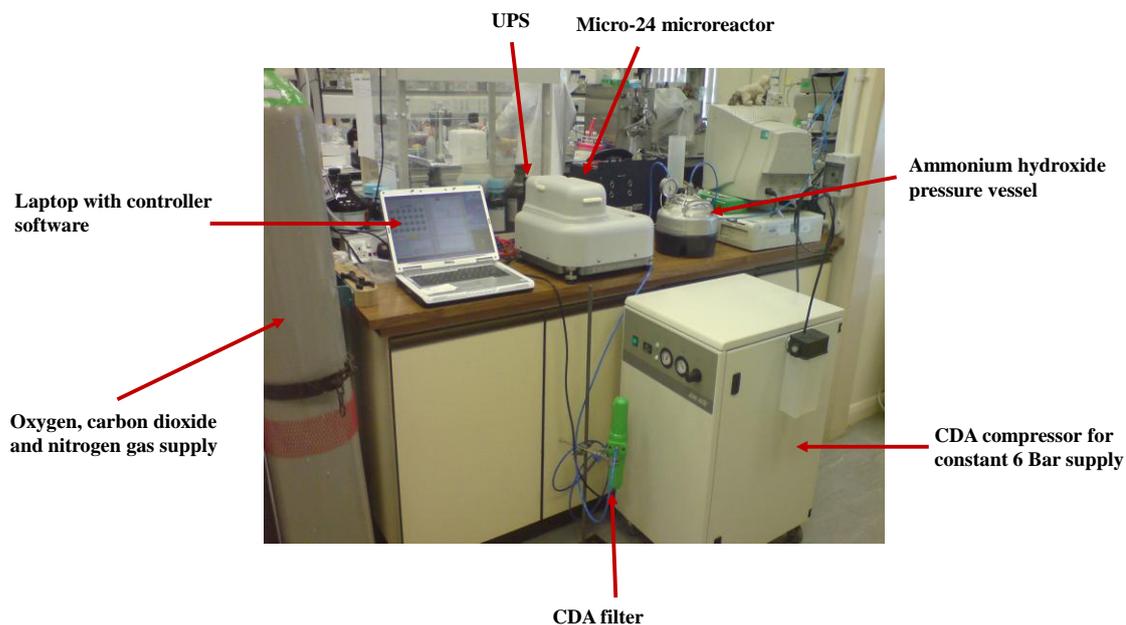


Figure 2.8 *Micro-24 microreactor set-up at Aston University. Photograph shows the components required for the Micro-24 to operate, including a laptop with controlling software, gas supplies, air compressor and a gas bubbler to generate ammonium hydroxide gas.*



Figure 2.9 *Micro-24 microreactor 24 well plate. The plate is clamped onto the shaking, vacuum platform of the instrument. The caps shown here are Type D, for use with high aeration rates (>40% DO).*

2.2.2.4.2. Antifoam and medium assessment in Micro-24 well plates

Prior to carrying out experiments, the medium and antifoam were tested in the well plates to ensure that no gas delivery membranes present in the bottom of each well would ‘wet’ out or that any blockages would occur because of the medium composition. Antifoam is required as a crucial component in fermentation processes. Foam or bubbles are often introduced during fermentation due to mixing in a system and therefore aeration. Severe foaming may introduce contamination and also reduce yields of a product. Mazu DF 204 (Sigma) was used as an antifoam agent. It is a proprietary organic agent that is widely used in the manufacturing, chemical and petroleum industries and is not harmful to micro-organisms. Figure 2.10 shows a medium and antifoam test incubation cassette set up. The test was run with 2× CBS defined medium, Mazu DF 204 antifoam and 2.5% DMSO. The Micro-24 Bioreactor was configured to provide the wells with a continuous flow of sterile CDA at 20 cm² per min and at 500 rpm orbital speed. Temperature, pH and DO control were switched off. This test was run for 10 min and observations were made for sufficient bubbling in the wells and also the bottom of the incubation cassette was checked for any wetness.

	1	2	3	4	5	6
A	2XCBS 200ppm Mazu DF 204	2XCBS 400ppm Mazu DF 204	2XCBS 600ppm Mazu DF 204	2XCBS 800ppm Mazu DF 204	2XCBS 1200ppm Mazu DF 204	2XCBS 1600ppm Mazu DF 204
B	2XCBS + 2.5% DMSO 200ppm Mazu DF 204	2XCBS + 2.5% DMSO 400ppm Mazu DF 204	2XCBS + 2.5% DMSO 600ppm Mazu DF 204	2XCBS + 2.5% DMSO 800ppm Mazu DF 204	2XCBS + 2.5% DMSO 1200ppm Mazu DF 204	2XCBS + 2.5% DMSO 1600ppm Mazu DF 204
C						
D						

Figure 2.10 Antifoam and media test set up of the Micro-24 microreactor 24 well plate.

The test showed that there was no detrimental effect of the medium and antifoam. It was decided that 600 ppm of Mazu DF 204 would be used for all proceeding experiments.

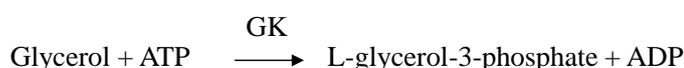
2.2.2.4.3. Performing the Micro-24 microreactor cultivations

The run conditions were programmed into the Micro-24 control software according to the DoE input factors required (temperature, pH and DO). A fourth input factor was also investigated, the presence and absence of 2.5% DMSO. The medium composition for wells that required DMSO was 111.78 mL 2× CBS, 0.125 mL 100% DMSO, 7.5 mL 40× glucose, 30 mL 10× drop out solution (for *S.cerevisiae* wild type and TM6* strains 10× drop out solution without uracil was used and for the *S.cerevisiae* BMS1 strain, 10× drop out solution without histidine was used), 0.3 mL trace elements solution and 0.3 mL vitamin solution. 5 mL of this medium was aliquotted into wells where DMSO was to be added. The medium composition for wells where that did not require DMSO was 111.90 mL 2× CBS, 7.5 mL 40× glucose, 30 mL 10× drop out solution (composition as above), 0.3 mL trace elements solutions and 0.3 mL vitamins solution. 5 mL of this medium was aliquotted into wells where DMSO was not present. The PI settings were altered to minimise oscillations after 2 h of culture growth. The PI settings that were used

for all yeast strains were pH (P = 8, I = 2) and DO (P = 2, I = 20). Cultures were grown and glucose levels were monitored via the use of Accu-Chek Active glucose analyser (Roche Diagnostics). The run was stopped when the glucose levels were between 10–30 mM for each well. The 5 mL cultures were centrifuged at $5000 \times g$, 4°C for 5 min. The cell pellets were frozen in liquid nitrogen and stored at -80°C for subsequent membrane fraction preparations.

2.2.3. Glycerol assay

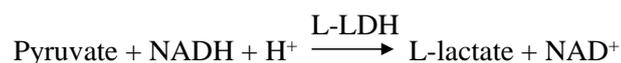
Residual glycerol in the culture supernatants was analysed using a glycerol quantitation kit (r-biopharm, Roche). The basis of the kit is that glycerol is phosphorylated by adenosine-5'-triphosphate (ATP) to L-glycerol-3-phosphate by glycerokinase (GK):



The adenosine-5'-diphosphate (ADP) formed in the above reaction is then re-converted into ATP by phosphoenolpyruvate (PEP) with the aid of pyruvate kinase (PK) and the formation of pyruvate:



In the presence of L-lactate dehydrogenase (L-LDH), pyruvate is reduced to L-lactate by reduced nicotinamide-adenine dinucleotide (NADH) and the oxidation of NADH to NAD^+ :



The amount of NADH oxidised to NAD^+ is stoichiometric with respect to the glycerol phosphorylated by GK and is measured by light absorption at 334, 340 or 365 nm. The kit includes 4 bottles: bottle 1; glycylglycine buffer (pH 7.4), 7 mg NADH, 22 mg ATP, 11 mg PEP-CHA, magnesium sulphate, bottle 2; 240 U of pyruvate kinase and 220U of L-lactate dehydrogenase, bottle 3; 34 U glycerokinase solution and bottle 4; glycerol assay positive control solution (known amount of glycerol). All samples to be tested were placed at 80°C for 15 min to de-activate any naturally-occurring enzymes in the sample. Bottle 1 contents were dissolved into 11 mL distilled water. The bottle was left to stand for 10 min at 25°C . For the blank sample, in a plastic disposable cuvette, 333.3 μL of bottle 1, 666.6 μL of water and 3.3 μL of bottle 2 were added, covered with parafilm, mixed by inverting and left to stand for 7 min. The absorbance at 340 nm was recorded and constituted the Absorbance 1 (A1) reading.

Then 3.3 μL bottle 3 was added to the same cuvette, mixed and left to stand for 10 min. The absorbance was recorded at 340 nm, which constituted Absorbance 2 (A2) reading. For the control or sample measurement, 33.3 μL bottle 4 was included in the first part of the reaction. The following equation was used to calculate the concentration of glycerol:

$$c = (V \times MW / \epsilon \times d \times v \ 1000) \times \Delta A$$

Where:

c = concentration of glycerol (g/L)

V = final volume (mL)

MW = molecular weight of glycerol (g/mol)

ϵ = extinction coefficient of NADH at 340 nm (6.3 L x per mmol x per cm)

d = light path (cm)

v = sample volume (mL)

$$\Delta A = (A1 - A2)_{\text{sample}} - (A1 - A2)_{\text{blank}}$$

2.2.4. Methanol assay

Residual methanol in culture supernatants was analysed using a Thermo Scientific FOCUS Gas Chromatograph. Appropriately-diluted culture supernatants (typical dilutions were 1:100) were filter sterilised with syringe filter to remove any debris. 1 μL of each sample was injected onto the column and the methanol peak was integrated using QuanLab[®] software. The sample was repeated to achieve a duplicate reading and was followed by a blank sample (filter-sterilised water) injection before the next sample was loaded. The mean value of the integrated peak area was used to estimate the residual methanol within the sample by comparison with methanol standards (0, 0.013, 0.065, 0.13, 0.25, 0.38, 0.5, 1, 2, 5 and 10%) and a standard curve (Figure 2.11).

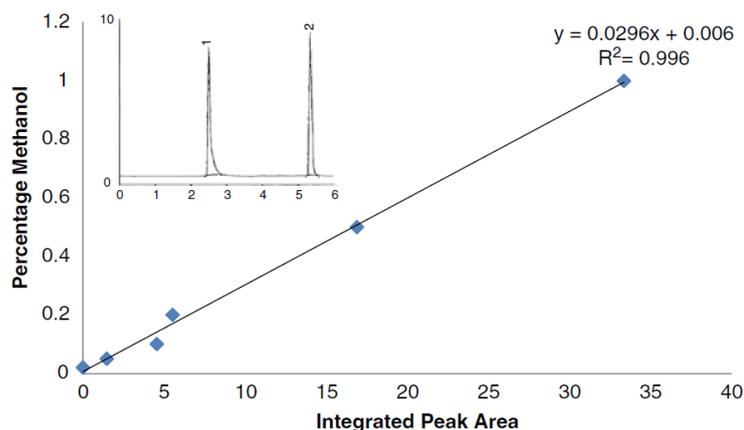


Figure 2.11 A representative calibration curve for determining the concentration of methanol. Inset is a representative gas chromatograph for one of the methanol standards from which the integrated peak areas are calculated (Bawa et al, 2012).

2.2.5. Green Fluorescent Protein (GFP) assay

100 μL supernatant was added to black 96 well plates followed by 50 μL 1M potassium phosphate buffer, pH 8.0. A recombinant GFP standard was used (Vector Laboratories); the standard curve ranged from 5 mg mL^{-1} to 20 mg mL^{-1} . The relative fluorescence units (RFU) were determined using a Spectramax Gemini plate reader with λ_{exe} of 397 nm and λ_{em} of 506 nm all measured at 25°C.

2.2.6. Small scale membrane preparations

Since yeast cell pellets retrieved from Micro-24 microreactor wells were relatively small (between 50-300 mg) it was appropriate to use the glass bead agitation method. Each pellet was mixed with 500 μL ice cold breaking buffer and 500 μL acid washed glass beads (Sigma) in a 2 mL breaking tube. A 1:2000 dilution of a protease inhibitor cocktail IV set (Calbiochem) was added to each tube. The caps of the tubes were secured and placed on ice. The tubes were agitated in a chilled Tissue Lyser LT (Qiagen) at 50 Hz for 10 min. The supernatant was removed from the tube by carefully creating a small hole at the bottom of the breaking tube with a sterile needle. The tubes were placed in larger 15 mL Falcon tubes with an adapted lid to hold the breaking tube in place. The tubes were centrifuged at $5000 \times g$ for 3 min. The recovered material was transferred to a new 1.5 mL Eppendorf tube and centrifuged at $15000 \times g$ for 15 min. The resultant supernatant was transferred to 1 mL ultra-centrifuge tubes and centrifuged at $100000 \times g$ for 1 h at 4°C. The supernatant was discarded and the pellet was re-suspended with 100 μL ice cold buffer A. The samples were stored at -20°C.

2.2.7. Large scale membrane preparations

For large scale membrane preparations, when at least 20 g of culture pellet was available for processing (typically from 35 L and 2 L bioreactors), homogenising pressure was applied instead of agitation to break the yeast cell walls. The cells were re-suspended in ice cold breaking buffer at a ratio of 2:1 buffer to cells. A 1:2000 dilution of a protease inhibitor cocktail IV set was added to cells and buffer. The cells were passed through an Emulsiflex-C3 cell disrupter (Avestin) fitted with a chilled heat exchanger for 20 min at a homogenising pressure of 30000 psi. Figure 2.12 shows the Emulsiflex-C3 set-up at Aston University. The cells were observed under a light microscope to check the extent of the cell breakage. Typically, more than 90% of the cells were disrupted. The sample was centrifuged at $10000 \times g$ for 30 min to remove the unbroken cells and cellular debris. The supernatant was transferred to ultra-centrifuge tubes and centrifuged at $100000 \times g$ for 1 h. The pellet was re-suspended in ice cold buffer A using a glass homogeniser at a ratio of 10 mL per gram of pellet. The samples were stored at -80°C .

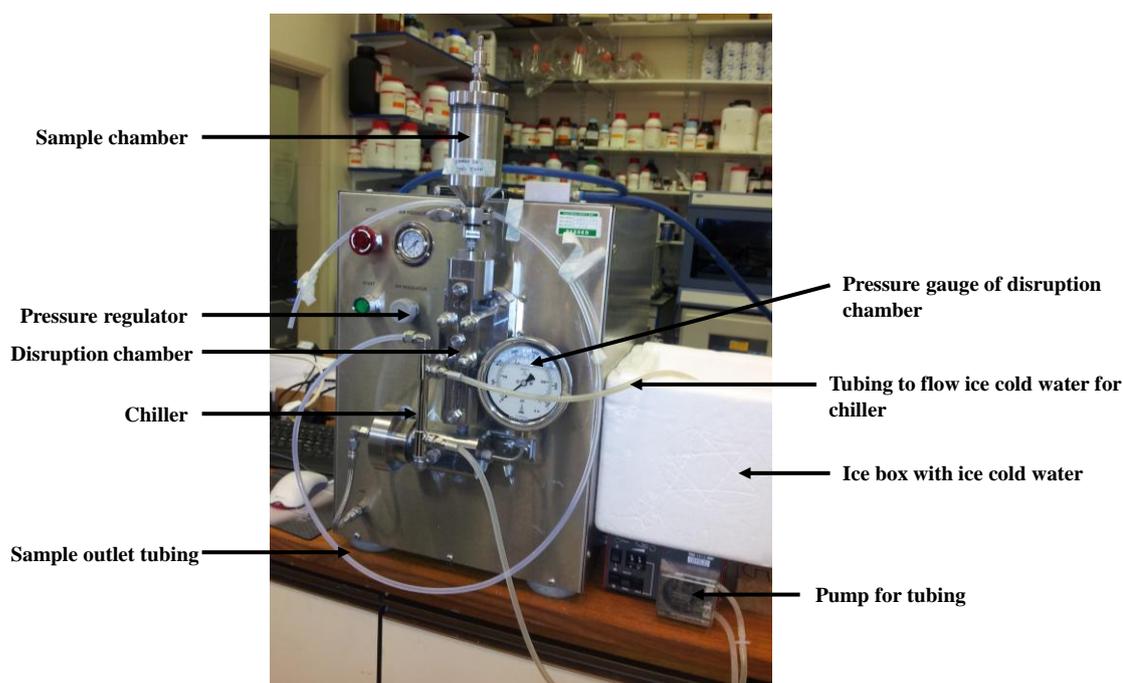


Figure 2.12 Emulsiflex-C3 (Avestin) pressure homogeniser. Photograph shows the set-up at Aston University and the associated parts including an ice box with ice-water mix for the chiller and external pump to pump chilled water for the sample chiller.

2.2.8. Protein quantification

Total protein quantification was determined by using a bicinchoninic acid (BCA) assay. A 1:50 mixture of BCA : copper (II) sulphate solution (4.9 mL : 0.1 mL) and bovine serum albumin standards were prepared (0-10 μg). The membrane fraction sample to be analysed was diluted appropriately (typically a dilution of 1:100 was made) and added to the BCA : copper (II)

sulphate solution. The final volume of either the standard or sample was 200 μ L. This was pipetted into a clear, flat-bottomed 96 well plate. After thorough mixing of the wells, the plate was incubated in the dark at 37°C for 15 min. The absorbance was read with a plate reader (BioTek Instruments) at 570 nm and the protein concentration was determined by plotting a standard curve of the BSA standards and extrapolating the membrane fraction value from the straight line equation ($y = mx + c$). Figure 2.13 shows an example standard curve generated from a BCA assay.

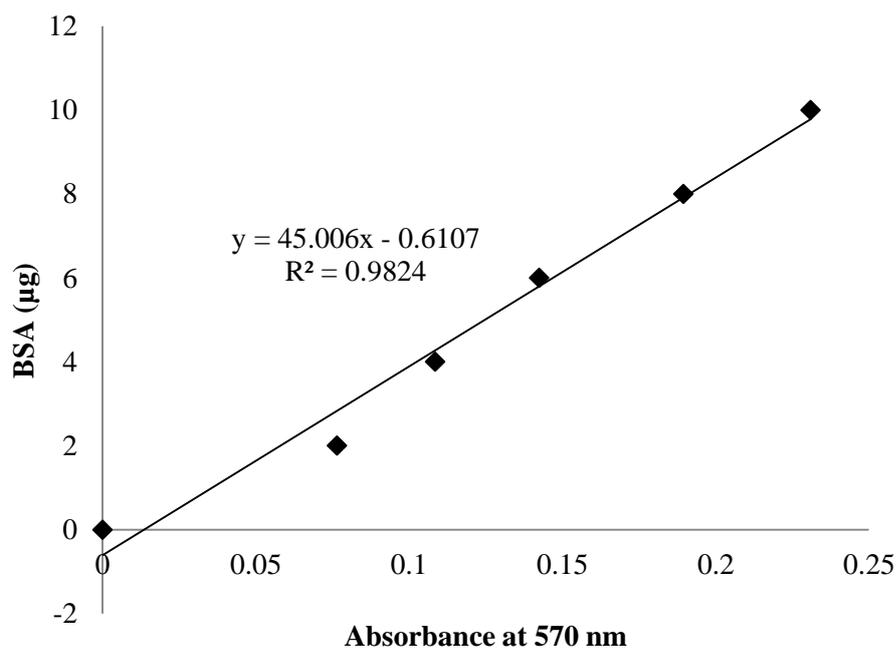


Figure 2.13 *BCA standard curve.* BCA reactions measured at 570 nm generated with BSA standards at 0, 2, 4, 6, 8 and 10 μ g. Data plotted and a straight line equation ($y = mx + c$) generated to derive unknown samples.

2.2.9. Immunoblot analysis

2.2.9.1. Sodium dodecyl sulphate polyacrylamide electrophoresis (SDS PAGE) gels

Sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) gels contained 10% polyacrylamide in the separating gel and 4% polyacrylamide in stacking gel. Table 2.2 and 2.3 show the components of these gels.

Table 2.2 10% separating gel. The APS was made fresh each time and the TEMED was always added last. Water saturated butanol was added on top of the gel to level it. It was washed off with distilled water and the space between the two glass plates were dried with filter paper.

10% separating gel	Per gel
30% polyacrylamide	1.9 mL
Water	2.2 mL
Tris-HCl 1.5M, pH 8.8	1.5 mL
10% SDS	60 μ L
20% ammonium persulphate (APS)	20 μ L
TEMED	4.5 μ L

Table 2.3 4% stacking gel. The APS was made fresh each time and the TEMED was always added last. A gel comb was added and the gel was left to set.

4% stacking gel	Per gel
30% polyacrylamide	0.3 mL
Water	1.5 mL
Tris-HCl 0.5M, pH 6.8	0.6 mL
10% SDS	25 μ L
20% ammonium persulphate (APS)	10 μ L
TEMED	2.5 μ L

Up to 15 μ L of sample was mixed with 5 μ L SDS Sample Buffer with β -mercaptoethanol. This was heated to 50°C for 5 min prior to loading onto the gel. 5 μ L of a protein marker (PageRuler pre-stained protein ladder, Fermentas) was also loaded in one well. The marker ranged from 10 kDa to 250 kDa. The gels were run in 1 \times SDS/tris/glycine running buffer (GeneFlow) at 100 V until the dye front reached the bottom of the gel.

2.2.9.2. Immunoblots

Once the SDS gel had run, the stacking gel was removed and discarded from the separating gel. The separating gel was placed into 1 \times tris/glycine/methanol transfer buffer (GeneFlow). Nitrocellulose transfer membrane (Whatman PROTRAN) and 6 sheets of filter paper (Whatman 3 mm chromatography paper) were cut to the same size as the separating gel. A fibre pad (pre-soaked in transfer buffer) and three filter papers were placed on the black side of a Bio-Rad Easy Lock Cassette. The separating gel was carefully placed on top of the filter papers followed

by the nitrocellulose membrane and the remaining three filter papers. Air bubbles were removed by rolling a plastic serological pipette over the nitrocellulose sandwich. A final fibre pad was added and the cassette was locked. The cassette was inserted into an electrophoretic blotting cell. An ice cooling unit was also added to the cell and filled with the transfer buffer. It was important to have the nitrocellulose membrane facing the anode of the electrophoretic cell. The transfer ran at 100 V for 1 h.

Once the transfer was complete, the nitrocellulose membrane was blocked with an aqueous solution of 5% dried milk in PBS for 1 h at room temperature on a rocker set at 20 rpm. A primary antibody, 6x His Monoclonal antibody (Serotec) at a dilution of 1:5000 was added to the 5% blocking solution and nitrocellulose membrane. This was placed at 4°C overnight. The membrane was washed twice in PBS 0.2% Tween-20 for 5 min. A secondary antibody, Anti-mouse IgG (Sigma) was added to a 5% blocking solution for 1 h at room temperature on a rocker set at 20 rpm. The membrane was washed twice with PBS 0.2% Tween-20 for 5 min. EZ-ECL Chemiluminescence (Geneflow) solution was used following the manufacturer's instructions. The membranes were air dried and exposed using a CCD camera (Uvitec) for 25 min and the image files saved.

2.2.10. Yeast membrane solubilisation with DDM

Solubilisation with n-dodecyl- β -D-maltoside (DDM) was performed with the following solubilisation buffer: 20 mM HEPES, pH 7.4, 50% glycerol, 250 mM NaCl, 1 μ L protease inhibitors, 5% (w/v) DDM and 0.5% (w/v) cholesteryl hemi-succinate (CHS). The membrane fraction added to the solubilisation buffer was 1:1 where the starting concentration was $\sim 20\text{mg mL}^{-1}$ and therefore the final concentration in the solubilisation was $\sim 10\text{mg mL}^{-1}$. After incubation with slow rotation at 4°C for 4 h, the sample was centrifuged at $100000 \times g$ for 1 h. The supernatant contained DDM solubilised hA_{2A}R and was stored at 4°C for further studies for a maximum of 3 days.

2.2.11. Yeast membrane solubilisation with styrene maleic acid (SMA) co-polymer

Solubilisation with SMA polymer (supplied by Dr. Yu-pin Lin and Dr. Mohammed Jamshad, University of Birmingham) was carried out in the following solubilisation buffer: 10% glycerol, 500 mM NaCl, 2.5% (w/v) SMA (prepared as in section 2.1.1.11.6.), 1% (w/v) DMPC and 50 mM Tris-HCl, pH 8 with a final protein concentration between 20 - 40 mg mL^{-1} . The slurry was stirred for 2 h at room temperature and then centrifuged at $100000 \times g$ for 1 h. The supernatant contained the styrene maleic acid lipid particle (SMALP) solubilised hA_{2A}R and was stored for no more than 3 days for further studies.

2.2.12. Yeast membrane solubilisation with poly (maleic anhydride-styrene) (PMAS)

Solubilisation with PMAS polymers (supplied by Dr. Anisa Mahomed, Aston University) was prepared as in section 2.1.1.11.4. and 2.1.1.11.5. The initial solubilisations were carried out as in a solubilisation buffer similar to that in 2.2.11. This was termed a ‘standard’ protocol and was compared to a protocol which only had the membrane fraction, the PMAS and the DMPC; this was termed the ‘simple’ protocol. Both versions were tested for the highest specific binding activity (Chapter 5) of hA_{2A}R. The effect of temperature (room temperature or 4°C) and the time of incubation (4 h or over-night) were also considered for the solubilisation protocol. It was found that the optimal protocol was adding a final concentration of 10-15 mg mL⁻¹ total protein to 1 % DMPC (final concentration) and 2.3 % PMAS (final concentration). The solubilisation mixture was placed on a rocker at 20 rpm at room temperature and rocked for at 16 h or over-night. This was then centrifuged at 100000 x g for 1 h and the supernatant contained the PMAS solubilised hA_{2A}R and was stored at 4°C for no more than 3 days for further studies.

2.2.13. Radio-ligand binding for hA_{2A}R

The radio-ligand binding assay is based on the use of a radioactively labelled compound which interacts with a receptor. Tritium [³H] is the most common isotope used to label these compounds due to its long half-life (12.3 years) (Lucas and Unterweger, 2000). All radio-ligand binding assays were performed using [³H]ZM241385 (ARC), a hA_{2A}R antagonist (Palmer et al., 1995).

The binding buffer used was 50 mM sodium phosphate, 0.5 M EDTA, pH 7.4 for the all types of radio-ligand binding assays performed. For all reactions, the tubes were mixed by inversion and incubated in a water bath at 30°C for 2 h. For the membrane bound reactions, the tubes were centrifuged at 14000 rpm for 5 min. The supernatant was discarded and the tube washed with tap water twice. The tubes were left open and inverted on a rack to dry for 1-2 h. 100 µL solouene was then added to each tube and placed in a 50°C oven for 2 h to dissolve the pellet. This was transferred to a scintillant vial where 4 mL of scintillant solution was added, mixed thoroughly and placed on the scintillation counter (Packard 1600TR Liquid Scintillation Analyser) for counting.

To analyse solubilised supernatants, 50 µL of binding reaction was applied to a Micro-BioSpin chromatography column (BioRad) and centrifuged at 3000 rpm for 2 min. These columns were previously equilibrated with 2 × washes of 500 µL binding buffer and centrifuged at 3000 rpm for 1 min. The eluate was transferred to a scintillant vial where 4 mL scintillant solution was

added, mixed thoroughly and placed on the scintillation counter (Packard 1600TR Liquid Scintillation Analyser) for counting.

2.2.13.1. Single-point binding assays for hA_{2A}R

These assays were used as the initial test for correctly folded hA_{2A}R to see if any binding was present. A single, high concentration, usually the concentration at the top end of the a saturation binding curve of [³H]ZM241385 (10 nM) was used to calculate specific binding activity (non-specific binding – total binding = specific binding). Table 2.4 shows the set-up for either membrane bound or solubilised hA_{2A}R single-point binding reactions.

Table 2.4 Single-point binding assay reaction set-up. Single-point assay for either membrane bound or solubilisation supernatant with the components of the total and non-specific reaction which include the tritiated ZM241385, adenosine deaminase, binding buffer, the amount of membrane or volume of solubilisation supernatant and unlabelled ZM241385 for the non-specific reactions.

Membrane (µg) or supernatant from solubilisation (µL)	Total binding reaction			Non-specific binding reaction			
	Final [³ H]ZM241385 concentration (nM)	Adenosine deaminase (U)	Binding buffer (µL)	Final [³ H]ZM241385 concentration (nM)	Adenosine deaminase (U)	Unlabelled ZM241385 concentration (µM)	Binding buffer (µL)
100 µg	10	0.1	up to 500	10	0.1	1	up to 500
50 µL	10	0.1	up to 200	10	0.1	1	up to 200

2.2.13.2. Saturation binding assays for hA_{2A}R

A saturation binding assay was used to measure specific binding (by subtracting the total binding from the non-specific binding) at various concentrations of [³H]ZM241385. The assay results in a curve. Non-linear regression is used to calculate the maximum number of binding sites that can be occupied (B_{max}) and also the equilibrium dissociation constant (K_d and pK_d (-log(K_d)); (Hulme and Trevethick, 2010). Table 2.5 and 2.6 show the reaction set-up for either membrane bound or for the solubilised supernatant saturation binding curves. The experimental method was as described in section 2.2.13.

Table 2.5 Saturation binding assay reaction set-up for membrane bound hA_{2A}R. Saturation binding assay for membrane bound hA_{2A}R with the components of the total and non-specific reaction which include the tritiated ZM241385 at different concentrations, adenosine deaminase, binding buffer, the amount of membrane or volume of solubilisation supernatant and unlabelled ZM241385 at one concentration for the non-specific reactions.

Membrane (μg)	Total binding reaction			Non-specific binding reaction			
	Final [³ H]ZM241385 concentration (nM)	Adenosine deaminase (U)	Binding buffer (μL)	Final [³ H]ZM241385 concentration (nM)	Adenosine deaminase (U)	Unlabelled ZM241385 concentration (μM)	Binding buffer (μL)
100	0.05	0.1	up to 500	0.05	0.1	1	up to 500
100	0.1	0.1	up to 500	0.1	0.1	1	up to 500
100	0.5	0.1	up to 500	0.5	0.1	1	up to 500
100	1	0.1	up to 500	1	0.1	1	up to 500
100	5	0.1	up to 500	5	0.1	1	up to 500
100	10	0.1	up to 500	10	0.1	1	up to 500

Table 2.6 Saturation binding assay reaction set-up for solubilised hA_{2A}R. Saturation binding assay for solubilised hA_{2A}R with the components of the total and non-specific reaction which include the tritiated ZM241385 at different concentrations, adenosine deaminase, binding buffer, the amount of membrane or volume of solubilisation supernatant and unlabelled ZM241385 at one concentration for the non-specific reactions.

Supernatant from solubilisation (μL)	Total binding reaction			Non-specific binding reaction			
	Final [³ H]ZM241385 concentration (nM)	Adenosine deaminase (U)	Binding buffer (μL)	Final [³ H]ZM241385 concentration (nM)	Adenosine deaminase (U)	Unlabelled ZM241385 concentration (μM)	Binding buffer (μL)
50	0.05	0.1	up to 200	0.05	0.1	1	up to 200
50	0.1	0.1	up to 200	0.1	0.1	1	up to 200
50	0.5	0.1	up to 200	0.5	0.1	1	up to 200
50	1	0.1	up to 200	1	0.1	1	up to 200
50	5	0.1	up to 200	5	0.1	1	up to 200
50	10	0.1	up to 200	10	0.1	1	up to 200

2.2.13.3. Competition binding assays for hA_{2A}R

A competition binding assay is used to measure specific binding (by subtracting the total binding from the non-specific binding) at various concentrations of unlabelled ZM241385 or other antagonists or agonists whilst the [³H]ZM241385 remains constant at a low concentration, in this instance, 2 nM. When a homologous competitive binding curve was carried out, the Cheng-Prusoff equation was applied to a one site binding model on the curve generated. From this equation the K_d and hence the pK_d could be calculated of the competitor:

$$K_d = K_i = EC_{50} - [\text{Radio-ligand}]$$

where K_i = dissociation constant and is equivalent to K_d for homologous curves; EC₅₀ = half the effective concentration of the unlabelled ZM241385 and [Radio-ligand] = concentration of the [³H]ZM241385.

When a heterologous competitive binding curve was carried out, the Cheng-Prusoff equation was again applied to a one site binding model on the curve generated. From this equation the K_i and hence the pK_i could be calculated of the competitor:

$$K_i = \frac{EC_{50}}{1 + [\text{Radioligand}]/K_d}$$

where K_i = dissociation constant; EC_{50} = half the effective concentration of the unlabelled agonist or antagonist to $hA_{2A}R$ and $[\text{Radio-ligand}]$ = concentration of the $[^3H]ZM241385$ (Motulsky, 1995).

These assays allowed the pharmacological characterisation of $hA_{2A}R$ after solubilisation with various polymers (Chapter 5). The agonist, NECA and the antagonists, theophylline, XAC and ZM2413835 were assayed. Table 2.7, 2.8, 2.9 and 2.10 shows the competition binding reaction set-up for each drug (ZM241385, theophylline, NECA and XAC, respectively). The experimental methods are described in section 2.2.13.

Table 2.7 Competition binding assay reaction set-up for membrane bound or solubilised $hA_{2A}R$ with ZM241385. Saturation binding assay for either membrane bound or solubilised $hA_{2A}R$ with the components of the total and non-specific reaction which include tritiated ZM241385 at one concentration of 2 nM, adenosine deaminase, binding buffer, the amount of membrane or volume of solubilisation supernatant and unlabelled ZM241385 at various concentration for the non-specific reactions.

Membrane (μg) Solubilised (μL)	Final $[^3\text{H}]ZM241385$ concentration (nM)	Adenosine deaminase (U)	Unlabelled ZM241385 concentration (nM)	Binding buffer 500 μL for membrane, 200 μL for solubilised
100 μg or 50 μL	2	0.1	1000	up to 500 or 200
100 μg or 50 μL	2	0.1	100	up to 500 or 200
100 μg or 50 μL	2	0.1	10	up to 500 or 200
100 μg or 50 μL	2	0.1	1	up to 500 or 200
100 μg or 50 μL	2	0.1	0.1	up to 500 or 200
100 μg or 50 μL	2	0.1	0.001	up to 500 or 200

Table 2.8 Competition binding assay reaction set-up for membrane bound or solubilised $hA_{2A}R$ with theophylline. Saturation binding assay for either membrane bound or solubilised $hA_{2A}R$ with the components of the total and non-specific reaction which include tritiated ZM241385 at one concentration of 2 nM, adenosine deaminase, binding buffer, the amount of membrane or volume of solubilisation supernatant and unlabelled theophylline at various concentration for the non-specific reactions.

Membrane (μg) Solubilised (μL)	Final [^3H]ZM241385 concentration (nM)	Adenosine deaminase (U)	Unlabelled theophylline concentration (mM)	Binding buffer 500 μL for membrane, 200 μL for solubilised
100 μg or 50 μL	2	0.1	10	up to 500 or 200
100 μg or 50 μL	2	0.1	5	up to 500 or 200
100 μg or 50 μL	2	0.1	1	up to 500 or 200
100 μg or 50 μL	2	0.1	0.5	up to 500 or 200
100 μg or 50 μL	2	0.1	0.1	up to 500 or 200
100 μg or 50 μL	2	0.1	0.05	up to 500 or 200
100 μg or 50 μL	2	0.1	0.01	up to 500 or 200

Table 2.9 Competition binding assay reaction set-up for membrane bound or solubilised $hA_{2A}R$ with NECA. Saturation binding assay for either membrane bound or solubilised $hA_{2A}R$ with the components of the total and non-specific reaction which include tritiated ZM241385 at one concentration of 2 nM, adenosine deaminase, binding buffer, the amount of membrane or volume of solubilisation supernatant and unlabelled NECA at various concentration for the non-specific reactions.

Membrane (μg) Solubilised (μL)	Final [^3H]ZM241385 concentration (nM)	Adenosine deaminase (U)	Unlabelled NECA concentration (μM)	Binding buffer 500 μL for membrane, 200 μL for solubilised
100 μg or 50 μL	2	0.1	10	up to 500 or 200
100 μg or 50 μL	2	0.1	5	up to 500 or 200
100 μg or 50 μL	2	0.1	1	up to 500 or 200
100 μg or 50 μL	2	0.1	0.5	up to 500 or 200
100 μg or 50 μL	2	0.1	0.1	up to 500 or 200
100 μg or 50 μL	2	0.1	0.05	up to 500 or 200
100 μg or 50 μL	2	0.1	0.01	up to 500 or 200

Table 2.10 Competition binding assay reaction set-up for membrane bound or solubilised hA_{2A}R with XAC. Saturation binding assay for either membrane bound or solubilised hA_{2A}R with the components of the total and non-specific reaction which include tritiated ZM241385 at one concentration of 2 nM, adenosine deaminase, binding buffer, the amount of membrane or volume of solubilisation supernatant and unlabelled XAC at various concentration for the non-specific reactions.

Membrane (µg) Solubilised (µL)	Final [³ H]ZM241385 concentration (nM)	Adenosine deaminase (U)	Unlabelled XAC concentration (µM)	Binding buffer 500 µL for membrane, 200 µL for solubilised
100 µg or 50 µL	2	0.1	100	up to 500 or 200
100 µg or 50 µL	2	0.1	50	up to 500 or 200
100 µg or 50 µL	2	0.1	10	up to 500 or 200
100 µg or 50 µL	2	0.1	5	up to 500 or 200
100 µg or 50 µL	2	0.1	1	up to 500 or 200
100 µg or 50 µL	2	0.1	0.1	up to 500 or 200
100 µg or 50 µL	2	0.1	0.05	up to 500 or 200
100 µg or 50 µL	2	0.1	0.01	up to 500 or 200

2.2.13.4. Competition binding curve on membrane bound hA_{2A}R from HEK cells

5 mL trypsin EDTA was added to a single T75 flask of HEK cells producing hA_{2A}R and was placed on a shaker for 5 min until the cells came loose. The cells were scraped and transferred to a 50 mL tube. 10 mL binding buffer was added to the tube and was probe sonicated (Polytron) for 15 s. 50 µL cells were added to each competition binding curve reaction and followed the same set-up as described in section 2.2.13.3. The experimental conditions are described in section 2.2.13., following the steps for the membrane bound reactions.

Chapter 3: The application of ‘Design of Experiments’ (DoE) to membrane protein production

Typically, maximising the yield of a functional membrane protein requires many ‘trial and error’ attempts, which may involve varying one factor at a time for a large number of variables, thereby amplifying the number of experiments and leading to high costs and increased times. Furthermore, there is no guarantee of establishing optimal conditions, even when all the factors have been tested because any interactions between these variables will not be tested; interactions between pH and temperature for example may be critical in any optimisation process (Isar et al., 2007, Shivam and Mishra, 2010). Statistical Design of Experiments (DoE) allows for a more rational approach that reduces the number of experiments required to determine what the key variables are and ultimately establishes a predictive model in the form of an equation for optimal protein yields. In this chapter, a DoE approach was used to investigate maximizing the yield of hA_{2A}R. In order to keep the number of experimental factors to a minimum, the first experiments were designed using the yeast *Saccharomyces cerevisiae* as the production host because an induction step could be avoided by using constitutive expression in this yeast species. The hA_{2A}R construct used encodes a full length, glycosylated protein with a haem-agglutinin tag (HA₃) and includes an alpha- factor signal sequence that was previously designed and cloned into pYX212 or pYX222 *TPI* vectors by Dr. Renaud Wagner, Université de Strasbourg, France. This was used because it was immediately available at Aston University.

The DoE process is normally sectioned into three parts: screening (where a large number of input factors are explored with the goal to reduce the number going forward for further investigation); process characterisation (identifies and quantifies key input factors by generating a predictive equation) and process optimisation (identifying previously unexplored responses that may require a different design) (Montgomery, 2006).

Table 3.1 defines the key terminology used in this chapter.

Table 3.1 Glossary of relevant DoE terminology. Adapted from (Antony, 2003).

Main effect	The effect of an independent variable on a dependent variable, averaging across the levels of any other independent variables
Interaction	A situation in which the simultaneous influence of two variables on a third is not additive
Contour	A function of two variables forming a curve along which the function has a constant value
Residual	The error in a result - the difference between an individual value and the mean value for a particular data set
R ²	The coefficient of determination, which indicates how well data points fit a line or curve
Adjusted R ² (R _{adj} ²)	A modified R ² value that adjusts for the number of explanatory terms in a model relative to the number of data points, this prevents the R ² value spuriously increasing when extra explanatory variables are added
Analysis of Variance (ANOVA)	A set of statistical models used to compare and analyse the differences between three or more mean values, looking for statistical significance
Degrees of freedom (DF)	The number of values or parameters in a system that are free to vary independently
Sum of squares (SS)	A measure of the total variability of a set of data around the mean, calculated by summing the squared deviations of the individual data points.
Adjusted mean square	An adjusted mean, corrected to compensate for data imbalances
F statistic (Fisher's test)	Derived from the difference between the mean responses at the level studied, it is analogous to a measure of signal to noise ratio for each factor.

Figure 3.1 outlines the steps for a DoE to be carried out. First, after the protein of interest is chosen, the input factors (also referred as parameters or variables) that the researcher wishes to interrogate need to be established. These are usually physical and environmental variables, such as temperature, pH, dissolved oxygen (DO) and medium composition that could have an effect on the yield of the protein at specific set-points levels. The simplest factor level design in a DoE is typically known as a 2^k , where k is the factor that is examined at 2 levels, usually a low and high level. In the design matrix, this is typically coded as -1 for the low level and +1 as the high level. This type of input factor is termed a 'numeric' or quantitative input factor. The input factor could also just be treated as an 'attribute' or qualitative factor where there are no levels associated with it. For instance, just testing the presence or absence of a reagent has an effect on protein yield (Montgomery, 2006, Antony, 2003).

The output response or responses must also be defined in order to run a DoE. This is usually the recombinant protein yield but the total cell biomass is also an example of an output response. The input and output factors are considered in the DoE set-up which define the experiments to be run based on this information. There are now fully developed DoE specific software programmes that facilitate DoE design and statistical analysis and these include Minitab[®], Modde[®], ECHIP[®] and Design - Expert[®]. The software programme returns several types of DoE. Each one has a specific 'design matrix' appropriate for the input factors and number of levels to be interrogated. The design matrix will define the number of replicates, the layout of the experiments and the type of relationships that will be examined between the factors (linear, interaction or quadratic). Based on this information, the most suitable DoE is selected and the experiments are carried out as outlined by the DoE (Bora et al., 2012).

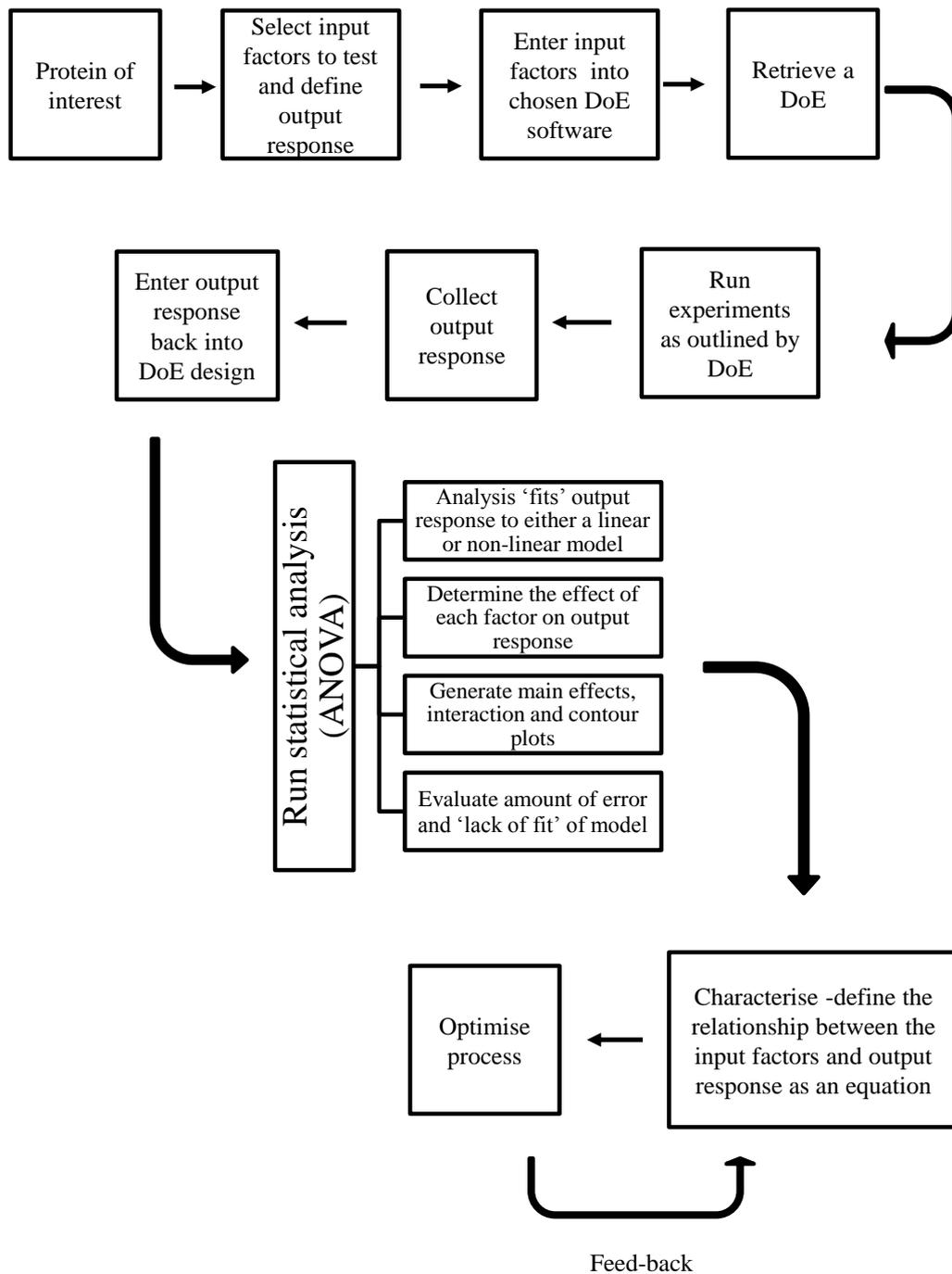


Figure 3.1 Carrying out a DoE- steps to follow in a bioprocess setting. The ‘road-map’ outlines a typical route to carrying out a DoE for recombinant protein production as described in detail in the text.

Following experimentation the output responses are collected and the values are entered back into the DoE. At this stage, statistical analysis is normally an ‘analysis of variance’ (ANOVA). The ANOVA is used to fit the output response to either a linear or non-linear model and calculate the effect of each input factor on this output response (Montgomery, 2006, Mandenius and Brundin, 2008).

The DoE software provides useful plots or graphs to visualise this. These typically include main effects, interaction and in some cases, contour plots. The main effects plots will show if there is an effect with each input factor when considered alone. For example, did a certain pH value have a significant effect on the protein yield? An interaction graph plots input factors against each other to see if a combination of them has an effect on the protein yield. For example, does only a specific temperature and pH value combination have a significant effect on the protein yield, but on their own, they do not? A contour plot is generated after a ‘response surface method’ DoE and will show how a response, visualised as contours of colour, changes in relation to two input factors plotted on the x and y axes (Montgomery, 2006).

The DoE model can be validated by examining the residuals. This is done in two ways: by examining the R^2 and the residual plots. The R^2 value is interpreted as the proportion of the variability in the data explained by the ANOVA model, where a high R^2 value indicates a better fit. A residual plot is a visual representation of the difference between individual values and the mean of the observable data set. A low R^2 may not represent a problem with the data itself, but rather could indicate that the model should be revised (Schmidt and Launsby, 1989). Figure 3.2 shows how residual plots can be an important diagnostic tool. Figure 3.2A illustrates data that fit a linear model well while Figure 3.2B shows data that do not fit this model well and may be more suited to a quadratic model.

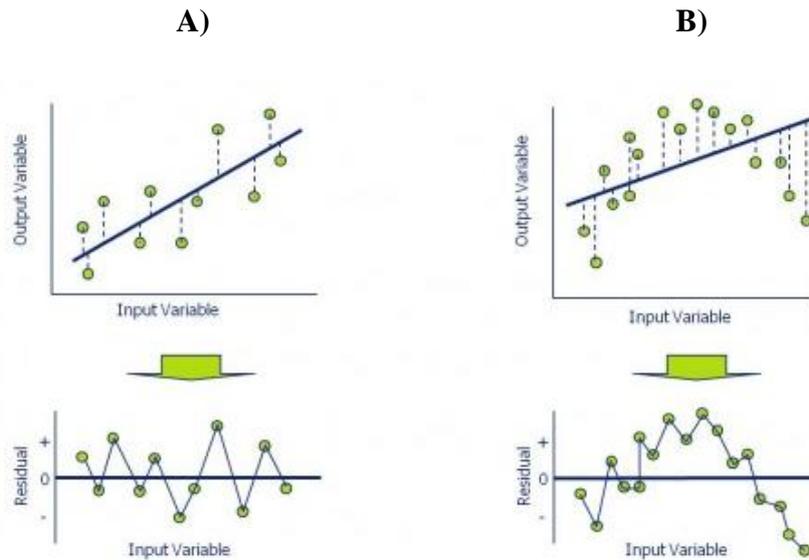


Figure 3.2. Illustration of residuals. A) in the upper graph, input variables are plotted against output variables and a line of fit is plotted; the lower graph shows the pattern of the residuals form and in this case the data fit the model B) in the upper graph, input variables are plotted against output variables and a line of fit is plotted showing; the lower graph shows the pattern of the residuals form and in this case data does not fit the model. Furthermore, the residuals can be referred to as unusual.

A review by Mandenius and Brundin in 2008 suggested that good models that are published in the literature tend to give R^2 values of >0.75 and poor models are usually <0.75 and should require re-evaluation (Mandenius and Brundin, 2008).

Following examination of the R^2 values and the residuals plots, the second phase of the DoE is process characterisation. Here, the goal is to identify and quantify the influence of key input factors on the output response. This requires generation of a predictive equation or model that can be applied under a range of operating conditions. In other words, if certain input factors were entered into the equation, a predicted protein yield would be given.

The third stage of the DoE method is process optimisation where previously unexplored responses are examined or ‘zoomed’ in on and usually a different DoE design is applied. For example this can be the application of a response surface method (RSM) that can be used to interrogate the non-linear behaviour of the response (such as quadratic relationships). The DoE method is an iterative process where the results from this can feed-back until the design is improved and the ‘model building’ is satisfactory.

Previously, DoEs and the use of a micro bioreactor system (Micro-24 microreactor, Pall Corporation) were used to ascertain optimal recombinant soluble green fluorescent protein

yields in *P. pastoris* (Holmes et al., 2009). In this chapter, hA_{2A}R was used as a model GPCR to investigate the application of a DoE to find optimal growth conditions for the highest yield of active hA_{2A}R produced in recombinantly in yeast.

3.1. Defining of DoE model for optimising hA_{2A}R yields in yeast

When taking the DoE approach, various design models are available for use and the most appropriate design is selected based on the number of input factors, type of input factors (quantitative and qualitative) and the number of levels for the quantitative factors. The main types of DoE include but are not limited to: full factorial, fractional factorial, Plackett -Burman, Taguchi, d-optimal and response surface methodology (Schmidt and Launsby, 1989). The classical experimental approach of investigating one factor at a time (OFAT) is easy to understand and straight forward to perform, however it is not an efficient method to study a large sample size. Also, it is practically impossible to address interactions between input factors and usually unnecessary data points are collected leading to costly and time-consuming experiments. Table 3.2 summarise the different types of DoE.

Table 3.2 Types of DoE design and their characteristics.

Type of DoE	Characteristics
Full-factorial	Take into account all possible combinations of input factors. Require many experimental runs but give main effects and interaction information. Linear response only (Fuller and Bisgaard, 1995).
Fractional factorial	Full factorial on selected factors, therefore fewer runs. Loses power when analysing interaction effects. Linear response only (Chen et al., 1993).
Plackett-Burman	2-level screening designs. Do not consider any interaction effects. Linear response only (Vindevogel and Sandra, 1991).
Taguchi	Highly fractional, estimate main effects using few runs. Not limited to 2-level designs. Do not consider some interaction effects. Linear response only (Rao et al., 2008).
d-optimal	Full factorial where number of runs are reduced via a computer derived method. Loses power when analysing interaction effects. Linear response only (de Aguiar et al., 1995).
Response surface methodology (RSM)	Examination of quadratic responses. Not suitable for high number of input factors. Non-linear response (Holmes et al., 2009).

3.1.1. Experimental set-up

The Micro-24 microreactor (Pall Corporation) used to conduct these experiments enabled the input conditions to be established as set-points. Figure 3.3 shows the set-up at Aston University. The instrument is controlled by software and the DO is controlled by gas cylinder oxygen and compressed sterile air. The pH is controlled by either addition of carbon dioxide or ammonia by bubbling nitrogen through an ammonium hydroxide bubbler. The instrument uses 24 modified deep wells or mini bioreactors as a plate format. These wells can be controlled individually for temperature, pH and DO and also be run in parallel. In the bottom of each well, DO tension, pH optical sensors, a sparging port and thermal conduction pads are present to provide control and monitoring. The pH and DO of the culture are controlled by a set point by sparging gas through the medium, in this instance, oxygen for DO and carbon dioxide or ammonium hydroxide gas for pH.

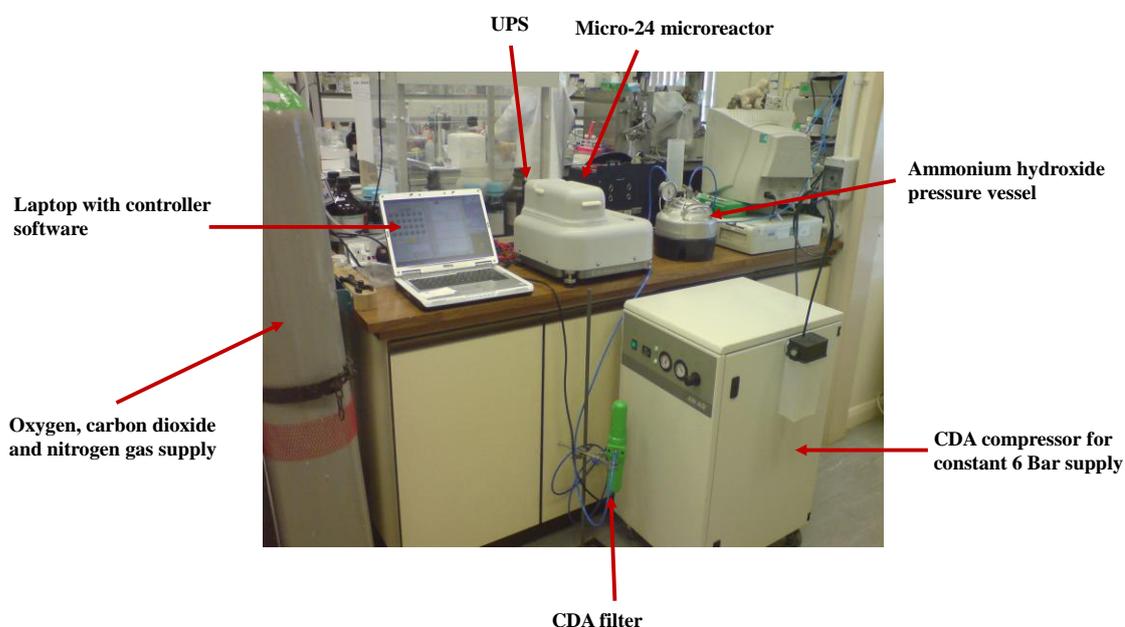


Figure 3.3 Micro-24 microreactor set-up at Aston University. Photograph shows the instrument and the components needed for its function including a lap-top with the software for control and monitoring, the gas supplies for DO and pH control, the ammonium hydroxide bubbler for ammonia production and a compressor to maintain a vacuum for the 24 well plates.

3.1.2. Input parameters

It is well known that the temperature, pH and the amount of DO in a culture will have an effect on yeast cell growth (Schubert et al., 1994). Other studies have also defined that these parameters will affect recombinant protein yields (Schmidt, 2005). Therefore these 3 factors were chosen as input factors for the DoE to test if optimal conditions could be achieved for hA_{2A}R production. Moreover, for the temperature, DO and pH, 3 levels (low, medium and high)

were interrogated in the DoE screen in order to detect any interaction relationships. The input factors selected in these experiments were based on previous studies of hA_{2A}R or other GPCRs recombinantly produced in yeast. For example, it is known that lower pH values (5.0-6.0) are favourable in yeast, therefore the input conditions for pH were 5.0, 5.5 and 6.0 (Fraser, 2006). It is also established that lower temperature values (~22°C) are favourable in yeast for hA_{2A}R production (Fraser, 2006) therefore the input conditions for temperature were 22, 24 and 26°C. The DO input conditions were chosen close to the normal range for a bioreactor set-up which is ~30% set-point. Therefore the input conditions for DO were 30, 40 and 50%. Finally, in the literature there have been reports that the addition of the additive, DMSO has a positive effect on recombinant GPCR yield in yeast (Andre et al., 2006, Lundstrom et al., 2006). Therefore this was an input factor that was interesting to include in this DoE. This input factor was however, included as a qualitative or attribute input factor, where it was either present or absent and did not vary in its level (with or without DMSO) and inclusion of this additional factor contributed to choosing a d-optimal design.

In this study, 4 input factors were initially investigated in a screening design where 3 of the factors were at 3 levels (low, medium and high; -1, 0 and +1 and therefore constituted a 3^k or 3^3 design) and one was at one level but was either present or absent in the experiments. With these input factors, the classical full factorial design was considered. However, to run it in full, 162 separate experiments were defined and this was too large to conduct. Therefore a computer derived design called a d-optimal design was selected as this took the 162 runs from the full factorial design and selected 24 experiments to be run, which was more practical. A d-optimal design is an option when trying to reduce the number of runs in a design when there are 3 levels and 3 or more factors to explore. However, there are compromises on this design such as the design being non-orthogonal.

Table 3.3 summarises the input factors that were entered into the Minitab[®] DoE and that generated the d-optimal design. In conjunction with these input factors, the output responses were also defined. These were cell biomass, total membrane protein yield and specific binding activity of hA_{2A}R.

Table 3.3 Input factors used in the DoE screen for hA_{2A}R production. Temperature (22, 24 and 26°C), pH (5.0, 5.5 and 6.0) and DO (30, 40 and 50%) of the yeast culture were examined. The addition of 2.5% DMSO is either present or absent in the culture media.

Input Factor	Input factor type	Low Level	Medium Level	High Level
Temperature (°C)	Quantitative	22	24	26
pH	Quantitative	5	5.5	6
dO (%)	Quantitative	30	40	50
DMSO (%)	Qualitative	0 (absent)	n/a	2.5 (present)

3.1.3. *Saccharomyces cerevisiae* strains

Screening experiments were carried out using a DoE designed to screen the best growth conditions for yeast to produce the highest hA_{2A}R yields. The DoE explored four input variables in three different *S. cerevisiae* yeast strains producing recombinant hA_{2A}R (Table 3.4). These strains were chosen as previous studies had shown they were high-yielding (Ferndahl et al., 2010, Bonander et al., 2009). The hA_{2A}R construct is a full length, glycosylated version with a haem-agglutinin tag (HA₃) and includes an alpha- factor signal sequence (Figure 3.4) that was previously designed and cloned into pYX212 or pYX222 *TPI* vectors by Dr. Renaud Wagner, Université de Strasbourg, France. Dr. Richard Darby, Aston University subsequently transformed these into a wild-type strain of *S. cerevisiae* (BY4741) in a pYX212 vector with uracil selection and into an over-expression mutant, *S. cerevisiae* *BMS1* (BY4741) in a pYX222 vector with histidine selection. The third strain was developed by Dr. Cecilia Ferndahl (Ferndahl et al., 2010) and this was *S. cerevisiae* TM6* (KOY-TM6*) and was transformed with the pYX212 vector with uracil selection. The strains are referred to as WT-hA_{2A}R, BMS1-hA_{2A}R and TM6-hA_{2A}R throughout this chapter. The origins of these strains are described in 1.2.2.3.

Table 3.4 *S. cerevisiae* strains producing hA_{2A}R used for DoE research.

Full name	<i>S. cerevisiae</i> strain	Vector	Selection	Short name
<i>S. cerevisiae</i> wild-type hA _{2A} R	BY4741	pYX212	Uracil	WT-hA _{2A} R
<i>S. cerevisiae</i> <i>BMS1</i> hA _{2A} R	BY4741 <i>yTHCBMS1</i>	pYX222	Histidine	BMS1-hA _{2A} R
<i>S. cerevisiae</i> TM6* hA _{2A} R	KOY-TM6*	pYX212	Uracil	TM6-hA _{2A} R

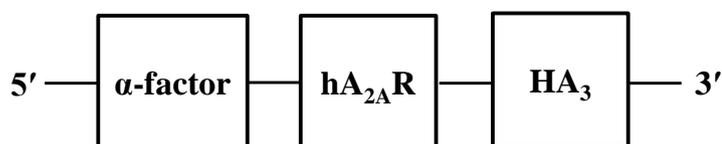


Figure 3.4 hA_{2A}R construct design schematic for *S. cerevisiae* transformations. Construct includes full-length hA_{2A}R, an alpha-factor signal sequence at the 5' end and a haem-agglutinin tag × 3 at the 3' end.

S. cerevisiae was used for these experiments to avoid the need for any media exchange during recombinant protein production. The strains chosen produce recombinant hA_{2A}R constitutively

in the pYX212 and 222 vectors, with glucose as its sole carbon source. Therefore for the purpose of these experiments, it was simpler to use *S. cerevisiae* strains rather than *P. pastoris*, where methanol feeding is required to induce recombinant production.

3.2. Execution of the d-optimal DoE design

When the input factors were entered into the Minitab[®] software, a d-optimal design was selected. The input factors were temperature (22, 24 and 26°C) pH (5.0, 5.5 and 6.0), DO (30, 40 and 50%) and the presence or absence of DMSO. The output responses that were measured were cell biomass, total membrane protein yield and specific binding activity of hA_{2A}R in three different *S. cerevisiae* strains: WT-hA_{2A}R, BMS1-hA_{2A}R and TM6-hA_{2A}R. This section describes the experimental results of the d-optimal DoE screen for all 3 *S. cerevisiae* strains producing recombinant hA_{2A}R.

Table 3.5 shows the output of the experiments that were defined by the d-optimal design and were to be performed in the Micro-24 microreactor plates. It can be seen that 24 different combinations of temperature, pH, DO and DMSO conditions from Table 3.3. The d-optimal design output gave 22 different experiment combinations. The grey highlighted rows in the table show repeats of certain combinations in the wells to increase the well usage to the full 24.

Table 3.5 d-optimal design of 4 input factors, 3 numeric factors and 1 attribute factor. The 3 numeric factors are temperature (22, 24 or 26°C), pH (5.0, 5.5 or 6.0) and DO (30, 40 or 50%). The attribute factor is presence/absence of DMSO. The highlighted rows indicate repeated input factor combinations. The total number of experiments is 24, to fit one whole Micro-24 plate.

Temperature (°C)	pH	dO (%)	2.5% DMSO
22	5.5	40	Absent
22	5.5	40	Absent
22	6	50	Present
22	6	50	Present
22	5	50	Present
22	6	30	Absent
22	5	40	Present
22	5.5	30	Absent
24	5	30	Absent
24	5.5	50	Present
24	5.5	50	Absent
24	6	40	Absent
24	5	30	Present
24	5	50	Absent
24	5.5	30	Present
24	6	40	Present
26	6	50	Absent
26	5.5	30	Present
26	5	40	Absent
26	6	30	Present
26	5.5	40	Present
26	5	50	Absent
26	6	30	Absent
26	5	40	Present

In DoE outputs, a run order is usually stated in the design output return. This means that the DoE design can state which order each experiment should be run in order to minimise the error due to the effect of lurking variables by using a random sequence (Lendrem et al., 2001).

However, in these experiments, the run order could not be adhered to due to instrument constraints. The temperature in the plate wells had to be ordered so that lowest temperatures were at the back of the plate and the highest temperatures were at the front of the plate. This was due to the coolant fans in the Micro-24 microreactor being positioned at the rear of the instrument (Figure 3.5). Each *S. cerevisiae* strain, WT-hA_{2A}R, BMS1-hA_{2A}R and TM6-hA_{2A}R was inoculated to a starting optical density of 1 from a 50 mL shake flask pre-culture in the Micro-24 microreactor well plate. WT-hA_{2A}R, BMS1-hA_{2A}R and TM6-hA_{2A}R were run on the Micro-24 microreactor in triplicate with the well plate set-up as illustrated in Figure 3.5, on

three different days. When the glucose levels reached between 10-30 mM, the cultures were harvested and the optical density noted at 600 nm (OD₆₀₀). The glucose levels were monitored by using Accu-Chek Active glucose analyser (Roche Diagnostics). Membrane preparations were carried out for each well and the total membrane protein quantified using a BCA assay. Finally, single-point radio-ligand binding assays were carried out with [³H]ZM241385 on each well for each strain.

	1	2	3	4	5	6
A	22°C, pH5, 50, Yes	22°C, pH6, 50, Yes	24°C, pH5, 30, No	24°C, pH5, 30, Yes	26°C, pH6, 50, No	26°C, pH5.5,40, Yes
B	22°C, pH5.5, 40. No	22°C, pH5.5, 30, No	24°C, pH5.5, 50, Yes	24°C, pH5, 50, No	26°C, pH5.5, 30, Yes	26°C, pH 5, 50, No
C	22°C, pH6, 30, No	22°C, pH6,50, Yes	24°C, pH 5.5, 50, No	24°C, pH5.5, 30, Yes	26°C, pH5, 40, No	26°C, pH6, 30, No
D	22°C, pH5,40, Yes	22°C, pH5.5, 40, No	24°C, pH6, 40, No	24°C, pH6, 40, Yes	26°C, pH6, 30, Yes	26°C, pH5, 40, Yes

Figure 3.5 Micro-24 microreactor plate set-up. Each well is shown with specified input factors shown in Table 3.3. Yes/No indicates the presence/absence of 2.5% DMSO.

The output data retrieved after the runs had finished were OD_{600} , total membrane protein yields and specific binding activity of $hA_{2A}R$ in WT- $hA_{2A}R$, BMS1- $hA_{2A}R$ and TM6- $hA_{2A}R$ strains. The OD_{600} for all the strains was approximately 5.0 ± 1.0 therefore the effects from the input factors were minimal (Figure 3.6).

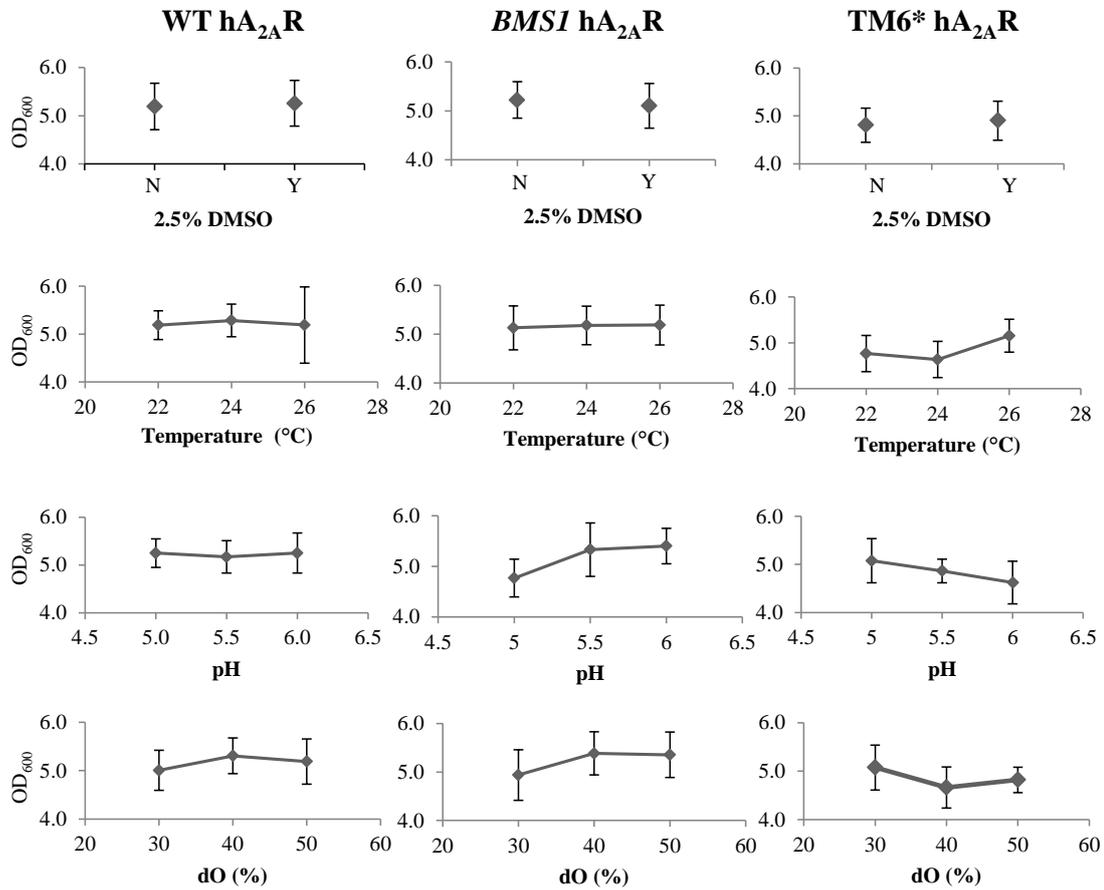


Figure 3.6 Main effects plots of OD_{600} values for each $hA_{2A}R$ strain, for each input factor at the end of the Micro-24 microreactor run. OD_{600} plots for WT- $hA_{2A}R$ are all on the left side, plots for BMS1- $hA_{2A}R$ are all in the centre and plots for TM6- $hA_{2A}R$ are all on the right side. Values are means of OD_{600} for triplicate readings of each strain for each input factor. \pm standard error of the mean (SEM) is displayed as error bars.

Total protein yield was also measured in membrane preparations and the SEM between replicates was high for some replicates and ranged between 4-79 μg total protein (Table 3.6).

Table 3.6 Total protein yields of membrane preparations from each input combination and *S. cerevisiae* strain. The input factors and the total protein yield as determined by BCA assay with bovine serum albumin as a standard are shown. The data are means from triplicate samples and the \pm SEM are displayed.

Conditions tested generated from d-optimal DoE screen				Total Protein Yield (μg)		
Temperature ($^{\circ}\text{C}$)	pH	dO (%)	2.5% DMSO	WT-hA _{2A} R	BMS1-hA _{2A} R	TM6-hA _{2A} R
22	5.0	50	Present	193 \pm 13	206 \pm 21	233 \pm 15
22	5.5	40	Absent	162 \pm 17	185 \pm 58	186 \pm 18
22	6.0	30	Absent	217 \pm 44	219 \pm 53	217 \pm 31
22	5.0	40	Present	236 \pm 28	217 \pm 19	205 \pm 19
22	6.0	50	Present	159 \pm 6	221 \pm 43	192 \pm 18
22	5.5	30	Absent	192 \pm 32	178 \pm 53	191 \pm 12
22	6.0	50	Present	198 \pm 21	210 \pm 35	190 \pm 4
22	5.5	40	Absent	200 \pm 29	213 \pm 47	173 \pm 37
24	5.0	30	Absent	162 \pm 8	193 \pm 37	197 \pm 18
24	5.5	50	Present	167 \pm 4	201 \pm 32	208 \pm 15
24	5.5	50	Absent	195 \pm 41	214 \pm 30	212 \pm 35
24	6.0	40	Absent	182 \pm 26	248 \pm 18	241 \pm 48
24	5.0	30	Present	198 \pm 5	199 \pm 36	172 \pm 14
24	5.0	50	Absent	182 \pm 41	185 \pm 43	195 \pm 17
24	5.5	30	Present	199 \pm 21	216 \pm 40	195 \pm 23
24	6.0	40	Present	223 \pm 56	204 \pm 14	195 \pm 32
26	6.0	50	Absent	168 \pm 15	177 \pm 43	195 \pm 37
26	5.5	30	Present	175 \pm 5	187 \pm 31	237 \pm 45
26	5.0	40	Absent	212 \pm 49	228 \pm 21	161 \pm 31
26	6.0	30	Present	204 \pm 11	233 \pm 20	181 \pm 30
26	5.5	40	Present	175 \pm 11	228 \pm 49	174 \pm 25
26	5.0	50	Absent	166 \pm 12	163 \pm 35	180 \pm 15
26	6.0	30	Absent	182 \pm 41	231 \pm 18	168 \pm 46
26	5.0	40	Present	207 \pm 34	209 \pm 79	193 \pm 60

Radio-ligand binding analyses of hA_{2A}R were the main output data of interest for this DoE screen, since the objective was to determine how the input factors had an effect or not on hA_{2A}R yield for each strain. Specific binding activity data were collected via radio-ligand binding assays using the hA_{2A}R antagonist, [³H] ZM241385 (Table 3.7). The range of specific binding activity was from 0- 500 fmol mg⁻¹ for all the strains and all the input conditions indicating the overall expression was low for these experiments.

Table 3.7 Specific binding activity from each input combination and *S. cerevisiae* strain. The input factors and specific binding activity as determined by a radio-ligand assay with the hA_{2A}R antagonist, [³H] ZM241385 . The data are means from triplicate samples and the ± SEM are displayed.

Conditions tested generated from d-optimal DoE screen				Specific binding (fmol mg ⁻¹)		
Temperature (°C)	pH	dO (%)	2.5% DMSO	WT-hA _{2A} R	BMS1-hA _{2A} R	TM6-hA _{2A} R
22	5	50	Present	7.9 ± 5.5	8.1 ± 3.6	4.5 ± 2.1
22	5.5	40	Absent	3.1 ± 1.6	13.8 ± 9.9	9.5 ± 3.0
22	6	30	Absent	320.6 ± 122.1	18.1 ± 6.0	4.6 ± 2.5
22	5	40	Present	9.3 ± 4.8	8.3 ± 2.6	3.9 ± 0.2
22	6	50	Present	28.8 ± 6.8	11.7 ± 4.2	5.9 ± 1.8
22	5.5	30	Absent	12.3 ± 6.4	14.7 ± 4.9	6.1 ± 1.2
22	6	50	Present	31.4 ± 2.4	16.6 ± 3.3	7.6 ± 1.3
22	5.5	40	Absent	25.8 ± 3.0	15.9 ± 2.4	7.6 ± 1.3
24	5	30	Absent	5.8 ± 5.5	9.4 ± 2.5	6.1 ± 0.6
24	5.5	50	Present	17.9 ± 3.9	19.5 ± 8.6	35.5 ± 12.6
24	5.5	50	Absent	14.1 ± 5.9	12.8 ± 3.8	4.6 ± 1.5
24	6	40	Absent	10.6 ± 6.4	45.6 ± 11.8	5.8 ± 2.3
24	5	30	Present	12.1 ± 4.9	12.3 ± 6.1	4.1 ± 3.0
24	5	50	Absent	18.3 ± 2.7	9.4 ± 2.5	4.2 ± 2.6
24	5.5	30	Present	13.4 ± 5.8	11.9 ± 3.5	3.9 ± 0.3
24	6	40	Present	22.8 ± 2.6	32.5 ± 8.6	3.6 ± 1.2
26	6	50	Absent	55.8 ± 24.3	10.5 ± 6.9	4.5 ± 0.7
26	5.5	30	Present	5.7 ± 8.7	13.4 ± 5.0	4.9 ± 3.5
26	5	40	Absent	7.3 ± 4.9	9.3 ± 3.7	2.9 ± 1.2
26	6	30	Present	7.7 ± 5.4	12.3 ± 5.4	5.6 ± 1.3
26	5.5	40	Present	14.7 ± 4.7	11.4 ± 4.3	4.1 ± 2.1
26	5	50	Absent	16.1 ± 4.6	18.4 ± 9.3	7.3 ± 1.1
26	6	30	Absent	22.6 ± 4.0	19.7 ± 5.7	2.7 ± 1.4
26	5	40	Present	21.7 ± 4.2	11.5 ± 3.9	3.2 ± 1.8

These values were entered into the d-optimal design. A linear regression analysis was performed on the data. Main effects and interaction plots were generated with the ANOVA analyses.

Figure 3.7 shows the main effects plot, the interaction plot and the ANOVA analysis output for the d-optimal design performed on WT-hA_{2A}R.

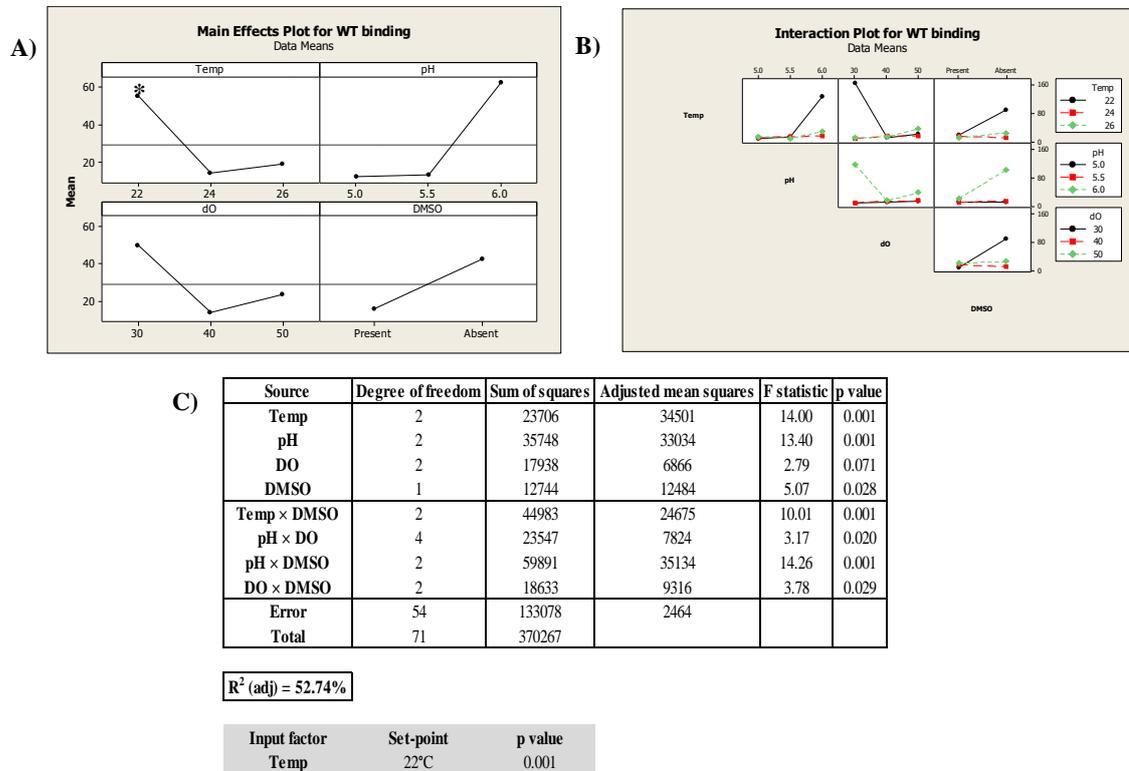


Figure 3.7 Main effects, interaction and ANOVA analysis for WT-hA_{2A}R. A) Main effects plot where the horizontal line indicates mean specific binding activity. B) Interaction plot with mean specific binding activity as the y axis. C) Analysis of Variance (ANOVA) for specific binding in WT-hA_{2A}R. The relationship between the specific binding activity and the input factors was analysed using ANOVA where the F statistic is the Fisher's F-test. The R^2 (adj) indicates 'goodness of fit' of the model. Input factors highlighted in grey show a significant p value (<0.05, 95% confidence) and a positive co-efficient and marked by * on the graphs (Appendix A4).

For the WT-hA_{2A}R strain, the highest yields of hA_{2A}R were produced when the culture medium was at a temperature of 22°C, pH 6.0, with 30% DO and in the absence of DMSO, according to the main effects plot (Figure 3.7A). Furthermore, the temperature effect alone of 22°C was found to give a significant increase (Figure 3.7C) in specific binding activity (p value = 0.001). When considering the interaction effects, the interaction plot (Figure 3.7B) showed that interaction were present such as a low temp, high pH interaction; a low temp, low DO interaction; a low temp, no DMSO interaction; a high pH, low DO interaction; a high pH, no DMSO interaction and finally a low DO, no DMSO interaction. These interactions gave higher yields of hA_{2A}R than any other interactions. The adjusted R^2 value for the WT-hA_{2A}R DoE was 52.74% (Figure 3.7C). Figure 3.7C also displays the ANOVA output for the input terms to determine if they were significant or not. As stated, the temperature at 22°C was significant (p

value < 0.05) and also the co-efficient was positive (Appendix A4 for linear regression co-efficient data).

Figure 3.8 shows the main effects plot, the interaction plot and the ANOVA analysis output for the d-optimal design performed on BMS1-hA_{2A}R.

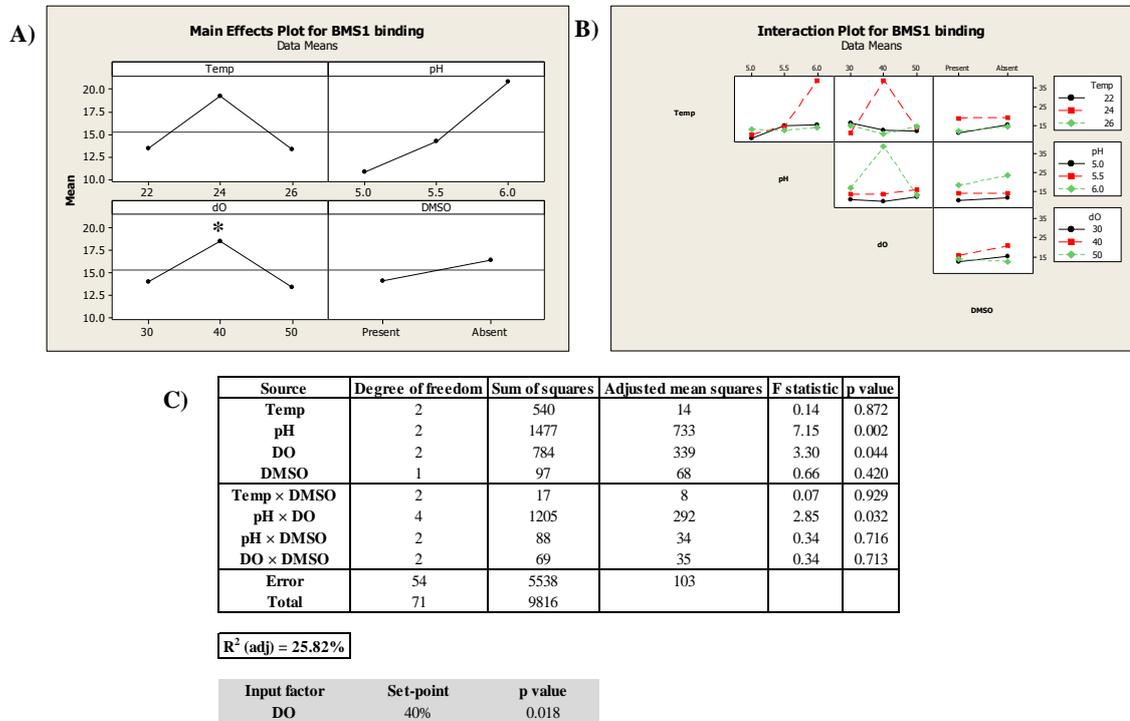


Figure 3.8 Main effects, interaction and ANOVA analysis for BMS1-hA_{2A}R. A) Main effects plot where the horizontal line indicates mean specific binding activity. B) Interaction plot with mean specific binding activity as the y axis. C) Analysis of Variance (ANOVA) for specific binding in BMS1-hA_{2A}R. The relationship between the specific binding activity and the input factors was analysed using ANOVA where the F statistic is the Fisher's F-test. The R² (adj) indicates 'goodness of fit' of the model. Input factors highlighted in grey show a significant p value (<0.05, 95% confidence) and a positive co-efficient and marked by * on the graphs (Appendix A4).

The main effects plot (Figure 3.8A) showed that the BMS1-hA_{2A}R strain, gave the highest yields of hA_{2A}R when the culture medium was at a temperature of 24°C, pH 6.0, with 40% DO and in the absence of DMSO. The 40% DO in the culture gave a significant increase in hA_{2A}R yield (p value = 0.018, Figure 3.8C). From the interaction plot (Figure 3.8B), interactions that gave higher yields of hA_{2A}R included a mid temperature, high pH interaction; a mid temperature, mid DO interaction; a high pH, mid DO interaction; a high pH, no DMSO interaction and finally a mid DO, no DMSO interaction. The adjusted R² value for the BMS1-hA_{2A}R DoE was 28.52% (Figure 3.8C).

Figure 3.9 shows the main effects plot, the interaction plot and the ANOVA analysis output for the d-optimal design performed on TM6-hA_{2A}R.

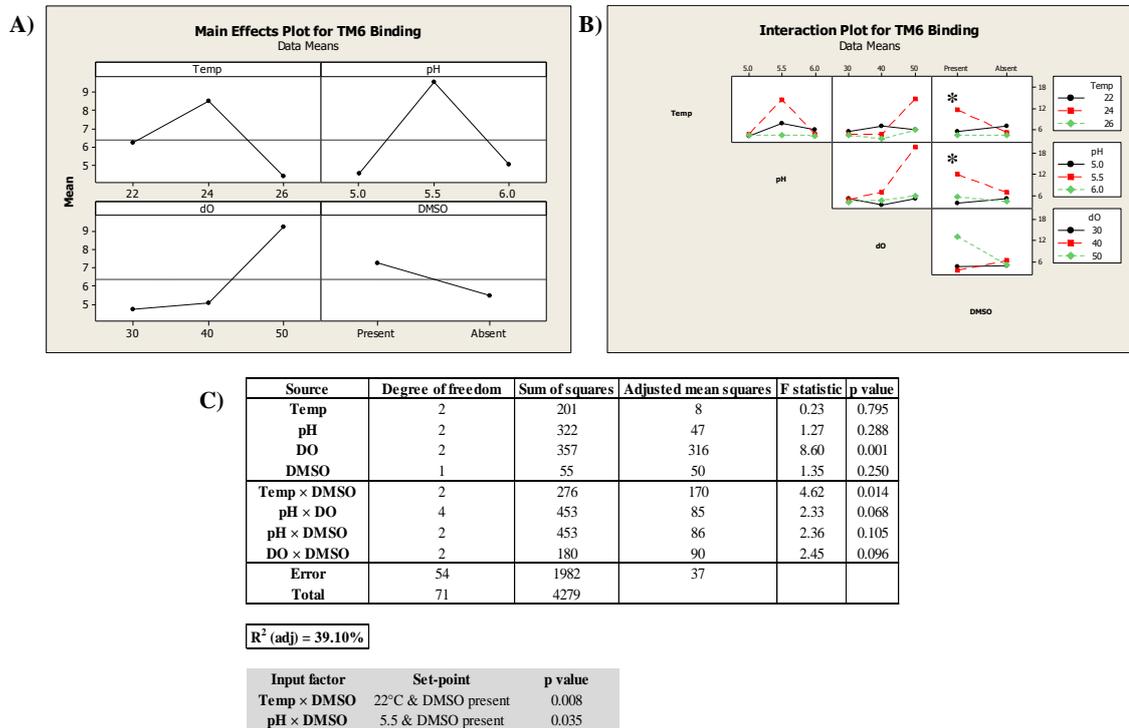


Figure 3.9 Main effects, interaction and ANOVA analysis for TM6-hA_{2A}R. A) Main effects plot where the horizontal line indicates mean specific binding activity. B) Interaction plot with mean specific binding activity as the y axis. C) Analysis of Variance (ANOVA) for specific binding in TM6-hA_{2A}R. The relationship between the specific binding activity and the input factors was analysed using ANOVA where the F statistic is the Fisher's F-test. The R² (adj) indicates 'goodness of fit' of the model. Input factors highlighted in grey show a significant p value (<0.05, 95% confidence) and a positive co-efficient and marked by * on the graphs (Appendix A4).

The main effects plot (Figure 3.9A) revealed that the TM6-hA_{2A}R strain, gave the highest yields of hA_{2A}R when the culture medium was at a temperature of 24°C, pH 5.5, with 50% DO and in the presence of DMSO. Contrastingly however, there were no significant singular effects for this strain. The interaction plot showed (Figure 3.9B) higher yields of hA_{2A}R were obtained with either a mid temp, mid pH interaction; a mid temperature, high DO interaction; a mid temperature, DMSO present interaction; a mid pH, high DO interaction; a mid pH, DMSO present interaction and finally a high DO, DMSO present interaction. Furthermore, the mid temperature, DMSO present and mid pH, DMSO present interactions were significant effects on hA_{2A}R yields and gave a p value of 0.008 and 0.035, respectively (Figure 3.9C). The adjusted R² value for the TM6-hA_{2A}R DoE was 39.10% (Figure 3.9C).

Residual plots were also retrieved from the analyses. Figure 3.10 displays all three residual plots for the yeast strains.

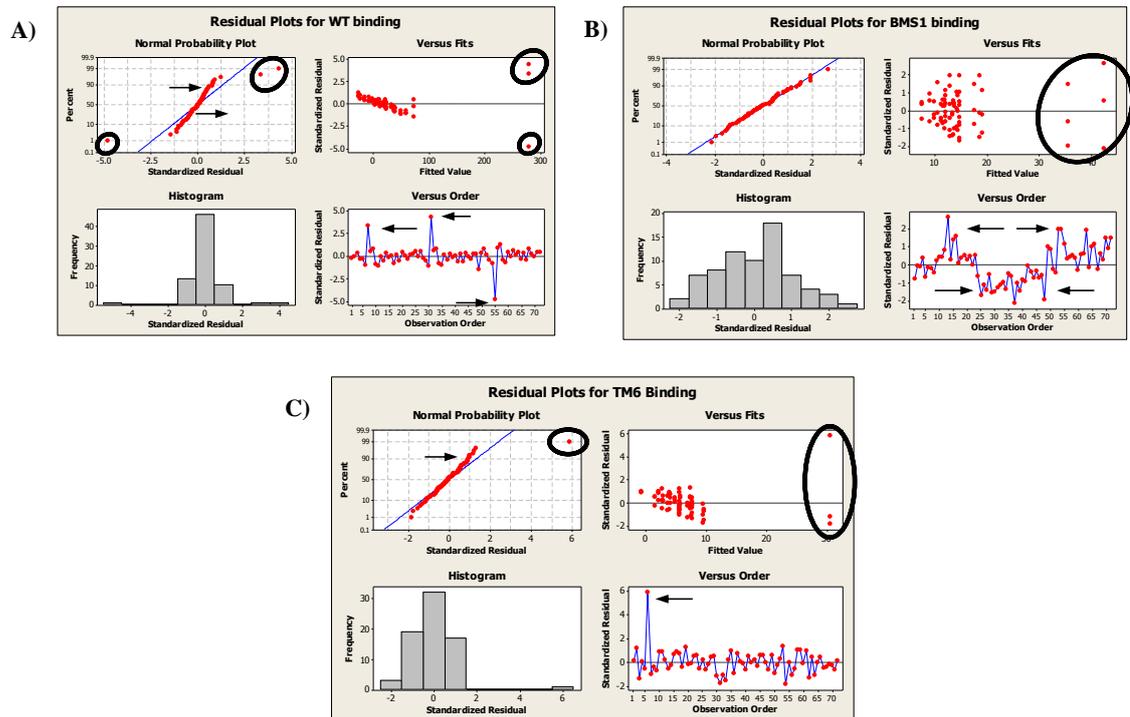


Figure 3.10 Residual plots for WT- $hA_{2A}R$, BMS1- $hA_{2A}R$ and TM6- $hA_{2A}R$ after linear regression analysis from *d*-optimal design. Visualised in four ways, as a normal probability plot, standardised versus fitted plot, a histogram and observed order of residuals. Black circles indicate outlier data and black arrows indicate outlier residuals and patterns or trends that do not fit the normal residuals.

The residual plots were examined to determine how well the data fit the model. It can be seen from Figure 3.10A when viewing the normal probability plot that the data points for the WT- $hA_{2A}R$ did not fit to the model. It can also be seen that there are 3 clear unusual residuals present in the versus fits, normal probability and versus order plot indicated by black arrows or black circles.

These points were replicates of the same input condition, 22°C, pH 6.0, 30% and DMSO absent and were also in the same position in the Micro-24 well plate for each replicate (Figure 3.11). Upon further investigation, the set-points of these input factors were checked. It was found that the temperature was maintained at 22°C (data not shown) as was the pH (6.0) but the DO did not reach the set-point of 30% and instead the DO was maintained at 60% (Figure 3.11). In general it was found that many of the DO set-points were not reached in the Micro-24 well plate and this is further discussed in section 3.4.1.

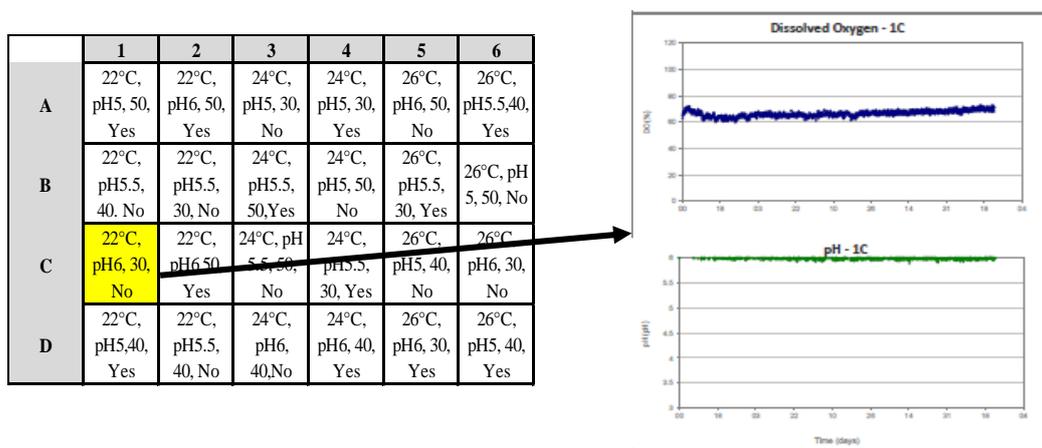


Figure 3.11 Actual set-points for DO and pH for WT-hA_{2A}R replicate 1. Data interrogated after residual plot revealed unusual residual for well 1C (highlighted in yellow) in Micro-24 plate well (left). The actual DO and pH data for well 1C are shown in the graph on the right.

This examination of the residuals was also performed for the BMS1-hA_{2A}R and the TM6-hA_{2A}R strain (Figure 3.10B and C). It was found that inspecting the residuals (black arrows and black circles) and tracing them back to the experiment and the actual set-points, the DO levels were not achieved. Figure 3.12 shows the well position of the input conditions that gave the high residuals from Figure 3.10, the input conditions and the actual set-points for the DO and pH for the BMS1-hA_{2A}R (Figure 3.12A) and the TM6-hA_{2A}R (Figure 3.12B) experiments.

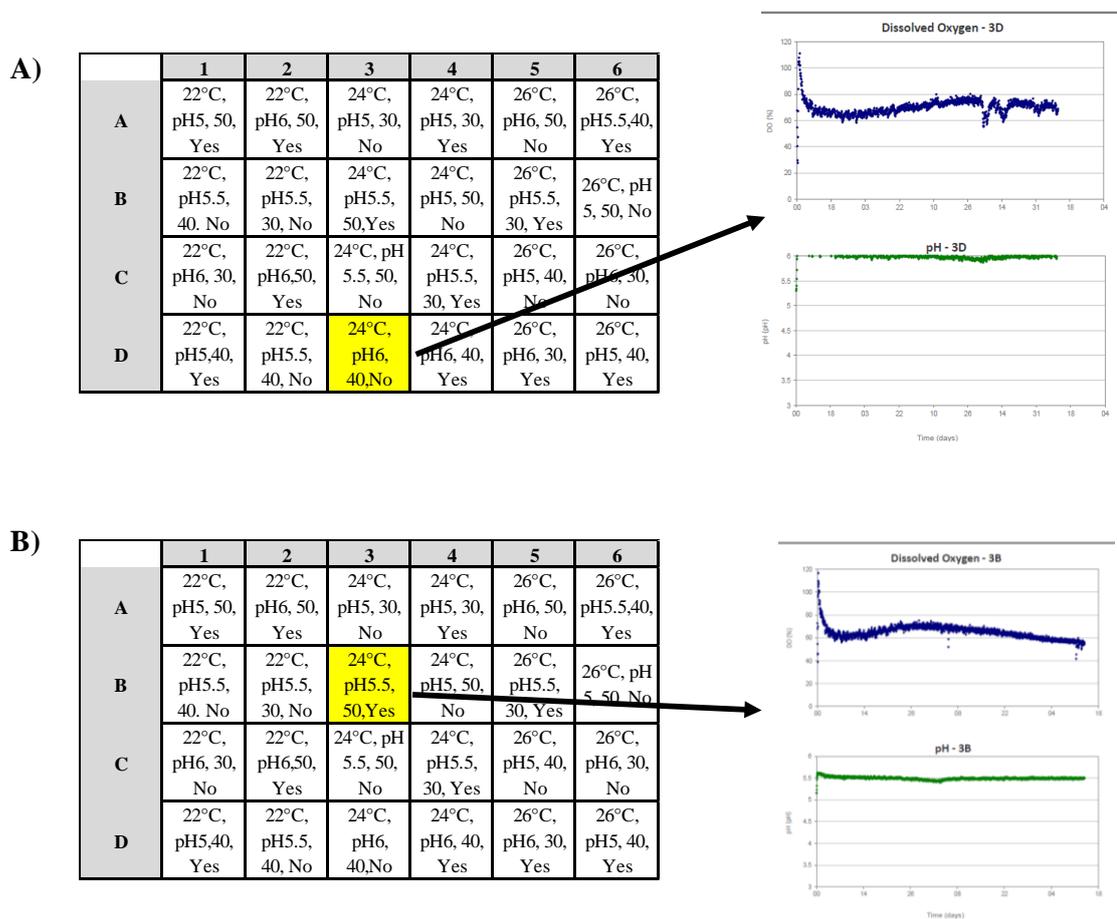


Figure 3.12 Actual set-points for DO and pH for BMS1-hA_{2A}R and TM6- hA_{2A}R replicate 1.
A) BMS1-hA_{2A}R data interrogated after residual plot revealed high residual for well 3D (highlighted in yellow) in Micro-24 plate well (left). The actual DO and pH data for well 3D are shown in the graph on the right. **B) TM6-hA_{2A}R data interrogated after residual plot revealed high residual for well 3B** (highlighted in yellow) in Micro-24 plate well (left). The actual DO and pH data for well 3B are shown in the graph on the right.

Figure 3.12A shows the Micro-24 well plate for BMS1-hA_{2A}R and the input conditions of 24°C, pH 6.0, 40% and DMSO absent and the actual DO and pH data. It can be seen that, not unlike the WT-hA_{2A}R strain, the pH was maintained at the set-point but the DO did not reach the set-point of 40%. For the TM6-hA_{2A}R (Figure 3.12B); input conditions (24°C, pH 5.5, 50% and DMSO present), the pH was maintained but the DO was not. These findings prompted further error investigation discussed in section 3.4.

The overall d-optimal analysis showed that specific input factor combinations gave rise to favourable conditions for optimal hA_{2A}R yields for all three strains. It was noted however that the R² values retrieved from the linear regression were low, overall instrument limitations introduced unanticipated error and that a linear model was not optimal.

3.3. Using response surface methodology (RSM) to improve the statistical outcome of the d-optimal data

Based upon inspection of the residuals data for all the strains the data were fitted to a quadratic model instead of a linear one and this was done using response surface methodology (RSM).

The goal was to improve upon the d-optimal model that was originally generated and used for the DoE experiments investigating hA_{2A}R production in 3 yeast strains. After applying a linear regression model to the data, it was evident that this was not the best model, since the ANOVA analysis revealed high residuals that warranted further investigation. A customised RSM analysis was run in Minitab[®] software on the data set (input factors and output response). Visual plots and residuals plot were also generated as before. It is worth noting that the RSM analysis carried out on the BMS1-hA_{2A}R and TM6-hA_{2A}R strains gave very poor R² (adj) values (15-20%), therefore it was decided to not continue with any further analysis for these strains. This is discussed further in the main discussion chapter (Chapter 6). Hence, the subsequent experiments described and discussed in this section are for WT-hA_{2A}R only.

A model was derived from the RSM analysis of the WT-hA_{2A}R data which gave an initial R² (adj) of 63.07%. The model was then reduced to improve the R² (adj) value by removing any insignificant terms. This was done by de-selecting a term that was not significant in the Minitab[®] software and re-running the analysis. It was found that the temperature [squared] effect was insignificant (p value > 0.05) and therefore was removed and re-analysed with ANOVA. The model improved to give a R² (adj) of 63.41%. Removal of other insignificant terms did not improve the model further. Table 3.8 shows the statistical output of the ANOVA analysis showing the terms that were analysed.

Table 3.8 ANOVA analysis output for WT-hA_{2A}R strain after RSM analysis. The relationship between the input factors and the output response (specific binding activity of hA_{2A}R) was calculated using analysis of variance (ANOVA). F statistic is the Fisher's test and p value was deemed significant if it was <0.05 (95% confidence level). The R² (adj) was 63.41% for this reduced model where (Temp*Temp) were removed. The full analysis is shown in Appendix A5.

Source	Degree of freedom	Adjusted Sum of squares	Adjusted mean squares	F statistic	p value
Regression	12	265920	22160	7.79	0.001
Linear	4	52711	13178	4.63	0.004
Temp	1	11798	11798	4.15	0.049
pH	1	43077	43077	15.14	0.001
DO	1	1779	1779	0.63	0.435
DMSO	1	382	382	0.13	0.716
Square	2	36288	18144	6.38	0.004
pH*pH	1	21091	21091	7.41	0.010
DO*DO	1	19010	19010	6.68	0.014
Interaction	6	155148	25858	9.09	0.001
Temp*pH	1	35284	35284	12.4	0.001
Temp*DO	1	27389	27389	9.62	0.004
Temp*DMSO	1	5809	5809	2.04	0.162
pH*DO	1	23895	23895	8.40	0.006
pH*DMSO	1	27188	27188	9.55	0.004
DO*DMSO	1	619	619	0.22	0.644

The analysis also generated co-efficients that could be applied as a predictive model in the form of a polynomial equation ($Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{12} X_1 X_2 + \beta_{23} X_2 X_3 + \beta_{13} X_1 X_3 + \beta_{14} X_1 X_4 + \beta_{24} X_2 X_4 + \beta_{34} X_3 X_4 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{44} X_4^2$) where β is the regression co-efficient from the ANOVA analysis and X is the input factor.

$$\text{Yield (fmol mg}^{-1}\text{)} = 1653.14 + (162.98 \times T) - (934.92 \times \text{pH}) - (56.47 \times \text{DO}) + (99.12 \times \text{DMSO}) + (227.77 \times \text{pH}^2) + (0.45 \times \text{DO}^2) - (49.93 \times T \times \text{pH}) + (2.50 \times T \times \text{DO}) + (10.83 \times T \times \text{DMSO}) - (7.35 \times \text{pH} \times \text{DO}) - (68.73 \times \text{pH} \times \text{DMSO}) + (0.56 \times \text{DO} \times \text{DMSO})$$

where T = temperature (°C), DO = dissolved oxygen (%), DMSO = DMSO present in culture.

From the RSM analysis, a response optimisation plot was generated (Figure 3.13) where a theoretical yield is entered and predicted input conditions are returned to achieve this yield. In this instance, the theoretical target yield was entered was 250 fmol mg⁻¹ (in order for the response optimiser to have a low and high limit within capable means). This graphic revealed that the optimal input level settings for maximum hA_{2A}R production in the WT strain were 22°C for temperature, pH 6.0, DO 30% in the absence of DMSO for a maximum predicted yield of 240.34 fmol mg⁻¹ which was close to the target yield. It was noted that the d-optimal screening design also generated the same optimal conditions.

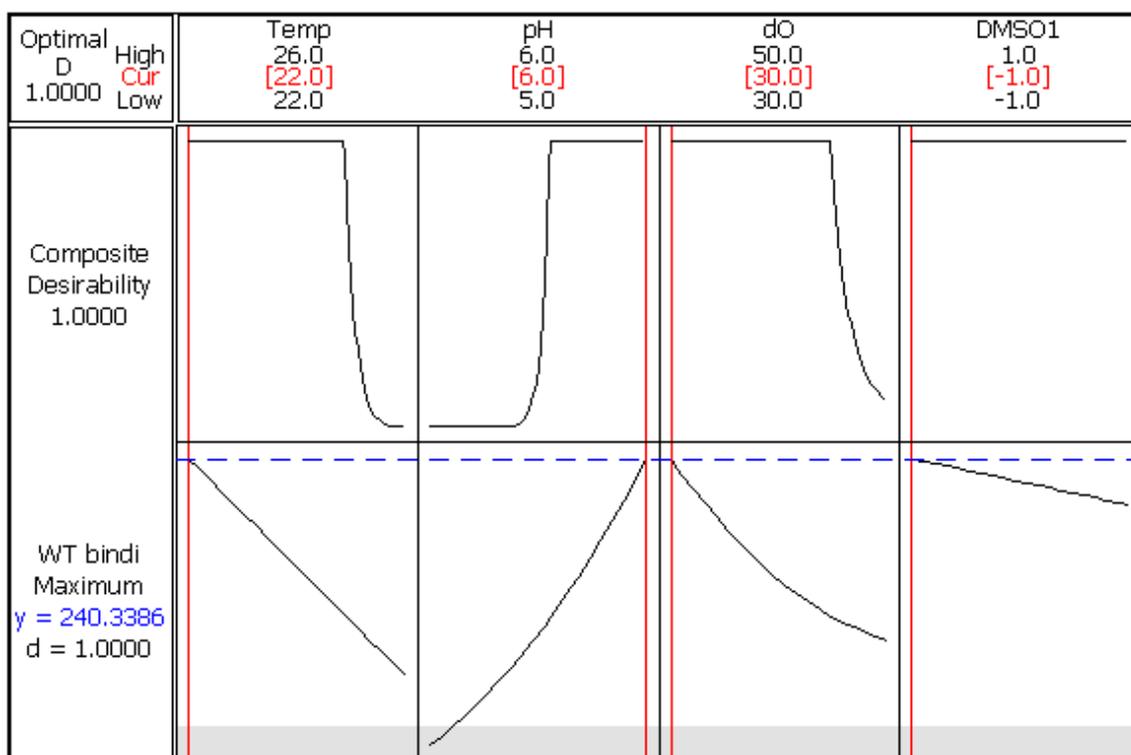


Figure 3.13 Response optimisation plot for WT-hA_{2A}R strain after RSM analysis. The plot is a visual tool from Minitab® which shows how factors affect predicted responses. Each column corresponds to an input factor. The input factor settings in red represent the predicted conditions for the maximum yield possible (blue value in left bottom). The red vertical line can be moved (in the software programme) and as the input factor is adjusted, the predicted WT-hA_{2A}R binding yield will change. The composite desirability is a measure of how good the prediction is and a target of 1.000 is desirable.

From the response optimisation plot, it was possible to generate a contour and wire-frame plot where the predicted optimal conditions were visualised (Figure 3.14). The contour plot (Figure 3.14A) shows the green colour deepen as the yield increases, therefore showing that optimal yields predicted are at low temperature, high pH, low DO and no DMSO present in the culture for the WT-hA_{2A}R strain. The surface plot shows a 3 dimensional representation of the contour plot with the same interpretation. These plots are useful as they also give an indication of further testing that would be required. For instance, the response area in the contour plot around pH 6.0 and temperature 22°C shows that further experiments could be carried out to explore input conditions lower than 22°C and higher than pH 6.0 to see if the yield could be improved further.

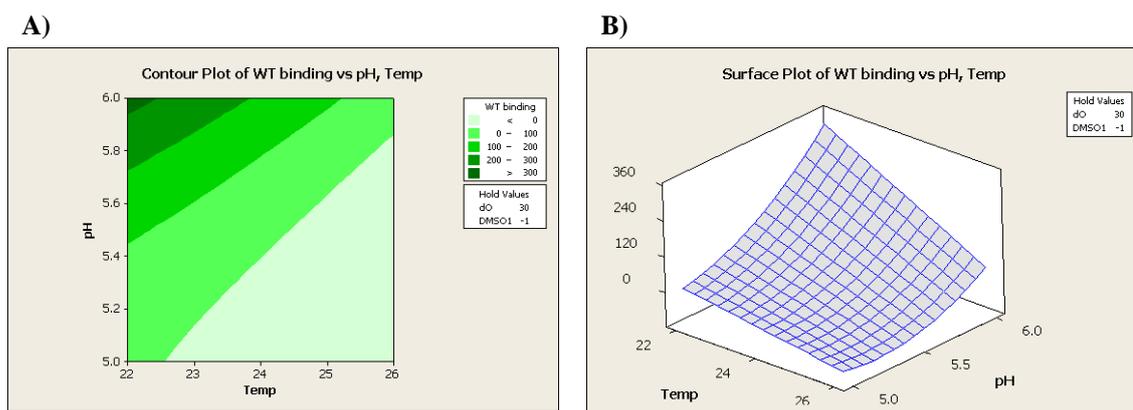


Figure 3.14 Contour and surface wire-frame plot for WT-hA_{2A}R strain after RSM analysis.
 A) Contour plot where the darker the green colour, the higher the desired yield. The x axis is temperature and the y axis is pH. The DO and DMSO are fixed at 30% and with no DMSO. B) Wire-frame or surface plot showing a more 3D view of the contour plot with temperature, pH and specific binding activity in the x, y and z axes. The hold values for DO and DMSO are 30% and no DMSO as for the contour plot.

3.3.1. Examination of residuals

A visual residual analysis after the RSM identified unusual points that did not fit the model as had been seen for the d-optimal design (section 3.2). Figure 3.15A shows the residuals plots for the RSM for the WT-hA_{2A}R strain. From the plots it can be seen that some points are unusually high or low. Also, there is a cyclic trend that is observed in the versus order plot (indicated by a black curve super-imposed). This plot illustrates the residuals observed for each experimental combination as it was run. The 4 data points that form the peaks and troughs of the curve were investigated to see if there was any unusual behaviour in the wells. The 4 data points were found to be 2 replicates of the same condition: 22°C, pH 6.0, 30% DO and no DMSO which was well 1C in the Micro-24 well plate and 22°C, pH 5.5, 30% DO and no DMSO which was well 2B in the Micro-24 well plate. This is a notable observation as these are the optimal predicted conditions from the RSM and the d-optimal linear analysis (22°C, pH 6.0, 30% DO and no DMSO) for high hA_{2A}R yield. Figure 3.15 shows 1 WT-hA_{2A}R replicate of the actual set-point data for pH and DO during the experiment. It is observed that for both wells (Figure 3.15B), the pH is maintained but the DO does not achieve the target set-points. This is discussed further in section 3.4.

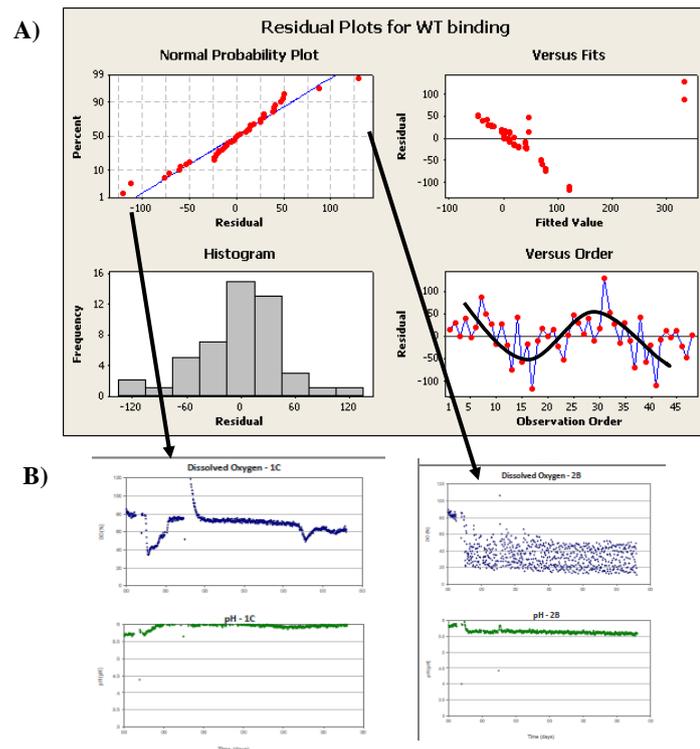


Figure 3.15 Residual plots for WT- $hA_{2A}R$ strain after RSM analysis showing actual DO and pH set-points for the whole experiment. A) the plots shown are normal probability, versus fits, histogram and versus order. Versus order plot has a black curve super-imposed to highlight cyclic behaviour of the observations B) 1C (left) and 2B (right) wells show the actual DO data (blue and top) and actual pH data (green and bottom).

The RSM analysis enabled a predictive equation for the model to be generated and assessed the optimal input conditions for $hA_{2A}R$ yield in the *S. cerevisiae* wild type strain, which were 22°C, pH 6.0, 30% DO in the absence of DMSO. The R^2 (adj) was improved to 63.41% but analysis of the residuals still revealed unusual points and a cyclic trend. The unusual residuals were associated with DO set-points never reaching the target input condition. Sources of error were therefore considered more generally to see if any can explain unusual behaviour in the system.

3.4. The importance of a robust experimental set-up to minimise error

For a DoE to be successful in predicting optimal conditions and generating a robust model, obtaining and using reliable data is crucial. This was considered briefly in sections 3.2. and 3.3.1. where actual data set-points of input parameters impacted the residual analysis and therefore the statistical robustness of the DoE model. This section explores further the influence of instrument and biological error on DoE modelling.

3.4.1. Instrument challenges

As with any piece of laboratory equipment or instrument, the Micro-24 comes with constraints and restrictions when it is used. This is to ensure that the instrument calibration and controls are

working correctly and are within the instrument specifications. However, when carrying out certain methods, such as DoE, these can be prohibitive. For example, when setting up the input factors for the DoE design, the DoE states which order to run an experiment. In this situation, this is difficult to follow, as the temperature parameter is a physical constraint for the Micro-24 instrument. In order for temperatures to reach the set-point as much as possible, the designs with different temperatures must be ordered in a specific way on the Micro-24 well plate. For example, in the case of the d-optimal design tested in section 3.2., the temperature input values were 22°C, 24°C and 26°C. In order for the temperatures to reach as close to the set point as possible, on the Micro-24 well plate, the lowest temperatures were assigned to the back of the plate and the highest at the front (Figure 3.5). This was because the coolant fans are positioned at the back of the instrument. The impact of this requirement on the DoE is large. This means that no randomisation can happen which contributes to the variation in the data when analysis is performed.

The Micro-24 plate wells each include a DO tension sensor, pH optical sensor, a sparging port and thermal conduction pads, all of which control and monitor the three parameters (temperature, pH and DO). The pH and DO of the culture are controlled by a set-point by sparging gas through the medium, in this instance, oxygen for the DO and carbon dioxide or ammonium hydroxide gas for the pH. Figure 3.16 show actual pH and DO data for one replicate of the WT-hA_{2A}R experiment for each well. The pH is shown as green points and the DO as blue and the graphs are laid out as the Micro-24 well plate would be. For the actual pH data during a run, only one well did not meet the expected set-point which was well 2D (indicated by red cross on the green pH axis Figure 3.16). For the DO however, most of the levels did not reach the set-points (red crosses on DO blue axes), in fact only 3 out of the 24 experiments had DO set-points that reached the desired level (indicated by black circles). Additionally, there were some wells (indicated by green circles) that eventually met the desired DO set-point but took most of the run time to achieve it or had high oscillations that spanned 30-50%. Figure 3.17 shows a successful calibration test of the DO sensors where oxygen is applied to each well (containing water only) for 15 min at a 90% set-point. The oxygen is then switched off and the actual DO levels returned to between 0-10% within 10 min and these pass the calibration specifications of the instrument. This shows that the instrument was calibrated correctly which suggests that the failure to reach DO set-points could be due to other factors associated with the experiment such as the DO consumption behaviour of the yeast cultures. Therefore the DO control was investigated by optimising the PI settings (proportional, integral) settings in the system. For future work, a new development from Pall will enable the oxygen uptake rate of the cells to be measured and this is an important feature to have for yeast cultures. This is because they rely on sufficient oxygen concentrations, especially strictly aerobic yeast

such as *P. pastoris* and hence monitoring the actual uptake rate would be a more meaningful parameter.

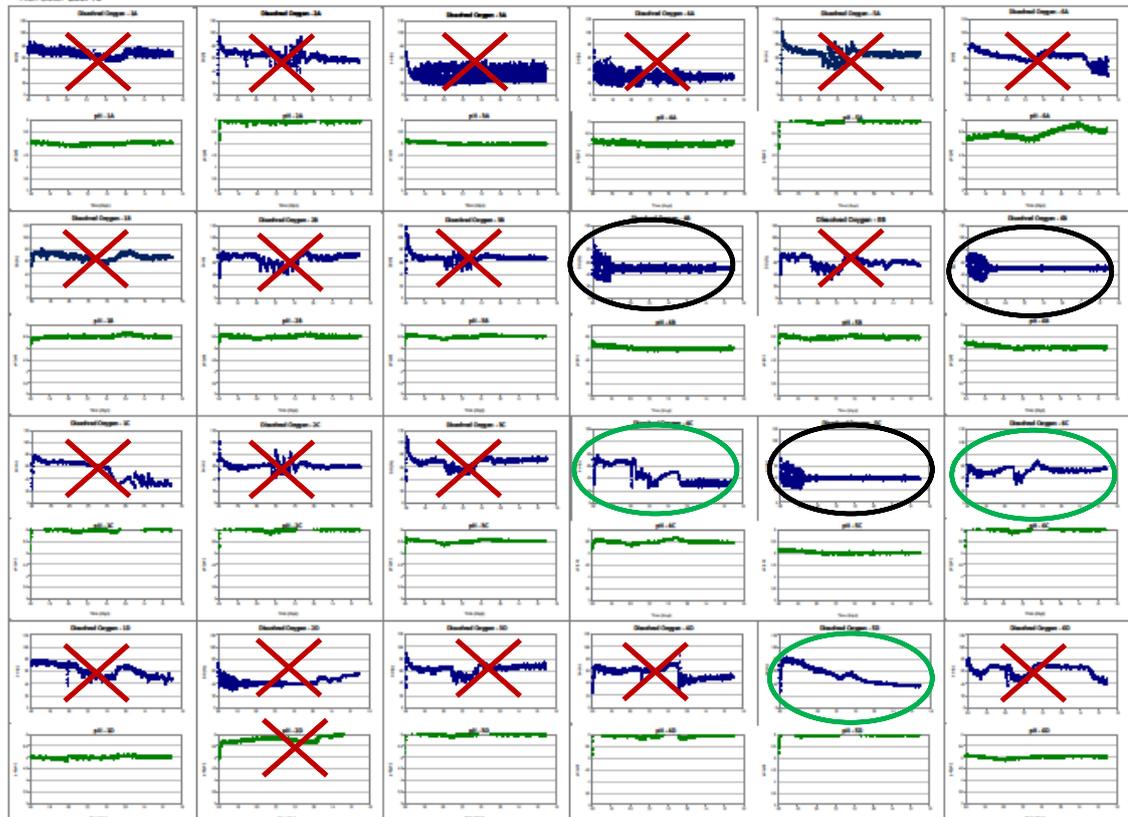


Figure 3.16 Actual DO and pH data for WT-hA_{2A}R experiment. Data shown are for one replicate experiment. The graphs are organised matching the Micro-24 well plate layout (wells 1A->6A across and 1A->1D down). Each graph shows the actual DO data (blue) and the actual pH (green) data. Black circles highlight wells that have reached the set-points for the DO, green circles indicate wells that eventually reached the set-point and red crosses indicate that the pH or DO did not reach the set-point.

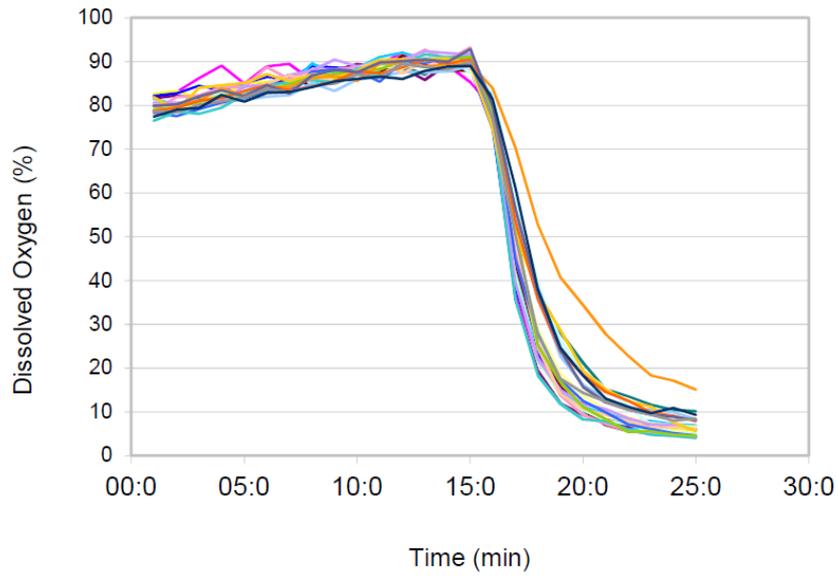


Figure 3.17 DO calibration test. 60%:40% oxygen:nitrogen gas mix applied at 2 bar pressure to Micro-24 microreactor well plate with water for 15 min at a 90% DO set-point. The oxygen supply was switched off and the DO dropped between 0-10% within 10 min. Each line represents the actual DO level for a single well in a Micro-24 well plate.

It is likely that these instrument effects have a very large impact on the DoE data and form part of the explanation for the unusual residuals and also the presence of different patterns or trends observed in the normal probability plots.

3.4.2. Biological sample error

For the analysis of hA_{2A}R, data measurements were not directly from the culture supernatants from each Micro-24 well. Instead, the assays were on membrane-bound protein and therefore membrane preparations need to be carried out to remove the yeast cell wall (Figure 3.18).

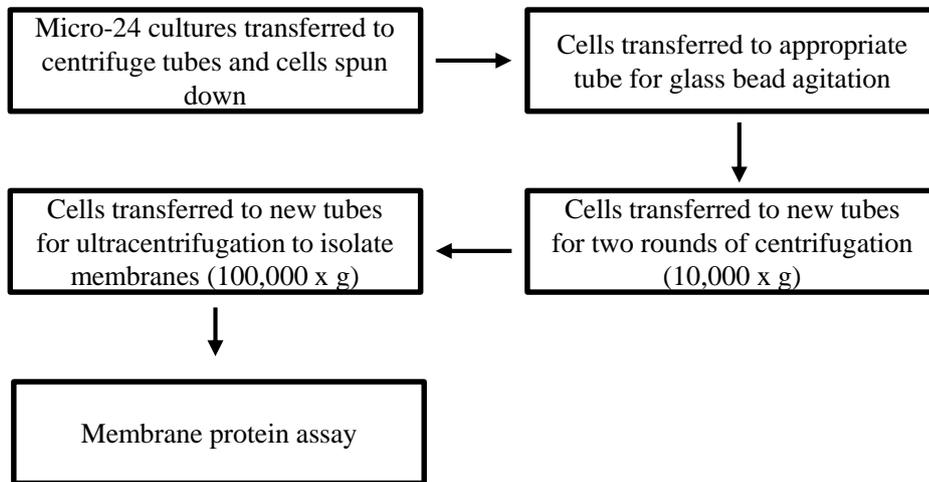


Figure 3.18. Membrane preparation process for Micro-24 yeast cultures. The flow diagram shows the steps required to isolate membranes in order to carry out an assay on the membrane protein, for example specific binding activity. See text for detailed description of steps.

It can be seen that several transfers of the cells to and from tubes are required depending on the step. The actual action of the glass bead agitation poses a potential issue with regards to membrane protein integrity. There is normally increased heat generation during this process and although safeguards (such as carrying out the procedure at 4°C) are in place, little is known as to how much the membrane protein is damaged during this process. This process must be considered as an additional source of error for the DoE as the output response, the specific binding activity, is not directly measured in the culture. For soluble proteins, this is not the case, as an assay can be directly applied to the culture supernatant after the Micro-24 run has completed.

3.5. Discussion

The purpose of the work performed in this chapter was to examine the effects of varying several input conditions on the specific binding activity of recombinant hA_{2A}R in yeast by employing a statistical Design of Experiments (DoE) approach.

DoE is widely used in industry to minimise the number of experiments carried out whilst simultaneously not compromising the quality of the data in an attempt to move away from trial

and error methods. In this chapter, a DoE screen examined the influence of temperature, pH, DO and the presence of an additive (2.5% DMSO) on recombinant hA_{2A}R yield. The input condition levels were temperature at 22°C, 24°C and 26°C; pH at 5.0, 5.5 and 6.0; DO at 30, 40 and 50% and the presence or absence DMSO. Therefore the type of design was a 3^k or 3³ plus one qualitative factor. These experiments were carried out in a small scale, high throughput bioreactor system called the Micro-24 microreactor (Pall Corporation).

The experiments were carried out in three *S. cerevisiae* strains producing hA_{2A}R, *S. cerevisiae* wild-type hA_{2A}R (WT-hA_{2A}R), *S. cerevisiae* BMS1 hA_{2A}R (BMS1-hA_{2A}R) and *S. cerevisiae* TM6* hA_{2A}R (TM6-hA_{2A}R). Using *S. cerevisiae* over *P. pastoris* was ideal for these experiments as there was no fed-batch type phase involved during the cultivation. Therefore it was not necessary to halt the Micro-24 microreactor cultivation run in order to add the feed manually to each well. Holmes and colleagues (Holmes et al., 2009) used *P. pastoris* producing GFP to run the DoE but had to halt the run in order to add more methanol feed and whilst this worked perfectly well, for the purpose of this research, the closed, batch fed system with glucose as the sole carbon source was ideal in order to simplify the experiment. The *S. cerevisiae* strains were selected due to prior work carried out by Bonander and colleagues for the BMS1-hA_{2A}R strain (Bonander et al., 2005; Bonander et al., 2009) and Otterstedt and colleagues for the TM6-hA_{2A}R strain (Otterstedt et al., 2004) where both were found to be high yielding for recombinant protein production.

A d-optimal design was selected, which reduced the required number of experimental runs when compared to the classical screening method (full factorial,) which would have demanded an impractically large number of experimental runs. The d-optimal design outlined 24 different experimental combinations of the 4 input factors at 3 different levels (except the DMSO addition which was either present or absent). The output response was defined as the specific binding activity of the hA_{2A}R. This was entered back into the d-optimal design. An ANOVA analysis was carried out to see if there were any input condition combinations that provided optimal hA_{2A}R yields as measured by specific binding activity. Significant effects from the input factors were shown and included the WT-hA_{2A}R strain where temperature at 22°C was found to give a significant increase in specific binding activity (p value = 0.001). For the BMS1-hA_{2A}R strain, DO at 40% was found to give a significant increase in specific binding activity (p value = 0.018). Finally, for the TM6-hA_{2A}R strain, it was found that temperature at 24°C and the presence of DMSO together gave a significant increase in the specific binding activity (p value = 0.008), as well as at pH 5.5 and the presence of DMSO, which also gave increased specific binding activity (p value = 0.035). The R² (adj) values for the WT-hA_{2A}R was 52.74%, BMS1-hA_{2A}R was 25.82% and TM6-hA_{2A}R was 39.10%. The residual plots

further indicated that the linear regression model was not the best fit for these experiments, drawing these conclusions into question.

Applying a quadratic model (RSM) (Montgomery, 2005) to the DoE data helped achieve an improved model with optimal conditions for increased hA_{2A}R production. This was done for the WT-hA_{2A}R strain only as the R² values for the BMS1-hA_{2A}R and the TM6-hA_{2A}R strains were low and therefore were not analysed further. For the WT-hA_{2A}R however, the optimal conditions were 22°C, pH 6.0 and DO 30% with no DMSO for hA_{2A}R production. The contour plot pointed to areas on the optimal range that could be further improved for optimisation studies, where the temperature could be tested lower than 22°C and the pH higher than 6.0. These conditions are not unexpected since hA_{2A}R production has been optimal in other studies at 22°C but in *P. pastoris* strains only (Fraser, 2006; Singh et al., 2008; Singh et al., 2010; André et al., 2006; Lundstrom et al., 2006). The pH for yeast cultures tends to be between 5.0-6.0 for optimal growth. Here, the ideal pH was 6.0 for hA_{2A}R production and further hinted at a higher pH from the contour plot. This tends to be at the buffered pH value in complex media such as BMGY and BMMY (André et al., 2006). With regards to DO, little work or no work has been done to determine the optimal DO for hA_{2A}R production. It is not clear why 30% DO was optimal for this strain, however, it could postulated that this wild type strain is not highly aerobic like *P. pastoris* (Cereghino and Cregg, 2000) or *S. cerevisiae* TM6* (Otterstedt et al., 2004; Ferndahl et al., 2010) and therefore the lower DO is more amenable to optimal hA_{2A}R production. In the case of DMSO, it was found by André and colleagues and Lundstrom and colleagues (André et al., 2006; Lundstrom et al., 2006) that when they carried out a GPCR screen, the binding activity increased in more than half the GPCRs when produced in DMSO. It is thought that the cell membrane permeability increases in the presence of DMSO (Murata et al., 2003). However, the DMSO addition had no effect on the hA_{2A}R production in these experiments. These conditions were determined from the RSM DoE and resulted in a predictive model or equation. However, it should be considered that the equation may not be very reliable since the error being high and especially not being able to maintain the DO adequately. The next step would be to carry out validation experiments to test if the predictive model is acceptable. Holmes and colleagues demonstrated this for soluble GFP where the line of parity for the experimental data versus the predicted data was R² = 0.57 (Holmes et al., 2009).

The causes of the high variation in the data were investigated globally for all the runs carried out by further inspecting the actual set-point values of the input factors in the Micro-24 instrument and whether they had been achieved. It was found that many of the DO set-points were not reached. Another cause of error was the manipulation of the cells before an assay could be applied. Carrying out individual membrane preparations on each well sample introduced user handling and the potential to lose sample or even affect the receptor protein.

These findings show that DoE statistical methods can be used for streamlining optimisation experiments. The impact of this is that many growth factor and recombinant protein production parameters may be investigated simultaneously. However, issues need to be addressed such as reducing the amount of variability within an experimental process. These can be done by optimising the downstream processing of the membrane protein preparation methods via automated processes.

The work derived from this chapter is published in a book chapter review (Appendix A9); (Bora et al., 2012) and a review article (Appendix A9); (Bawa et al., 2011).

Chapter 4: An investigation into the induction phase of *P. pastoris* cultures: production of hA_{2A}R as a case study

This chapter describes an investigation of the induction protocol in *P. pastoris* cultures producing hA_{2A}R at both laboratory and pilot scale. The construct used in this chapter was selected in collaboration with the industrial sponsor, AstraZeneca Ltd, on account of its immediate availability at Aston University; in Chapter 5, a second, multiply-tagged construct is investigated that was designed with AstraZeneca's input. A major finding of the work described in Chapter 4 is the detection of pre-induction recombinant protein production, which was also confirmed for soluble green fluorescent protein (GFP). When glucose, which has been shown to repress *AOX* expression, was the pre-induction carbon source, GFP was still produced in the pre-induction phases. GFP yields were higher and biomass yields were lower than in an equivalent glycerol-grown culture. Both hA_{2A}R and GFP were also produced in methanol-free cultivations; functional protein yields were maintained or increased after depletion of the carbon source. Analysis of the pre-induction phases of 10 L pilot scale cultivations also demonstrated that pre-induction yields were at least maintained after methanol induction, even in the presence of cytotoxic concentrations of methanol.

4.1. Analysis of growth rates during methanol induction in pilot-scale bioreactors

This section describes the work performed to investigate methanol induction feeding regimes in hA_{2A}R producing *P. pastoris* in 35 L pilot-scale bioreactors.

4.1.1. Expression strain

The human _{2A} adenosine receptor (hA_{2A}R) construct considered for this study was designed by Dr. Niall Fraser, University of Glasgow, United Kingdom (Fraser, 2006). A de-glycosylated version of the sequence was used where the N-linked glycosylation site at Asn154 was mutated to Gln; it is referred to as dG-hA_{2A}R in this chapter. The hA_{2A}R sequence is tagged with FLAG and His₁₀ tags on the N-terminus of the receptor and includes an α -factor leader sequence (Figure 4.1).

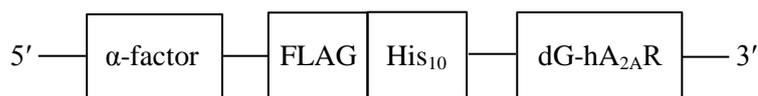


Figure 4.1 Schematic diagram for dG - hA_{2A}R construct. Construct design schematic for dG-hA_{2A}R designed by Dr. Niall Fraser. Design includes a FLAG and His₁₀ tag and α -factor leader sequence.

The sequence was cloned into the pPICZ α A vector (Life Technologies Corporation) and transformed into the *P. pastoris* X33 strain by electroporation; and the transformed cells were grown on increasing concentrations of zeocin containing plates. Three colonies (clones 26, 27, 28) were selected from 250 $\mu\text{g mL}^{-1}$ zeocin selection plates and were cultured in YPD broth with 250 $\mu\text{g mL}^{-1}$. When the cells reached exponential growth, glycerol stocks of the cells were prepared for storage at -80°C . This work was previously carried out by Dr. Richard Darby, Aston University, United Kingdom. These clones (26, 27 and 28) were screened for dG-hA_{2A}R expression by culturing them in BMGY flasks (50 mL culture volume) overnight at 30°C . The cells were then transferred to BMMY (0.5% methanol) media shake flasks and cultured to anOD₆₀₀ of 1. Samples were taken at 24 h and 48 h post-induction and membrane preparations and immuno-blot analyses were performed. Figure 4.2 shows expression of dG-hA_{2A}R from all the clones selected; clone 26 at 48 h post-induction in BMMY medium showed the strongest signal at about 50 kDa along with other fainter bands demonstrating a mixture of degraded forms of the receptor and where the Kex2 protease was unable to cleave the alpha mating factor signal sequence.

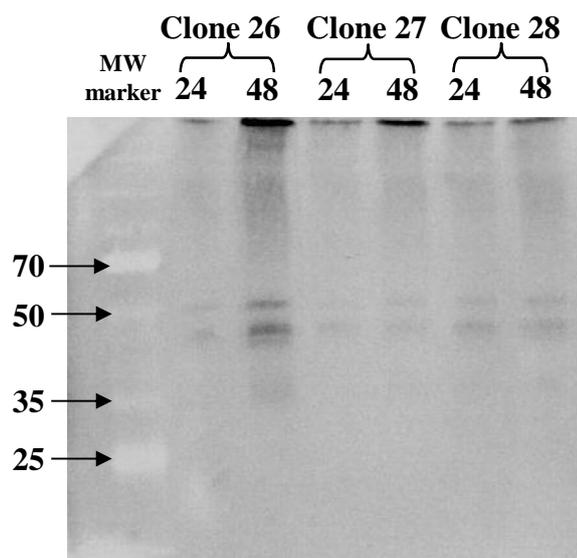


Figure 4.2 Immuno-blot from X33-dG-hA_{2A}R colony screen. Immuno-blot of dG-hA_{2A}R production from colony screening of BMMY cultures (0.5% methanol) in shake flasks. Lanes labelled Clone 26 (24 and 48) are membrane fraction preparations for clone 26 *P. pastoris* cells harvested at 24 h and 48 h from the BMMY cultures. Lanes labelled Clone 27 (24 and 48) are membrane fraction preparations for clone 26 *P. pastoris* cells harvested at 24 h and 48 h from the BMMY cultures. Lanes labelled Clone 28 (24 and 48) are membrane fraction preparations for clone 26 *P. pastoris* cells harvested at 24 h and 48 h from the BMMY cultures. Ladder is 250 kDa with 70, 50, 35 and 25 kDa indicated on figure. The primary antibody used in the method was anti-His antibody (Serotec).

Since colony 26 gave the strongest band, this colony was selected for subsequent cultivations. Before the strain was used for a 35 L pilot-scale cultivation, the strain was cultured using a 2 L

bench-top bioreactor to ensure growth in (BSM) minimal medium, to retrieve information about growth rates for this strain and also to test methanol induction and hence, hA_{2A}R production at a larger scale. Figure 4.3 shows a 2 L bioreactor cultivation for the X33-dG-hA_{2A}R colony 26 strain. The bioreactor set-points were 30°C, pH 5.0, 30% DO and stirrer speed at 500 rpm. The cultivation shows the glycerol batch phase (I), the glycerol fed-batch phase (II), the transition phase (III) and the methanol induction phase (IV) where the phases are separated by blue vertical lines. The batch phase (I) started off with 40 g L⁻¹ glycerol in the bioreactor and as that was consumed by the cells, the DO (green line) started to decrease and the carbon dioxide (red line) started to increase, indicating that the cells were growing. This was also reflected in the optical density measurements (black line). After about 20 h, the fed- batch phase (II) was started by adding a constant feed of 50% glycerol at 12 mL h⁻¹ for 4 h. The transition or starvation phase was for about 2 h, where the temperature was lowered to 22°C, since it is established that lowering the temperature to 22°C during induction increases the functional yield of hA_{2A}R (Singh et al, 2008; Singh et al, 2010 and Fraser, 2006) and the fed-batch phase was stopped. Within this phase, a DO spike was detected indicating that the glycerol had been consumed and therefore methanol induction could begin. The methanol induction phase (IV) was run for the duration of the cultivation at 2 mL h⁻¹ at 50% methanol. However, at about 45 h, the DO levels started to rise (cells not demanding oxygen) and the carbon dioxide levels started to decrease, indicating that the cells had stopped growing. This also showed that the methanol feed rate probably exceeded the rate of consumption of the cells. Samples were taken throughout the cultivation and stored at -80°C for further use.

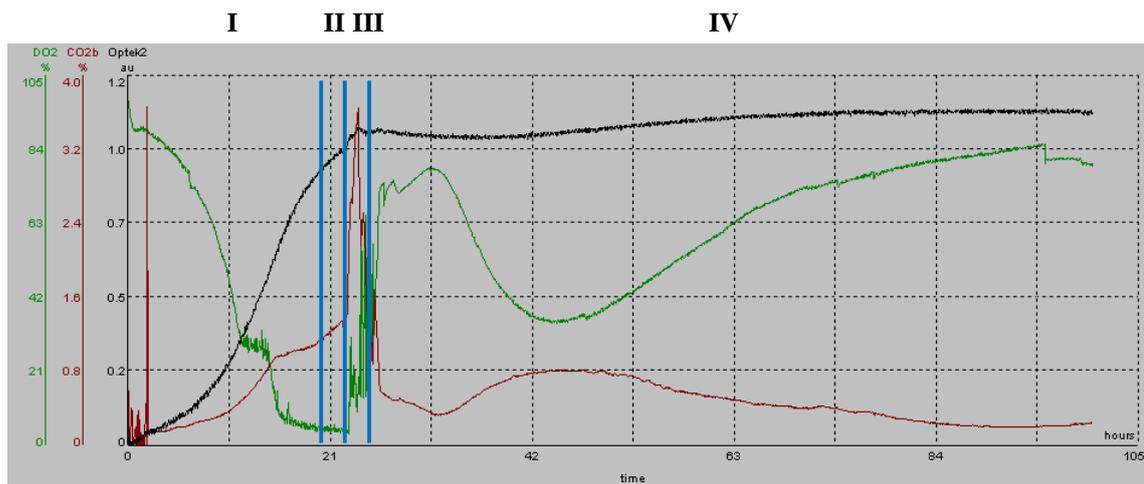


Figure 4.3 Bioreactor cultivation of *P. pastoris* X33 dG-hA_{2A}R. Optical density (black line), carbon dioxide off-gas (red line) and DO (green line) are shown for *P. pastoris* X33 dG-hA_{2A}R cultivation in BSM media with PTM₁ salts. The batch phase (I) was with 40 g L⁻¹ glycerol, the fed-batch phase (II) was with 50% glycerol constant feed of 12 mL h⁻¹ and the methanol induction phase (IV) was carried out with 50% methanol at 2 mL h⁻¹. The phases are marked by vertical blue lines. The bioprocess set-points were at 30°C, pH 5.0 and 30% DO. The

temperature was lowered to 22°C for the induction phase at the beginning of the transition phase (III).

From the optical density measurements, the specific growth rate (μ) was calculated for the second half of the glycerol batch phase (I) since the growth was exponential; $\mu = 0.21 \text{ h}^{-1}$. Two of the samples were taken at 42 h and at 90 h during the methanol induction phase (IV) where membrane preparations and immuno-blot analysis was performed. Figure 4.4 shows expression of dG-hA_{2A}R from clone 26 from a minimal medium bioreactor run with a band at about 50 kDa for both time-points. The blot shows fainter, larger bands which could suggest forms of the receptor where the Kex2 protease was unable to cleave the alpha factor sequence. This could be confirmed by extracting this band from the gel and after suitable purification and performing mass spectrometry analysis. A single point radio-ligand binding assay was performed on clone 26 to give specific activity of $4.0 \pm 0.2 \text{ pmol mg}^{-1}$ which confirmed expression of correctly-folded dG-hA_{2A}R.

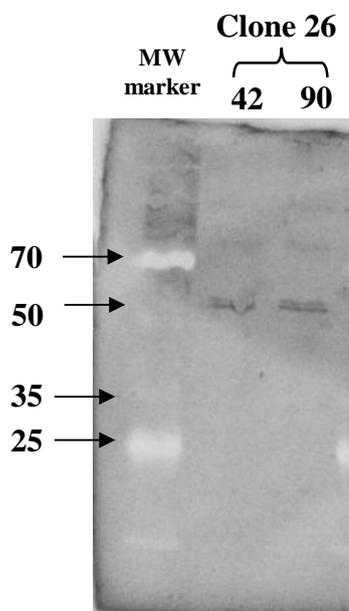


Figure 4.4 Immuno-blot from X33-dG-hA_{2A}R 2 L bioreactor cultivation (clone 26). Immuno-blot of dG-hA_{2A}R from X33-dG-hA_{2A}R 2 L bioreactor cultivation (clone 26). Lanes are for Clone 26 membrane fraction preparations of *P. pastoris* cells harvested at 42 h and 90 h during the methanol induction phase. Ladder is 250 kDa with 70, 50, 35 and 25 kDa indicated on figure. The primary antibody used in the method was anti-His antibody (Serotec).

From these cultivations, clone 26 of the *P. pastoris* X33-dG-hA_{2A}R strain was used for the subsequent 35 L pilot-scale bioreactor studies at the AstraZeneca fermentation facility.

4.1.2. Methanol feed regimes

Selecting optimal culture conditions such as temperature, pH, and the amount of DO in the culture medium as well as the addition of chemical additives and ligands for optimal GPCR

production in *P. pastoris* cultivations is often desired, since achieving high yields is a requirement for downstream structural analysis (Andre et al., 2006, Mattanovich et al., 2009, Cereghino and Cregg, 2000, Cos et al., 2006b). Furthermore, Holmes and colleagues showed that along with culture conditions influencing protein production, the induction feeding strategy of *P. pastoris* is also important (Holmes et al., 2009) and this is particularly relevant when growing cultures at a large scale (10 – 2000 L bioreactors). Detailed studies examining specific growth rates (μ) during the methanol induction phase of *P. pastoris* cultures have shown that μ influences recombinant protein yields and cell biomass (Potvin et al., 2012). In particular, this could be achieved by applying exponential feeding strategies to maintain cell growth at a constant specific and desired growth rate (μ_{set}) (Jahic, 2002). This is a common strategy employed in industry where large bioreactors are used for recombinant protein production and where controlled addition of the nutrient feed is essential in order to minimise ‘over-flow’ metabolism (side metabolites) which could affect the quality of the product (Lee et al., 1999; Wlashin et al., 2006). Exponential feeding follows the empirical model of cell growth to regulate the feeding rate i.e. in the correct nutrient and operating conditions; the cells grow exponentially thereby achieving high cell biomass (Aulicino et al., 2010). This method has been used in previous studies and is termed ‘feed forward strategy’ since a theoretical specific growth rate (μ_{set}) of an organism can be achieved by feeding the cells in an exponential manner and hence will influence recombinant protein production and cell biomass (Potvin et al., 2012). This is described as an equation and can therefore be programmed into specialised software to deliver the exponential feed via the bioreactor controllers and feed pumps:

$$F(t) = F_0 e^{\mu_{set}t}$$

$F(t)$ is the feed rate ($g\ h^{-1}$) at time, t (h); F_0 is the initial feed rate and μ_{set} is the desired theoretical constant specific growth rate (h^{-1}).

Several studies have been published on the influence of μ_{set} for soluble proteins (Çelik et al., 2009; Çelik et al., 2010; Jungo et al., 2007; Kobayashi et al., 2000; Zhang et al., 2000), however, no detailed studies are available for membrane protein production and feed forward strategies. This chapter therefore describes the effect of exponential feeding of methanol at two different μ_{set} on hA_{2A}R production in *P. pastoris* in 35 L pilot-scale cultivations. The two μ_{set} were $0.01\ h^{-1}$ which was termed μ_{low} and $0.03\ h^{-1}$ which was termed μ_{high} . These μ_{set} were chosen as the literature had suggested that μ_{set} of $0.03\ h^{-1}$ for soluble proteins such as human EPO and human growth hormone had resulted in high yields (Çelik et al., 2009; Çelik et al., 2010). A low μ_{set} was also selected ($0.01\ h^{-1}$) as studies have also shown that for soluble proteins the productivity decreased but the cell biomass increased at lower μ_{set} than $0.03\ h^{-1}$. It was therefore interesting to see if this was the same for membrane protein production.

4.1.3. Pilot-scale cultivations

Two simultaneous pilot scale bioreactor cultivations (10 L starting volume) of *P. pastoris* X33 strain expressing dG - hA_{2A}R were compared following the same batch, fed-batch and transition feeding regimes. During the induction phase, different methanol feed profiles were applied; a low methanol feed profile was devised to give $\mu_{\text{set}} = 0.01 \text{ h}^{-1}$, while a higher methanol feed profile was devised to achieve $\mu_{\text{set}} = 0.03 \text{ h}^{-1}$ followed by 0.05 h^{-1} . Cultures were monitored for optical density at 600 nm (OD₆₀₀), dry cell weight (DCW), residual glycerol and methanol concentrations in the culture supernatant, total membrane protein yield and the yield of correctly folded dG - hA_{2A}R, as measured by radio-ligand binding with the tritiated antagonist, ZM241385.

Table 4.1 shows the phases of the two simultaneous bioprocesses examined in this study. A total cultivation time of 91.2 h included the following phases: batch (phase I), fed-batch (phase II), transition (phases IIIA and IIIB) and induction (phase IV and V). An identically conducted batch phase (I) on 10 g L^{-1} glycerol lasted for 19.1 h for both bioprocesses. The IIA phase had an exponential glycerol feed of $\mu_{\text{set}} = 0.15 \text{ h}^{-1}$ and IIB was at $\mu_{\text{set}} = 0.03 \text{ h}^{-1}$. The transition (III) phase consisted of a short starvation period (IIIA) followed by a brief pre-induction constant methanol feed at $4 \text{ g L}^{-1} \text{ h}^{-1}$ (IIIB), for both bioreactors. During the induction (IV) phase, one cultivation (μ_{low}) was subjected to a 46 h exponential methanol feed rate, predicted to achieve $\mu_{\text{set}} = 0.01 \text{ h}^{-1}$. The second cultivation (μ_{high}) was subjected to $\mu_{\text{set}} = 0.03 \text{ h}^{-1}$ for 41 h (phase IV) and then $\mu_{\text{set}} = 0.05 \text{ h}^{-1}$ for a further 5 h (phase V). At this point, the characteristic DO spike indicated full consumption of the glycerol carbon source and the stirrer speed decreased as result of the DO control (Fig. 4.5A). In the subsequent 20 h fed-batch phase (II), the same exponential, growth rate-limiting glycerol feed profile of $4 \text{ g L}^{-1} \text{ h}^{-1}$ was applied to both bioprocesses, increasing exponentially at a rate of 0.15 h^{-1} for 10 h and then 0.03 h^{-1} until the end of this phase. The transition phase had two different sections (IIIA and IIIB): in phase IIIA, no feed was applied to the cultivations for 1 h; in phase IIIB a constant methanol feed of $4 \text{ g L}^{-1} \text{ h}^{-1}$ was applied. At the end of phase IIIB, the induction phase commenced differently for the two cultivations.

Table 4.1 Bioprocess events during two simultaneous *P. pastoris* cultivations producing recombinant dG - hA_{2A}R with different exponential methanol feed rates (μ_{set}). I denotes batch phase; IIA and B denotes fed-batch phase; IIIA and B denotes transition phase and IV denotes the induction phase. The μ_{high} cultivation has another induction phase, V where the μ_{set} was adjusted to 0.05 h^{-1} .

Age of Cultivation (h)	Feed Event		Phase		Temperature (°C)
0.0 14.4 16.4 18.4 19.1	↑ Glycerol Batch (10 g) ↓		I		↑ 30 ↓
20.4 22.4 23.6	Glycerol Fed-batch (Exponential feed at $\mu_{set} = 0.15\text{ h}^{-1}$)		IIA		
28.8 38.4 40.6	Glycerol Fed-batch (Exponential feed at $\mu_{set} = 0.03\text{ h}^{-1}$)		IIB		
42.2	Transition (No feed)		IIIA		
42.9 44.4	↑ Transition (Methanol pre-induction constant feed at $46\text{ g L}^{-1}\text{ h}^{-1}$) ↓		IIIB		
45.4 46.4 47.4 62.4 64.4 66.4 68.4 70.4 86.4 86.4 87.4 88.9 90.4 91.2	↑ μ_{low} ↓	↑ μ_{high} ↓	μ_{low}	μ_{high}	↑ 22 ↓
	Induction (Methanol exponential feed at $\mu_{set} = 0.01\text{ h}^{-1}$)	Induction (Methanol exponential feed at $\mu_{set} = 0.03\text{ h}^{-1}$)	IV	IV	
		Induction (Methanol exponential feed at $\mu_{set} = 0.05\text{ h}^{-1}$)	IV	V	

4.2. A low methanol induction feed profile increases biomass yield leading to improved overall volumetric dG-hA_{2A}R yields

Applying a μ_{set} of 0.01 h^{-1} during the methanol induction phase for a pilot-scale cultivation of *P. pastoris* gave an overall increase in cell biomass and recombinant dG-hA_{2A}R yield when compared to the cultivation with a μ_{set} of 0.03 h^{-1} during the methanol induction phase (Table 4.2).

4.2.1. The μ_{low} *P. pastoris* cultivation yields increased biomass when compared to the μ_{high} *P. pastoris* cultivation

Correct control of the required bioprocess settings is verified in Figure 4.5. The DO set-point was maintained at 30%, the pH set-point was maintained at 5 and the temperature set-points were maintained at 30°C for phases I, II and III and 22°C for phases IV and V for both the μ_{low} (Figure 4.5A) and μ_{high} (Figure 4.5B) cultivations. Stirrer speed was set in cascade mode; for the μ_{low} cultivation it increased to its maximum value as the DCW increased, demonstrating an increasing cellular demand for aeration (Figure 4.5A). In contrast, the stirrer speed decreased as the cells ceased to grow in phase IV of the μ_{high} cultivation (Figure 4.5B); in phase V, the stirrer speed decreased at a slightly lower rate as the DCW increased in response to a change in μ_{set} to 0.05h^{-1} .

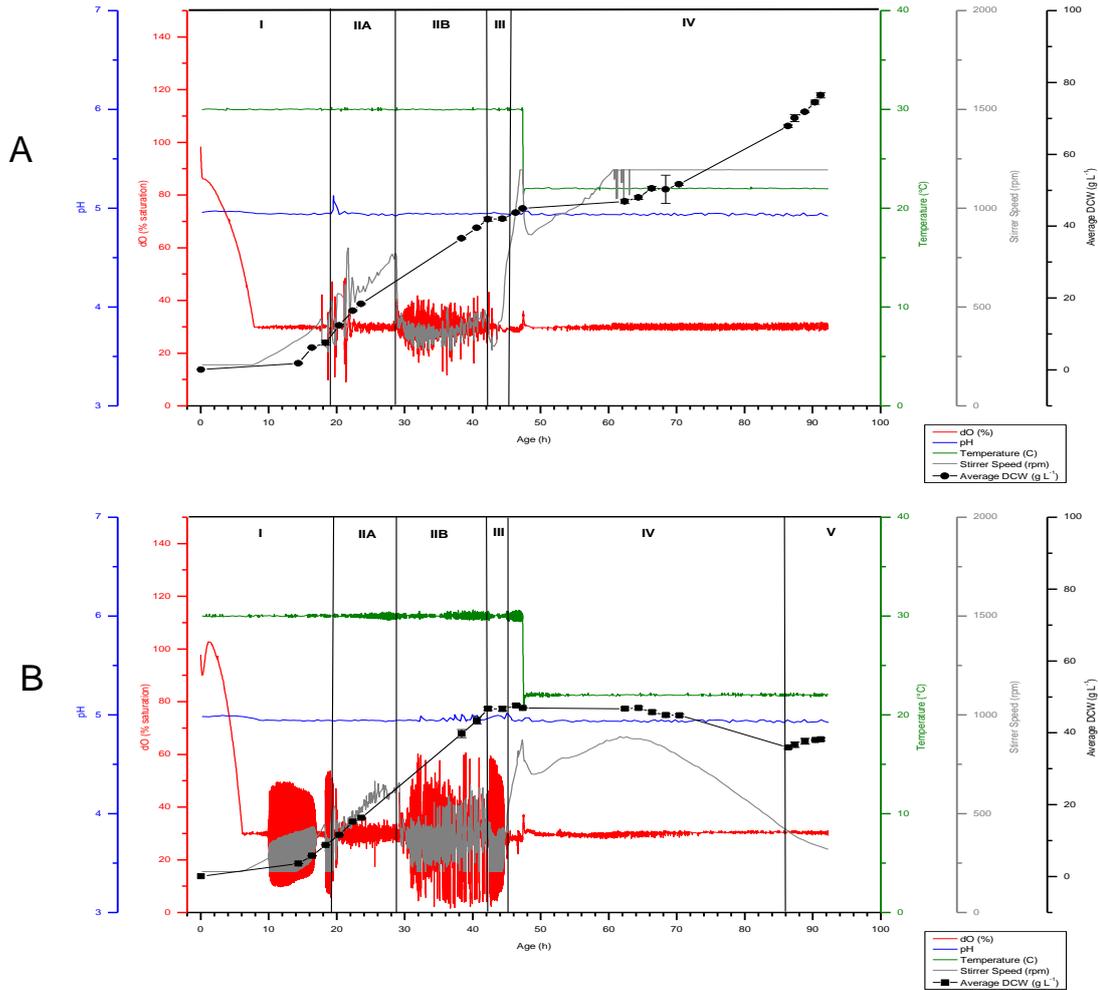


Figure 4.5 Bioprocess parameters for the μ_{low} and μ_{high} cultivations. The bioprocess parameters, DO (red), pH (blue), temperature (green) and stirrer speed (grey), were monitored during the entire duration of the μ_{low} (A) and μ_{high} (B) cultivations. DO and pH were maintained at 30% and 5, respectively; the temperature set-points were maintained as defined in Table 4.1. Stirrer speed (grey) increased as the dry cell weight (DCW; closed circles) increased for the μ_{low} cultivation during phase IV (A). A decrease in stirrer speed for the μ_{high} cultivation was observed as the DCW (closed squares) decreased in phases IV and V (B).

Figure 4.6A shows the OD₆₀₀ and DCW data throughout both cultivations. During phases I and II, the biomass yield remained almost identical, as expected, while differences were apparent in the induction phase (IV and V), most notably around 70 h duration. DCW and OD₆₀₀ increased throughout the methanol feeding phase for the μ_{low} cultivation ($\mu_{set} = 0.01\text{h}^{-1}$). For the μ_{high} cultivation ($\mu_{set} = 0.03\text{h}^{-1}$), the DCW and OD₆₀₀ values increased slightly in the first half of phase IV then decreased after 70 h duration. When μ_{set} was increased to 0.05h^{-1} , DCW and OD₆₀₀ increased slightly. At 91 h, the final biomass yields were 1,057.2 g for the μ_{low} cultivation and 589.8 g for the μ_{high} cultivation. This is apparent in harvest flasks after centrifugation to sediment the cells (Figure 4.6B).

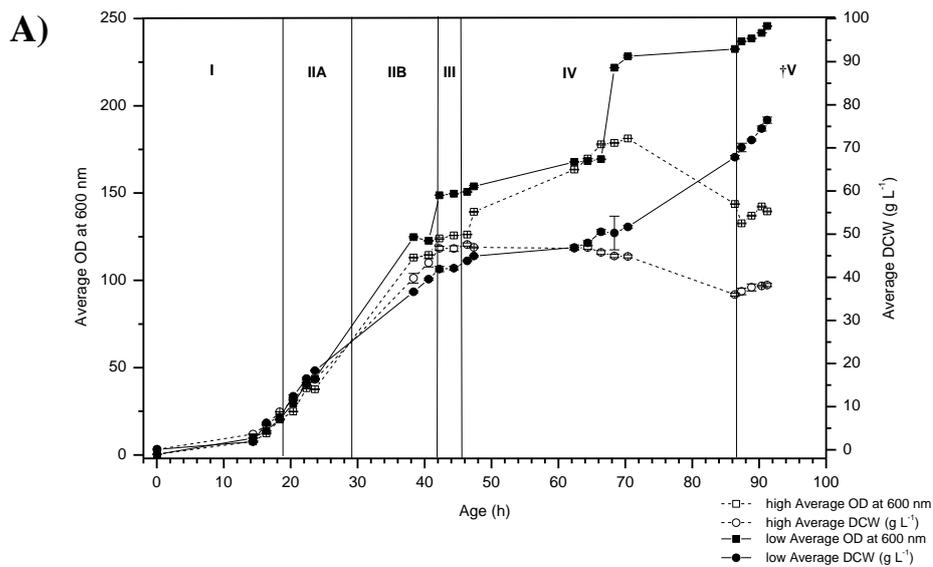


Figure 4.6 Biomass yield of the μ_{low} and μ_{high} cultivation. A) Cell biomass (OD_{600} ; squares and DCW; circles) was monitored throughout the μ_{low} (closed symbols) and μ_{high} (open symbols) cultivations for all phases. Measurements were made in triplicate. †Phase V indicates a μ_{set} setting change for the μ_{high} cultivation only. At 91 h, cell biomass for the μ_{low} cultivation was approximately double that of the μ_{high} cultivation. B) Photograph of centrifuge flasks with 1 L culture showing *P. pastoris* cells for μ_{high} (left) and μ_{low} (right) at the end of the cultivations.

4.2.2. Increased volumetric dG-hA_{2A}R yields are achieved in the μ_{low} cultivation

Table 4.3 shows B_{max} estimates and the amount of biomass generated (DCW and total membrane protein) for both cultivations during the transition and induction phases. Overall, when the cell biomass is taken into account, the μ_{low} cultivation produces improved volumetric yields of dG - hA_{2A}R. In fact, it the μ_{low} cultivation yields 7 times more active receptor in 1 L of

P. pastoris culture when compared to the μ_{high} cultivation (38522 pmol versus 5573 pmol). This finding suggests that the slower feeding of methanol has a positive effect on cell biomass generation and hence increased production of cell membranes leading to increased production of the desired recombinant receptor.

Table 4.2 The yield of dG - hA_{2A}R is higher in the μ_{low} cultivation than the μ_{high} cultivation. The total yield of dG - hA_{2A}R (pmol) and the specific yield of dG - hA_{2A}R per gram of DCW (pmol g⁻¹) were derived from the B_{max} estimates and total membrane protein measurements. Values are calculated per L of culture for the transition and induction phases of the μ_{low} and μ_{high} cultivations and are the mean of triplicate determinations, with the standard error of the mean. Triplicate measurements were from three different membrane preparations. B_{max} estimates were calculated from single-point radio-ligand binding analysis using [³H]ZM241385.

	Phase	Age (h)	DCW (g L ⁻¹)	Total membrane protein yield in 1 L culture (mg)	B_{max} estimate (pmol mg ⁻¹)	Total dG-hA _{2A} R yield in 1 L culture (pmol)	Specific yield (pmol g ⁻¹)
μ_{low}	III A	42.2	41.91 ± 0.6	764.4	4.4 ± 0.1	3363.6	80.3
	III B	44.4	42.10 ± 0.4	2791.1	7.1 ± 0.6	19816.9	470.7
	IV	66.4	50.53 ± 0.5	4084.4	4.5 ± 0.2	18380.0	363.7
	IV	89.0	71.82 ± 0.2	7704.4	5.0 ± 0.2	38522.2	536.4
μ_{high}	III A	42.2	46.71 ± 0.5	813.3	4.4 ± 0.2	3578.7	76.6
	III B	44.4	46.66 ± 0.7	3060.0	4.0 ± 0.0	12240.0	262.3
	IV	66.4	45.76 ± 0.5	3277.8	3.8 ± 0.1	12455.5	272.2
	V	89.0	37.64 ± 0.8	1797.8	3.1 ± 0.1	5573.1	148.1

4.3. Increased dG-hA_{2A}R yields correlate with low residual methanol in the cultivation but cytotoxic levels of residual methanol maintain baseline levels of active dG-hA_{2A}R

Table 4.3 shows the binding activity of dG - hA_{2A}R during the transition (III A and III B) and induction (IV and V) phases for both the μ_{low} and μ_{high} cultivations. The amount of residual methanol and the cumulative addition of methanol in the cultivations is also shown. Membranes isolated from the μ_{low} cultivation had higher binding activity than those from the μ_{high} cultivation; notably, the highest binding activity (7.1 ± 0.6 pmol mg⁻¹) was measured in membranes of cells that had been subjected to constant methanol feeding (phase III; 4 g L⁻¹ h⁻¹). The residual methanol concentration in the μ_{low} cultivation remained below 5 g L⁻¹. Furthermore, the activity was at its highest only when the residual methanol was at 0.5 g L⁻¹ in the system. The activity then decreased from 7.1 ± 0.6 pmol mg⁻¹ to 4.5 ± 0.2 pmol mg⁻¹ when the residual methanol increased by another 0.5 g L⁻¹. Interestingly, the activity increased slightly to 5.0 ± 0.2 pmol mg⁻¹ as the entire residual methanol in the system was consumed by the cells after 23 h of cultivation. For the μ_{high} cultivation, the highest level of binding activity reached

was $4.4 \pm 0.2 \text{ pmol mg}^{-1}$ during the transition phase (IIIA) where no methanol feeding had occurred and the cells were in starvation phase. The binding activity decreased gradually as the residual methanol in the system increased. During phases IV and V, the residual methanol levels remained above 5 g L^{-1} which is considered to be cytotoxic (Guarna et al., 1997). Moreover, the residual methanol reached levels that were about 17 times the cytotoxic concentration (84.7 g L^{-1}). While it is not unexpected that the binding activity was low, it was surprising that any activity was present at all. More methanol was added and at a faster rate for the μ_{high} cultivation when considering the cumulative addition of the methanol to the bioreactor. Methanol was added at a slower rate for the μ_{low} culture and hence less was added in total at the end of the cultivation. At 66.4 h when the exponential feeding began (phase IV), for the μ_{low} culture 0.8 kg methanol was added but for the μ_{high} culture 1.3 kg was added and by 89 h, the μ_{low} cultivation had 2.2 kg added whereas for the μ_{high} cultivation had 3.8 kg of methanol added by the end of the run (89 h), (Table 4.3). Figure 4.7 illustrates the difference in rate of addition of the methanol for both cultivations. The trend lines show that an exponential addition is achieved in both cultivations.

Table 4.3 *dG-hA_{2A}R activity, residual methanol and cumulative methanol values for μ_{low} and μ_{high} cultivations. Values are for B_{max} estimates which were calculated from single-point radioligand binding analysis using [³H]ZM241385, residual methanol analysis and cumulative methanol measurements for the transition phase of carbon source starvation (IIIA), transition phase of constant methanol feeding of $4 \text{ g L}^{-1} \text{ h}^{-1}$ (IIIB) and the methanol induction phase of either μ_{low} or μ_{high} (IV and V). Measurements are the mean of triplicate determinations, with the standard error of the mean. Cumulative methanol addition values are from a single reading.*

	Phase	Age (h)	Estimated B_{max} (pmol mg^{-1})	Residual methanol (g L^{-1})	Cumulative methanol added (kg)
μ_{low}	IIIA	42.2	4.4 ± 0.1	0 ± 0.0	0
	IIIB	44.4	7.1 ± 0.6	0.5 ± 0.1	0.2
	IV	66.4	4.5 ± 0.2	1.0 ± 0.1	0.8
	IV	89	5.0 ± 0.2	0 ± 0.0	2.2
μ_{high}	IIIA	42.2	4.4 ± 0.2	0 ± 0.0	0
	IIIB	44.4	4.0 ± 0.0	2.1 ± 0.2	0.2
	IV	66.4	3.8 ± 0.1	39.1 ± 9.0	1.3
	V	89	3.1 ± 0.1	84.7 ± 0.5	3.8

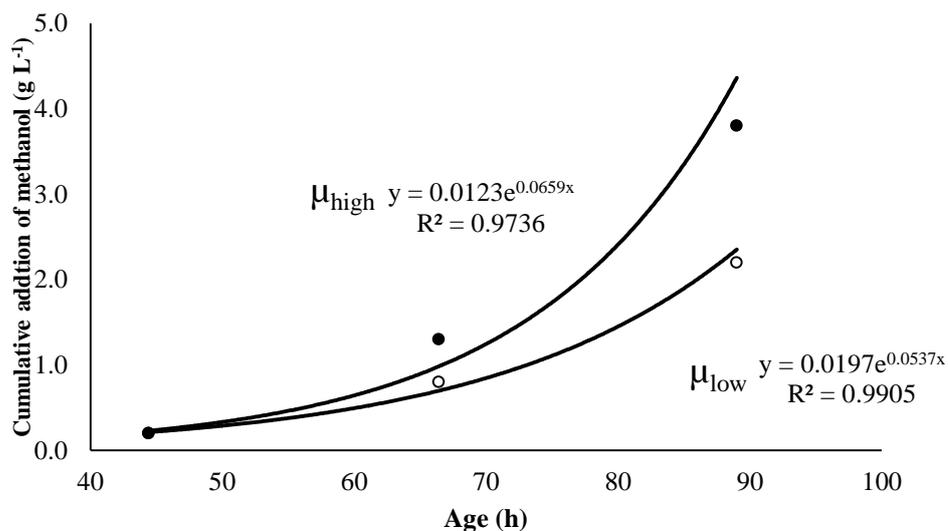


Figure 4.7 Cumulative addition of methanol to μ_{low} and μ_{high} cultivations. Cumulative methanol addition values are from a single reading and fit an exponential curve (solid trend line). For μ_{low} the exponent is 0.0537 (open circles) and for μ_{high} the exponent is 0.0659 (closed circles).

4.4. Yield co-efficients differ between μ_{low} and μ_{high} cultivations during induction

Table 4.4 shows the calculated yield co-efficient for each phase for both the μ_{low} and μ_{high} cultivations. For phases I, IIA, IIB and III the values were similar, as expected. In the induction phase (phase IV), however, the yield co-efficient (0.23) was higher for the μ_{low} cultivation than the corresponding value (0.02) for the μ_{high} cultivation; the yield co-efficient for the μ_{high} cultivation increased (to 0.30) in phase V.

Table 4.4 Calculated yield co-efficients for the μ_{low} and μ_{high} cultivations. Yield co-efficients for each phase of the μ_{low} and μ_{high} cultivations. There was no phase V in the μ_{low} cultivation.

Phase	Age (h)	Yield co-efficients	
		μ_{low}	μ_{high}
I	18.4	0.51	0.59
IIA	28.8	0.17	0.17
IIB	40.6	1.60	1.27
III	42.2	0.75	0.76
IV	86.4	0.23	0.02
V	91.2	—	0.30

4.5. Calculated specific growth rates versus set growth rates for μ_{low} and μ_{high} cultivations

To determine whether the theoretical constant specific growth rates had been achieved, the actual specific growth rates were calculated (Table 4.5). As expected, both cultivations had similar specific growth rates in the batch and fed-batch phases (I, IIA and IIB). In the transition (III) and induction (IV) phases, the specific growth rate of the μ_{low} cultivation was 0.012 h^{-1} , in agreement with $\mu_{set} = 0.01 \text{ h}^{-1}$. However, the μ_{high} cultivation achieved a negative specific growth rate of -0.001 h^{-1} , which was unexpected as $\mu_{set} = 0.03 \text{ h}^{-1}$. During phase V of the cultivation for the μ_{high} cultivation, $\mu_{set} = 0.05 \text{ h}^{-1}$ and the specific growth rate increased to 0.012 h^{-1} .

Table 4.5 Calculated specific growth rates for the μ_{low} and μ_{high} cultivations. Specific growth rates (h^{-1}) were calculated from the DCW data. The R^2 value shows the fit to the rate equation. There was no phase V in the μ_{low} cultivation.

Phase	μ_{low}		μ_{high}	
	Specific Growth Rate (h^{-1})	R^2	Specific Growth Rate (h^{-1})	R^2
I	0.246	0.97	0.256	0.99
IIA	0.122	0.99	0.108	0.96
IIB	0.035	0.99	0.041	0.99
III & IV	0.012	0.93	-0.001	0.47
V	—	—	0.012	0.92

4.6. Pre-induction activity in *P. pastoris* cultivations

Observations of pre-induction activity in the μ_{low} and μ_{high} cultivations prompted further investigation into these unexpected findings (Table 4.6).

4.6.1. dG - hA_{2A}R binding activity is present in the pre-induction glycerol phases of the μ_{low} and μ_{high} cultivations

Even in the absence of methanol, dG - hA_{2A}R receptor expression (i.e. binding activity) was 4.4 pmol mg^{-1} for both μ_{low} (SEM = $\pm 0.1 \text{ pmol mg}^{-1}$) and μ_{high} (SEM = $\pm 0.2 \text{ pmol mg}^{-1}$) cultivations (Table 4.6). For the μ_{low} cultivation, in phase IIB, where a constant methanol feed was initiated and glycerol was still present at 1.1 g L^{-1} , binding activity was present. In phases IV and V of the μ_{low} cultivation, no glycerol was present and binding activity was present. For the μ_{high} cultivation during the transition and induction phases (IIB and IV), glycerol was present in the culture medium and binding activities were 4.0 and 3.8 pmol mg^{-1} , respectively. In phase V, there was no residual glycerol in the culture medium and binding activity was 3.1 pmol mg^{-1} .

4.6.2. dG - hA_{2A}R binding activity is also present in the pre-induction glycerol phases of a 2 L bench-top bioreactor cultivation

Following on from the 35 L pilot-scale work, 2 L bench-top cultivations were carried out for further investigations of pre-induction binding activity. The influence of glycerol on the yield of dG - hA_{2A}R, in a 2 L cultivation was analysed throughout phases I-IV. Figure 4.8 shows the residual glycerol measurements, DCW and B_{max} estimates for the recombinant *P. pastoris* culture expressing dG - hA_{2A}R. Residual glycerol peaked during phase II at ~ 2 g L⁻¹ and was not present during the transition and induction phases (III and IV). Binding activity between 1.1 and 3.1 pmol mg⁻¹ was measured during the batch and fed-batch phases (I and II), suggesting leaky expression. During the transition phase (III), binding activity increased to 4.1 pmol mg⁻¹ and during the induction phase (IV) it was 3.7 – 4.4 pmol mg⁻¹. The measured specific growth rates during this cultivation were 0.08 h⁻¹ during phases I and II and zero during phases III and IV. The specific yield from the 1L cultivation was 122.2 pmol g⁻¹, comparable to that of the μ_{high} cultivation (148.1 pmol g⁻¹) rather than that of the μ_{low} cultivation (536.4 pmol g⁻¹). Notably, the impact of the methanol feed was minimal since the pre-induction specific yield was not substantially increased (Figure 4.8). hA_{2A}R was also produced in all phases of a methanol-free cultivation of glycerol-grown *P. pastoris* (Figure 4.9). During the batch phase (I), glycerol was present at 1.9 g L⁻¹ and the recombinant hA_{2A}R yield was 1.1 pmol mg⁻¹, reaching a final yield of 1.6 pmol mg⁻¹ at the end of the cultivation. The specific yield from the 1L cultivation was 90.3 pmol g⁻¹ and the total yield was 5,598.3 pmol. This was lower than the yield achieved in the corresponding induced culture (Figure 4.8; 122.2 pmol g⁻¹ and 12,986.3 pmol at the end of phase IV) indicating the positive impact of the methanol feed on total yield. The DCW reached a maximum of 62.0 g L⁻¹ in contrast to that of 106.3 g L⁻¹ for the corresponding induced culture (Figure 4.8).

Table 4.6 Residual glycerol in the culture medium does not repress dG - hA_{2A}R expression. B_{max} estimates and residual glycerol values were measured. Samples were tested from the transition (III) and induction (IV and V) phases for both the μ_{low} and μ_{high} cultivations. †Phase V indicates μ_{set} setting change for μ_{high} cultivation only. Values are the means of triplicate determinations, with the standard error of the mean (SEM) given in parentheses. Student's t-test was used to compare corresponding B_{max} estimates from the μ_{low} and μ_{high} cultivations; * = P ≤ 0.01 and ** = P ≤ 0.001.

Cultivation phase	Age (h)	B _{max} estimate (pmol mg ⁻¹)		Residual glycerol (g L ⁻¹)	
		μ _{low}	μ _{high}	μ _{low}	μ _{high}
IIIA	42.2	4.4 (0.1)	4.4 (0.2)	0 (0)	2.9 (0.1)
IIIB	44.4	7.1 (0.6) **	4.0 (0.0)	1.1 (0.2)	4.2 (0.2)
IV	66.4	4.5 (0.2) *	3.8 (0.1)	0 (0)	1.1 (0.2)
†V	88.9	5.0 (0.2) **	3.1 (0.1)	0 (0)	0 (0)

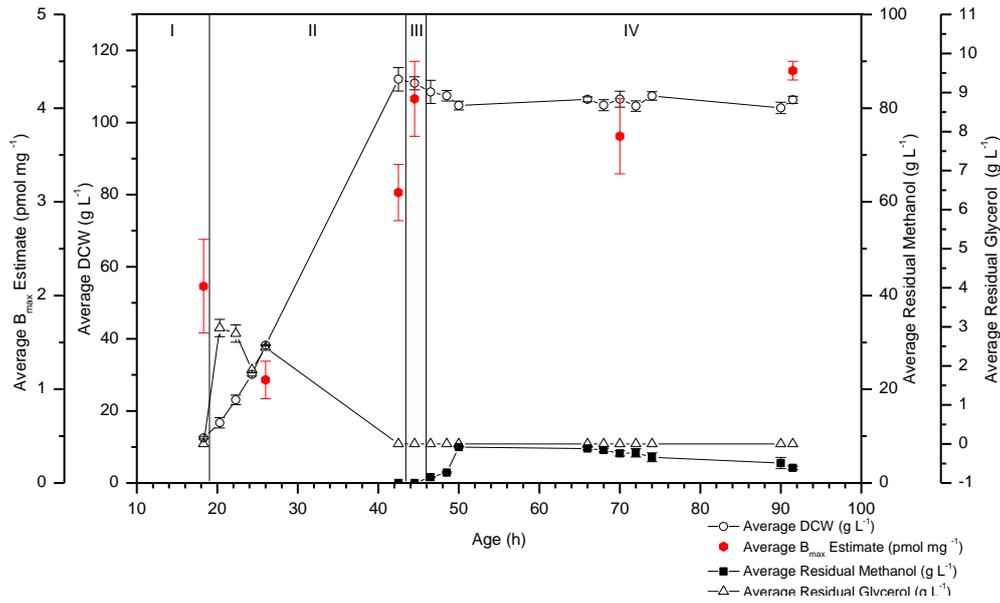


Figure 4.8 *dG - hA_{2A}R binding activity is present in all phases of a subsequent 2 L cultivation.* A 2 L cultivation was analysed for *dG - hA_{2A}R* binding activity (pmol mg⁻¹; red symbols), DCW (g; circles), residual glycerol concentration (g L⁻¹; triangles) and residual methanol concentration (g L⁻¹; squares). Measurements were made in triplicate. *dG - hA_{2A}R* binding activity was measured in all the phases, including pre-induction phases I and II. The residual glycerol concentration was at its highest during phase II, dropping to zero in phase II and for the duration of the cultivation. DCW increased during phases I and II and plateaued during phases III and IV. Residual methanol ranged from 1.25 – 7.66 g L⁻¹ during the induction phase.

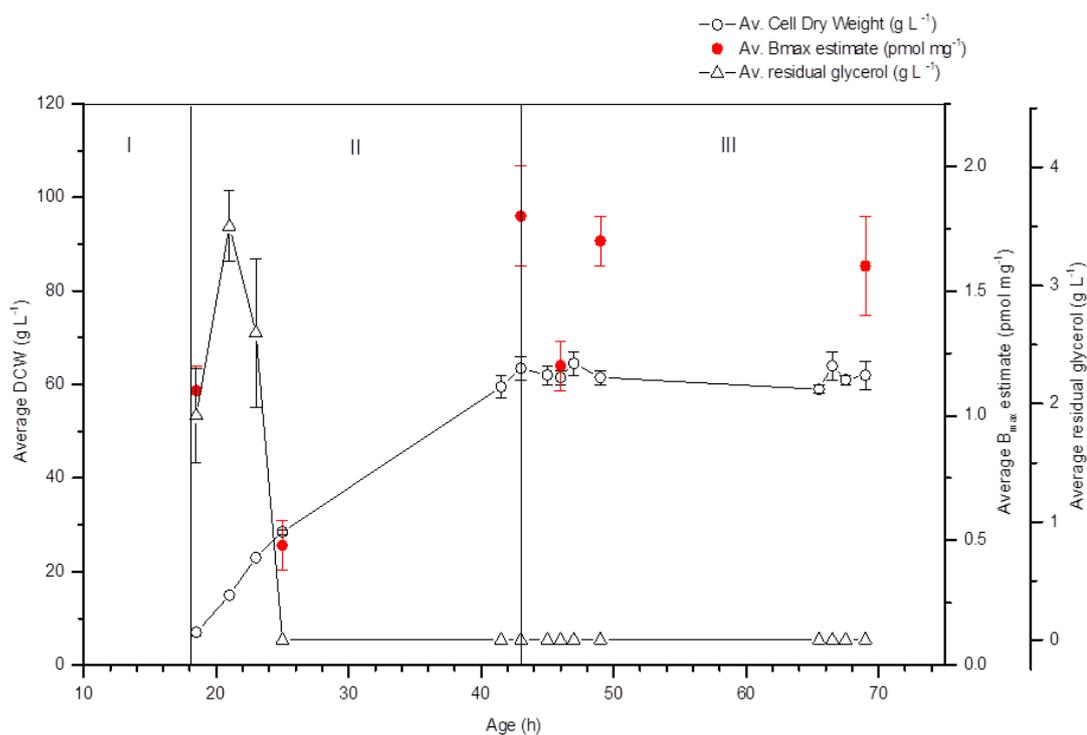


Figure 4.9 dG - hA_{2A}R binding activity is present in all phases in a methanol-free cultivation. A 2 L cultivation was analysed for the same strain and construct cultured previously (Figure 4.8). dG - hA_{2A}R binding activity (pmol mg⁻¹; red symbols), DCW (g; circles) and residual glycerol concentration (g L⁻¹; triangles) are shown. Measurements were made in triplicate. dG - hA_{2A}R binding activity was measured in all the phases, glycerol batch phase (I); glycerol fed-batch phase (II) and carbon source starvation phase (III). The residual glycerol concentration was at its highest during phase I, dropping to zero in phase II and for the duration of the cultivation. DCW increased during phases I and II and plateaued during phases III.

4.6.3. Soluble GFP is produced in all phases of *P. pastoris* bioreactor cultivations

In order to assess whether these observations were specific to dG - hA_{2A}R as the recombinant target protein, the production of recombinant soluble GFP was also examined. Figure 3.9A shows a 2 L bioreactor cultivation with glycerol as the pre-induction carbon source. During the batch phase (I), glycerol was present at 3.0 g L⁻¹ and the recombinant GFP yield was 2.6 mg L⁻¹ prior to methanol addition. This increased to 2.7 – 3.5 mg L⁻¹ in the fed-batch phase (II; the glycerol was consumed 3 h into this fed-batch phase), reaching a final yield of 4.7 mg L⁻¹ in the induction phase (IV). The low impact of the methanol feed was confirmed in a subsequent experiment: when glycerol replaced methanol in the induction phase, the final yield was 3.6 mg L⁻¹.

As glucose has been shown to repress *AOX1* even in the presence of methanol (Potvin et al, 2012), a second cultivation containing glucose as the pre-induction carbon source was analysed

(Figure 4.9B). During the batch phase (I), when the glucose concentration was 6 mM, the yield of recombinant GFP was 2.5 mg L⁻¹ prior to methanol addition. This increased to 2.5 – 3.5 mg L⁻¹ in the fed-batch phase (II; the glucose was consumed 3 h into this fed-batch phase), reached a plateau in the transition phase (III) and then increased during the induction phase (IV) to a maximum value of 13.6 mg L⁻¹. Methanol-free induction (with glucose replacing methanol in the induction phase) in a subsequent experiment gave a final yield of 2.2 mg L⁻¹. This confirmed the higher impact of methanol induction on glucose-grown cells.

Following methanol-induction, DCW values for glycerol-grown cells were approximately 5 times higher than for glucose-grown cells in the equivalent pre-induction phase, while glucose-grown cells produced 3 times the yield of GFP compared with glycerol-grown cells. Since GFP was present in the pre-induction phases of both glycerol- and glucose-grown *P. pastoris* bioreactor cultivations (Fig. 4.10), its production was analyzed in the absence of a methanol induction step (Fig. 4.11). GFP yield was measured in the culture supernatant in all phases of two 1 L *P. pastoris* cultivations, one grown on glycerol (Fig. 4.11A) and one grown on glucose (Fig. 4.11B). In these cultivations, the transition phase (III) was extended from 2 h to 30 h. During the batch phase (I) of the glycerol-grown culture (Fig. 4.11A), glycerol was present at 2.9 g L⁻¹ and the recombinant GFP yield was 2.1 mg L⁻¹ in line with the earlier glycerol-grown cultivations (2.6 mg L⁻¹ prior to methanol addition, Fig. 4.10A). The total GFP yield remained stable in the fed-batch phase (II) and reached a final yield of 8.4 mg L⁻¹ at the end of the cultivation. This was higher than the yield achieved in the corresponding induced cultures (4.7 mg L⁻¹ at the end of phase IV; Fig. 4.10) indicating the negative impact of the methanol feed on total yield. The DCW for the methanol free cultivation reached a maximum of 92.5 g L⁻¹ (Fig. 4.11A) in contrast to that of 175 g L⁻¹ for the corresponding induced culture (Fig. 4.10).

During the batch phase (I) of the glucose-grown culture (Fig. 4.11B), glucose was present at 3.4 mM and the recombinant GFP yield was 1.8 mg L⁻¹, which was lower than that for the earlier glucose-grown cultivation (2.5 mg L⁻¹; Fig. 4.10B). The total GFP yield increased in the fed-batch phase (II), remained stable in the transition phase and reached a final yield of 5.6 mg L⁻¹ at the end of the cultivation. This was lower than the yield achieved in the corresponding induced culture (13.6 mg L⁻¹ at the end of phase IV; Fig. 4.10B) indicating the positive impact of the methanol feed on total yield. The DCW reached a maximum of 46.5 g L⁻¹ in contrast to that of 30 g L⁻¹ for the corresponding induced culture (Fig. 4.10B).

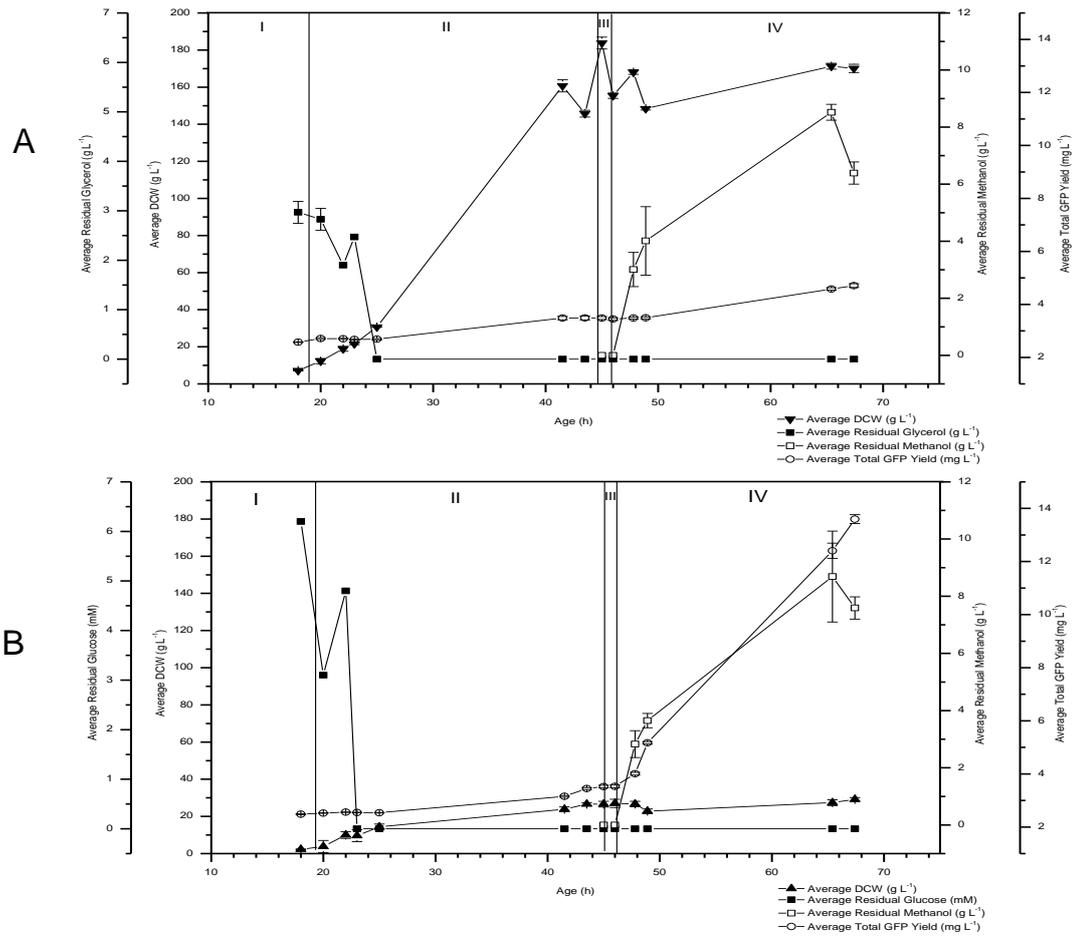


Figure 4.10 Recombinant GFP is produced in the pre-induction phases of two independent 2 L cultivations. GFP yield (mg L⁻¹; circles) was measured in the culture supernatant in all phases of two 2 L *P. pastoris* cultivations, one grown on glycerol (A) and one grown on glucose (B). The residual glycerol concentration (A; squares) was at its highest during phase I, eventually dropping to zero in phase II. The DCW (A; triangles) for this cultivation reached a maximum of 175 g L⁻¹. For the glycerol cultivation, residual methanol (B; open squares) ranged from 3.01 to 8.52 g L⁻¹. For the glucose cultivation, the residual methanol (B; open squares) ranged from 2.83 to 8.69 g L⁻¹. The residual glucose concentration (B; squares) was at its highest during phase I, eventually dropping to zero in phase II. The DCW (B; triangles) for this cultivation reached a maximum of 30 g L⁻¹. All measurements were made in triplicate.

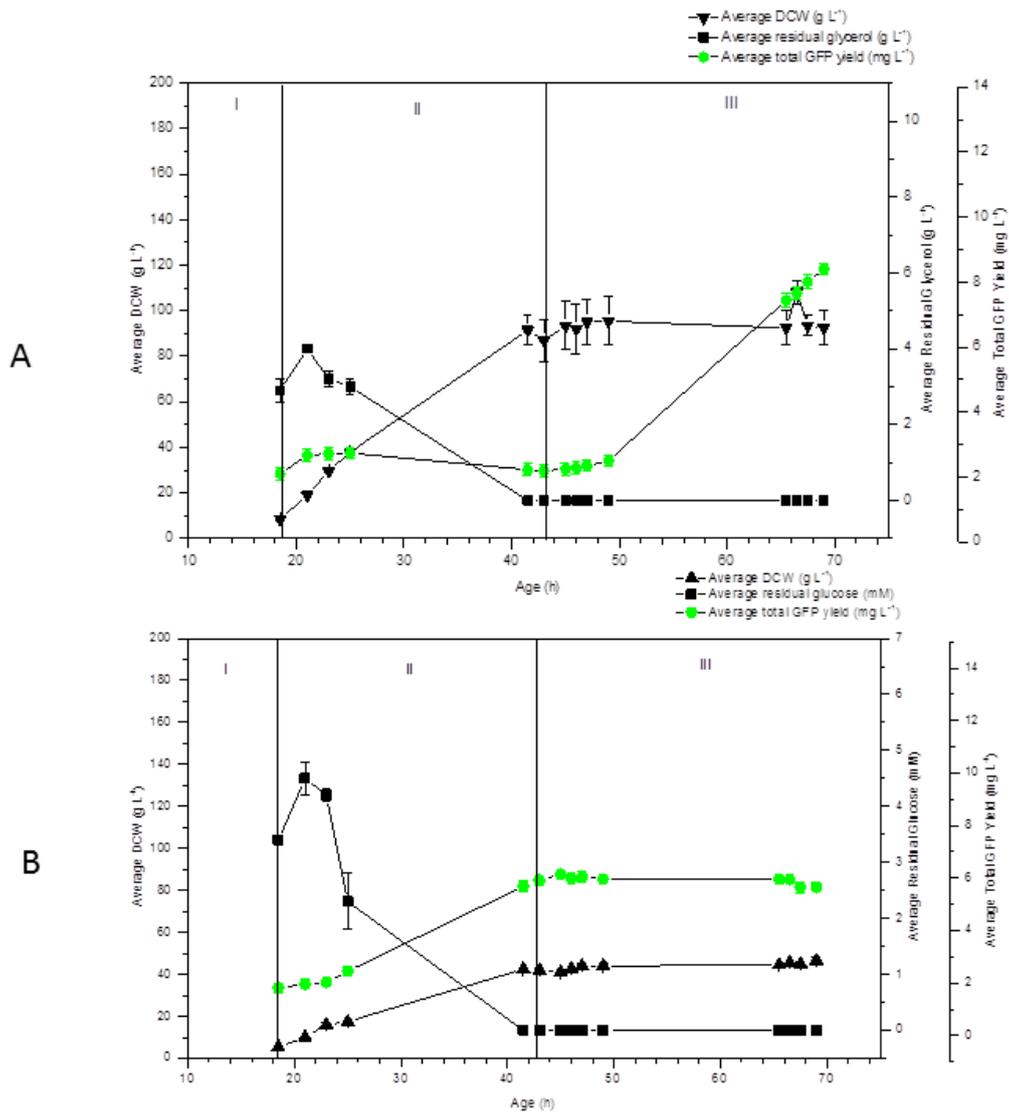


Figure 4.11 Recombinant GFP is produced in methanol-free cultivations of *P. pastoris* grown on either glycerol or glucose as the carbon source. GFP yield (mg L⁻¹; circles) was measured in the culture supernatant in all phases of two 1 L *P. pastoris* cultivations, one grown on glycerol (A) and one grown on glucose (B). The residual glycerol concentration (A; g L⁻¹; squares) was at its highest during phase I, eventually dropping to zero in phase II. The DCW (A; g L⁻¹; triangles) for this cultivation reached a maximum of 92.5108 g L⁻¹. The residual glucose concentration (B; mM; squares) eventually dropped to zero in phase II. The DCW (B; g L⁻¹; triangles) for this cultivation reached a maximum of 46.5 g L⁻¹. All measurements were made in triplicate.

4.7. Discussion

The experiments performed in this chapter addressed the effect of applying two methanol feed profiles during induction on the cell biomass and recombinant hA_{2A}R yields in *P. pastoris*. The study used 35 L pilot-scale bioreactors with software that enabled pre-programming a desired feeding regime; previous studies had shown that applying exponential feeding of methanol to *P. pastoris* cells to control the specific growth rate can be beneficial for recombinant protein production. The cultivations were regulated to reach a theoretical specific growth rate (μ_{set}) by controlling the feeding strategy at the methanol induction phase. The μ_{set} chosen were 0.01 h⁻¹ termed μ_{low} and 0.03 h⁻¹ termed μ_{high} . Two separate *P. pastoris* cultivations were set-up to recombinantly produce a de-glycosylated version of the human adenosine A_{2A} receptor (dG-hA_{2A}R) with either the μ_{low} or the μ_{high} set growth rate applied at the methanol induction phase.

A central finding in this study was that recombinant hA_{2A}R yields in *P. pastoris* could be improved by controlling the methanol feeding regime. The μ_{low} cultivation gave overall hA_{2A}R yields of 536.4 pmol g⁻¹ compared to the μ_{high} cultivation, which yielded 148.1 pmol g⁻¹. This was attributed to the doubling in cell biomass in the μ_{low} cultivation (71.82 g L⁻¹) when compared to the μ_{high} cultivation (37.64 g L⁻¹). In the μ_{low} and μ_{high} cultivations, the methanol feed influenced the biomass yield and particularly increased the total membrane protein yield in the μ_{low} cultivation (7,704 mg for μ_{low} compared to 1,798 mg for μ_{high}). Notably, a similar increase in total membrane protein yield has previously been reported to increase total volumetric yields of GPCRs in a respiratory yeast strain of *S. cerevisiae* (Ferndahl et al., 2010). As a consequence of increased total membrane protein yield, the yield per unit volume of the μ_{low} cultivation was approximately 40 times higher than shake flask cultivations of the same hA_{2A}R construct (Fraser, 2006); total membrane protein yields from shake flasks are 200 mg compared with almost 7704 mg in bioreactors. This yield improvement compares very favourably with that achieved by Singh and colleagues using an optimised hA_{2A}R construct (Singh et al., 2010): on transferring from shake flasks to bioreactors a yield improvement of approximately 25 times was achieved on account of both an increase in specific productivity (25 pmol mg⁻¹ in a shake flask; 100 pmol mg⁻¹ in a bioreactor) and biomass yield (OD₆₀₀ = 13 in a shake flask; OD₆₀₀ = 80 in a bioreactor). Furthermore, this increase in yield in the μ_{low} cultivation suggested that a slowed metabolism may be desirable in maximising yields. This is consistent with previous reports demonstrating that improvements in GPCR yield can be achieved by lowering the culture temperature (André et al., 2006; Singh et al., 2008; Lundstrom et al., 2006; Fraser., 2006); under low temperature conditions, cells have a reduced flux through the TCA cycle, reduced levels of oxidative stress proteins and lower levels of molecular chaperones (Dragosits et al., 2010).

When considering the B_{\max} estimate data for hA_{2A}R, (where overall biomass is not considered) there is a significant but less dramatic increase in pmol of hA_{2A}R per milligram of total protein (for μ_{low} 7.1 ± 0.6 pmol mg⁻¹ and for μ_{high} 4.0 ± 0.0 pmol mg⁻¹ at phase IIIB and 5.0 ± 0.2 pmol mg⁻¹ for μ_{low} and 3.1 ± 0.1 pmol mg⁻¹ for μ_{high} at the end of phase V). From these findings, it could be hypothesised that growing the cells at a higher density may have had a negative effect on the hA_{2A}R being produced at the cellular level. Jahic and colleagues (Jahic et al., 2003) showed that high cell density *P. pastoris* cultures exhibit increased proteolysis and that cultures between OD₆₀₀ 100-500 are at risk of this. There is further evidence in the literature that high yeast culture cell densities tend to be under osmotic stress (Mattanovich et al., 2004). Osmotic stress occurs when the salt concentrations in the media formulations rise and fall during consumption by the yeast cells. In particular relation to membrane proteins, it is thought that the yeast cell surface assembly is sensitive to osmotic stress (Gasch et al., 2000) and that major structural changes occur such as adjustment of the cell wall organisation during hyperosmotic conditions (Mager and Siderius, 2002). For different protein targets, different residual methanol concentrations have been found for optimal productivity, ranging from 0.4-30 g L⁻¹ (Hellwig et al., 2001, Zhang et al., 2000). Overall, this suggests that applying an exponential methanol feeding regime at a low set growth rate yields high biomass for *P. pastoris* and leads to higher membrane protein yields than are achieved in a high set growth rate.

The slight increase in dG- hA_{2A}R activity for the μ_{low} cultivation was probably due to the lack of residual methanol that was present in the cultivation and that it never exceeded the reported cytotoxic level of 5 g L⁻¹. However in stark contrast, the μ_{high} cultivation contained residual methanol levels of 84.70 g L⁻¹ which gave rise to a cytotoxic environment. The effect of this was reflected in the cell biomass yields as they did not increase once this (μ_{high}) methanol feeding regime was applied. Interestingly however, dG-hA_{2A}R activities were only slightly lower than the activities achieved from the μ_{low} cultivation and therefore it was quite unexpected to achieve any activity in such a cytotoxic methanol environment.

Another notable finding from this study was the presence of dG-hA_{2A}R activity prior to methanol induction. For both the μ_{low} and μ_{high} cultivations, binding activity was present at 4.4 pmol mg⁻¹ prior to any induction. Furthermore, it was found that residual glycerol was still present in some phases indicating that glycerol did not repress the *AOX1* gene. This finding warranted further investigation and therefore several 2 L bench-top bioreactor cultivations were performed on *P. pastoris* and the pre-induction phases were studied in more detail.

In one bioreactor experiment, dG- hA_{2A}R was recombinantly produced in *P. pastoris* and it was found that the activity was present regardless of the presence of glycerol.

This was assessed further by carrying out bench-top bioreactor cultivations for another recombinant protein, in order to rule out any protein specific phenomena. Soluble GFP was recombinantly produced in *P. pastoris* cultivations with either glycerol or glucose as the batch and fed-batch carbon source. This was done to test if the carbon source was also a factor. It was found that there was indeed a pre-induction production of GFP with either glycerol or glucose present in the cultivation. This could suggest that the leaky expression phenomena are present in all of the tested systems in this chapter but would require further testing. This is reported in the literature for some *P. pastoris AOX1* leaky promoter behaviour (Lombardi et al., 2010) and also in *E. coli* systems (Ham et al., 2006, Guzman et al., 1995).

These findings show that against the general consensus, pre-induction production of hA_{2A}R and GFP are present. The impact of this finding is that this adds insight into the importance of controlling and optimising culture conditions during all the phases of *P. pastoris* cultivations and not just the induction phase. There are however unanswered questions such as: is this true for other protein targets? Future work would therefore be to select several more recombinant protein targets and perform the same bioreactor and comparative shake flask cultivations to see if this is true. This information could then be useful when planning *P. pastoris* recombinant protein cultivations.

This chapter forms the basis of an original research manuscript currently in submission and also a published book chapter titled 'Optimising *P. pastoris* induction' by Zharain Bawa and Richard Darby (Bawa, 2012) ;(Appendix A9).

Chapter 5: The use of novel polymers to extract recombinant hA_{2A}R from *P. pastoris* membranes

Extracting membrane proteins from their native membranes to enable structural studies is challenging due to their heavy dependence on the membrane lipid bilayer for their integrity. The first step of extracting the membrane protein from the membrane is termed solubilisation, followed by purification of the protein to enable further biophysical studies such as circular dichroism (CD), analytical ultra-centrifugation (AUC) and dynamic light scattering (DLS). Eventually, x-ray crystallography techniques are used to achieve crystal structures. The solubilisation process typically uses compounds such as surfactants or detergents as discussed in Chapter 1. These may cause denaturation of the membrane protein or formation of aggregates if not optimised. In this chapter solubilisation of hA_{2A}R from *P. pastoris* membranes using poly (maleic anhydride-styrene) polymers (PMAS) is described with the goal of making solubilisation a quick and cheap alternative to employing traditional detergents. The hA_{2A}R construct described in Chapter 5 was designed in collaboration with the industrial sponsor, AstraZeneca Ltd, in order to use their standard tagging protocol. This was done to render the hA_{2A}R construct amenable to different downstream applications within AstraZeneca that would not be possible for the hA_{2A}R construct used in Chapter 4.

5.1. Design of a novel hA_{2A}R construct

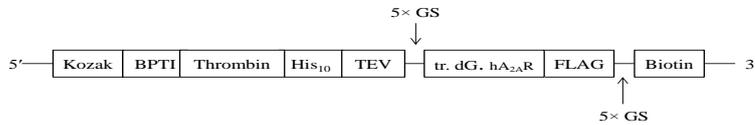
An improved construct design was investigated for hA_{2A}R. The goal was to use a truncated version of the hA_{2A}R sequence. Fraser (Fraser, 2006) used a full length version of the sequence even though it was mutated for de-glycosylation. The new construct illustrated in Figure 5.1 kept the de-glycosylation but also introduced a truncated version of the protein (A316) as in Singh et al., 2010.

5.1.1. Construct design

The new hA_{2A}R construct was designed with multiple tags which included 10× histidine, FLAG and biotin tags. It was also designed with a bovine pancreatic trypsin inhibitor (BPTI) sequence (Parekh et al., 1995) which replaces the commonly used alpha factor leader peptide sequence in yeast vector-construct designs (Waters et al., 1988, Tashiro et al., 1993). A thrombin cleavage site was designed into the construct where thrombin proteases are highly specific for cleavage between Arg and Gly residues (Vu et al., 1991). Another cleavage site included was the TEV cleavage site. Two Gly-Ser linkers were designed either side the hA_{2A}R ORF. Finally a Kozak sequence was included at the beginning of the construct (Kozak, 1999). A synthesised *Xba*I was included to allow flexibility of the construct to be used in other projects. Figure 5.1 shows the design, DNA and amino acid sequence for the new construct with the new name: MT-hA_{2A}R.

The construct was synthesised and cloned into the pPICZB vector (Life Technologies) by GeneArt Life Technologies Corporation (Figure 5.2).

A)



B) Coding sequence for new construct:

```
GAAACGATGAGAAATGAAGGTTTTGATCGTTTTTGTGGCTATCTTCGCTGCTTTGCCATTG
GCTTTGGCTTTGGTTCCAAGAGGTTCATCACCATCACCATCACCACCACAGAG
AACTTGACTTCCAAGGTTCTGGTTCCGGTTCGGATCTGGTTCCATGCCAATTATGGGT
TCCTCCGTTTACATCACTGTTGAGTTGGCTATCGCTGTTTTGGCTATCTGGGTAACGTT
TTGGTTGTTGGGCTGTTGGTTGAACTCCAACCTGCAGAACGTTACAACTACTTCGTT
GTTTCCTGGCTGCTGCTGACATTGCTGTTGGAGTTTTGGCTATCCATTCCGTATCACT
ATCTCCACTGGTTTCTGTGCTGCTTGTACCGGTTGTTGTTTCATTGCTTGTTCGTTTTG
GTTTTGACTCAGTCCATCATCTCCTTCCTGTTGGCTATGCTATCGACAGATATATCGCT
ATCAGAAATCCCATTGAGATACAACGGTTTGGTTACTGGTACTAGAGCTAAGGGTATTATC
GCTATCTGTTGGGTTTTGTCCCTCGCTATCGGTTTGGCTCCATGTTGGGTTGGAACAAC
TGTGGTCAGCCAAAAGAGGGTAAGCAACACTCTCAAGGTTGTTGGTGAGGGTCAGGTTGCT
TGTTTTGTCGAGGACGTTGTTCCAATGAACACTACATGGTTACTTCAACTTTTTTCGCTTGT
GTTTTGGTTCCTTTGTTGTTGATGTTGGGTTTACTTGAGAATCTTCTGGCTGCTAGA
AGACAGTTGAAGCAGATGGAATCCCAGCCATTGCCAGGTGAAGAGCTAGATCCACATTG
CAGAAAGAGGTTACCGTGTAAAGTCCCTGGCTATCATCGTTGGTTTTGTTCCGCTTTGTTG
TGGTTGCCATTGCACATCACTCAACTGTTTTACTTTCTTCTGTCTGACTGTTCCACGCT
CCATTGTTGGTTGATGACTTGGCTATCGTTTTGTCCACACTAACTCCGTTGTTAACCCA
TTCATCTACGCTTACAGAATCAGAGAGTTCAGACAGACTTTCAGAAAGATCATCAGATCC
CACGTTTTGAGACAGCAAGAGCCATTCAAGGCTGACTACAAGGATGATGACGACAAGGGT
TCCGGATCAGGTTCTGGATCAGGATCAGGTC AATTCCGGTGGTGGTACTGGTGGTGCTCCA
GCTCCAGCTGCTGGTGGTGGTGGTAAAGCTGGTGAAGGTGAAATCCAGCTCCA
TTGGCTGGTACTGTTTTCCAAGATCTTTGGTTAAGGAAGGTGACACTGTTAAGGCTGGTCCAG
ACTGTTTTGGTTTTGGAGGCTATGAAGATGGAACACTGAGATCAACGCTCCAACCTGACCGT
AAGGTTGAGAAGGTTTTGGTCAAAGAAGAGATGCTGTTTCAGGGAGGTCAGGTTTTGATT
AAGATCGGTTCTAGA
```

C) Translated amino acid sequence with same colour code key as in B):

```
ETMRMKVLIVLLAIFAALPLALALVPRGSHHHHHHHHHHENLYFQGS GSGSGSGS MPIMGSSVYITVELAIA
VLAILGNVLCWAVWLNLSNLQNVNTNYFVVS LAAADIAVGVLAIPFAITISTGFCAACHGCLFIACFVLVLTQ
SSIFSLAIAIDRYAIRIPLRYNGLVTGTRAKGIIAICWVLSFAIGLTPMLGWNNCGQPKEGKQHSQCGEGQ
VACL FEDV VPMNYMVYFNFFACVLP LLLMLGVYLRIFLAARRQLKQMESQPLPGERARSTLQKEVHAAK
SLAII VGLFALCWLP LHIINCFTFFCPDCSHAPLWMLYLAIVLSHTNSV VNPFIYAYRIREFRQTFRKIIRSHVLR
QQEPFKA DYKDDDDK GSGSGSGSGSQFGGGTGGAPAPAAGGAGAGKAGEGEIPAPLAGTVSKILVKEGD
TVKAGQTVLVLEAMKMETEINAPTDGKVEKVLVKERDAVQGGQGLIKIGSR
```

Figure 5.1 MT-hA_{2A}R construct design and sequence. A) Schematic diagram of a multi-tag, truncated (tr.), de-glycosylated (dG) hA_{2A}R construct, named MT-hA_{2A}R. B) Coding DNA sequence of MT-hA_{2A}R that was cloned into pPICZB vector (Life Technologies) using EcoRI/NotI cloning sites present in multiple cloning region of vector backbone. C) Amino acid sequence translated from DNA coding region.

Colour code key:

Kozak sequence BPTI signal (bovine pancreatic trypsin inhibitor) Thrombin cleavage site 10x histidine tag TEV cleavage site Gly-Ser linker (5x) FLAG tag (with internal enterokinase cleavage site) Truncated (A316) (Singh et al., 2010) and de-glycosylated (N154Q) hA_{2A}R (Fraser, 2006) Biotin tag Synthesised XbaI site

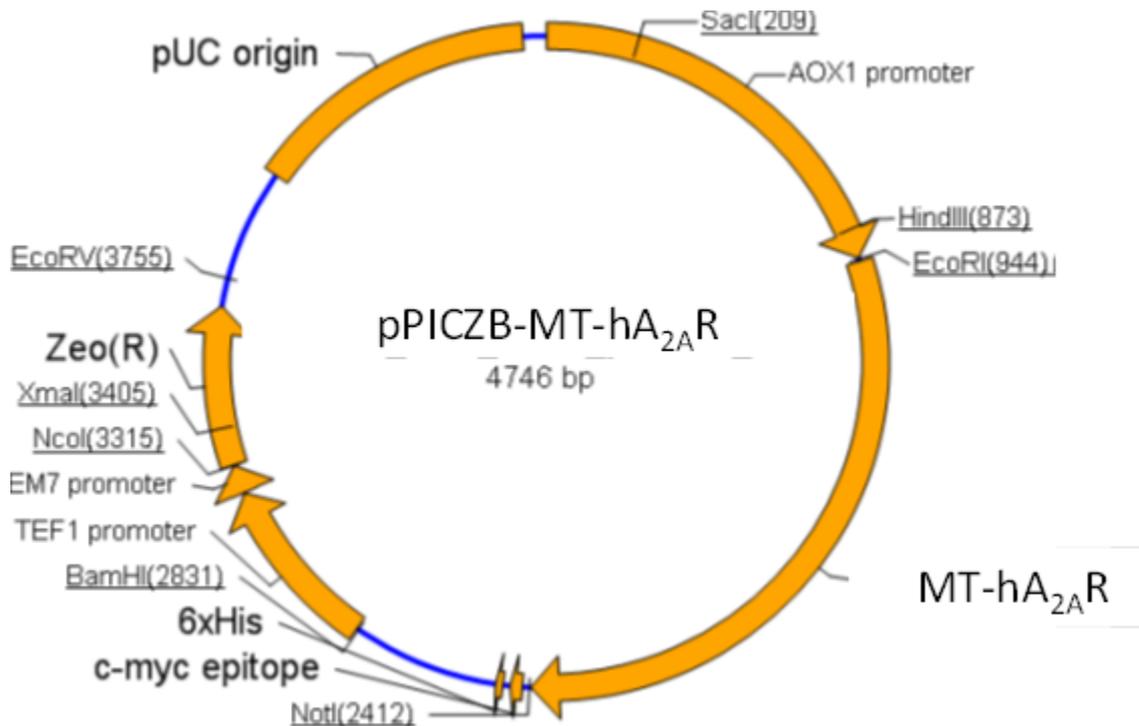


Figure 5.2 pPICZB-MT-hA_{2A}R vector map. Vector map of the MT-hA_{2A}R cloned into the pPICZB vector multiple cloning sites using EcoRI/NotI. Notable features of the vector include the AOX1 promoter region used for inducing recombinant protein production with methanol and the Zeocin resistance gene (Zeo (R)), used for colony selection after transformations into the desired strain are performed. Appendix A2 and A3 contains the alignment sequencing performed at Life Technologies for the pPICZB-MT-hA_{2A}R vector.

5.1.2. Validation of construct design

The MT-hA_{2A}R construct was tested further and sub-cloned out of the pPICZB vector and cloned into a mammalian vector (pCDNA3.1) and transfected into human embryonic kidney (HEK) cells. This work was performed and analysed by Dr. Sarah Routledge, Aston University, United Kingdom. The purpose of this investigation was to observe if any of the tags in the MT-hA_{2A}R construct interfered with the specific binding activity by comparing it to HEK cells transfected with His tagged hA_{2A}R and an untagged construct of hA_{2A}R. Figure 5.3 shows the competition binding curves performed where cold ZM241385 competed with [³H] ZM241385. The pK_d values were derived from the K_i values where the untagged hA_{2A}R construct gave the highest pK_d of 8.2. Both the His-tagged and MT-hA_{2A}R construct gave a pK_d of 8.9 suggesting good affinity and that the tags did not interfere with hA_{2A}R binding.

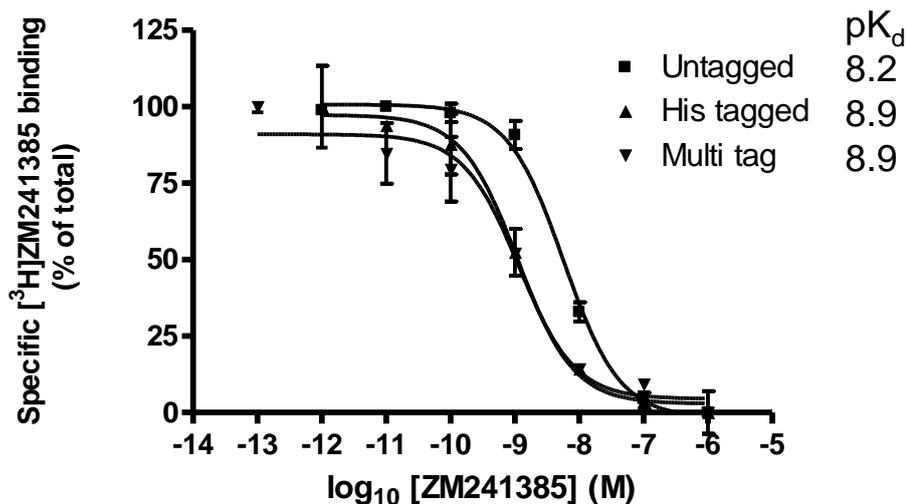


Figure 5.3 Competition binding curves for untagged, His- tagged and MT-hA_{2A}R from HEK cell membranes. Competition binding curves were performed where cold ZM241385 competed with [³H] ZM241385. The curve was fitted to a one site binding model and data points were triplicate over 3 separate experiments with the standard error of the mean (SEM) shown. pK_d values were derived from the EC₅₀ values from the curves.

The novel hA_{2A}R construct (MT- hA_{2A}R) was therefore transformed into a protease-deficient strain (SMD1163) and was verified for growth and expression. This allowed it to be used for studies of novel hA_{2A}R solubilisations. Subsequent experiments used the SMD1163- MT- hA_{2A}R grown in FM22 minimal medium.

5.1.3. Generation of a high yielding *P. pastoris* strain for production of recombinant MT-hA_{2A}R

The pPICZB-MT-hA_{2A}R vector was transformed into *E. coli* XL 10-Gold® ultra-competent cells (Stratagene) and the vector purified using a mini-prep method. Figure 5.4 shows the agarose gel of the purified vector from the propagated XL 10-Gold® ultra-competent *E. coli* cells and the vector after linearisation with *PmeI*. The vector was linearised so that it could integrate with the *P. pastoris* genome via homologous recombination.

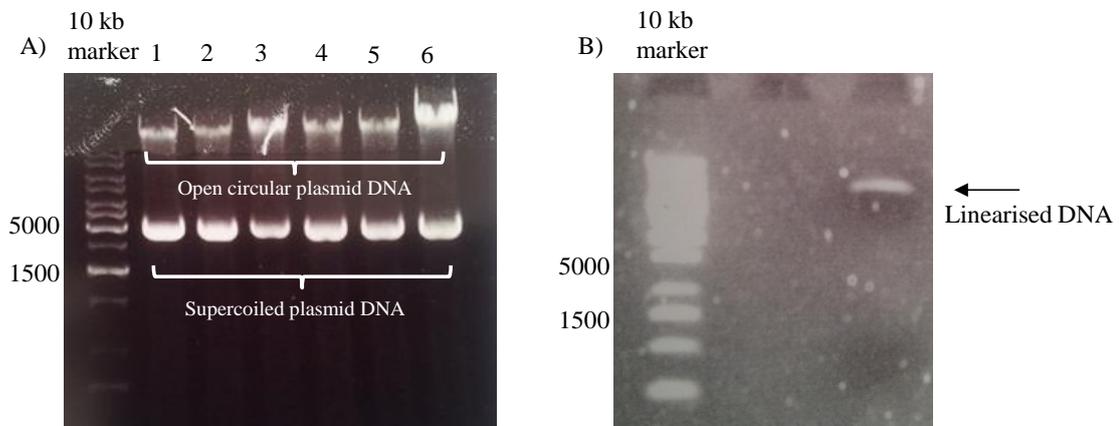


Figure 5.4 Agarose gel (1%) showing purified pPICZB-MT-hA_{2A}R before and after linearization. A) pPICZB-MT-hA_{2A}R purified after mini-preparation technique, lanes 1-6. Supercoiled plasmid DNA is indicated at ~ 5000 kb. Open circular plasmid DNA is also indicated in the upper part of the gel. B) pPICZB-MT-hA_{2A}R after linearization with *PmeI* with a size estimated between 10-20 kb.

Once the pPICZB-MT-hA_{2A}R DNA was linearized and the sequence verified, the chosen *P. pastoris* strain for transformation was prepared. The SMD1163 (*his4 pep4 prb1*) protease deficient strain (White et al., 1995 and Brierley, 1998) was selected and was the kind gift of Dr. Shweta Singh, Imperial College, London, United Kingdom. After electroporation of the pPICZB-MT-hA_{2A}R DNA into the SMD1163 *P. pastoris* cells, they were grown on zeocin selection plates of increasing concentrations. Figure 5.5 illustrates the colony selection process for expression screening.

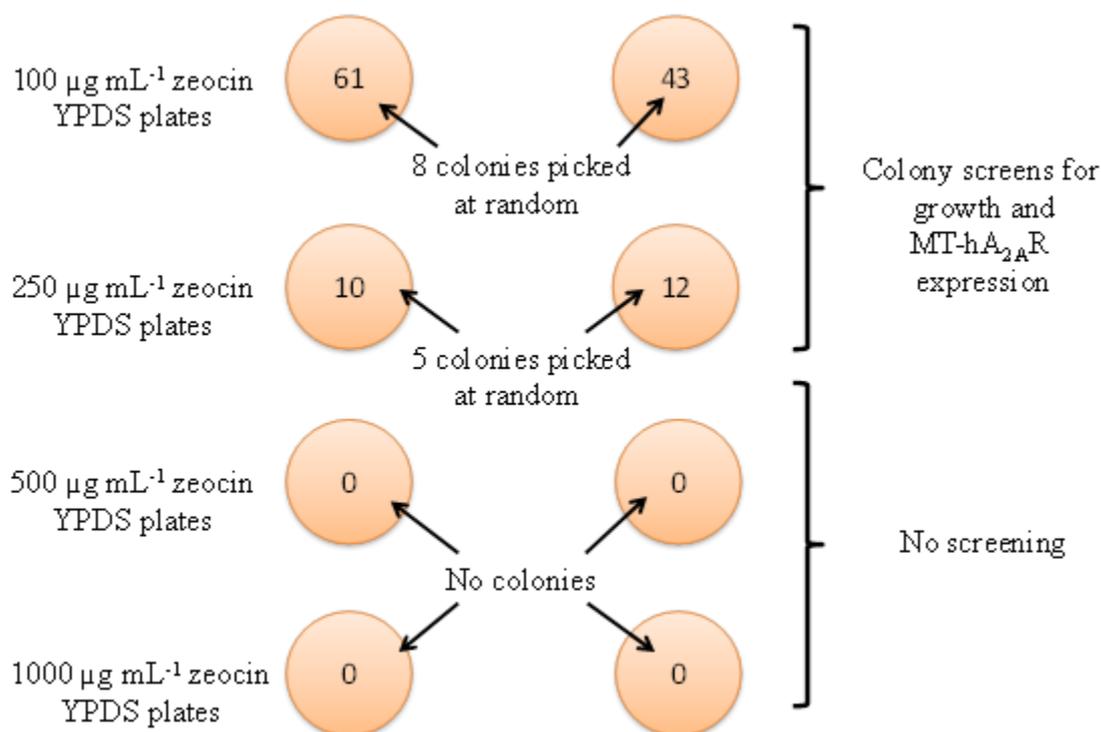


Figure 5.5 Scheme for screening MT-hA_{2A}R expression. Diagram shows YPDS culture plates with 4 concentrations of zeocin. Colonies grew on 100 $\mu\text{g mL}^{-1}$ and 250 $\mu\text{g mL}^{-1}$ plates and were taken forward for expression screens. No colonies were observed on the 500 $\mu\text{g mL}^{-1}$ and 1000 $\mu\text{g mL}^{-1}$ plates.

Before the expression screen began, a few colonies from the 100 $\mu\text{g mL}^{-1}$ and 250 $\mu\text{g mL}^{-1}$ plates were grown in BMGY complex medium in shake flasks to determine their growth rates and hence to determine at which time point the induction of recombinant MT-hA_{2A}R should approximately begin (around late log phase) by changing the medium to BMMY during the expression screen experiments. The mean growth rate (μ) of the colonies was 0.36 h^{-1} and their doubling time for growth was about 67 min. These calculations were done at the logarithmic phase of growth. The late log phase occurred at 5-10 h and therefore induction started at this time (Figure 5.6).

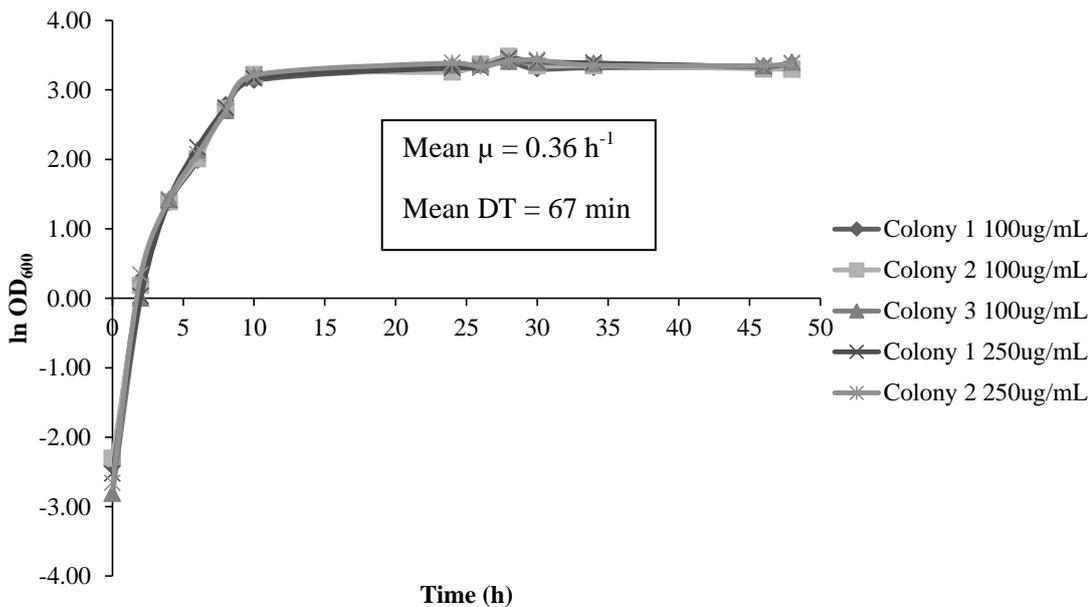


Figure 5.6 Growth curves for several colonies of SMD1163- MT-hA_{2A}R in BMGY cultured in shake flasks. 5 colonies (3 from 100 μg mL⁻¹ plates and 2 from 250 μg mL⁻¹ plates) were cultured in BMGY in shake flasks and grown at 30°C for 48 h. Optical density measurements were taken at 600 nm. The specific growth rates (μ) were calculated from the exponential part of the curve for each colony and averaged. The doubling times (DT) were calculated from μ . The optimal time for starting induction is between 5-10 h.

Immuno-blot analysis showed expression of the MT-hA_{2A}R from a selection of colonies from both concentrations of zeocin (colony A3 and A7 from the 100 µg mL⁻¹ plates and colony B1 and B4 from the 250 µg mL⁻¹ plates). The main band was approximately 38-40 kDa (Figure 5.7). There were other larger bands seen in the immuno-blot and this may be attributed to uncleaved thrombin signal peptides and the smaller bands may be attributed to degraded products of the receptor..

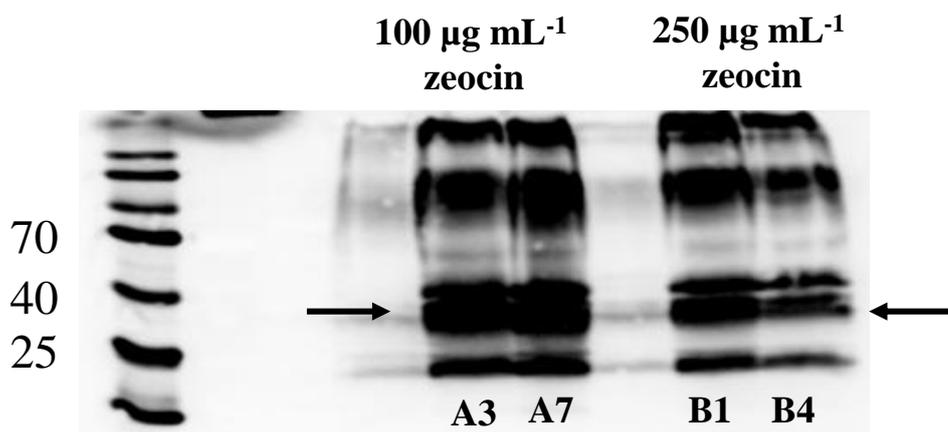


Figure 5.7 Immuno-blot from SMD1163-MT-hA_{2A}R colony screen. Immuno-blot of MT-hA_{2A}R from colony screening in BMMY shake flasks. Lanes labelled A3 and A7 correspond to colonies A3 and A7 from the 100 µg mL⁻¹ zeocin YPDS plates. Membrane preparations were made from cells that were harvested at the end of methanol (0.5%) induction after 24 h from the BMMY shake flask experiments. Lanes labelled B1 and B4 correspond to colonies B1 and B4 from the 250 µg mL⁻¹ zeocin YPDS plates. Membrane preparations were also from cells that were harvested at the end of methanol (0.5%) induction after 24 h from the BMMY shake flask experiments. The primary antibody used in the method was anti-His antibody (Serotec).

5.1.4. Analysis of functional recombinant MT-hA_{2A}R in *P. pastoris*

Radio-ligand binding analysis was carried out on all the colonies selected (Figure 5.8). Single-point binding analysis was performed on the membrane preparations of the colonies. It can be seen that colony A7 gave the highest specific binding (~ 7.0 pmol mg⁻¹) to the hA_{2A}R antagonist [³H] ZM241385 hence this colony was selected for subsequent experiments. SMD1163 *P. pastoris* transformed with the pPICZB vector only were also grown on the zeocin selection plates. Colonies grew on the 100 µg mL⁻¹, 250 µg mL⁻¹ and 500 µg mL⁻¹ zeocin plates and were selected for single-point radio-ligand binding. There was no activity detected in these cells with the vector only control.

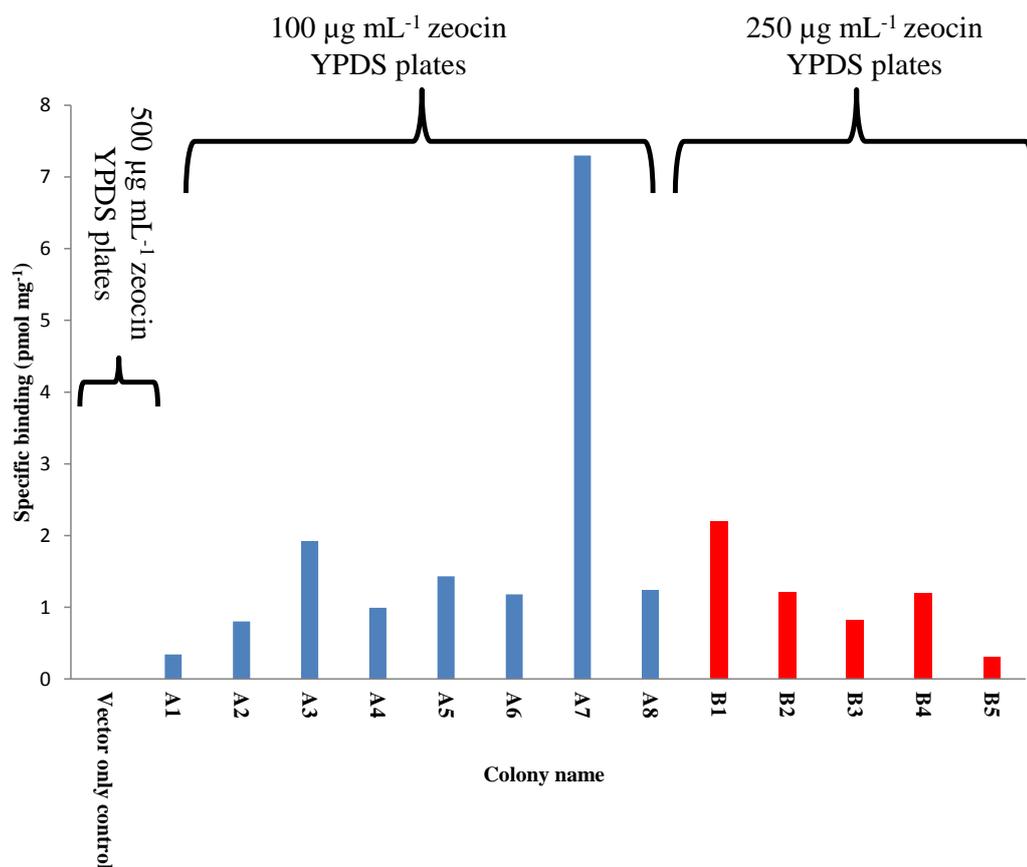


Figure 5.8 Specific binding activities from SMD1163 MT-hA_{2A}R colony expression screen. Single-point radio-ligand binding analysis using [³H] ZM241385 on membrane preparations of 8 selected colonies from 100 µg mL⁻¹ zeocin YPDS plates (after growth in BMGY shake flask and induction with BMMY) and 5 selected colonies from 250 µg mL⁻¹ zeocin YPDS plates (after growth in BMGY shake flask and induction with BMMY). Colony A7 gave the highest specific binding activity and was chosen for all subsequent experiments. A vector only control was also grown on the 100, 250 and 500 µg mL⁻¹ zeocin YPDS plates where the pPICZB vector was transformed into SMD1163 P. pastoris cells with no insertion of the construct. The no binding activity was present for all colonies and the figure shows this for the 500 µg mL⁻¹ zeocin only.

As the purpose of the work described in this chapter was to study novel solubilising compounds, it was necessary to generate large amounts MT-hA_{2A}R membranes to enable this research. The most efficient and controlled method was to run a 2 L bioreactor of SMD1163 MT- hA_{2A}R cells. This also tested if the cells could be scaled-up from 25 mL culture volume to 2 L culture volume and also to see if the cells could grow on minimal medium (FM22 media) and not just complex medium (BMGY and BMMY). The results were compared to a 2 L bioreactor run for the X33-dG -hA_{2A}R in FM22 media. Membrane preparations were carried out at the final induction time-point (90 h) for both constructs. For both construct types and strains, the total wet cell biomass was recorded (Figure 5.9, inset table). The wet cell biomass was similar in all situations, where the X33-dG -hA_{2A}R wet cell mass was 125 g and SMD1163 MT- hA_{2A}R was 120 g at the end of the runs. The specific growth rate was calculated for the exponential part of the growth curves for both constructs and they were similar in value, 0.086 for X33-dG -hA_{2A}R

and 0.082 for SMD1163 MT- hA_{2A}R. Radio-ligand binding analysis with [³H] ZM241385, a saturation binding curve, was carried out for both constructs and the B_{max} and pK_d values calculated (Figure 5.9). Both constructs gave similar pK_d values of 8.3 (X33-dG -hA_{2A}R) and 8.5 for (SMD1163 MT- hA_{2A}R). However, the B_{max} was slightly increased for SMD1163 MT- hA_{2A}R (6.4 pmol mg⁻¹) when compared to X33-dG -hA_{2A}R (4.4 pmol mg⁻¹). This suggested that the SMD1163 MT- hA_{2A}R strain yielded more expression of the receptor than the X33-dG -hA_{2A}R strain-construct combination.

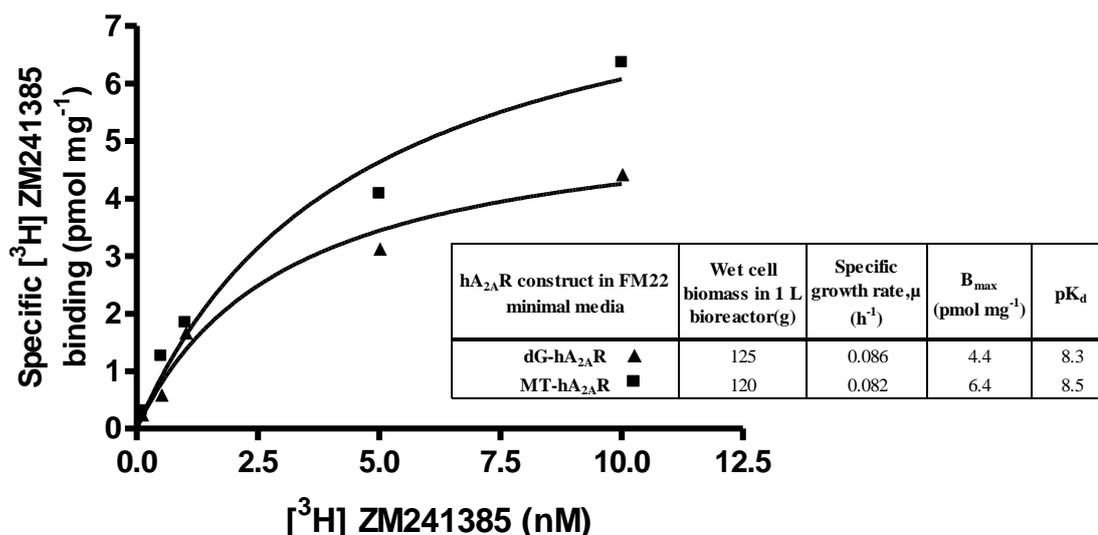


Figure 5.9 Saturation binding curve and summary of wet cell weight, specific growth rates (μ), B_{max} and pK_d data for X33-dG - hA_{2A}R and SMD1163-MThA_{2A}R in FM22 minimal media. Saturation binding of [³H] ZM241385 to membrane bound dG - hA_{2A}R grown in FM22 media (closed triangles), MT-hA_{2A}R grown in FM22 media (closed squares) at a single time-point during the induction phase (91h). Data were all from triplicate sampling. Inset table shows wet cell biomass weighed at end of bioreactor runs (91 h). The specific growth rates (μ) were derived from the exponential section of the growth curve for both constructs. The B_{max} and pK_d values were calculated from the saturation binding curve using non-linear regression. Radio-ligand binding assays were carried out on membrane preparations of cells harvested at the final time-point of 91 h during the induction phase for both constructs.

5.2. Solubilisation of MT-hA_{2A}R from SMD1163 *P. pastoris* membranes with the detergent, n-dodecyl-β-d-maltopyranoside (DDM)

Typically, n-dodecyl-β-d-maltopyranoside (DDM) is used for solubilising hA_{2A}R from native membranes (yeast or mammalian); (le Maire et al., 2000). Figure 5.10 shows the saturation binding profile for DDM solubilised MT- hA_{2A}R from SMD1163 *P. pastoris* cells. The B_{max} value is 5.6 ± 2.3 pmol mg⁻¹ and the pK_d is 8.3 ± 0.1 which was comparable to the membrane bound MT- hA_{2A}R and gave values of B_{max} = 6.4 ± 0.5 pmol mg⁻¹ and pK_d = 8.5 ± 0.1. This confirmed that the SMD1163 MT- hA_{2A}R strain was able to be solubilised via detergents and hence the PMAS solubilisations were initiated.

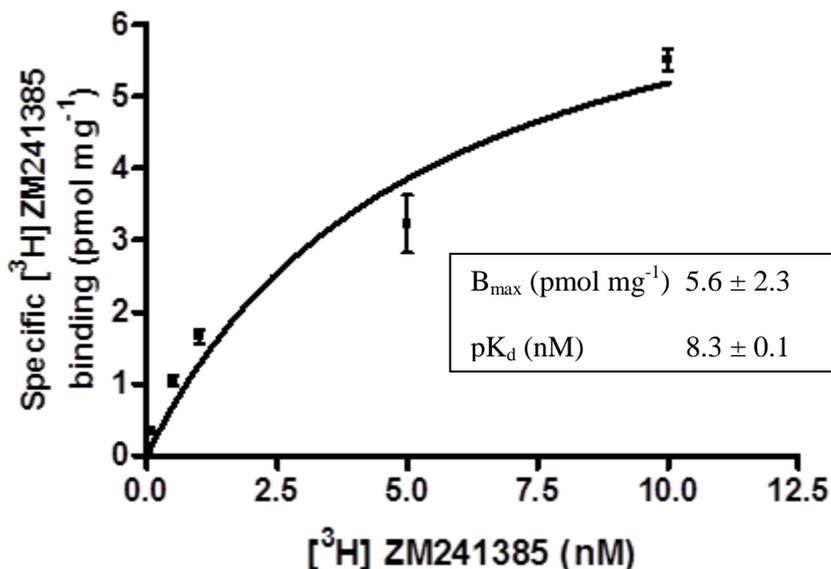


Figure 5.10 Saturation binding curve for DDM solubilised MThA_{2A}R from SMD1163 *P. pastoris* cells. Saturation binding of [³H] ZM241385 to DDM solubilised MThA_{2A}R. Data are the mean of triplicate experiments ± SEM. The B_{max} and pK_d values were derived from non-linear regression of the curve.

5.3. Solubilisation of MT-hA_{2A}R from SMD1163 *P. pastoris* membranes with responsive hydrophobically associating polymers

In an alternative approach, MT-hA_{2A}R was solubilised from SMD1163 *P. pastoris* membranes with responsive hydrophobically associating polymers, in particular, poly (maleic anhydride-styrene) or PMAS. Figure 5.11 shows the generic chemical structure of poly (maleic anhydride-styrene) (Figure 5.11A) and poly (maleic acid-styrene) (Figure 5.11B) after hydrolysis (addition of NaOH) which opens the anhydride rings to form maleic acid molecules in order to make the polymer soluble. The styrene molecules form the hydrophobic regions of the polymer and maleic acid is the hydrophilic part. The number of styrene to maleic anhydride/acid can vary. As there are many forms of poly (maleic anhydride-styrene), a sub-set of PMAS was chosen with different molecular weights in order to test whether these variations improved the solubilisations. For the same reason, one PMAS was esterified with methanol.

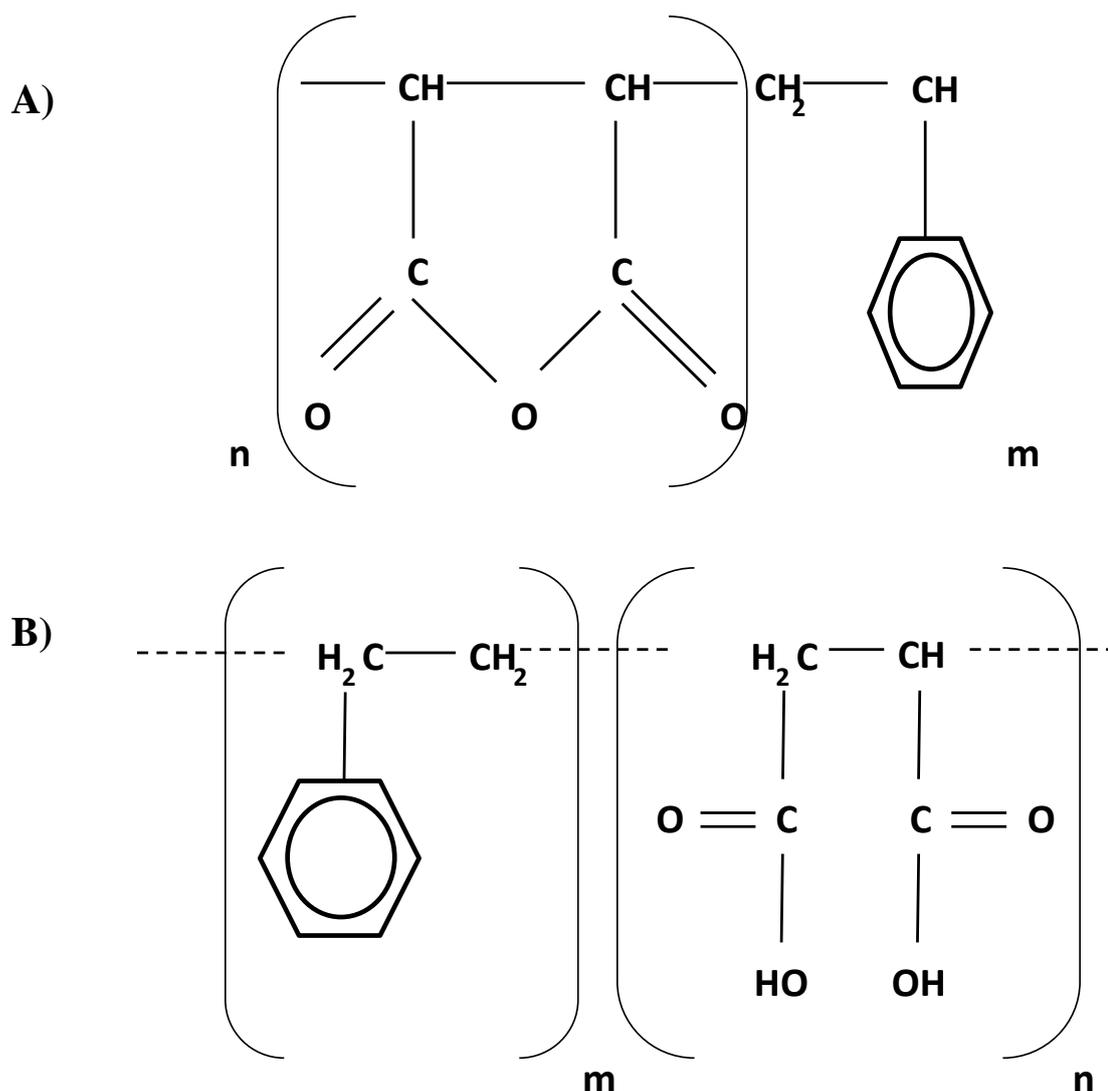


Figure 5.11 Chemical structure of poly (maleic anhydride-styrene) and poly (maleic acid-styrene) A) is the chemical structure of poly (maleic anhydride-styrene) showing the styrene benzene ring and the anhydride group in the polymer. The number of styrene molecules and anhydride groups can vary denoted by m and n . B) shows poly (maleic acid-styrene) after hydrolysis with NaOH. When the pH reaches above 6 when the NaOH is added, the carboxyl (COOH) group is ionised to COO^- with Na^+ in solution.

Detergents have always been a popular choice for solubilising membrane proteins (Seddon et al., 2004) however; they come with their challenges including the formation of unstable protein-detergent micelles, loss of membrane protein function due to aggregation and the disordered nature of detergent molecules leading to the lack of crystal lattice formation (Alguel et al., 2010). Therefore, research into finding more robust agents such as modified detergents and polymers for solubilising membrane proteins is a current objective (Jamshad et al., 2011). It was recognised by Tonge and Tighe (Tonge and Tighe, 2001) that PMAS (mixed with a phospholipid), have properties that can be exploited in biomedical science, such as drug

delivery, to target areas of the body including the lung (Tonge and Tighe, US Patent number 6,436,905) and also as a contact lens fluid to prevent dry eye syndrome (Tighe et al., 2013, Patent application number GB 1312343.5, Appendix A6). It was later recognised that by mixing the PMAS with a phospholipid and cell membranes, membrane proteins were solubilised in a polymer-membrane protein-lipid complex (Knowles et al., 2009). Moreover, the typical DDM solubilisation protocol of membrane proteins from *P. pastoris* membranes was modified with PMAS as the solubilising agent instead of DDM and DMPC instead of CHS as the supplementary lipid and the buffers and protease inhibitors remained unchanged in the solubilisation mixture. This initial work was performed by Knowles and colleagues and Jamshad and colleagues with the membrane proteins, PagP and bacteriorhodopsin. Furthermore, the results led to successful biophysical experiments (Knowles et al., 2009; Jamshad et al., 2011). In these studies, the polymer used was called styrene maleic acid (SMA) which is the same as PMAS 2000P used in this thesis (Chapter 1); (Tighe et al., 2013, Patent application number GB 1312343.5, Appendix A6). The aim was to look for improvements on 2000P by changing its structure and also by optimising the actual solubilisation protocol.

Table 5.1 shows the PMAS that were used in this study. These were chosen based on commercial availability of the raw material. 2000P has a molecular weight of approximately 7500 g mol^{-1} and consists of a chain of two styrene to every one maleic anhydride/acid. It is the largest polymer that was tested. In terms of size, 1000F was next largest (5500 g mol^{-1}) followed by the smallest, 1600 and 1600ME (both 1600 g mol^{-1}). These PMAS were all 1:1 for styrene : maleic anhydride/acid. The four PMAS solutions were prepared using hydrolysis and reflux reactions by Dr. Anisa Mahomed, Aston University, United Kingdom. Reflux reactions were carried out for the higher molecular weight and more hydrophobic PMAS where the anhydride rings were more difficult to open and hydrolysis was not sufficient.

Table 5.1 Four PMAS solutions used for testing solubilisation of hA_{2A}R in *P. pastoris* membranes. The number of styrene : maleic anhydride/acid in the polymer chain were 2 : 1 for 2000P. All others were 1:1. For 1600ME, esterification of the 1600 PMAS was performed with methanol. The molecular weights differed for each PMAS. P denotes powder of the raw material, F denotes flakes of the raw material.

PMAS name	m(number of styrenes)	n(number of maleic anhydride)	Approximate molecular weight (g mol ⁻¹)
2000P	2	1	7500
1000F	1	1	5500
1600	1	1	1600
1600ME	1	1 + methyl group	1600

Each PMAS solution (final concentration of 3%) was prepared at pH 7.0 and pH 11.0 and was used for the solubilising experiments. These pHs were selected because pH 11.0 is the final pH once the hydrolysis procedure is complete while pH 7.0 is more physiological pH. Therefore, pH 11.0 was tested to see if the polymer could be used directly to solubilise the membrane protein. The actual pH was also measured when the membrane and lipid were added to the solubilisation mixture to see how much it altered. It was found that when any of the pH 7.0 PMAS were added to the *P. pastoris* SMD1163-MT-hA_{2A}R membranes plus the DMPC lipid, the pH altered in the range of 7.1-7.2. For the pH 11.0 PMAS the pH altered in the range of 8.2-8.5.

5.3.1. Preliminary optimisation of PMAS solubilisation conditions

Initially work was done with SMA *co*-polymer (PMAS 2000P equivalent) and *P. pastoris* expressing hA_{2A}R (unpublished results) in collaboration with Dr. Yu-pin Lin, University of Birmingham, United Kingdom using Knowles and colleagues specified protocol (Knowles et al., 2009). However, there was an opportunity to improve upon the solubilisation protocol and therefore the main solubilisation conditions (solubilisation buffer composition, time period of solubilisation incubation and temperature of the solubilisation) were tested for the PMAS 2000P. The optimal conditions retrieved were used for all subsequent experiments.

Knowles and colleagues had included a buffer in the solubilisation mixture similar to that used in DDM solubilisations (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10% glycerol, 10% DMPC and membrane (40-80 mg mL⁻¹)); (Knowles et al., 2009). Therefore, this was an opportunity to investigate whether this buffer was necessary in the solubilisation as the *P. pastoris* membranes were already suspended in 50 mM Tris-HCl, 5% glycerol, 2mM EDTA and 0.2% protease inhibitors. If the solubilisation reaction could be kept simple, without the buffer, it would be more user-friendly and less expensive. In this section the solubilisation reactions are therefore referred to as 'standard' or 'simple'. The standard reaction includes: MT-hA_{2A}R SMD1163 *P. pastoris* membranes; 10% DMPC, 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10% glycerol and the PMAS. The simple reaction includes: MT-hA_{2A}R SMD1163 *P. pastoris* membranes; 10% DMPC, and the PMAS. Table 5.2 summarises the results from testing the type of solubilisation, the time of incubation and the temperature on the B_{max} estimate. From this test, it can be seen that the best condition was a simple solubilisation, incubated overnight at room temperature.

Table 5.2 Summary of solubilisation condition tests using PMAS 2000P pH 7.0 and DDM. Table shows B_{max} estimate values of solubilised MT-hA_{2A}R from SMD1163 *P. pastoris* membranes using PMAS 2000P pH 7.0 or DDM. The simple versus standard buffer recipe was tested, the time of the incubation and the temperature of the incubation. RT denotes room temperature and O/N denotes overnight (~16 h). Data are from at least 2 replicates and SEM are shown \pm .

	Time of incubation	Temperature of incubation	B_{max} estimate (pmol mg ⁻¹)	
			DDM	PMAS 2000P pH 7.0
Standard solubilisation buffer	1 h	RT	2.4 \pm 0.6	2.1 \pm 0.1
Simple solubilisation buffer	1 h	RT	1.2 \pm 0.8	3.9 \pm 0.3
Simple solubilisation buffer	O/N	RT	1.1 \pm 1.3	4.9 \pm 0.1
Simple solubilisation buffer	O/N	4°C	5.3 \pm 0.3	4.7 \pm 0.1

Standard solubilisation buffer for DDM solubilisations are 20 mM HEPES, pH 7.4, 50% glycerol, 250 mM NaCl, 1 μ L protease inhibitors, 5% (w/v) DDM and 0.5% (w/v) cholesteryl hemi-succinate (CHS) plus membrane fraction. Simple solubilisation buffer for DDM solubilisations are 5% (w/v) DDM and 0.5% (w/v) CHS plus membrane fraction. Standard solubilisation buffer for PMAS 2000P pH 7.0 solubilisations are 10% glycerol, 500 mM NaCl, 2.3% (w/v) PMAS, 1% (w/v) DMPC, 50 mM Tris-HCl, pH 8.0 and membrane fraction. Simple solubilisation buffer for PMAS 2000P pH 7.0 solubilisations are 2.3% (w/v) PMAS, 1% (w/v) DMPC plus membrane fraction as outlined in Chapter 2 section 2.2.10 and 2.2.12.

Figure 5.12 displays the solutions in the solubilisation protocol at each step. It is important to note the clarity of the final solubilisation mixture as this gives a preliminary indication if any hyper-coiled structures have been formed (Tonge and Tighe, 2001) and hence if the solubilisation has worked. It can be seen that the membrane and the membrane plus the DMPC are quite cloudy when mixed together. But when the PMAS is added to the mixtures, some became clear. It was noted that some would turn clear immediately but some took longer and hence the overnight rocking incubation period was necessary.

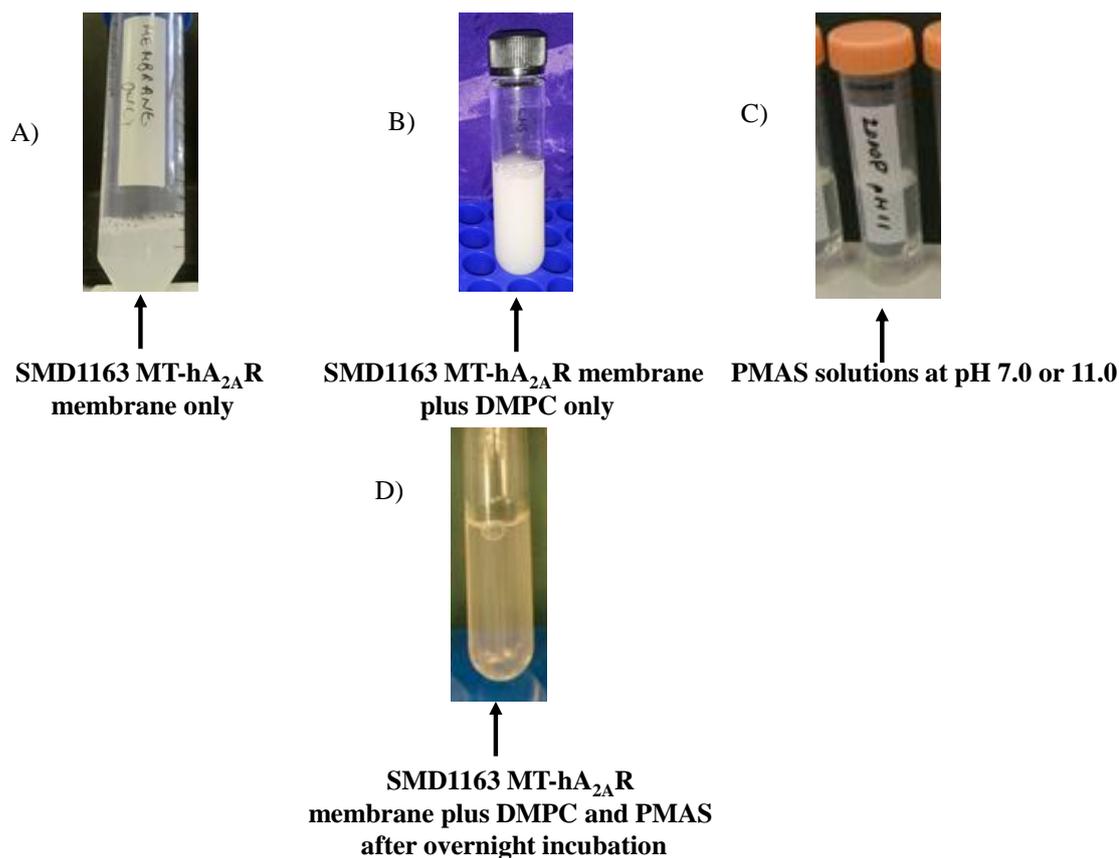


Figure 5.12 Clarity of solutions in solubilisation experiments. A) Shows the SMD1163 *P. pastoris* membrane with MT-hA_{2A}R only. B) Shows the SMD1163 *P. pastoris* membrane with MT-hA_{2A}R only with DMPC. C) Shows the 2000P pH 11.0 PMAS solution only. D) Shows the SMD1163 *P. pastoris* membrane with MT-hA_{2A}R plus DMPC and PMAS after overnight, room temperature incubation.

5.3.2. Screening PMAS solutions for solubilisation of MT-hA_{2A}R from SMD1163 *P. pastoris* membranes

This section describes the screening process carried out to observe which PMAS were the best at solubilising functional MT-hA_{2A}R from SMD1163 *P. pastoris* membranes and therefore, which ones were suitable to move forward with more detailed experiments. Radio-ligand binding assays were carried out on the supernatant fraction obtained via ultra-centrifugation (solubilised material). Table 5.3 summarises the B_{max} and pK_d data (if applicable). In the first instance, single point radio-ligand binding was performed with the [³H]ZM241385 and if no saturation binding curves were performed due to poor performance of the PMAS, these data were considered as B_{max} estimates. It can be seen that the 1000F (pH 7.0 and 11.0) and the 1600 (pH 7.0 and 11.0) PMAS solubilisations did not give any radio-ligand binding activity, suggesting that they failed to solubilise any intact MT-hA_{2A}R at all. Therefore, these polymers were eliminated from any further study. It is interesting to point out that the 1000F pH 11.0 solubilisation mix cracked and broke the ultracentrifugation tube after being subjected to a

100000 \times g spin. This occurred on two occasions but did not happen in the 1000F pH 7.0 samples (Figure 5.13). It was unclear why this occurred but it renders the polymer unsuitable for further study.

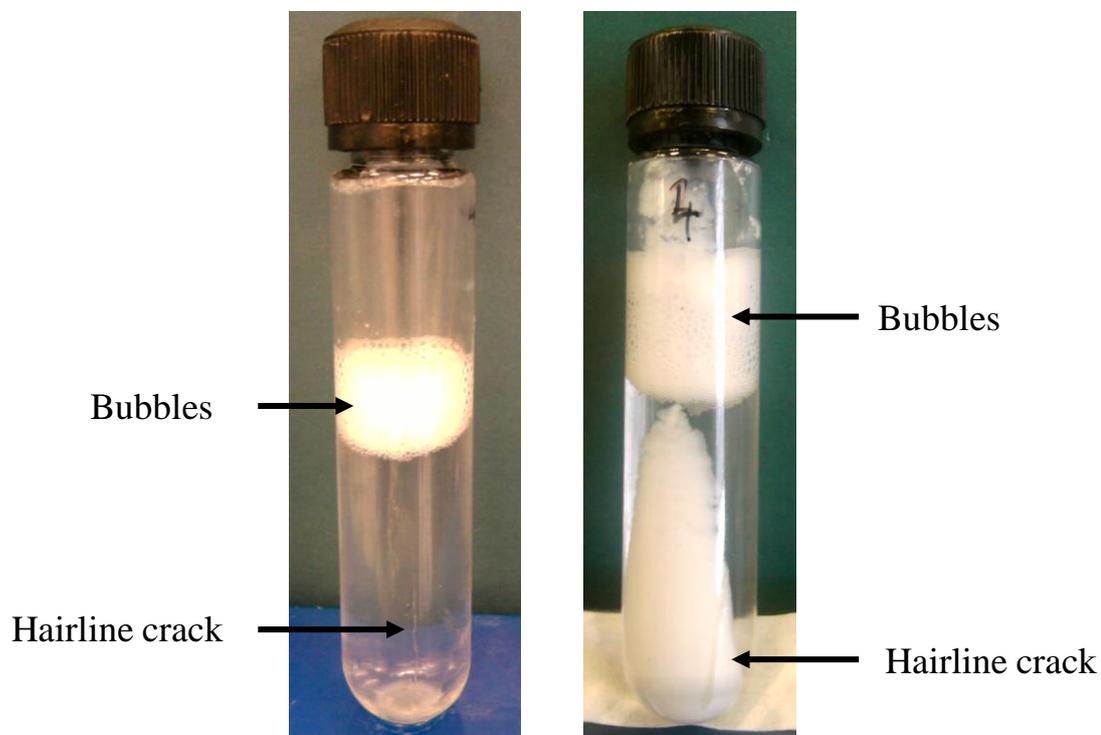


Figure 5.13 Cracked ultracentrifugation tubes with 1000F pH 11.0 solubilisation mixture after 1 h spin at 100000 \times g. Tubes are from 2 separate experiments.

The pK_d of the MT-hA_{2A}R was evaluated for remaining solubilisations (DDM, 2000P pH 7.0 and 11.0 and 1600ME pH 7.0 and pH 11.0) and were compared to the membrane-bound MT-hA_{2A}R via saturation or competition binding curves. Table 5.3 shows that there was no significant difference between the pK_d calculations suggesting that the receptor did not alter in affinity as a consequence of the solubilisations. The B_{max} values were different when compared to the membrane-bound MT-hA_{2A}R which gave the highest B_{max} value of 6.4 ± 0.5 pmol mg⁻¹. The other solubilisations (DDM, 2000P pH 7.0, 1600ME pH 7.0 and pH 11.0) gave lower B_{max} or B_{max} estimates ranging from 4.9-5.6 pmol mg⁻¹ which were lower than membrane-bound MT-hA_{2A}R.

Table 5.3 B_{max} and pK_d values for DDM and PMAS solubilisations. Saturation binding curves were performed on DDM solubilisations and membrane-bound MT-hA_{2A}R previously and the B_{max} and pK_d values were calculated from non-linear regression and recorded in this table. For the PMAS solubilisations, single point radio-ligand binding assays were carried out in the first instance with 10 nM [³H]ZM241385 and if no binding activity was observed, no further radio-ligand binding analysis was performed, as was the case for the 1000F and 1600 PMAS. For the remaining PMAS (2000P pH 7.0 and 11.0 and 1600ME pH 7.0 and pH 11.0), competition binding curves with ZM241385 were constructed and the pK_d values were derived from the EC_{50} values once the curve was best fit to a one site binding model. All data are means from at least 3 separate experiments including the \pm SEM and nd indicates 'not determined'. † indicates B_{max} estimate from homologous competition binding curves. $p < 0.05$ when compared to native membrane (ANOVA with Bonferroni's post hoc test).

Solubilisation reagent	B_{max} or B_{max} estimate† (pmol mg ⁻¹)	pK_d
None, native SMD1163-MT-hA_{2A}R membrane	6.4 ± 0.5	8.5 ± 0.1
DDM pH 7.0 plus CHS	5.6 ± 2.3	8.3 ± 0.4
2000P pH 7.0 plus DMPC	$4.9^\dagger \pm 1.0$	8.3 ± 0.2
2000P pH 11.0 plus DMPC	$6.4^\dagger \pm 0.7$	8.9 ± 0.4
1000F pH 7.0 plus DMPC	0.0	nd
1000F pH 11.0 plus DMPC	0.0	nd
1600 pH 7.0 plus DMPC	0.0	nd
1600 pH 11.0 plus DMPC	0.0	nd
1600ME pH 7.0 plus DMPC	$5.4^\dagger \pm 1.6$	8.3 ± 0.2
1600ME pH 11.0 plus DMPC	$5.2^\dagger \pm 1.8$	8.4 ± 0.3

This activity data seemed to relate with the physical clarity of the solubilisation solution for the DDM and PMAS experiments. Figure 5.14 shows a representation of clarity of the solubilisation solution versus B_{max} values and relative clarity of the solution. It can be seen that the very cloudy solutions of 1000F and 1600 PMAS, at both pH values tested, gave zero B_{max} values, whereas the other solutions were clearer and correlated with increased B_{max} values.

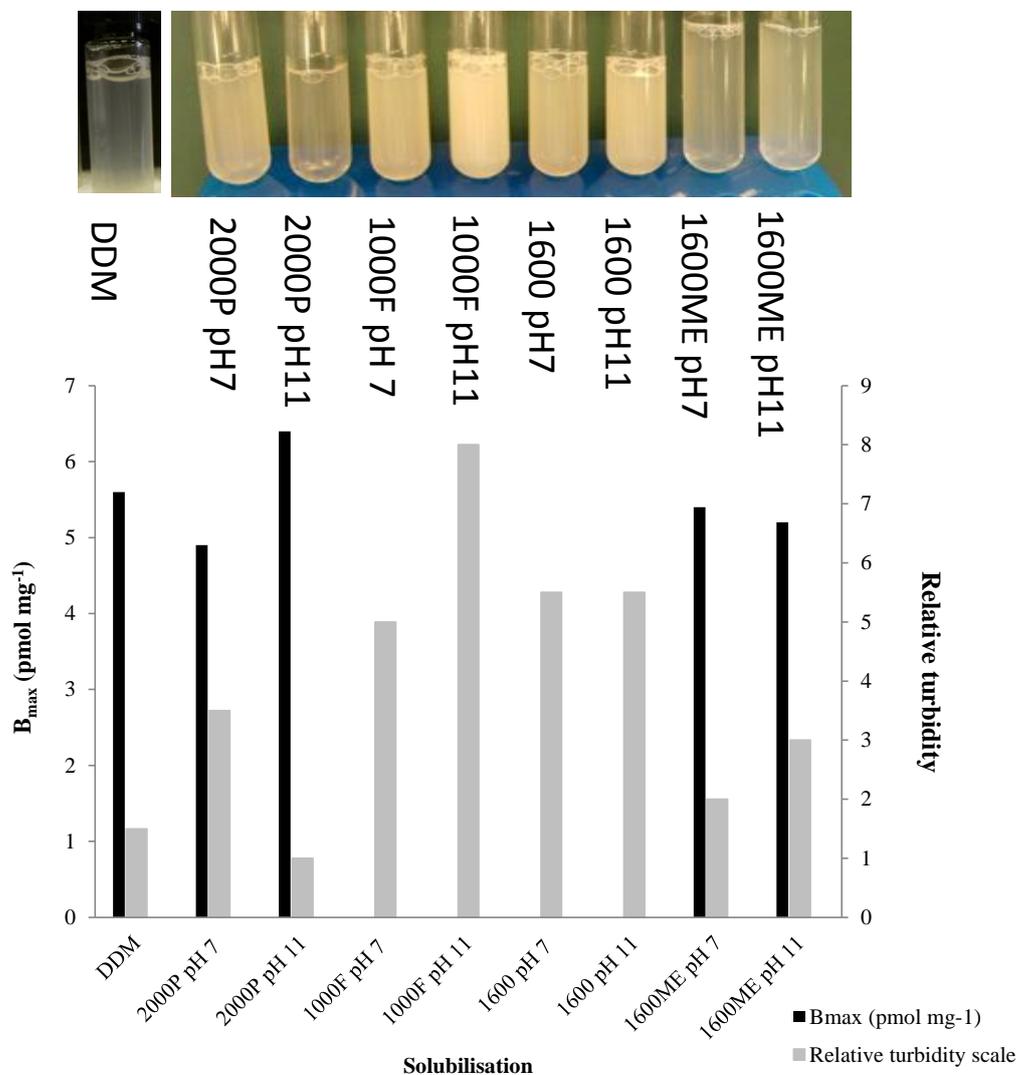


Figure 5.14 Graphical representation of clarity of solubilisation solution versus $MT-hA_{2A}R$ activity. The B_{max} values versus relative clarity of the solution are plotted. The relative clarity values were calculated using ImageJ software where a PMAS solution with no membrane or lipid added was used as background to calculate the intensity values. A photograph of the PMAS and DDM solubilisations is shown above the bar graph. Tubes are shown after overnight incubation at room temperature and before ultra-centrifugation. Visual comparison of clarity versus B_{max} shows a trend where the clearer the solubilisation, the better the activity and vice versa.

The final analysis into the PMAS screen was to calculate percent recovery of MT-hA_{2A}R from each solubilisation. Therefore, single-point binding assays were performed on the supernatant (expected solubilised material) and also the pellet (expected non-solubilised material) after the ultra-centrifugation step. Figure 5.15 shows that 1000F and 1600 pH 7.0 and 11.0 polymers solubilised zero or almost zero MT-hA_{2A}R as the activity remained in the pellet. 1600ME pH 7 and 11 gave higher activity in the supernatant (32.5-46.1%) than in the pellet (29.4-29.9%). For 2000P pH 7.0, a higher amount of activity was retained in the pellet (62.6%) than in the supernatant (31.2%) which was noteworthy. This was the opposite for 2000P pH 11.0 as most of the activity was found in the supernatant (72.0%) rather than in the pellet (23.3%) and was comparable to the DDM recovery (supernatant, 70.9%, pellet 15.3%). Compared to all the other PMAS, 2000P pH 11.0 performed the best at this screening stage. For all samples however, there was activity that was lost and unaccounted for, suggesting loss of sample during the process of solubilisation ranging from 3.0-48.2%.

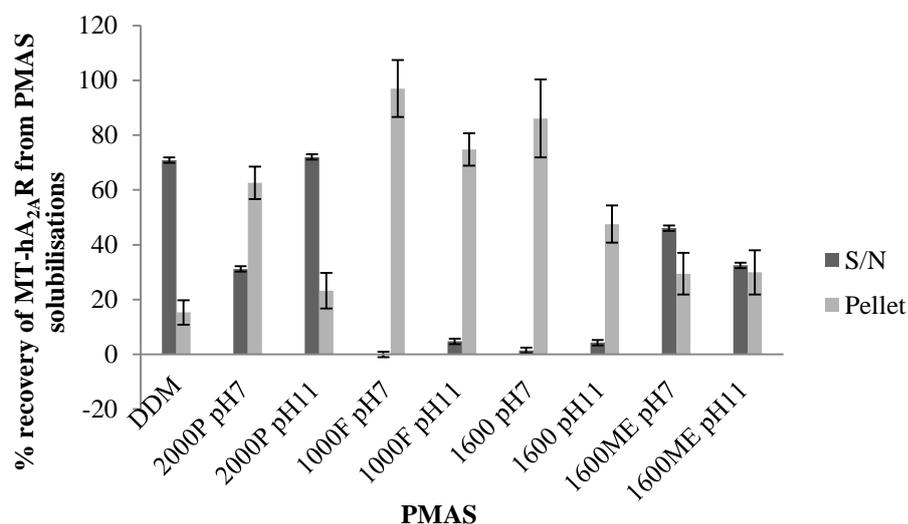


Figure 5.15 Percent recovery of MT-hA_{2A}R from PMAS solubilisations. MT-hA_{2A}R activity measured in the supernatant and pellet fractions of the solubilisation mixtures after ultra-centrifugation. Dark grey bars are % recovery values for supernatant fraction of the solubilisation mixture, light grey bars are for the pellet fraction of the solubilisation mixture. All values calculated are from means from at least 3 experiments and their SEM shown in the graph.

5.3.3. Comparative pharmacology of PMAS solubilised MT-hA_{2A}R

It is important with any receptor research, that its pharmacology is examined when investigating techniques that could alter it using competition radio-ligand binding assays. In this instance, the MT- hA_{2A}R from the PMAS and DDM solubilisations was interrogated with established agonists and antagonists to hA_{2A}R. The molecules used in the assays were the agonist N-ethylcarboxamidoadenosine (NECA) and the antagonists, xanthine amine congener (XAC); theophylline and ZM241385. Table 5.4 shows the pK_i values for each assay derived using the Cheng-Prusoff equation. For all solubilisations the pK_i values showed the general order of affinity for each agonist and antagonists was ZM241385 > XAC and NECA >theophylline which is in agreement with the literature (Fraser, 2006; Singh et al., 2010). There were some significant differences detected upon statistical examination (ANOVA analysis) of the pK_i values from the various solubilisations when compared to membrane-bound MT- hA_{2A}R. For the ZM241385, the competition binding curve showed that the solubilisation with the 1600ME pH 7.0 PMAS gave a pK_i that was significantly lower (p value <0.05) than the membrane bound MT- hA_{2A}R for ZM241385 and XAC (p value <0.01). The solubilisation with the 2000P pH 7.0 PMAS gave a pK_i that was significantly lower (p value <0.05) than the membrane bound MT- hA_{2A}R for NECA. There were no significant differences with any other pK_i values when compared to the membrane bound MT- hA_{2A}R suggesting no alterations in the receptor during the solubilisations. Figures 5.16 shows the full competition binding curves achieved from these assays. The figure shows the four curves (panel A, B, C and D) for each molecule (ZM241385, XAC, theophylline and NECA) for each solubilisation that passed the screening process carried out in section 5.3.2. Therefore the data are for the DDM, 2000P pH 7.0, 2000P pH 11.0, 1600ME pH 7.0 and 1600ME pH 11.0 solubilisations against each molecule. The membrane-bound MT- hA_{2A}R competition binding curve is also included.

Table 5.4 Summary of pK_i values as an indication of the pharmacological properties of MT-hA_{2A}R after PMAS and DDM solubilisations. The pK_i values are means of two separate experiments and their SEM shown. Competition binding assays were performed using [³H]ZM241385 as the competitor to ZM241385, NECA, XAC and theophylline. Data were fit to a one site binding model and the K_i (and pK_i) values were calculated from the EC_{50} values and using the Cheng-Prusoff equation. $p < 0.05$ indicated by * and $p < 0.01$ indicated by ** when compared to native membrane (ANOVA with Dunnett's post hoc test)

Solubilisation type of MT-hA _{2A} R	pK_i			
	ZM241385	XAC	NECA	Theophylline
Membrane-bound MT-hA_{2A}R	8.4 ± 0.2	6.5 ± 0.1	6.1 ± 0.1	4.3 ± 0.3
DDM	8.2 ± 0.2	6.6 ± 0.1	5.9 ± 0.1	4.5 ± 0.3
2000P pH 7.0	7.8 ± 0.2	6.3 ± 0.2	5.7 ± 0.1*	4.5 ± 0.3
2000P pH 11.0	8.2 ± 0.4	6.9 ± 0.1	6.1 ± 0.1	4.8 ± 0.3
1600ME pH 7.0	7.5 ± 0.1*	5.8 ± 0.1**	6.4 ± 0.1	3.8 ± 0.3
1600ME pH 11.0	7.6 ± 0.1	6.4 ± 0.1	6.2 ± 0.1	4.0 ± 0.2

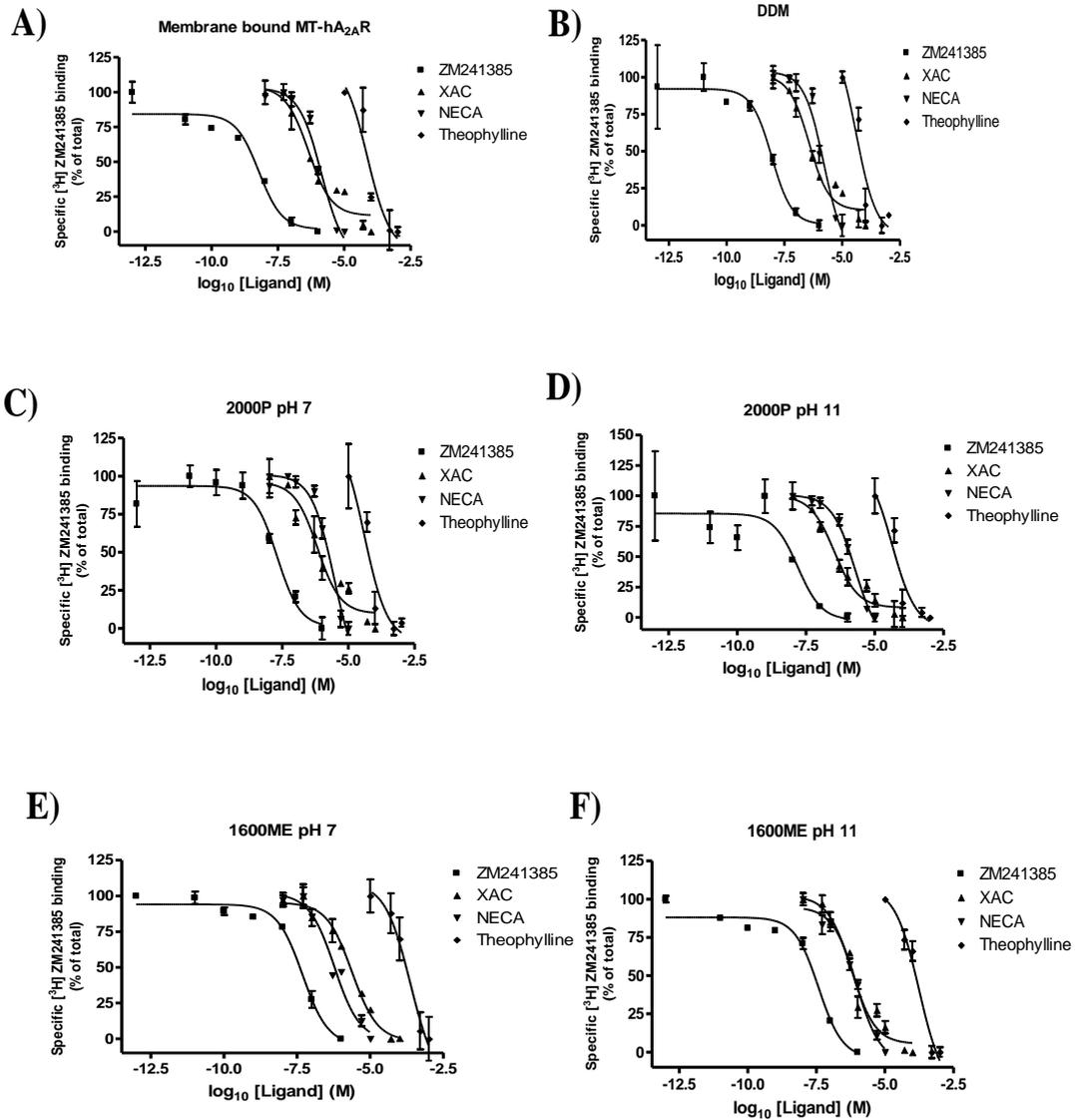


Figure 5.16 Competition binding curves for MT-hA_{2A}R (membrane bound, DDM solubilised, 2000P pH 7.0 solubilised, 2000P pH 11.0 solubilised, 1600ME pH 7.0 solubilised and 1600ME pH 11.0 solubilised) with ZM241385, XAC, NECA and theophylline agonist and antagonists. Binding curves of NECA, ZM241385, XAC and theophylline were determined using [³H]ZM241385 as the competitor for A)membrane bound MT-hA_{2A}R, B)DDM solubilised, C)2000P pH 7.0 solubilised, D)2000P pH 11.0 solubilised, E)1600ME pH 7.0 solubilised and F)1600ME pH 11.0 solubilised MT-hA_{2A}R. The general trend of affinity for all the drugs and the solubilisations were ranked as ZM241385 > XAC and NECA > theophylline. The data were fit to a one site binding model and the pK_i values were determined from the EC₅₀ values from the curves. Data shown are from 2 separate experiments. ANOVA analysis with Dunnett's post hoc tests were performed on the pK_i values and the data summarised in Table 5.6.

5.3.4. Investigation of the lipid composition in PMAS solubilisation mixtures

A typical DDM detergent solubilisation of hA_{2A}R includes cholesteryl hemisuccinate (CHS) as it has been shown to stabilise the receptor (Jaakola et al., 2008). In the PMAS solubilisations and the previous SMA solubilisations, the phospholipid, 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) is present in the solubilisation mixture. In previous work, no explanation was given as to why this phospholipid was used instead of CHS. Tests were therefore carried out on whether using CHS or DMPC in the solubilisation mixture has an effect on the recovery of functional protein. DDM and 2000P (pH 7.0 and pH 11.0) solubilisations were carried out using the simple buffer format described in section 5.3.1. CHS or DMPC were added to the solubilisations. The DDM solubilisation was tested further with a mixture of DMPC and CHS and another reaction where no lipid was added at all. After the incubation and the ultra-centrifugation steps, single-point binding assays were carried out and the B_{max} values were estimated. Binding was observed for the DDM plus CHS solubilisation (5.6 ± 2.3 pmol mg⁻¹), which was expected. However, when the lipid was switched to DMPC, no binding was observed at all. Furthermore, the DDM solubilisation without any lipid had less of a detrimental effect as some binding was observed (0.1 ± 0.1 pmol mg⁻¹). A mixture of DMPC and CHS gave a little more binding (0.2 ± 0.1 pmol mg⁻¹). The B_{max} estimates for the 2000P pH 7.0 and pH 11.0 gave different results. The DMPC gave expected binding results (2000P pH 7.0; 4.9 ± 1 pmol mg⁻¹ and 2000P pH 11.0; 6.4 ± 0.7 pmol mg⁻¹). When CHS was added instead of DMPC, the binding was decreased slightly (2000P pH 7.0; 4.2 ± 1.8 pmol mg⁻¹ and 2000P pH 11.0; 6.1 ± 1.3 pmol mg⁻¹) but not as dramatically as in the case of the DDM experiments suggesting that CHS has a minimal impact on the PMAS solubilised MT-hA_{2A}R (Table 5.5).

Table 5.5 Summary of lipid (DMPC and CHS) composition in PMAS and DDM solubilisation experiments. Data shown are for B_{max} estimates from single-point binding assays. DDM solubilisations were carried out with CHS, with DMPC, with CHS and DMPC or with no additional lipids. 2000P pH 7.0 and pH 11.0 solubilisations were carried out with CHS or with DMPC. Data are means from triplicate separate experiments with the SEM shown.

Type of solubilisation	Estimated B _{max} (pmol mg ⁻¹)	SEM ±
DDM with CHS	5.6	2.3
DDM with DMPC	0	0
DDM with CHS and DMPC	0.2	0.1
DDM with no added lipid	0.1	0.1
2000P pH 7 with DMPC	4.9	1
2000P pH 7 with CHS	4.2	1.8
2000P pH 11 simple DMPC	6.4	0.7
2000P pH 11 simple CHS	6.1	1.3

The 2000P pH 11.0 solubilisation was explored further since it gave the higher B_{\max} result when CHS was used as the lipid (Table 5.5) and when it was compared to the 2000P pH 7.0 solubilisation. A homologous competition binding curve to [^3H] ZM241385 was carried out to determine the pK_d values when a solubilisation was carried out with CHS and DMPC and compared to the DDM plus CHS solubilisation (Figure 5.17). The 2000P pH 11.0 solubilisation with CHS gave a higher pK_d of 9.5 ± 0.6 when compared to the solubilisation with DMPC which gave a pK_d of 8.9 ± 0.4 and the DDM with CHS solubilisation which gave 8.2 ± 0.2 . This was an interesting result and suggests that the presence of CHS in the 2000P pH 11.0 solubilisations increases the affinity of the MT-hA_{2A}R receptor.

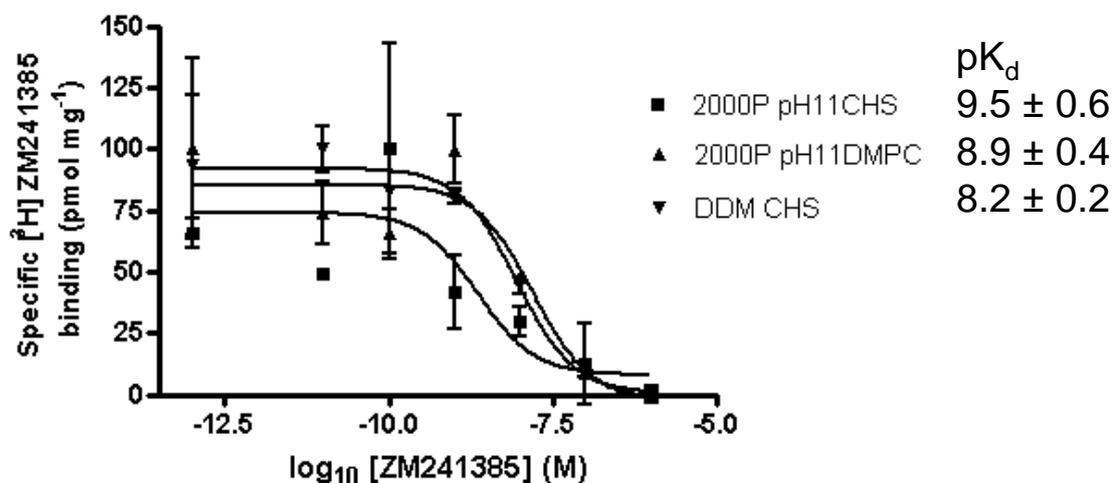


Figure 5.17 Competition binding curve after solubilisation in (i) 2000P pH 11.0 with CHS and DMPC and (ii) DDM with CHS. Binding curves of ZM241385 for 2000P pH 11.0 CHS or DMPC and DDM with CHS solubilisations were determined using [^3H]ZM241385 as the competitor. The data were fit to a one site binding model and the pK_d values were determined from the EC_{50} values from the curves. Data shown on curve are from 3 separate experiments.

5.3.5. Preliminary purification tests of MT-hA_{2A}R after 2000P pH 11.0 solubilisation compared to DDM solubilised membrane.

After solubilisation, the next step is to purify the isolated receptor. This is often difficult as the receptor may change structure or fold differently once the lipid-like environment has been removed (Serebryany et al., 2012). Therefore, the purification procedure itself requires optimisation. In this section, a preliminary purification test was carried out on a 2000P pH 11.0 solubilised MT-hA_{2A}R sample and a DDM solubilised sample. Table 5.6 illustrates the percentage activity calculated from the starting native membrane material derived from pmols per 1 L culture for each step of the solubilisation and manual purification process. The data show similar solubilisation and purification profiles for both the DDM and 2000P pH 11.0. For the DDM solubilisation, the total eluted material comprised 41.72% of the total input of the native membrane compared to 36.51% for the 2000P pH 11.0 solubilised material. For [³H]ZM241385 binding, the B_{max} value of the elution for the DDM solubilisation was 7.9 ± 0.8 pmol mg⁻¹ and 7.9 ± 1.0 nmol mg⁻¹ for the 2000P pH 11.0 solubilisation. The theoretical hA_{2A}R specific activity maximum is 24 nmol mg⁻¹ and therefore this purification corresponded to only 0.03% of the theoretical maximal value. This suggests that most of the receptor population was not active for both solubilisations where protein aggregation or denaturation may have occurred during the purification process. It is possible that technical issues with the protocol were responsible for these results. Future experimentation would enable protocol optimisation using the DDM solubilisation as a positive control. Due to technical issues, a silver stain gel of the purified samples was not successful and therefore not shown.

Table 5.6 MT-hA_{2A}R activity of native membrane, solubilised with either DDM or 2000P pH 11.0 and at each purification step. Data are expressed as percentage of total volumetric pmol where starting material (native membrane) is 100% activity. SEM are from duplicate samples.

Treatment and purification steps	DDM		2000P pH 11.0	
	%	SEM	%	SEM
Native membrane	100.00	0.32	100.00	0.32
Solubilised	77.96	1.04	69.06	3.22
Flow-through	35.14	0.34	32.16	0.31
Wash 1	0.00	0.00	0.00	0.00
Wash 2	0.00	0.00	0.00	0.00
Elution 1	0.02	0.00	0.01	0.00
Elution 2	0.24	0.01	0.26	0.01
Elution 3	20.71	0.44	16.98	0.36
Elution 4	20.75	0.00	19.26	0.00

5.4. Discussion

The research in this chapter is concerned with solubilising hA_{2A}R from *P. pastoris* membranes and investigating the use of novel compounds as replacements for detergents in that process.

It was first necessary to design and develop a novel, more updated, working hA_{2A}R construct for these studies. The construct designed was a de-glycosylated, truncated (A316) multi-tag version of hA_{2A}R, named MT-hA_{2A}R that was transformed into a protease deficient *P. pastoris* strain (SMD1163). It was compared to a previously used construct (a full-length, de-glycosylated form with two tags (dG-hA_{2A}R) in a non-protease deficient strain (X33)) for activity. The specific binding activity for the MT-hA_{2A}R was 6.4 pmol mg⁻¹ and for the dG-hA_{2A}R was 4.4 pmol mg⁻¹. The pK_d values were comparable (8.5 for MT-hA_{2A}R and 8.3 for dG-hA_{2A}R). The MT-hA_{2A}R construct was also cloned into a mammalian vector and transfected into HEK cells. An untagged version of the hA_{2A}R construct was compared for activity, and it was verified that the multiple tags did not interfere with the binding assays. The MT-hA_{2A}R construct in SMD1163 was therefore suitable to be used in the subsequent solubilisation experiments. The hA_{2A}R construct used in the design terminated at A316 which removed the intracellular C-terminal tail and the intracellular loop 3 (ICL3). This version was used for the structural determination of hA_{2A}R by Jaakola and colleagues in 2008 by modifying it with a T4 lysozyme (T4L) to make the structure more stable (Jaakola et al., 2008) as it was previously reported that the C-terminal tail degrades during recombinant production (Weiß and Grisshammer, 2002, Singh et al., 2008, Singh et al., 2010). This design was therefore perfect for the needs of this thesis research. The hA_{2A}R A316 was further modified to contain no glycosylation site via a mutation (N154Q) (Fraser, 2006). The construct included three tags to allow flexibility when detecting or purifying the construct. The tags were His₁₀, FLAG and a biotin tag. These tags are excellent for detection (with antibodies for immunoblot studies) and also for purification of recombinant proteins (Terpe, 2003). It has been debated however if these tags cause altered solubility or increase aggregation or even hinder crystallisation (Derewenda, 2004). Therefore the general method is to cleave off the tags (Carson et al., 2007), and hence in this design, a tobacco etch virus (TEV) protease cleavage site and an enterokinase cleavage site present in the FLAG tag sequence were included. A bovine pancreatic trypsin inhibitor (BPTI) was used as a signal sequence instead of an α -factor sequence in the construct which could be cleaved off by the presence of a thrombin cleavage site.

Expression screens of the new construct were successful and when compared to a previously used hA_{2A}R strain, gave good specific binding activity and the affinity of the receptor was similar. This was further elaborated upon where the construct was sub-cloned into mammalian

vector. Here the goal was to test if the tags were an issue and therefore specific activity was tested against an unlabelled hA_{2A}R. The results showed that the tags did not interfere with the binding activity. As discussed earlier, the tags may influence the binding activity and even downstream crystallisation and therefore it was necessary to design cleavage sites in the construct.

A set of four poly (maleic anhydride/acid-styrene) (PMAS) solutions prepared at pH 7.0 and pH 11.0 were tested as solubilising agents on MT-hA_{2A}R SMD1163 membranes and compared with the detergent, DDM. These varied in molecular weight, the ratio of styrene : maleic anhydride/acid and esterification. It was found that after optimisation, the most favourable solubilisation conditions were room temperature, at least 16 h or overnight and using a simple combination of membrane, DMPC lipid and PMAS. The MT-hA_{2A}R SMD1163 membranes were subjected to DDM solubilisations as a control. For the membrane-bound MT-hA_{2A}R, the B_{max} was 6.4 ±0.5 pmol mg⁻¹ and the pK_d was 8.5 ±0.1 and the results of the solubilisations were compared to the membrane bound receptor. The pK_d did not differ significantly for PMAS when compared to the membrane-bound MT-hA_{2A}R. The PMAS solubilisations which gave no or little activity, were 1000F pH 7.0 and 11.0 and also 1600 pH 7.0 and 11.0, therefore they were not tested further. It was also noteworthy that the clarity of the solubilisation solution was related to the binding activity of the MT-hA_{2A}R. The PMAS which gave the highest recovery most % recovered material in the correct fraction % of the solubilisation (supernatant) was 2000P pH 11.0 (72.0%) which was comparable to the DDM solubilisation (70.9%).

The PMAS that were included for further testing were 2000P pH 7.0 and 11.0 and 1600ME pH 7.0 and 11.0, along with DDM and membrane-bound MT-hA_{2A}R. The receptor after each of these solubilisations was pharmacologically characterised with hA_{2A}R agonist and antagonists. For all solubilisations, the pK_i values gave a rank order of affinity for the agonist and antagonists of ZM241385 > XAC and NECA >theophylline which were in agreement with literature findings (Fraser, 2006) and was the case for each solubilisation and membrane-bound MT-hA_{2A}R. Furthermore, it was found that for the 1600ME pH 7.0 PMAS solubilisation, the pK_i was significantly lower (p value < 0.05) for ZM241385 and XAC (p value <0.01) when compared to the MT-hA_{2A}R. Also, the 2000P pH 7.0 PMAS solubilisation gave a pK_i that was significantly lower (p value <0.05) than the membrane bound MT- hA_{2A}R for NECA. However there were no further significant differences with any other solubilisations when competition binding curves were performed, suggesting no major changes in hA_{2A}R receptor binding. The type of lipid present in the solubilisation reaction was further investigated. It was found that the presence of CHS in the solubilisation mixture of 2000P pH 11.0 gave higher pK_d result than with DMPC as the lipid. This suggests that the presence of CHS increases the affinity of the

MT-hA_{2A}R with the 2000P pH 11.0 solubilisation, and is in line with a CHS dependence on activity for other GPCRs (Jaakola et al., 2008, Rasmussen et al., 2007).

In general, the PMAS agents are more beneficial to use over detergent methods because they are inexpensive to prepare, a range of PMAS with differing properties may be prepared easily, DMPC or CHS can be used as the lipid in the solubilisation mixture and they are simple to use. With further optimisation, the aim would be to exceed the functional yield obtained with classic detergents such as DDM but in a more native environment for hA_{2A}R since the polymer retains the native lipid. These findings show that PMAS agents can be used for solubilising MT-hA_{2A}R and the data are comparable with DDM solubilisation. The impact of this work using this class of polymers has provided an insight into different system to detergents for membrane protein solubilisation. For future work, some unanswered questions need to be addressed: would this system work for any other membrane proteins? Is the polymer suitable for downstream biophysical work and crystallisation trials? Therefore it would be useful to perform experiments to investigate these questions.

The results in the chapter are currently in preparation for publication as an original research article and these results have also contributed to a filed patent (application number GB 1312343.5, Appendix A6).

Chapter 6: General discussion

The objectives of this thesis were to increase recombinant hA_{2A}R production in yeast by investigating bioprocess conditions at large (pilot-scale) and small (micro-bioreactors) scale and also extract functional receptor from *P. pastoris* membranes using novel polymers.

A volumetric yield increase in hA_{2A}R produced in *P. pastoris* was achieved by investigating the methanol feeding regime of *P. pastoris*. By applying a slow, exponential feed of methanol during induction in 35 L pilot-scale bioreactors. This type of bioprocess improvement (feed forward strategy) has not been applied or tested for any GPCR work prior to the submission of this thesis. Small-scale bioreactor work (Micro-24 microreactor) showed that a statistical DoE could be applied to screen for bioprocess conditions such as temperature, pH and DO to give optimal hA_{2A}R yields in yeast. Optimal conditions could be retrieved but data with high error needed to be addressed.

Another focus of this thesis was the extraction of hA_{2A}R from the yeast membrane with novel polymers. Here, PMAS was successful in solubilising hA_{2A}R from *P. pastoris* membranes as well as or better (in some cases) than the DDM detergent. Furthermore, the 2000P pH 11.0 PMAS worked well as a solubilising agent with either DMPC or CHS. Table 6.1 outlines the thesis objectives, how the objectives were met and in which chapter of the thesis they were addressed. The table also summarises the major outcomes and other interesting findings from this thesis.

Table 6.1 Thesis objectives, optimal conditions and major findings.

Thesis objectives proposed	How objective met and in which thesis chapter	Optimal conditions	Major findings
Increase hA_{2A}R production by investigating bioprocess conditions at a large scale (pilot-scale bioreactors)	Large (pilot-scale) <i>P. pastoris</i> X33-dG-hA _{2A} R cultures; controlled exponential feeding (μ_{set}) of methanol during induction phase (Chapter 4)	$\mu_{\text{set}} = 0.01 \text{ h}^{-1}$ for methanol induction at 22°C for large (pilot-scale) <i>P. pastoris</i> X33-dG-hA _{2A} R cultures	Cytotoxic levels of methanol in the cultivation is not a hindrance to maintaining protein production levels; pre-induction levels of recombinant protein are present in a methanol inducible system
Increase hA_{2A}R production by investigating bioprocess conditions at a small scale (micro-bioreactors)	Small (micro bioreactors) <i>S. cerevisiae</i> WT-hA _{2A} R cultures, temperature, pH, DO and additive presence tested via DoE; response surface methodology (RSM) for optimal conditions and predictive equation generated (Chapter 3)	22°C, pH 6.0, 30% DO, no DMSO culture conditions for small (micro bioreactors) <i>S. cerevisiae</i> WT-hA _{2A} R cultures	High error suggests complete process requires optimisation for membrane protein research
Extract the hA_{2A}R from <i>P. pastoris</i> membranes using novel polymers	Medium (bench-top) <i>P. pastoris</i> SMD1163-MT-hA _{2A} R cultures; Poly (maleic anhydride-styrene) PMAS (Chapter 5)	2000P pH 11.0 PMAS plus DMPC for solubilisation of MT-hA _{2A} R for large (bench-top) <i>P. pastoris</i> SMD1163-MT-hA _{2A} R cultures	CHS with 2000P pH 11.0 PMAS and esterification of PMAS increases MT-hA _{2A} R affinity, suggesting both may have a stabilising effect of the receptor

Ever since scientists began to work with recombinant proteins, the field has strived to increase target protein yields in order for more quality product to be available for commercial and therapeutic use. This has been the case for soluble proteins such as hEPO, insulin and hGH (Schmidt, 2004) In the case of membrane proteins, especially GPCRs, achieving crystal structures is the main goal (Tate, 2001). Once this goal is met, GPCR structure-function studies invariably lead to drug discovery programmes which interest the pharmaceutical and clinical industries greatly. Therefore, increasing the yield and extracting the membrane protein without disturbing the integrity of the protein is highly desired by researchers.

6.1. Controlling methanol feeds during membrane protein induction in *P. pastoris* to increase yields

It is established that both the concentration and rate of addition of methanol to *P. pastoris* cultures can substantially affect recombinant soluble protein yields (Potvin et al., 2012, Celik et al., 2009, Celik et al., 2010, Jungo et al., 2007, Kobayashi et al., 2000, Zhang et al., 2000). In contrast, the effect on recombinant membrane protein yields, particularly on GPCRs, had not been examined as systematically. Therefore in this study, the application of two methanol-phase feed profiles on the yield of recombinant hA_{2A}R was analysed in two pilot-scale 35 L *P. pastoris* cultivations, μ_{low} and μ_{high} . The set-growth rates were applied via exponential feeding of methanol to the *P. pastoris* cells in order to reach a desired specific set growth rate (μ_{set}). These were 0.01 h⁻¹ for the μ_{low} and 0.03 h⁻¹ for the μ_{high} . The major findings of this study were that the μ_{low} cultivation yielded higher amounts of hA_{2A}R compared to the μ_{high} cultivation when the amount of biomass was taken into account (536.4 pmol g⁻¹ for the μ_{low} cultivation compared to 148.1 pmol g⁻¹ for the μ_{high} cultivation). Moreover, hA_{2A}R levels were still present in cytotoxic residual methanol concentrations and pre-induction levels of hA_{2A}R (and another recombinant protein from an additional cultivation) were detected.

6.1.1. Cytotoxic levels of methanol in the cultivation are not completely detrimental to hA_{2A}R activity

The residual methanol concentration in the μ_{low} cultivation did not exceed 3 g L⁻¹ and was zero for most of the induction phase. In contrast, the residual methanol concentration in the μ_{high} cultivation increased throughout the bioprocess to > 80 g L⁻¹. Nonetheless, pre-induction hA_{2A}R levels were maintained even under these cytotoxic methanol concentrations, however this is not fully understood. In the literature, although detailed research on protein production in cytotoxic methanol conditions is lacking, there have been a few studies examining the benefits of high methanol concentrations in *P. pastoris* cultivations on recombinant protein production. One notable study was by Khatri & Hoffman, where they examined the impact of methanol

concentration on single-chain antibody fragment (scFv) in oxygen-limited *P. pastoris* cultures. They found that high methanol concentrations (30 g L⁻¹) were required in order to fully induce recombinant protein production. However, this was in an oxygen-limited system and also the temperature of the cultivation was reduced. The decrease in temperature is thought to reduce proteolysis. Moreover, they found that accumulation of degraded versions of their product (scFv) was less prevalent in the high methanol concentrated cultures, thereby simplifying downstream processing of their protein (Khatri and Hoffmann, 2006). Similar findings were established with Katakura and colleagues with their work on human β 2-glycoprotein I domain V, where production of that recombinant protein was increased in *P. pastoris* at methanol concentrations of 30 g L⁻¹ (Katakura et al., 1998).

Most of the literature however, demonstrated that optimal protein production is in the lower range for residual methanol concentrations of 2-3 g L⁻¹. Examples include, Cunha and colleagues (Cunha et al., 2004) and Schenk and colleagues (Schenk et al., 2008) where they showed that residual methanol concentrations between 2 g L⁻¹ and 3.5 g L⁻¹ were optimal for soluble protein production, while methanol concentrations between 3.7 g L⁻¹ and 20 g L⁻¹ were cytotoxic and inhibited growth (Cunha et al., 2004; Schenk et al., 2007), consistent with our findings. In a recent study, Barrigón and colleagues carried out cultures for *Rhizopus oryzae* lipase production in *P. pastoris*. They found that biomass yields and productivities were very low for methanol concentrations lower than 2 g L⁻¹ and that they had a significant increase in yield and product when the methanol was between 2-3 g L⁻¹. This was the case for the μ_{low} cultivation, however, biomass still increased in the presence of no methanol which was contrasting to Barrigón' data (Barrigón et al., 2013).

Concomitantly, it is becoming established that mixed feeding (methanol together with another carbon source) is the preferred *P. pastoris* induction regime. It was previously reported that residual methanol levels below 2.98 g L⁻¹ gave optimal yields of recombinant GFP (3.74 g) when following a mixed 60% methanol and 40% sorbitol induction regime. In contrast, a 100% methanol induction regime, where the residual methanol reached a maximum level of 182.25 g L⁻¹, gave a poor recombinant GFP yield (0.09 g); (Holmes et al., 2009). More recently, Jordà and colleagues were able to develop a ¹³C-labelled metabolic flux analysis system where they characterised the central metabolism of *P. pastoris* when grown on a mixed feed of methanol and glucose. They were able to ascertain that during recombinant production of *Rhizopus oryzae* lipase that a significant redistribution of carbon fluxes occurred in the central carbon metabolism that adds stress on certain pathways such as the glycolytic, TCA cycle and methanol dissimilation rates. Furthermore, in this study, methanol played a role as a supporting substrate to compensate for the increased energy demands of the recombinant protein production

secretion processes. This could explain why mixed feeding can increase protein production and reduce metabolic burden in *P. pastoris* (Jordà et al., 2012).

6.1.2. Pre-induction activity is present in pilot-scale and bench-top bioreactors

A significant finding emerging from this work is that hA_{2A}R and, subsequently GFP, were produced prior to induction with methanol. In both the μ_{low} and μ_{high} cultivations, hA_{2A}R binding activity was already measurable before induction with methanol had begun. This is consistent with a previous observation by Singh and colleagues during their bioreactor study using minimal medium (Singh et al., 2008). In a bioreactor study by Çelik and colleagues, recombinant hEPO appeared to be produced prior to the onset of the methanol feed, a finding that was not further elaborated (Çelik et al., 2009). In contrast, pre-induction hA_{2A}R binding activity was not apparent in shake flasks using complex medium in studies by Singh and colleagues (Singh et al., 2012) or by Fraser (Fraser, 2006). The demonstration of protein production prior to induction is especially noteworthy due to the presence of glycerol, which is a known repressor of the *AOX* promoter. Glycerol, glucose, ethanol and acetate have all been shown to support growth of *P. pastoris* cells without inducing the *AOX* promoter (Inan & Meagher, 2001). For example, Hellwig and colleagues (Hellwig et al., 2001) demonstrated glycerol in the culture medium inhibited production of a recombinant single-chain antibody by causing ethanol and acetate to accumulate.

The data in Chapter 4 further demonstrated that methanol induction did not substantially increase the pre-induction specific yield, measured in the transition phase, for cells initially cultured on glycerol. In both the μ_{low} and μ_{high} cultivations, the total yield of hA_{2A}R was at least maintained once the methanol feed commenced during the induction phase (IV). However, the increase in total yield was not substantial, being only approximately double that of the pre-induction yield of the μ_{low} cultivation. A subsequent *P. pastoris* cultivation also produced recombinant GFP in the pre-induction phases, with similar results following induction with methanol. These increases in yield are similar to those seen by Çelik and colleagues in their hEPO cultivations (Çelik et al., 2009).

Previously, it was demonstrated that high yields of recombinant GFP (3.74 g) but low biomass yields ($OD_{595} = 64$) were produced when *P. pastoris* was grown at a low induction phase growth rate ($\mu = 0.006 \text{ h}^{-1}$). In contrast, a higher induction phase growth rate ($\mu = 0.015 \text{ h}^{-1}$) resulted in a higher biomass yield ($OD_{595} = 74$), but a lower total GFP yield (0.98 g) (Holmes et al., 2009). Çalık and colleagues (Çalık et al., 2010) applied three different growth rates in the methanol induction phase of *P. pastoris* cultures expressing soluble human growth hormone: 0.02, 0.03 and 0.04 h^{-1} using a mixed sorbitol and methanol feed. The highest biomass yield (48 g L^{-1}) was obtained at a growth rate of 0.04 h^{-1} , but the highest yield of recombinant human growth hormone (270 mg L^{-1}) was produced at 0.03 h^{-1} . The authors suggested that this might be due to the lower growth rate producing lower biomass and therefore lower extracellular proteases being present in the cultivation. Similar trends were also observed for GFP, hEPO, human serum albumin, heavy-chain fragment C of *Botulinum* neurotoxin, serotype A and avidin (Çelik et al., 2009; Jungo et al., 2007; Kobayashi et al., 2000; Zhang et al., 2000).

In order to examine the balance between increased biomass yields and specific productivity, the glucose as a pre-induction carbon source was examined. Glucose has previously been shown to support growth of *P. pastoris*, but at a lower specific growth rate and yielding lower amounts of biomass than that achieved by cells grown on glycerol (Inan & Meagher, 2001). Following induction with methanol, DCW values for glycerol-grown cells were approximately 5 times higher than for glucose-grown cells in the equivalent pre-induction phase. However, the glucose-grown cells produced 3 times the GFP yield of the glycerol-grown cells. The data in this thesis suggests that specific productivity may not necessarily be solely dependent on methanol induction and can be increased at the expense of improvements in biomass yields in the methanol induction phase.

In summary, the application of a low methanol feed profile to a pilot-scale *P. pastoris* cultivation resulted in higher biomass and hA_{2A}R yields ($536.4 \text{ pmol g}^{-1}$) than a cultivation to which a high feed profile had been applied ($148.1 \text{ pmol g}^{-1}$). Production of hA_{2A}R (and GFP) was detected prior to methanol induction. Furthermore, the transition phase yield was not substantially increased following methanol induction. Using glucose as a pre-induction carbon source demonstrated that in the subsequent methanol induction phase, specific productivity can be increased at the expense of improvements in biomass yield. These data provide a platform to rationalise the production of recombinant proteins, such as GPCRs, in bioreactor cultivations. To further investigate the exponential feeding work set in Chapter 4, it would be interesting to see if a lower μ_{set} would yield higher biomass and therefore total hA_{2A}R yields. Furthermore, would the biomass have continued to increase further after 92 h since it was still rising for the

0.01 h⁻¹ μ_{low} growth rate at the end of that cultivation? Finally, it would be useful for these growth rate experiments to be tested on other GPCRs.

6.2. Using novel polymers to extract hA_{2A}R from *P. pastoris* membranes

The hA_{2A}R receptor and all other membrane proteins are not soluble in water or aqueous buffers due to their hydrophobic regions when correctly folded. Therefore, detergents are usually used to extract membrane proteins from the lipid bilayer whilst maintaining their native structure.

n-dodecyl- β -D-maltopyranoside (DDM) has been widely used for the extraction of GPCRs which have been heterologously produced in eukaryotic systems and some have had crystal structures resolved. Examples include the model GPCR used for this thesis research, the human adenosine A_{2A} GPCR (Jaakola et al., 2008); the human β 2 adrenergic GPCR (Cherezov et al., 2007), the mammalian voltage-dependent Shaker family K⁺ channel (Long et al., 2005) and 2 recently resolved structures from *P. pastoris* hosts, the human histamine H₁ GPCR complex with doxepin (Shimamura et al., 2011) and the human M₂ muscarinic acetylcholine GPCR bound to an antagonist (Haga et al., 2012).

The issue with DDM and other detergents is that many membrane proteins will either aggregate or denature during the solubilisation process, thereby leading to unsuccessful structural downstream studies; it is thought that membrane proteins that remain stable in detergent micelles tend to render better ordered crystals (Sonoda et al., 2011; Prive, 2007, le Maire et al., 2000). More specifically, as discussed in Chapter 1, detergents can only provide an approximate environment to that of the natural lipid bilayer. Neutron-scattering studies have shown that the membrane consists of layers that run perpendicular to each other (Wiener & White, 1992), the constituents of the membrane layers themselves are highly variable in phospholipid head-groups and acyl chains (Debnath et al., 2011, Damianoglou et al., 2010) and the theory of lipid rafts, states that certain lipids help maintain the function of the membrane protein in localised areas of the membrane bilayer (Lee, 2011a, Coskun and Simons, 2011, Grossmann et al., 2006). Moreover, the aim has been to mitigate the harmful effect of stripping away lipid from intimate contact with proteins. One approach has been to solubilise with a normal detergent but then to replace the detergent with a less aggressive surfactant. This may help restore some function, but it cannot replace important lipids. Therefore there has been a drive to move away from detergent solubilisation towards using other compounds. Research is being carried out using modified detergents such as calix[4]arene based detergents (C4C[n]). These were briefly investigated as part of a collaboration (Calixar SAS, France) where successful solubilisation of the hA_{2A}R was carried out with a range of C4C[n] compounds. However, the results showed that further optimisation was required as the data

were not as good as for DDM solubilisations and the purification methods needed to be further optimised. As a consequence, this research is currently on-going.

In this thesis however, the focus was to investigate non-detergent, novel compounds called PMAS in the solubilisation of hA_{2A}R from *P. pastoris* membranes. PMAS experience modifications in structural conformation when the pH changes in the environment. The alternating co-polymers of maleic anhydride, which is hydrolysed to maleic acid, and styrene changes its shape from an extended chain to a secondary structure and the hydrophobic and hydrophilic groups shift to separate domains when below or above a critical pH. This is known as hydrophobic association or hyper-coiling which results in a more compact formation of the polymer (Tonge and Tighe, 2001). The properties of these types of polymers include adsorption at the air-water interface, a reduction of the surface tension of water and as a surface active agent or surfactant. The co-polymer itself can self-associate above a certain concentration to form a micellar structure where the carboxylic pendant groups appear at the outer surface and the aromatic styrene pendant groups cluster within the inner core of spherical micelles (Tonge and Tighe, 2001).

Four PMAS solutions at pH 7.0 and pH 11.0 were used to solubilise the MT-A_{2A}R produced in SMD1163 *P. pastoris* cells. The PMAS included 2000P (molecular weight approximately 7500 g mol⁻¹, with 2 styrenes : 1 maleic anhydride); 1000F (molecular weight approximately 5500 g mol⁻¹, with 1 styrene : 1 maleic anhydride); 1600 (molecular weight approximately 1600 g mol⁻¹, with 1 styrene : 1 maleic anhydride); 1600ME (molecular weight approximately 1600 g mol⁻¹, with 1 styrene : 1 maleic anhydride and modified with methyl group), all at pH 7.0 and pH 11.0 (Patent pending, application number GB 1312343.5, Appendix A6).

Initial tests were carried out to determine the optimal solubilisation conditions for MT-hA_{2A}R from *P. pastoris* (3% PMAS, 1% DMPC, 20 mg mL⁻¹ membrane, at room temperature and for at least 16 h). 1000F and 1600 PMAS at both pHs, gave little or no specific binding activity (and the most turbid solutions) for MT-hA_{2A}R when compared to the other PMAS, DDM-solubilised material and membrane bound MT-hA_{2A}R. It is unclear why these two PMAS did not give solubilised material. One reason could be related to the molecular weights. The 1000F PMAS is lower in mass compared to 2000P (5500 g mol⁻¹ vs 7500 g mol⁻¹), so the responsive hyper-coiled structure may not be large enough for the membrane protein to be solubilised. Alternatively, further optimisation may be required for this PMAS. The 1600 PMAS molecular weight (≈1600 g mol⁻¹) is also low when compared to the 2000P PMAS. However, the 1600ME molecular weight is almost the same as the 1600 yet the 1600ME PMAS gave better specific binding activity results, as discussed below.

The PMAS that were carried forward for further investigations were the 2000P and 1600ME at both pHs. The 2000P has the same chemical structure as the styrene maleic acid (SMA) copolymer (Tonge and Tighe, 2001; Knowles et al., 2009; Rajesh et al., 2010; Jamshad et al., 2011) used as a solubilising agent by others. The protocol in previous studies relies on a pH change from acidic (pH 6.5) to basic (pH 8.0) and it is postulated that a doughnut like structure forms around the membrane protein and the lipid bilayer. However, this speculation is unlikely as it is more likely to hypercoil forming a more irregular shape.

The 2000P PMAS was prepared at two pHs, 7.0 and 11.0. The pH 11.0 solution is the pH after hydrolysis with 1 M NaOH. Therefore, tests were carried out to see if the pH 11.0 PMAS could be used directly for successful membrane protein solubilisation without the need to adjust the pH first. The pH 7.0 PMAS version was also examined to provide a solution more typical for biological experiments. By measuring the actual pH in the solubilisation mixture, it was found that it dropped from 11.0 to 8.2 (PMAS pH 11.0, membrane and DMPC lipid). For the pH 7.0 PMAS, the pH did not alter very much (7.0 to 7.1). Therefore for the pH 11.0 PMAS, the pH remained in the basic region even after the membrane and DMPC were added. The receptor was active and furthermore gave the best results from all the PMAS tests (B_{\max} value of 6.4 ± 0.7 pmol mg^{-1} , pK_d of 8.9 ± 0.4 and 72% recovery of solubilised MT-hA_{2A}R). Interestingly, the 2000P PMAS at pH 7.0 gave results with a lower B_{\max} of 4.9 ± 1.0 pmol mg^{-1} . Moreover, pharmacological characterisation of the receptor showed that the solubilised MT-hA_{2A}R in 2000P pH 11.0 had activity comparable to both the membrane-bound receptor and previous literature results: NECA > ZM241385 > XAC > theophylline (Singh et al, 2010; Fraser, 2006). The affinity for NECA was significantly increased for all PMAS and DDM solubilisations when compared to the membrane-bound MT-hA_{2A}R and for theophylline, the 2000P pH 11.0 was the only PMAS to increase the affinity of the MT-hA_{2A}R. This suggests that the 2000P pH 11.0 did not reduce the affinity of the hA_{2A}R to any of the antagonists or agonist.

The 1600ME (pH 7.0 = 5.4 ± 1.6 pmol mg^{-1} and pH 11 = 5.2 ± 1.8 pmol mg^{-1}) gave B_{\max} results not dissimilar to the DDM solubilisation (5.6 ± 2.3 pmol mg^{-1}). This PMAS was the lowest in molecular weight (~1600) and partial esterification was performed where the anhydride rings were opened with methanol and NaOH. The esterification of the PMAS (1600ME pH 7.0 and pH 11.0) made a notable difference in the solubilisation capability in these experiments as without the functionalising with methanol; the PMAS (1600 pH 7.0 and pH 11.0) was not able to solubilise the membrane. This was an expected result as the introduction of methyl ester moieties increased the number of hydrophobic regions of the polymer chain. Consequently, it is thought that more hydrophobic binding sites for the lipids are present and therefore can destabilise the lipid assemblies within the cell membrane more readily (Ladaviere et al., 2002, Thomas and Tirrell, 2000). Longer polymer chain lengths are also thought to increase the

hydrophobic regions and thereby increase the chance of disrupting the cell membrane lipids. However, from these data, it can be seen that the addition of methyl ester groups had more of an effect (1600ME, molecular weight $\approx 1600 \text{ g mol}^{-1}$) than chain length 1000F (no methyl ester groups, molecular weight $\approx 5500 \text{ g mol}^{-1}$). The 2000P however, did not have any esterification modifications, yet gave the best (2000P pH 11.0) solubilising results. This could be due to the ratio of styrene to maleic anhydride, where the ratio is 2 styrene : 1 maleic anhydride, therefore there are twice as many hydrophobic moieties than hydrophilic ones in the polymer chain. Hence, these polymers could have more hydrophobic binding sites for the lipids and therefore more chance of disruption of the cell membrane (Ladaviere et al., 2002, Thomas and Tirrell, 2000).

Investigating the type of lipid in the 2000P pH 11.0 PMAS solubilisations revealed that DMPC and CHS gave differing affinities for the MT-hA_{2A}R receptor (2000P pH 11.0 plus DMPC $pK_d = 8.9 \pm 1.4$ and 2000P pH 11.0 plus CHS $pK_d = 9.5 \pm 0.6$) but the B_{max} values were similar (2000P pH 11.0 plus DMPC $B_{max} = 6.4 \text{ pmol mg}^{-1}$ and 2000P pH 11.0 plus CHS $B_{max} = 6.1 \text{ pmol mg}^{-1}$). The lipid type was tested because previously (Knowles et al, 2009; Rajesh et al, 2010; Jamshad et al, 2011) no explanation was given why DMPC was chosen as the lipid to supplement the solubilisation reaction. Also in DDM solubilisations, the cholesterol derivative, cholesteryl hemisuccinate (CHS) is routinely used. Cholesterol is known to interact with membrane proteins directly (Paila and Chattopadhyay, 2010). In mammalian cells, it is a major component of the outer cell membranes constituting about half of the total lipids (Lee, 2011). It is known that cholesterol affects the ligand binding to many GPCRs and also increase their thermal stability. Hanson and colleagues (2008) resolved the structure of the β_2 -adrenergic receptor with a specific cholesterol binding site where 2 cholesterol molecules were bound in a cleft. The side of the cleft comprised helices I and IV and the back of the cleft was made of helices II and III (Hanson and Stevens, 2009). Contrastingly in the same year, Jaakola and colleagues resolved the structure of the human adenosine A_{2A} receptor and showed that phospholipid was bound in the same area instead of cholesterol (Jaakola et al., 2008). However, cholesterol is still deemed to be a major stabiliser for hA_{2A}R as a molecular dynamics simulation revealed that the binding of cholesterol in the cleft affects function by changing the conformation of helix II (Lyman et al., 2009). Yet its precise interaction with GPCRs is ambiguous; it is not clear whether it is attributed to indirect bilayer effects or to specific receptor binding and putative non-annular binding sites (Paila and Chattopadhyay, 2010). The literature however suggests that the presence of cholesterol modifies GPCR activity and the associated specific response appears to be receptor dependent. Additionally, some experiments have shown that both up-regulation and down-regulation are possible as are both direct and indirect actions (Oates and Watts, 2011). The experimental evidence is unclear as a number of

GPCRs can be produced in host systems that do not produce native cholesterol such as *E. coli* (Weiß and Grisshammer, 2002, Attrill et al., 2009), the GPCRs are stable, as determined by ligand binding and some are able to activate G proteins (Grisshammer and Hermans, 2001). This was certainly the case for the data in this thesis where correctly folded MT-hA_{2A}R was produced in *P. pastoris* membrane in the absence of cholesterol (ergosterol is present instead in the *P. pastoris* membranes) and moreover, the receptor remained stable during PMAS extraction supplemented with only phospholipid (DMPC). This was also the case with PMAS and CHS as well, however, the pK_d increased from 8.9 ± 1.4 (PMAS plus DMPC) to 9.5 ± 0.6 (PMAS plus CHS) suggesting the affinity of the receptor had increased and therefore an alteration had occurred during this solubilisation. Contrastingly, when the DDM plus DMPC solubilisation was explored, B_{max} was zero for MT-hA_{2A}R, whereas DDM plus CHS gave a B_{max} of 5.2 pmol mg⁻¹ and a pK_d of 8.2 ± 0.2. Here the data suggest that the detergent-lipid and polymer-lipid combinations are influencing the stability of the receptor and furthermore the PMAS solubilisation seems to be more flexible in terms of using either DMPC or CHS as the supplementing lipid for MT-hA_{2A}R solubilisations.

For future considerations regarding this work, it would be important to optimise the purification of 2000P pH 11.0 for down-stream biophysical analysis such as circular dichroism (CD), analytical ultra-centrifugation (AUC) and NMR studies. It would be valuable to retrieve transmission electron microscopy (TEM) images, to observe exactly the types of structures that are formed with the hA_{2A}R, the lipid and PMAS. Further work on the type of lipid (phospholipid or cholesterol or both) for the PMAS solubilisations and testing whether this is a receptor-dependent phenomenon would be fruitful. Moreover, developing modified PMAS with lipid molecules already present may even simplify and improve the solubilisation further. For example, functionalising the 2000P PMAS with a cholesterol molecule may improve the solubilisation further. There is also a necessity for performing thermostability experiments for the polymers and the solubilisation mixtures themselves.

6.3. Using DoE for hA_{2A}R production improvement in yeast

In industry, the DoE approach is widely used in order to improve upon product quality and process efficiency (Mandenius and Brundin, 2008). Recently DoE has been increasingly applied to the biotechnology industries especially concerning bioprocess development - specifically recombinant protein production. Although the approach has made the transition to recombinant protein production, so far it has been applied to soluble proteins only (Bora et al., 2012) and to date, there is a lack of evidence in the literature that membrane proteins have been studied in this way.

This thesis applied DoE methodology to improve production of the membrane protein, hA_{2A}R in *S. cerevisiae* strains. The main difference between soluble proteins and membrane proteins is that a direct assay can be used to measure the soluble protein product, while membrane proteins such as hA_{2A}R are assayed indirectly (membrane fractions must be prepared). DoE relies on an accurate output response to be fed back into the model (Montgomery, 2006). Each extra stage between production and assay has the potential to introduce error into the output response, which is a challenge when working with membrane proteins and perhaps this is a reason why DoE has previously been confined to soluble proteins research only.

However, work in this thesis demonstrated that DoE can be applied to membrane protein research. Optimal conditions were retrieved for hA_{2A}R production in a wild type strain of *S. cerevisiae* that were 22°C, pH 6.0, 30% DO and no DMSO additive in the culture. These conditions are typical growth and induction conditions for hA_{2A}R production (Fraser, 2006; Singh et al., 2008) and it was interesting that the input conditions did not deviate from these standard levels despite the significant errors associated with the DoE approach that were identified in this thesis. This work should be improved upon to in order to reduce the error and therefore generate a more robust DoE model.

Some future considerations to enable this would be to minimise the user handling as much as possible. This could be achieved with the use of automation and robotic platforms. Such systems are already in place for other biological assays in the genomics field for DNA microarray work (Illumina Inc. and Affymetrix Inc.). Of course, one limitation is that these systems require sophisticated engineering and are very expensive. However, the cost of the experiment may be equal to or even cheaper than repeating highly variable experiments. Figure 6.1 shows a proposed schematic of what an automated system should be capable of doing for optimal, high throughput screening of membrane protein production in yeast.

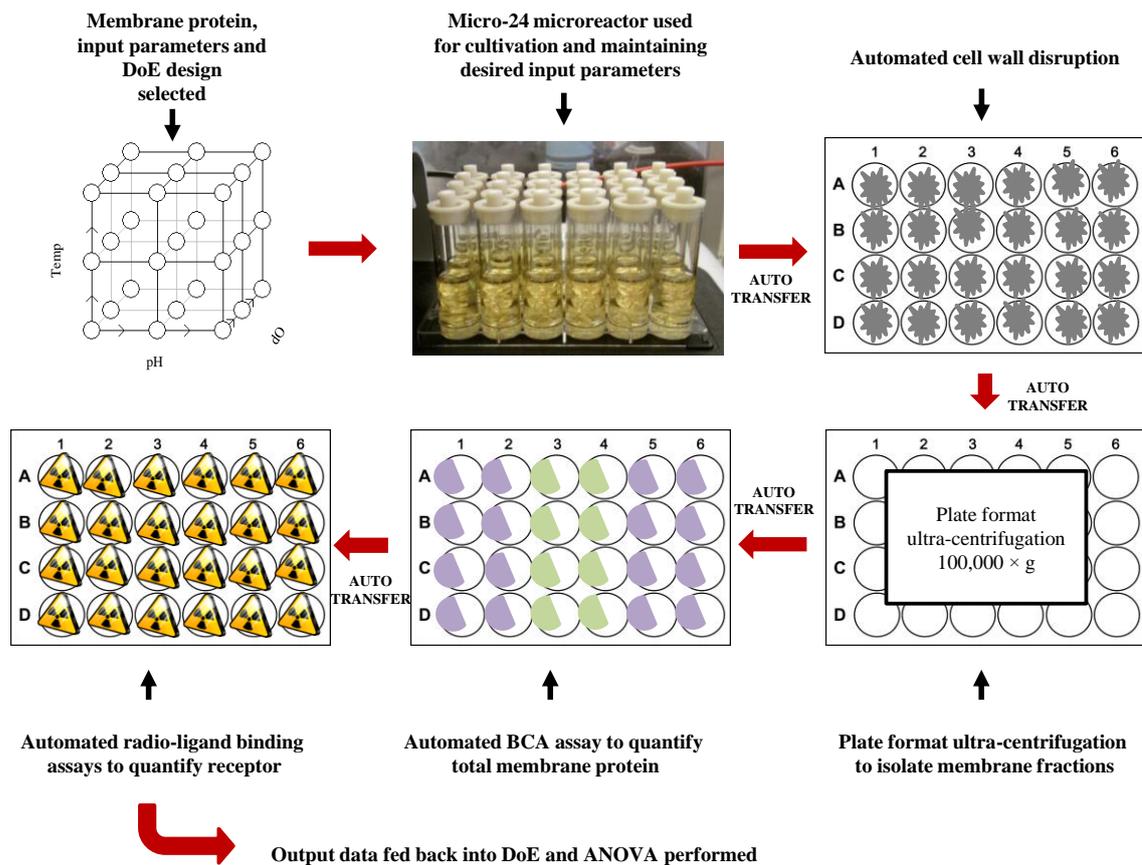


Figure 6.1 Schematic for proposed automation on of membrane protein production using DoE and Micro-24 microreactor. The scheme is a flow diagram of all the major processes required from the concept of DoE to obtaining the output radio-ligand binding assay. These include the choice of membrane protein, the choice of input factors, to running the Micro-24 microreactor, carrying out the membrane preparations, total protein quantification and radio-ligand binding assays all with automated capabilities.

A further improvement for the Micro-24 microreactor, making it more comparable to larger scale bioreactors, would be to introduce fed-batch systems of feeding via a multiple pump. This would be specifically applicable to *P. pastoris* for feeding methanol into the wells at a known amount and rate. Figure 6.2 shows a sketch of a possible set-up. The caps of the Micro-24 wells would be fitted with micro-tubing whilst maintaining their valve capacity. The tubing could be attached to one or several pumps which are connected to the desired feeds in separate vessels. This approach negates the need to stop a run and manually feed each well in a batch style.

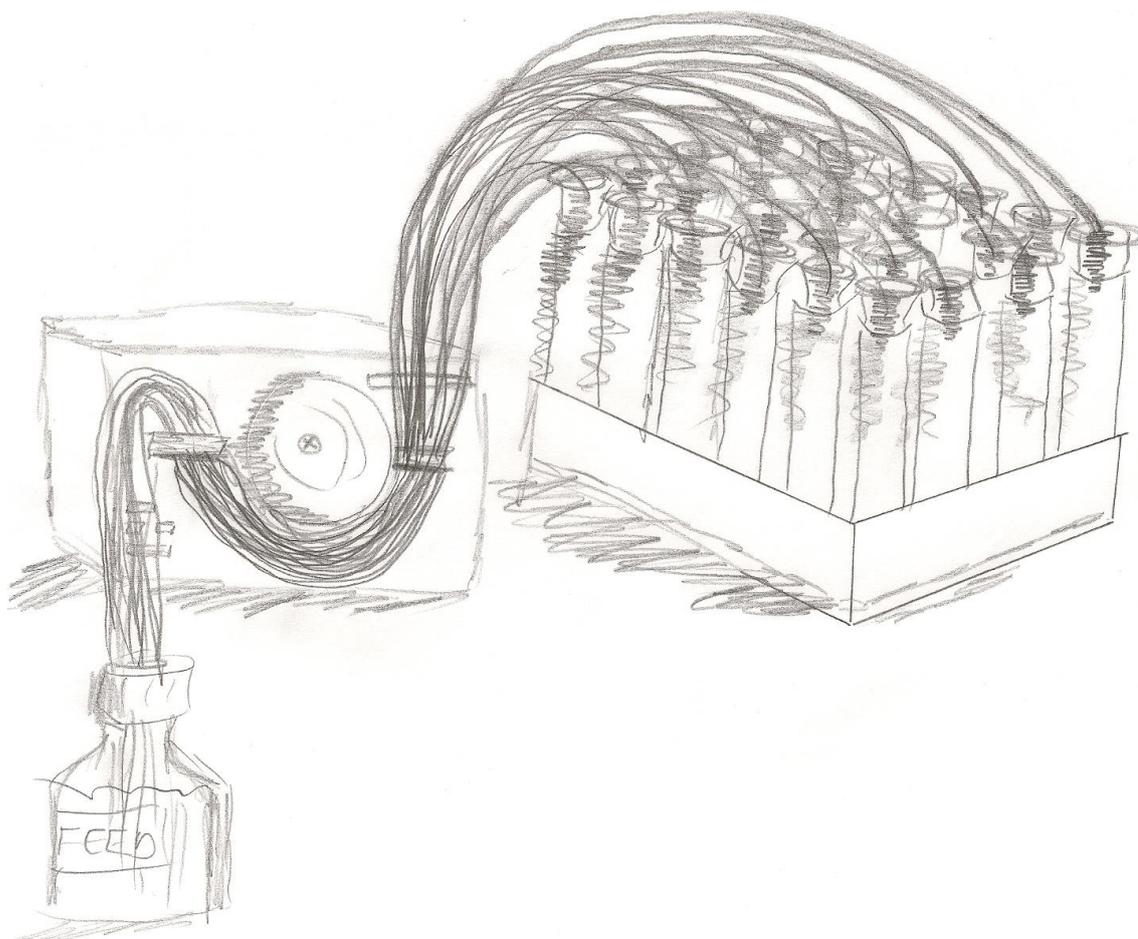


Figure 6.2 Sketch of Micro-24 well plate with modifications. Each well cap would have an inserted micro-tubing for the correct volume. This will be connected to a variable pump speed which would pump the feed into wells without the need to open the wells.

In conclusion, this thesis met the objectives set out for increasing hA_{2A}R yields in two species of yeast, *P. pastoris* and *S. cerevisiae*, by investigating large and small scale cultivations and also examining the use of novel polymers for extracting hA_{2A}R in a stable manner.

7. References

- ABAD, S., NAHALKA, J., BERGLER, G., ARNOLD, S. A., SPEIGHT, R., FOTHERINGHAM, I., NIDETZKY, B. & GLIEDER, A. 2010. Stepwise engineering of a *Pichia pastoris* D-amino acid oxidase whole cell catalyst. *Microb Cell Fact*, 9, 24.
- ADINARAYANA, K., ELLAIAH, P. & PRASAD, D. S. 2003. Purification and partial characterization of thermostable serine alkaline protease from a newly isolated *Bacillus subtilis* PE-11. *AAPS PharmSciTech*, 4, 440-448.
- AHUJA, S. & SMITH, S. O. 2009. Multiple switches in G protein-coupled receptor activation. *Trends Pharmacol Sci*, 30, 494-502.
- AL-MHANNA, N. M. M. 2010. Observation of Crabtree Effect and Diauxic Behaviour of Yeast by Using Absorption. *CHEMICAL ENGINEERING*, 21.
- ALGUEL, Y., LEUNG, J., SINGH, S., RANA, R., CIVIERO, L., ALVES, C. & BYRNE, B. 2010. New tools for membrane protein research. *Current Protein & Peptide Science*, 11, 156.
- ANDRE, N., CHEROUATI, N., PRUAL, C., STEFFAN, T., ZEDER-LUTZ, G., MAGNIN, T., PATTUS, F., MICHEL, H., WAGNER, R. & REINHART, C. 2006. Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen. *Protein Sci*, 15, 1115-26.
- ANTONY, J. 2003. *Design of experiments for engineers and scientists*, Butterworth-Heinemann.
- ARSLAN, G., KULL, B. & FREDHOLM, B. B. 1999. Signaling via A2A adenosine receptor in four PC12 cell clones. *Naunyn-Schmiedeberg's archives of pharmacology*, 359, 28-32.
- ASADA, H., UEMURA, T., YURUGI-KOBAYASHI, T., SHIROISHI, M., SHIMAMURA, T., TSUJIMOTO, H., ITO, K., SUGAWARA, T., NAKANE, T., NOMURA, N., MURATA, T., HAGA, T., IWATA, S. & KOBAYASHI, T. 2011. Evaluation of the *Pichia pastoris* expression system for the production of GPCRs for structural analysis. *Microb Cell Fact*, 10, 24.
- ATTRILL, H., HARDING, P. J., SMITH, E., ROSS, S. & WATTS, A. 2009. Improved yield of a ligand-binding GPCR expressed in *E. coli* for structural studies. *Protein expression and purification*, 64, 32-38.
- ATTWOOD, T. & FINDLAY, J. 1994. Fingerprinting G-protein-coupled receptors. *Protein engineering*, 7, 195-203.
- AULICINO, J., HERMIDA, M., MEDINA, R., SABAROTS, L., DUCREY, L., ORTI, E. & ROGERS, F. 2010. Developing an automatically controlled feeding process in an *E. coli* fermentation process for recombinant protein production.
- BAKER, M. 2010. Making membrane proteins for structures: a trillion tiny tweaks. *Nat Methods*, 7, 429-34.

- BARA-JIMENEZ, W., SHERZAI, A., DIMITROVA, T., FAVIT, A., BIBBIANI, F., GILLESPIE, M., MORRIS, M., MOURADIAN, M. & CHASE, T. 2003. Adenosine A2A receptor antagonist treatment of Parkinson's disease. *Neurology*, 61, 293-296.
- BARRIGÓN, J. M., MONTESINOS, J. L. & VALERO, F. 2013. Searching the best operational strategies for *Rhizopus oryzae* lipase production in *Pichia pastoris* Mut+ phenotype: Methanol Limited or Methanol Non-Limited Fed-Batch cultures? *Biochemical Engineering Journal*.
- BAUMANN, K., CARNICER, M., DRAGOSITS, M., GRAF, A. B., STADLMANN, J., JOUHTEN, P., MAAHEIMO, H., GASSER, B., ALBIOL, J., MATTANOVICH, D. & FERRER, P. 2010. A multi-level study of recombinant *Pichia pastoris* in different oxygen conditions. *BMC Syst Biol*, 4, 141.
- BAWA, Z., BLAND, C. E., BONANDER, N., BORA, N., CARTWRIGHT, S. P., CLARE, M., CONNER, M. T., DARBY, R. A., DILWORTH, M. V., HOLMES, W. J., JAMSHAD, M., ROUTLEDGE, S. J., GROSS, S. R. & BILL, R. M. 2011. Understanding the yeast host cell response to recombinant membrane protein production. *Biochem Soc Trans*, 39, 719-23.
- BAWA, Z. D., RAJ. 2012. Optimising *Pichia pastoris* induction. In: BILL, R. M. (ed.) *Recombinant protein production*. London: Humana Press.
- BAYBURT, T. H. & SLIGAR, S. G. 2003. Self-assembly of single integral membrane proteins into soluble nanoscale phospholipid bilayers. *Protein Sci*, 12, 2476-81.
- BAYBURT, T. H. & SLIGAR, S. G. 2010. Membrane protein assembly into Nanodiscs. *FEBS Lett*, 584, 1721-7.
- BLAXALL, H., MURPHY, T., BAKER, J., RAY, C. & BYLUND, D. 1991. Characterization of the alpha-2C adrenergic receptor subtype in the opossum kidney and in the OK cell line. *Journal of Pharmacology and Experimental Therapeutics*, 259, 323-329.
- BOCKAERT, J. & PHILIPPE PIN, J. 1999. Molecular tinkering of G protein-coupled receptors: an evolutionary success. *The EMBO journal*, 18, 1723-1729.
- BOISON, D. 2005. Adenosine and epilepsy: from therapeutic rationale to new therapeutic strategies. *The neuroscientist*, 11, 25-36.
- BOISON, D. 2008. Adenosine as a neuromodulator in neurological diseases. *Current opinion in pharmacology*, 8, 2-7.
- BONANDER, N. & BILL, R. M. 2012. Optimising yeast as a host for recombinant protein production (review). *Methods Mol Biol*, 866, 1-9.
- BONANDER, N., DARBY, R. A., GRGIC, L., BORA, N., WEN, J., BROGNA, S., POYNER, D. R., O'NEILL, M. A. & BILL, R. M. 2009. Altering the ribosomal subunit ratio in yeast maximizes recombinant protein yield. *Microb Cell Fact*, 8, 10.
- BONANDER, N., HEDFALK, K., LARSSON, C., MOSTAD, P., CHANG, C., GUSTAFSSON, L. & BILL, R. M. 2005. Design of improved membrane protein production experiments: quantitation of the host response. *Protein Sci*, 14, 1729-40.
- BONANDER, N., JAMSHAD, M., OBERTHUR, D., CLARE, M., BARWELL, J., HU, K., FARQUHAR, M. J., STAMATAKI, Z., HARRIS, H. J., DIERKS, K., DAFFORN, T. R., BETZEL, C., MCKEATING, J. A. & BILL, R. M. 2013. Production, purification and characterization of recombinant, full-length human claudin-1. *PLoS One*, 8, e64517.

- BORA, N. 2012. Large-scale production of secreted proteins in *Pichia pastoris*. In: BILL, R. M. (ed.) *Recombinant protein production in yeast*. London: Humana Press.
- BORA, N., BAWA, Z., BILL, R. M. & WILKS, M. D. 2012. The implementation of a design of experiments strategy to increase recombinant protein yields in yeast (review). In: BILL, R. M. (ed.) *Recombinant Protein Production in Yeast*. London: Humana Press.
- BOTSTEIN, D. C., SA AND CHERRY, MJ 1997. Yeast as a Model Organism. *Science*, 277, 1259-1260.
- BRETTTHAUER, R. K. & CASTELLINO, F. J. 1999. Glycosylation of *Pichia pastoris*-derived proteins. *Biotechnology and Applied Biochemistry*, 30, 193-200.
- BURDICK, R. K., BORROR, C. M. & MONTGOMERY, D. C. 2005. *Design and Analysis of Gauge R and R Studies: Making Decisions with Confidence Intervals in Random and Mixed ANOVA Models*, SIAM.
- BURGERS, P. M. & PERCIVAL, K. J. 1987. Transformation of yeast spheroplasts without cell fusion. *Analytical biochemistry*, 163, 391-397.
- ÇALIK, P., BAYRAKTAR, E., İNANKUR, B., SOYASLAN, E. Ş., ŞAHİN, M., TAŞPINAR, H., AÇIK, E., YILMAZ, R. & ÖZDAMAR, T. H. 2010. Influence of pH on recombinant human growth hormone production by *Pichia pastoris*. *Journal of Chemical Technology and Biotechnology*, 85, 1628-1635.
- CARSON, M., JOHNSON, D. H., MCDONALD, H., BROUILLETTE, C. & DELUCAS, L. J. 2007. His-tag impact on structure. *Acta Crystallogr D Biol Crystallogr*, 63, 295-301.
- CELIK, E., CALIK, P. & OLIVER, S. G. 2009. Fed-batch methanol feeding strategy for recombinant protein production by *Pichia pastoris* in the presence of co-substrate sorbitol. *Yeast*, 26, 473-84.
- CELIK, E., CALIK, P. & OLIVER, S. G. 2010. Metabolic flux analysis for recombinant protein production by *Pichia pastoris* using dual carbon sources: Effects of methanol feeding rate. *Biotechnol Bioeng*, 105, 317-29.
- CEREGHINO, J. L. & CREGG, J. M. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS microbiology reviews*, 24, 45-66.
- CHA, J.-H. J. 2000. Transcriptional dysregulation in Huntington's disease. *Trends in neurosciences*, 23, 387-392.
- CHAE, P. S., RASMUSSEN, S. G., RANA, R. R., GOTFRYD, K., CHANDRA, R., GOREN, M. A., KRUSE, A. C., NURVA, S., LOLAND, C. J., PIERRE, Y., DREW, D., POPOT, J. L., PICOT, D., FOX, B. G., GUAN, L., GETHER, U., BYRNE, B., KOBILKA, B. & GELLMAN, S. H. 2010. Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods*, 7, 1003-8.
- CHARALAMBOUS, K., MILLER, D., CURNOW, P. & BOOTH, P. J. 2008. Lipid bilayer composition influences small multidrug transporters. *BMC biochemistry*, 9, 31.
- CHEN, J., SUN, D. & WU, C. 1993. A catalogue of two-level and three-level fractional factorial designs with small runs. *International Statistical Review/Revue Internationale de Statistique*, 131-145.

- CHEN, Z. M., LI, Q., LIU, H. M., YU, N., XIE, T. J., YANG, M. Y., SHEN, P. & CHEN, X. D. 2010. Greater enhancement of *Bacillus subtilis* spore yields in submerged cultures by optimization of medium composition through statistical experimental designs. *Appl Microbiol Biotechnol*, 85, 1353-60.
- CHEREZOV, V., ROSENBAUM, D. M., HANSON, M. A., RASMUSSEN, S. G., THIAN, F. S., KOBILKA, T. S., CHOI, H.-J., KUHN, P., WEIS, W. I. & KOBILKA, B. K. 2007. High-resolution crystal structure of an engineered human β 2-adrenergic G protein-coupled receptor. *science*, 318, 1258-1265.
- CHUN, E., THOMPSON, A. A., LIU, W., ROTH, C. B., GRIFFITH, M. T., KATRITCH, V., KUNKEN, J., XU, F., CHEREZOV, V. & HANSON, M. A. 2012. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure*, 20, 967-976.
- CLARE, J. J., ROMANES, M. A., RAYMENT, F. B., ROWEDDER, J. E., SMITH, M. A., PAYNE, M. M., SREEKRISHNA, K. & HENWOOD, C. A. 1991. Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies. *Gene*, 105, 205-212.
- CONGREVE, M. & MARSHALL, F. 2010. The impact of GPCR structures on pharmacology and structure-based drug design. *British journal of pharmacology*, 159, 986-996.
- COS, O., RAMON, R., MONTESINOS, J. L. & VALERO, F. 2006a. Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: a review. *Microb Cell Fact*, 5, 17.
- COS, O., RAMON, R., MONTESINOS, J. L. & VALERO, F. 2006b. A simple model-based control for *Pichia pastoris* allows a more efficient heterologous protein production bioprocess. *Biotechnol Bioeng*, 95, 145-54.
- COSKUN, U. & SIMONS, K. 2011. Cell membranes: the lipid perspective. *Structure*, 19, 1543-8.
- CREGG, J. B., KJ. HESSLER, AY. MADDEN, KR 1985. *Pichia pastoris* as a Host System for Transformations. *MOLECULAR AND CELLULAR BIOLOGY*, 3376-3385.
- CUNHA, A. E., CLEMENTE, J. J., GOMES, R., PINTO, F., THOMAZ, M., MIRANDA, S., PINTO, R., MOOSMAYER, D., DONNER, P. & CARRONDO, M. J. 2004. Methanol induction optimization for scFv antibody fragment production in *Pichia pastoris*. *Biotechnol Bioeng*, 86, 458-67.
- D'ANJOU, M. C. & DAUGULIS, A. J. 2000. Mixed-feed exponential feeding for fed-batch culture of recombinant methylotrophic yeast. *Biotechnology letters*, 22, 341-346.
- DALY, R. & HEARN, M. T. 2005. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *J Mol Recognit*, 18, 119-38.
- DAMIANOGLU, A., RODGER, A., PRIDMORE, C., R DAFFORN, T., A MOSELY, J., M SANDERSON, J. & R HICKS, M. 2010. The synergistic action of melittin and phospholipase A2 with lipid membranes: development of linear dichroism for membrane-insertion kinetics. *Protein and Peptide Letters*, 17, 1351-1362.
- DARBY, R. A., CARTWRIGHT, S. P., DILWORTH, M. V. & BILL, R. M. 2012. Which yeast species shall I choose? *Saccharomyces cerevisiae* versus *Pichia pastoris* (review). *Methods Mol Biol*, 866, 11-23.

- DARBY, R. A. J., M. GRGIC, L. HOLMES, WJ. BILL, RM. 2010. Production of recombinant G protein-coupled receptor in yeast for structural and functional analysis. *In: POYNER, D. R. W., M. (ed.) G Protein-coupled receptors essential methods*. Wiley-Blackwell.
- DE AGUIAR, P. F., BOURGUIGNON, B., KHOTS, M., MASSART, D. & PHAN-THAN-LUU, R. 1995. D-optimal designs. *Chemometrics and Intelligent Laboratory Systems*, 30, 199-210.
- DE DEKEN, R. 1966. The Crabtree effect and its relation to the petite mutation. *Journal of general Microbiology*, 44, 157-165.
- DE FORESTA, B., LE MAIRE, M., ORLOWSKI, S., CHAMPEIL, P., LUND, S., MOELLER, J. V., MICHELANGELI, F. & LEE, A. G. 1989. Membrane solubilization by detergent: use of brominated phospholipids to evaluate the detergent-induced changes in calcium-ATPase/lipid interaction. *Biochemistry*, 28, 2558-2567.
- DE JONG, L. A. A., GRÜNEWALD, S., PIET FRANKE, J., UGES, D. R. A. & BISCHOFF, R. 2004. Purification and characterization of the recombinant human dopamine D2S receptor from *Pichia pastoris*. *Protein Expression and Purification*, 33, 176-184.
- DE SCHUTTER, K., LIN, Y. C., TIELS, P., VAN HECKE, A., GLINKA, S., WEBER-LEHMANN, J., ROUZE, P., VAN DE PEER, Y. & CALLEWAERT, N. 2009. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat Biotechnol*, 27, 561-6.
- DEBNATH, D. K., BASAIAWMOIT, R. V., NIELSEN, K. L. & OTZEN, D. E. 2011. The role of membrane properties in Mistic folding and dimerisation. *Protein Engineering Design and Selection*, 24, 89-97.
- DEMAIN, A. L. & VAISHNAV, P. 2009. Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv*, 27, 297-306.
- DEREWENDA, Z. S. 2004. Rational protein crystallization by mutational surface engineering. *Structure*, 12, 529-535.
- DI GUANA, C., LIB, P., RIGGSA, P. D. & INOUYEB, H. 1988. Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene*, 67, 21-30.
- DÍAZ-RUIZ, R., AVÉRET, N., ARAIZA, D., PINSON, B., URIBE-CARVAJAL, S., DEVIN, A. & RIGOULET, M. 2008. Mitochondrial Oxidative Phosphorylation Is Regulated by Fructose 1, 6-Bisphosphate A POSSIBLE ROLE IN CRABTREE EFFECT INDUCTION? *Journal of Biological Chemistry*, 283, 26948-26955.
- DORAN, P. M. 1995. *Bioprocess Engineering Principles*, Academic Press.
- DORE, A. S., ROBERTSON, N., ERREY, J. C., NG, I., HOLLENSTEIN, K., TEHAN, B., HURRELL, E., BENNETT, K., CONGREVE, M., MAGNANI, F., TATE, C. G., WEIR, M. & MARSHALL, F. H. 2011. Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine. *Structure*, 19, 1283-93.
- DOUGLASS, A. D. & VALE, R. D. 2005. Single-molecule microscopy reveals plasma membrane microdomains created by protein-protein networks that exclude or trap signaling molecules in T cells. *Cell*, 121, 937-950.

- DRAGOSITS, M., STADLMANN, J., GRAF, A., GASSER, B., MAURER, M., SAUER, M., KREIL, D. P., ALTMANN, F. & MATTANOVICH, D. 2010. The response to unfolded protein is involved in osmotolerance of *Pichia pastoris*. *BMC Genomics*, 11, 207.
- DREW, D., SLOTBOOM, D. J., FRISO, G., REDA, T., GENEVAUX, P., RAPP, M., MEINDL-BEINKER, N. M., LAMBERT, W., LERCH, M. & DALEY, D. O. 2005. A scalable, GFP-based pipeline for membrane protein overexpression screening and purification. *Protein science*, 14, 2011-2017.
- DUJON, B. 2010. Yeast evolutionary genomics. *Nature Reviews Genetics*, 11, 512-524.
- DUKE, E. & JOHNSON, L. 2010. Macromolecular crystallography at synchrotron radiation sources: current status and future developments. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 466, 3421-3452.
- DUPLAY, P., BEDOUELLE, H., FOWLER, A., ZABIN, I., SAURIN, W. & HOFNUNG, M. 1984. Sequences of the malE gene and of its product, the maltose-binding protein of *Escherichia coli* K12. *Journal of Biological Chemistry*, 259, 10606-10613.
- DUQUESNE, K. & STURGIS, J. N. 2010. Membrane protein solubilization. *Heterologous Expression of Membrane Proteins*. Springer.
- FERNDAHL, C., BONANDER, N., LOGEZ, C., WAGNER, R., GUSTAFSSON, L., LARSSON, C., HEDFALK, K., DARBY, R. A. & BILL, R. M. 2010. Increasing cell biomass in *Saccharomyces cerevisiae* increases recombinant protein yield: the use of a respiratory strain as a microbial cell factory. *Microb Cell Fact*, 9, 47.
- FLOWER, D. R. 1999. Modelling G-protein-coupled receptors for drug design. *Biochimica et Biophysica Acta*, 1422, 207-234.
- FOORD, S. M., BONNER, T. I., NEUBIG, R. R., ROSSER, E. M., PIN, J.-P., DAVENPORT, A. P., SPEDDING, M. & HARMAR, A. J. 2005. International Union of Pharmacology. XLVI. G protein-coupled receptor list. *Pharmacological reviews*, 57, 279-288.
- FRASER, N. J. 2006. Expression and functional purification of a glycosylation deficient version of the human adenosine 2a receptor for structural studies. *Protein Expr Purif*, 49, 129-37.
- FREDHOLM, B. B., IJZERMAN, A. P., JACOBSON, K. A., KLOTZ, K.-N. & LINDEN, J. 2001. International Union of Pharmacology. XXV. Nomenclature and classification of adenosine receptors. *Pharmacological reviews*, 53, 527-552.
- FREDRIKSSON, R., LAGERSTRÖM, M. C., LUNDIN, L.-G. & SCHIÖTH, H. B. 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Molecular pharmacology*, 63, 1256-1272.
- FULLER, H. & BISGAARD, S. 1995. A comparison of dispersion effect identification methods for unreplicated two-level factorials. *Center for Quality and Productivity Improvement*.
- GARCÍA-ARRAZOLA, R., SIU, S. C., CHAN, G., BUCHANAN, I., DOYLE, B., TITCHENER-HOOKER, N. & BAGANZ, F. 2005. Evaluation of a pH-stat feeding strategy on the production and recovery of Fab' fragments from *E. coli*. *Biochemical engineering journal*, 23, 221-230.

- GASCH, A. P., SPELLMAN, P. T., KAO, C. M., CARMEL-HAREL, O., EISEN, M. B., STORZ, G., BOTSTEIN, D. & BROWN, P. O. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Science Signaling*, 11, 4241.
- GELLISSEN, G. 2000. Heterologous protein production in methylotrophic yeasts. *Applied microbiology and biotechnology*, 54, 741-750.
- GERNGROSS, T. U. 2004. Advances in the production of human therapeutic proteins in yeasts and filamentous fungi. *Nature biotechnology*, 22, 1409-1414.
- GIETZ, R. D. & WOODS, R. A. 2002. Screening for Protein-Protein Interactions in the Yeast Two-Hybrid System in Embryonic Stem Cells. *Embryonic Stem Cells*. Springer.
- GLOVER, K. J., WHILES, J. A., VOLD, R. R. & MELACINI, G. 2002. Position of residues in transmembrane peptides with respect to the lipid bilayer: a combined lipid NMR and water chemical exchange approach in phospholipid bicelles. *Journal of Biomolecular NMR*, 22, 57-64.
- GOFFEAU, A., BARRELL, B., BUSSEY, H., DAVIS, R., DUJON, B., FELDMANN, H., GALIBERT, F., HOHEISEL, J., JACQ, C. & JOHNSTON, M. 1996. Life with 6000 genes. *Science*, 274, 546-567.
- GRISSHAMMER, R. & HERMANS, E. 2001. Functional coupling with $G\alpha_s$ and $G\alpha_{i1}$ protein subunits promotes high-affinity agonist binding to the neurotensin receptor NTS-1 expressed in *Escherichia coli*. *FEBS letters*, 493, 101-105.
- GROSSMANN, G., OPEKAROVA, M., NOVAKOVA, L., STOLZ, J. & TANNER, W. 2006. Lipid raft-based membrane compartmentation of a plant transport protein expressed in *Saccharomyces cerevisiae*. *Eukaryot Cell*, 5, 945-53.
- GUARNA, M., LESNICKI, G., TAM, B., ROBINSON, J., RADZIMINSKI, C., HASENWINKLE, D., BORASTON, A., JERVIS, E., MACGILLIVRAY, R. & TURNER, R. 1997. On-line monitoring and control of methanol concentration in shake-flask cultures of *Pichia pastoris*. *Biotechnology and bioengineering*, 56, 279-286.
- GUDERMANN, T., NÜRNBERG, B. & SCHULTZ, G. 1995. Receptors and G proteins as primary components of transmembrane signal transduction. *Journal of molecular medicine*, 73, 51-63.
- GUTIÉRREZ-DE-TERÁN, H., MASSINK, A., RODRÍGUEZ, D., LIU, W., HAN, G. W., JOSEPH, J. S., KATRITCH, I., HEITMAN, L. H., XIA, L. & IJZERMAN, A. P. 2013. The Role of a Sodium Ion Binding Site in the Allosteric Modulation of the α_2A Adenosine G Protein-Coupled Receptor. *Structure*, 21, 2175-2185.
- GUZMAN, L.-M., BELIN, D., CARSON, M. J. & BECKWITH, J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of bacteriology*, 177, 4121-4130.
- HAGA, K., KRUSE, A. C., ASADA, H., YURUGI-KOBAYASHI, T., SHIROISHI, M., ZHANG, C., WEIS, W. I., OKADA, T., KOBILKA, B. K. & HAGA, T. 2012. Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature*, 482, 547-551.
- HAM, T. S., LEE, S. K., KEASLING, J. D. & ARKIN, A. P. 2006. A tightly regulated inducible expression system utilizing the fim inversion recombination switch. *Biotechnology and bioengineering*, 94, 1-4.

- HANSON, M. A. & STEVENS, R. C. 2009. Discovery of new GPCR biology: one receptor structure at a time. *Structure*, 17, 8-14.
- HELENIUS, A. & SIMONS, K. 1975. Solubilization of membranes by detergents. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, 415, 29-79.
- HELLWIG, S., EMDE, F., RAVEN, N. P., HENKE, M., VAN DER LOGT, P. & FISCHER, R. 2001. Analysis of single-chain antibody production in *Pichia pastoris* using on-line methanol control in fed-batch and mixed-feed fermentations. *Biotechnology and bioengineering*, 74, 344-352.
- HENRICSSON, C., DE JESUS FERREIRA, M. C., HEDFALK, K., ELBING, K., LARSSON, C., BILL, R. M., NORBECK, J., HOHMANN, S. & GUSTAFSSON, L. 2005. Engineering of a novel *Saccharomyces cerevisiae* wine strain with a respiratory phenotype at high external glucose concentrations. *Appl Environ Microbiol*, 71, 6185-92.
- HERSKOWITZ, I. 1988. Life cycle of the budding yeast *Saccharomyces cerevisiae*. *Microbiological Reviews*, 52, 536.
- HINO, T., ARAKAWA, T., IWANARI, H., YURUGI-KOBAYASHI, T., IKEDA-SUNO, C., NAKADA-NAKURA, Y., KUSANO-ARAI, O., WEYAND, S., SHIMAMURA, T., NOMURA, N., CAMERON, A. D., KOBAYASHI, T., HAMAKUBO, T., IWATA, S. & MURATA, T. 2012. G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody. *Nature*, 482, 237-40.
- HOLMES, W. J., DARBY, R. A., WILKS, M. D., SMITH, R. & BILL, R. M. 2009. Developing a scalable model of recombinant protein yield from *Pichia pastoris*: the influence of culture conditions, biomass and induction regime. *Microb Cell Fact*, 8, 35.
- HUANG, X.-Y., MORIELLI, A. D. & PERALTA, E. G. 1993. Tyrosine kinase-dependent suppression of a potassium channel by the G protein-coupled m1 muscarinic acetylcholine receptor. *Cell*, 75, 1145-1156.
- HULME, E. C. & TREVETHICK, M. A. 2010. Ligand binding assays at equilibrium: validation and interpretation. *Br J Pharmacol*, 161, 1219-37.
- INAN, M. & MEAGHER, M. M. 2001. Non-repressing carbon sources for alcohol oxidase (AOX1) promoter of *Pichia pastoris*. *Journal of bioscience and bioengineering*, 92, 585-589.
- ISAR, J., AGARWAL, L., SARAN, S., KAUSHIK, R. & SAXENA, R. K. 2007. A statistical approach to study the interactive effects of process parameters on succinic acid production from *Bacteroides fragilis*. *Anaerobe*, 13, 50-6.
- JAAKOLA, V. P., GRIFFITH, M. T., HANSON, M. A., CHEREZOV, V., CHIEN, E. Y., LANE, J. R., IJZERMAN, A. P. & STEVENS, R. C. 2008. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science*, 322, 1211-7.
- JACOBSON, K. A., STILES, G. L. & JI, X.-D. 1992. Chemical modification and irreversible inhibition of striatal A2a adenosine receptors. *Molecular pharmacology*, 42, 123-133.
- JAHIC, M., WALLBERG, F., BOLLOK, M., GARCIA, P. & ENFORS, S.-O. 2003. Temperature limited fed-batch technique for control of proteolysis in *Pichia pastoris* bioreactor cultures. *Microbial Cell Factories*, 2, 6.

- JAHIC, M. R.-M., J.MARTINELLE, M. HULT, K. ENFORS S-O. 2002. Modeling of growth and energy metabolism of *Pichia pastoris* producing a fusion protein. *Bioprocess and Biosystems Engineering*, 24, 385-393.
- JAMSHAD, M., LIN, Y. P., KNOWLES, T. J., PARSLow, R. A., HARRIS, C., WHEATLEY, M., POYNER, D. R., BILL, R. M., THOMAS, O. R., OVERDUIN, M. & DAFFORN, T. R. 2011. Surfactant-free purification of membrane proteins with intact native membrane environment. *Biochem Soc Trans*, 39, 813-8.
- JIDENKO, M., NIELSEN, R. C., SØRENSEN, T. L.-M., MØLLER, J. V., LE MAIRE, M., NISSEN, P. & JAXEL, C. 2005. Crystallization of a mammalian membrane protein overexpressed in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 11687-11691.
- JORDÀ, J., JOUHTEN, P., CÁMARA, E., MAAHEIMO, H., ALBIOL, J. & FERRER, P. 2012. Metabolic flux profiling of recombinant protein secreting *Pichia pastoris* growing on glucose: methanol mixtures. *Microb Cell Fact*, 11, 57.
- JUNGO, C., SCHENK, J., PASQUIER, M., MARISON, I. W. & VON STOCKAR, U. 2007. A quantitative analysis of the benefits of mixed feeds of sorbitol and methanol for the production of recombinant avidin with *Pichia pastoris*. *J Biotechnol*, 131, 57-66.
- KALIPATNAPU, S. & CHATTOPADHYAY, A. 2005. Membrane protein solubilization: recent advances and challenges in solubilization of serotonin1A receptors. *IUBMB Life*, 57, 505-12.
- KATAKURA, Y., ZHANG, W., ZHUANG, G., OMASA, T., KISHIMOTO, M., GOTO, Y. & SUGA, K.-I. 1998. Effect of methanol concentration on the production of human β 2-glycoprotein I domain V by a recombinant *Pichia pastoris*: A simple system for the control of methanol concentration using a semiconductor gas sensor. *Journal of fermentation and bioengineering*, 86, 482-487.
- KHATRI, N. K. & HOFFMANN, F. 2006. Impact of methanol concentration on secreted protein production in oxygen-limited cultures of recombinant *Pichia pastoris*. *Biotechnology and bioengineering*, 93, 871-879.
- KIM, S.-J. 2009. Optimization of the Functional Expression of *Coprinus cinereus* Peroxidase in *Pichia pastoris* by Varying the Host and Promoter. *Journal of Microbiology and Biotechnology*, 19, 966-971.
- KITSON, S. M., MULLEN, W., COGDELL, R. J., BILL, R. M. & FRASER, N. J. 2011. GPCR production in a novel yeast strain that makes cholesterol-like sterols. *Methods*, 55, 287-92.
- KNOWLES, T. J., FINKA, R., SMITH, C., LIN, Y.-P., DAFFORN, T. & OVERDUIN, M. 2009. Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *Journal of the American Chemical Society*, 131, 7484-7485.
- KOBAYASHI, K., KUWAE, S., OHYA, T., OHDA, T., OHYAMA, M. & TOMOMITSU, K. 2000. High level secretion of recombinant human serum albumin by fed-batch fermentation of the methylotrophic yeast, *Pichia pastoris*, based on optimal methanol feeding strategy. *Journal of bioscience and bioengineering*, 90, 280-288.
- KOLAKOWSKI JR, L. F. 1994. GCRDb: a G-protein-coupled receptor database. *Receptors & channels*, 2, 1.

- KOZAK, M. 1999. Initiation of translation in prokaryotes and eukaryotes. *Gene*, 234, 187-208.
- KRAGH-HANSEN, U., LE MAIRE, M., NOEL, J. P., GULIK-KRZYWICKI, T. & MOELLER, J. V. 1993. Transitional steps in the solubilization of protein-containing membranes and liposomes by nonionic detergent. *Biochemistry*, 32, 1648-1656.
- KULISEVSKY, J., BARBANOJ, M., GIRONELL, A., ANTONIJOAN, R., CASAS, M. & PASCUAL-SEDANO, B. 2002. A double-blind crossover, placebo-controlled study of the adenosine A2A antagonist theophylline in Parkinson's disease. *Clinical neuropharmacology*, 25, 25-31.
- KUSUMI, A., SAKO, Y. & YAMAMOTO, M. 1993. Confined lateral diffusion of membrane receptors as studied by single particle tracking (nanovid microscopy). Effects of calcium-induced differentiation in cultured epithelial cells. *Biophysical journal*, 65, 2021-2040.
- LADAVIERE, C., TRIBET, C. & CRIBIER, S. 2002. Lateral organization of lipid membranes induced by amphiphilic polymer inclusions. *Langmuir*, 18, 7320-7327.
- LAVALLIE, E. R., DIBLASIO, E. A., KOVACIC, S., GRANT, K. L., SCHENDEL, P. F. & MCCOY, J. M. 1993. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm. *Nature Biotechnology*, 11, 187-193.
- LE MAIRE, M., CHAMPEIL, P. & MØLLER, J. V. 2000. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1508, 86-111.
- LEACH, K. V., C. SEXTON, P.M. CHRISTOPOULOS, A. 2010. Measurement of ligand - G protein-coupled receptor interactions. In: POYNER, D. R. W., M. (ed.) *G Protein-coupled receptors*. Wiley-Blackwell.
- LEBON, G., WARNE, T., EDWARDS, P. C., BENNETT, K., LANGMEAD, C. J., LESLIE, A. G. & TATE, C. G. 2011. Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature*, 474, 521-5.
- LEE, A. G. 2011a. Biological membranes: the importance of molecular detail. *Trends in biochemical sciences*, 36, 493-500.
- LEE, A. G. 2011b. Biological membranes: the importance of molecular detail. *Trends Biochem Sci*, 36, 493-500.
- LENDREM, D., OWEN, M. & GODBERT, S. 2001. DOE (design of experiments) in development chemistry: potential obstacles. *Organic Process Research & Development*, 5, 324-327.
- LI, H., SETHURAMAN, N., STADHEIM, T. A., ZHA, D., PRINZ, B., BALLEW, N., BOBROWICZ, P., CHOI, B.-K., COOK, W. J. & CUKAN, M. 2006. Optimization of humanized IgGs in glycoengineered *Pichia pastoris*. *Nature biotechnology*, 24, 210-215.
- LIN, S.-H. & GUIDOTTI, G. 2009. Purification of membrane proteins. *Methods in enzymology*, 463, 619-629.
- LINDHOLM, L. H., IBSEN, H., DAHLÖF, B., DEVEREUX, R. B., BEEVERS, G., DE FAIRE, U., FYHRQUIST, F., JULIUS, S., KJELDSSEN, S. E. & KRISTIANSSEN, K. 2002. Cardiovascular morbidity and mortality in patients with diabetes in the Losartan Intervention

- For Endpoint reduction in hypertension study (LIFE): a randomised trial against atenolol. *The Lancet*, 359, 1004-1010.
- LINGWOOD, D. & SIMONS, K. 2010. Lipid rafts as a membrane-organizing principle. *science*, 327, 46-50.
- LIU, W., CHUN, E., THOMPSON, A. A., CHUBUKOV, P., XU, F., KATRITCH, V., HAN, G. W., ROTH, C. B., HEITMAN, L. H. & IJZERMAN, A. P. 2012. Structural basis for allosteric regulation of GPCRs by sodium ions. *Science*, 337, 232-236.
- LODOLO, E. J., KOCK, J. L. F., AXCELL, B. C. & BROOKS, M. 2008. The yeast *Saccharomyces cerevisiae*— the main character in beer brewing. *FEMS Yeast Research*, 8, 1018-1036.
- LOMBARDI, A., BURSOMANNO, S., LOPARDO, T., TRAINI, R., COLOMBATTI, M., IPPOLITI, R., FLAVELL, D. J., FLAVELL, S. U., CERIOTTI, A. & FABBRINI, M. S. 2010. *Pichia pastoris* as a host for secretion of toxic saporin chimeras. *The FASEB Journal*, 24, 253-265.
- LONG, S. B., CAMPBELL, E. B. & MACKINNON, R. 2005. Crystal structure of a mammalian voltage-dependent Shaker family K⁺ channel. *Science*, 309, 897-903.
- LONGO, V. D., SHADEL, G. S., KAEBERLEIN, M. & KENNEDY, B. 2012. Replicative and chronological aging in *Saccharomyces cerevisiae*. *Cell Metab*, 16, 18-31.
- LUCAS, L. & UNTERWEGER, M. 2000. Comprehensive review and critical evaluation of the half-life of Tritium. *JOURNAL OF RESEARCH-NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY*, 105, 541-550.
- LUIJK, B., VAN DEN BERGE, M., KERSTJENS, H., POSTMA, D., CASS, L., SABIN, A. & LAMMERS, J. W. 2008. Effect of an inhaled adenosine A_{2A} agonist on the allergen-induced late asthmatic response. *Allergy*, 63, 75-80.
- LUNDSTROM, K., WAGNER, R., REINHART, C., DESMYTER, A., CHEROUATI, N., MAGNIN, T., ZEDER-LUTZ, G., COURTOT, M., PRUAL, C., ANDRE, N., HASSAINE, G., MICHEL, H., CABBILLAU, C. & PATTUS, F. 2006. Structural genomics on membrane proteins: comparison of more than 100 GPCRs in 3 expression systems. *J Struct Funct Genomics*, 7, 77-91.
- LYMAN, E., HIGGS, C., KIM, B., LUPYAN, D., SHELLY, J. C., FARID, R. & VOTH, G. A. 2009. A role for a specific cholesterol interaction in stabilizing the Apo configuration of the human A_{2A} adenosine receptor. *Structure*, 17, 1660-8.
- MACAULEY-PATRICK, S., FAZENDA, M. L., MCNEIL, B. & HARVEY, L. M. 2005. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*, 22, 249-70.
- MAGER, W. H. & SIDERIUS, M. 2002. Novel insights into the osmotic stress response of yeast. *FEMS yeast research*, 2, 251-257.
- MANCIA, F., PATEL, S. D., RAJALA, M. W., SCHERER, P. E., NEMES, A., SCHIEREN, I., HENDRICKSON, W. A. & SHAPIRO, L. 2004. Optimization of protein production in mammalian cells with a coexpressed fluorescent marker. *Structure*, 12, 1355-60.
- MANDENIUS, C. F. & BRUNDIN, A. 2008. Bioprocess optimization using design-of-experiments methodology. *Biotechnology progress*, 24, 1191-1203.

- MATTANOVICH, D., GASSER, B., HOHENBLUM, H. & SAUER, M. 2004. Stress in recombinant protein producing yeasts. *J Biotechnol*, 113, 121-35.
- MATTANOVICH, D., GRAF, A., STADLMANN, J., DRAGOSITS, M., REDL, A., MAURER, M., KLEINHEINZ, M., SAUER, M., ALTMANN, F. & GASSER, B. 2009. Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microb Cell Fact*, 8, 29.
- MATTANOVICH, D. B., P. DATO, L. GASSER, B. SAUER, M. AND PORRO, D. 2012. Recombinant protein production in yeasts. In: LORENCE, A. (ed.) *Recombinant Gene Expression*. Humana Press.
- MIKATA, K. & YAMADA, Y. 1995. *Ogataea kodamae*, a new combination for a methanol-assimilating yeast species, *Pichia kodamae* van der Walt et Yarrow. *Res Commun Inst Ferment*, 17, 99-101.
- MINNING, S., SERRANO, A., FERRER, P., SOLÁ, C., SCHMID, R. D. & VALERO, F. 2001. Optimization of the high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris*. *Journal of biotechnology*, 86, 59-70.
- MONTGOMERY, D. C. 2006. *Design and analysis of experiments*, Wiley. com.
- MONTGOMERY, D. C. & MYERS, R. H. 1995. Response surface methodology: process and product optimization using designed experiments. *Raymond H. Meyers and Douglas C. Montgomery. A Wiley-Interscience Publications*.
- MORAES, I., EVANS, G., SANCHEZ-WEATHERBY, J., NEWSTEAD, S. & STEWART, P. D. S. 2014. Membrane protein structure determination—The next generation. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1838, 78-87.
- MOTULSKY, H. 1995. *The GraphPad guide to analysing radioligand binding data*.
- MURATA, Y., WATANABE, T., SATO, M., MOMOSE, Y., NAKAHARA, T., OKA, S. & IWAHASHI, H. 2003. Dimethyl sulfoxide exposure facilitates phospholipid biosynthesis and cellular membrane proliferation in yeast cells. *J Biol Chem*, 278, 33185-93.
- NING, D., JUNJIAN, X., XUNZHANG, W., WENYIN, C., QING, Z., KUANYUAN, S., GUIRONG, R., XIANGRONG, R., QINGXIN, L. & ZHOUYAO, Y. 2003. Expression, purification, and characterization of humanized anti-HBs Fab fragment. *Journal of biochemistry*, 134, 813-817.
- OATES, J. & WATTS, A. 2011. Uncovering the intimate relationship between lipids, cholesterol and GPCR activation. *Current opinion in structural biology*, 21, 802-807.
- OGATA, K. N., HIDEO. OHSUGI MASAHIRO 1969. A yeast capable of utilizing methanol. *Agricultural and biological chemistry*, 33, 1519-1520.
- OLDHAM, W. M. & HAMM, H. E. 2006. Structural basis of function in heterotrimeric G proteins. *Quarterly reviews of biophysics*, 39, 117-166.
- OTTERSTEDT, K., LARSSON, C., BILL, R. M., STAHLBERG, A., BOLES, E., HOHMANN, S. & GUSTAFSSON, L. 2004. Switching the mode of metabolism in the yeast *Saccharomyces cerevisiae*. *EMBO Rep*, 5, 532-7.

- PAILA, Y. D. & CHATTOPADHYAY, A. 2010. Membrane cholesterol in the function and organization of G-protein coupled receptors. *Subcell Biochem*, 51, 439-66.
- PAL, Y., KHUSHOO, A. & MUKHERJEE, K. 2006. Process optimization of constitutive human granulocyte-macrophage colony-stimulating factor (hGM-CSF) expression in *Pichia pastoris* fed-batch culture. *Applied microbiology and biotechnology*, 69, 650-657.
- PALCZEWSKI, K. 2000. Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science*, 289, 739-745.
- PALMER, T. M., POUCHER, S. M., JACOBSON, K. A. & STILES, G. L. 1995. 125I-4-(2-[7-amino-2-[2-furyl][1, 2, 4] triazolo [2, 3-a][1, 3, 5] triazin-5-yl-amino] ethyl) phenol, a high affinity antagonist radioligand selective for the A2a adenosine receptor. *Molecular pharmacology*, 48, 970-974.
- PAREKH, R., FORRESTER, K. & WITTRUP, D. 1995. Multicopy overexpression of bovine pancreatic trypsin inhibitor saturates the protein folding and secretory capacity of *Saccharomyces cerevisiae*. *Protein expression and purification*, 6, 537-545.
- POPOT, J.-L. 2010. Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annual review of biochemistry*, 79, 737-775.
- POTVIN, G., AHMAD, A. & ZHANG, Z. 2012. Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: A review. *Biochemical Engineering Journal*, 64, 91-105.
- POUCHER, S., KEDDIE, J., SINGH, P., STOGGALL, S., CAULKETT, P., JONES, G. & COLLIS, M. 1995. The in vitro pharmacology of ZM 241385, a potent, non-xanthine, A2a selective adenosine receptor antagonist. *British journal of pharmacology*, 115, 1096-1102.
- PRABHA, L., GOVINDAPPA, N., ADHIKARY, L., MELARKODE, R. & SASTRY, K. 2009. Identification of the dipeptidyl aminopeptidase responsible for N-terminal clipping of recombinant Exendin-4 precursor expressed in *Pichia pastoris*. *Protein expression and purification*, 64, 155-161.
- PRASHER, D. C. 1995. Using GFP to see the light. *Trends in Genetics*, 11, 320-323.
- PRICE, L. A., STRNAD, J., PAUSCH, M. H. & HADCOCK, J. R. 1996. Pharmacological characterization of the rat A2a adenosine receptor functionally coupled to the yeast pheromone response pathway. *Molecular pharmacology*, 50, 829-837.
- PRITCHETT, J. & BALDWIN, S. A. 2004. The effect of nitrogen source on yield and glycosylation of a human cystatin C mutant expressed in *Pichia pastoris*. *Journal of Industrial Microbiology and Biotechnology*, 31, 553-558.
- PRIVE, G. G. 2007. Detergents for the stabilization and crystallization of membrane proteins. *Methods*, 41, 388-97.
- RAO, R. S., KUMAR, C. G., PRAKASHAM, R. S. & HOBBS, P. J. 2008. The Taguchi methodology as a statistical tool for biotechnological applications: a critical appraisal. *Biotechnology journal*, 3, 510-523.
- RASMUSSEN, S. G., CHOI, H.-J., ROSENBAUM, D. M., KOBILKA, T. S., THIAN, F. S., EDWARDS, P. C., BURGHAMMER, M., RATNALA, V. R., SANISHVILI, R. &

- FISCHETTI, R. F. 2007. Crystal structure of the human β_2 adrenergic G-protein-coupled receptor. *Nature*, 450, 383-387.
- REPASI, J. 2013. Small is beautiful – the future of SMEs in pharmaceutical development. *New Generation Pharmaceutical*. New York: GDS Publishing Ltd.
- RESINA, D., COS, O., FERRER, P. & VALERO, F. 2005. Developing high cell density fed-batch cultivation strategies for heterologous protein production in *Pichia pastoris* using the nitrogen source-regulated FLD1 Promoter. *Biotechnol Bioeng*, 91, 760-7.
- RHODES, P. 1985. *An Outline History of Medicine.*, London, Butterworths.
- ROUTLEDGE, S. J. & BILL, R. M. 2012. The effect of antifoam addition on protein production yields. *Methods Mol Biol*, 866, 87-97.
- ROUTLEDGE, S. J., HEWITT, C. J., BORA, N. & BILL, R. M. 2011. Antifoam addition to shake flask cultures of recombinant *Pichia pastoris* increases yield. *Microb Cell Fact*, 10, 17.
- ROUTLEDGE, S. J. C., M. 2012. Setting up a bioreactor for recombinant protein production in yeast. In: BILL, R. M. (ed.) *Recombinant protein production in yeast*. London: Humana Press.
- RYFF, J.-C. 1993. To treat or not to treat?: The judicious use of interferon- α -2a for the treatment of chronic hepatitis B. *Journal of hepatology*, 17, S42-S46.
- SACKMANN, E. 1996. Supported membranes: scientific and practical applications. *Science-AAAS-Weekly Paper Edition*, 271, 43-48.
- SALAZAR, O. & ASENJO, J. A. 2007. Enzymatic lysis of microbial cells. *Biotechnology letters*, 29, 985-994.
- SANCHEZ, M., MISULOVIN, Z., BURKHARDT, A. L., MAHAJAN, S., COSTA, T., FRANKE, R., BOLEN, J. & NUSSENZWEIG, M. 1993. Signal transduction by immunoglobulin is mediated through Ig alpha and Ig beta. *The Journal of experimental medicine*, 178, 1049-1055.
- SANDERS, C. R. & PRESTEGARD, J. 1990. Magnetically orientable phospholipid bilayers containing small amounts of a bile salt analogue, CHAPSO. *Biophysical journal*, 58, 447-460.
- SARDET, C., TARDIEU, A. & LUZZATI, V. 1976. Shape and size of bovine rhodopsin: a small-angle x-ray scattering study of a rhodopsin-detergent complex. *Journal of molecular biology*, 105, 383-407.
- SARRAMEGNA, V., DEMANGE, P., MILON, A. & TALMONT, F. 2002. Optimizing functional versus total expression of the human mu-opioid receptor in *Pichia pastoris*. *Protein Expr Purif*, 24, 212-20.
- SCHAPPER, D., ALAM, M. N., SZITA, N., ELIASSON LANTZ, A. & GERNAEY, K. V. 2009. Application of microbioreactors in fermentation process development: a review. *Anal Bioanal Chem*, 395, 679-95.
- SCHENK, J., BALAZS, K., JUNGO, C., URFER, J., WEGMANN, C., ZOCCHI, A., MARISON, I. W. & VON STOCKAR, U. 2008. Influence of specific growth rate on specific productivity and glycosylation of a recombinant avidin produced by a *Pichia pastoris* Mut+ strain. *Biotechnol Bioeng*, 99, 368-77.

- SCHIOTH, H. B. & FREDRIKSSON, R. 2005. The GRAFS classification system of G-protein coupled receptors in comparative perspective. *Gen Comp Endocrinol*, 142, 94-101.
- SCHMIDT, F. R. 2004. Recombinant expression systems in the pharmaceutical industry. *Appl Microbiol Biotechnol*, 65, 363-72.
- SCHMIDT, F. R. 2005. Optimization and scale up of industrial fermentation processes. *Appl Microbiol Biotechnol*, 68, 425-35.
- SCHMIDT, S. R. & LAUNSBY, R. G. 1989. *Understanding industrial designed experiments*, Air Academy Press.
- SCHNEIDER, J. C. & GUARENTE, L. 1991. [25] Vectors for expression of cloned genes in yeast: Regulation, overproduction, and underproduction. *Methods in enzymology*, 194, 373-388.
- SCHUBERT, J., SIMUTIS, R., DORS, M., HAVLIK, I. & LÜBBERT, A. 1994. Bioprocess optimization and control: Application of hybrid modelling. *Journal of Biotechnology*, 35, 51-68.
- SEDDON, A. M., CURNOW, P. & BOOTH, P. J. 2004. Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta*, 1666, 105-17.
- SEREBRYANY, E., ZHU, G. A. & YAN, E. C. 2012. Artificial membrane-like environments for in vitro studies of purified G-protein coupled receptors. *Biochim Biophys Acta*, 1818, 225-33.
- SERRANO-VEGA, M. J., MAGNANI, F., SHIBATA, Y. & TATE, C. G. 2008. Conformational thermostabilization of the β 1-adrenergic receptor in a detergent-resistant form. *Proceedings of the National Academy of Sciences*, 105, 877-882.
- SERRANO-VEGA, M. J. & TATE, C. G. 2009. Transferability of thermostabilizing mutations between beta-adrenergic receptors. *Mol Membr Biol*, 26, 385-96.
- SHEN, S., SULTER, G., JEFFRIES, T. W. & CREGG, J. M. 1998. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene*, 216, 93-102.
- SHI, F., XU, Z. & CEN, P. 2006. Optimization of γ -polyglutamic acid production by *Bacillus subtilis* ZJU-7 using a surface-response methodology. *Biotechnology and Bioprocess Engineering*, 11, 251-257.
- SHIMAMURA, T., SHIROISHI, M., WEYAND, S., TSUJIMOTO, H., WINTER, G., KATRITCH, V., ABAGYAN, R., CHEREZOV, V., LIU, W., HAN, G. W., KOBAYASHI, T., STEVENS, R. C. & IWATA, S. 2011. Structure of the human histamine H1 receptor complex with doxepin. *Nature*, 475, 65-70.
- SHIVAM, K. & MISHRA, S. 2010. Response surface optimization of the critical medium components for the production of alpha-galactosidase from *Aspergillus parasiticus* MTCC-2796. *Folia biologica*, 56, 78.
- SIDDIQUI, M. A. A. & PERRY, C. M. 2006. Human papillomavirus quadrivalent (types 6, 11, 16, 18) recombinant vaccine (Gardasil®). *Drugs*, 66, 1263-1271.
- SINGH, S., GRAS, A., FIEZ-VANDAL, C., RUPRECHT, J., RANA, R., MARTINEZ, M., STRANGE, P. G., WAGNER, R. & BYRNE, B. 2008. Large-scale functional expression of

WT and truncated human adenosine A2A receptor in *Pichia pastoris* bioreactor cultures. *Microb Cell Fact*, 7, 28.

SINGH, S., HEDLEY, D., KARA, E., GRAS, A., IWATA, S., RUPRECHT, J., STRANGE, P. G. & BYRNE, B. 2010. A purified C-terminally truncated human adenosine A(2A) receptor construct is functionally stable and degradation resistant. *Protein Expr Purif*, 74, 80-7.

SINGH, S., ZHANG, M., BERTHELEME, N., KARA, E., STRANGE, P. G. & BYRNE, B. 2012. Radioligand binding analysis as a tool for quality control of GPCR production for structural characterization: adenosine A(2a)R as a template for study. *Curr Protoc Protein Sci*, Chapter 29, Unit 29 3.

SITKOVSKY, M. V., LUKASHEV, D., APASOV, S., KOJIMA, H., KOSHIBA, M., CALDWELL, C., OHTA, A. & THIEL, M. 2004. Physiological control of immune response and inflammatory tissue damage by hypoxia-inducible factors and adenosine a2a receptors*. *Annu. Rev. Immunol.*, 22, 657-682.

SMITH, D. B. & JOHNSON, K. S. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene*, 67, 31-40.

SPIRA, F., MUELLER, N. S., BECK, G., VON OLSHAUSEN, P., BEIG, J. & WEDLICH-SOLDNER, R. 2012. Patchwork organization of the yeast plasma membrane into numerous coexisting domains. *Nat Cell Biol*, 14, 640-8.

STAGG, J. & SMYTH, M. 2010. Extracellular adenosine triphosphate and adenosine in cancer. *Oncogene*, 29, 5346-5358.

STANBURY, P. F. 1988. Fermentation technology. *Extraction*, 2, 1.1.

STEVENS, R. C., YOKOYAMA, S. & WILSON, I. A. 2001. Global efforts in structural genomics. *Science*, 294, 89-92.

TASHIRO, K., TADA, H., HEILKER, R., SHIROZU, M., NAKANO, T. & HONJO, T. 1993. Signal sequence trap: a cloning strategy for secreted proteins and type I membrane proteins. *Science*, 261, 600-603.

TATE, C. G. 2001. Overexpression of mammalian integral membrane proteins for structural studies. *FEBS Letters*, 504, 94-98.

TATE, C. G. & SCHERTLER, G. F. 2009. Engineering G protein-coupled receptors to facilitate their structure determination. *Current opinion in structural biology*, 19, 386-395.

TAYLOR, M., BUSHARA, H., CAPRON, A., XU, S. & BUTTERWORTH, A. 1994. Laboratory and field evaluation of defined antigen vaccines against *Schistosoma bovis* and *S. japonicum* in animals and of defined antigens for the immunodiagnosis of human *S. japonicum* infection. *Science and technology for development-health-second programme (1987-1991). Summaries of the final reports of the research contracts. Parasitology European Commission DGXII, A*, 189-200.

TERPE, K. 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol*, 60, 523-33.

THOMAS, J. L. & TIRRELL, D. A. 2000. Polymer-induced leakage of cations from dioleoyl phosphatidylcholine and phosphatidylglycerol liposomes. *Journal of controlled release*, 67, 203-209.

- THORPE, E. D., D'ANJOU, M. C. & DAUGULIS, A. J. 1999. Sorbitol as a non-repressing carbon source for fed-batch fermentation of recombinant *Pichia pastoris*. *Biotechnology letters*, 21, 669-672.
- TIERNEY, K. J., BLOCK, D. E. & LONGO, M. L. 2005. Elasticity and phase behavior of DPPC membrane modulated by cholesterol, ergosterol, and ethanol. *Biophysical journal*, 89, 2481-2493.
- TONGE, S. & TIGHE, B. 2001. Responsive hydrophobically associating polymers: a review of structure and properties. *Advanced drug delivery reviews*, 53, 109-122.
- TUCKER, J. & GRISSHAMMER, R. 1996. Purification of a rat neurotensin receptor expressed in *Escherichia coli*. *Biochem. J*, 317, 891-899.
- VINDEVOGEL, J. & SANDRA, P. 1991. Resolution optimization in micellar electrokinetic chromatography: use of Plackett-Burman statistical design for the analysis of testosterone esters. *Analytical Chemistry*, 63, 1530-1536.
- VISSER, W., SCHEFFERS, W. A., BATENBURG-VAN DER VEGTE, W. H. & VAN DIJKEN, J. P. 1990. Oxygen requirements of yeasts. *Applied and environmental microbiology*, 56, 3785-3792.
- VU, T.-K. H., HUNG, D. T., WHEATON, V. I. & COUGHLIN, S. R. 1991. Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell*, 64, 1057-1068.
- WAGNER, S., BADER, M. L., DREW, D. & DE GIER, J. W. 2006. Rationalizing membrane protein overexpression. *Trends Biotechnol*, 24, 364-71.
- WAGNER, S., KLEPSCH, M. M., SCHLEGEL, S., APPEL, A., DRAHEIM, R., TARRY, M., HÖGBOM, M., VAN WIJK, K. J., SLOTBOOM, D. J. & PERSSON, J. O. 2008. Tuning *Escherichia coli* for membrane protein overexpression. *Proceedings of the National Academy of Sciences*, 105, 14371-14376.
- WALKER, G. M. 1998. *Yeast physiology and biotechnology*, Wiley. com.
- WALLIN, E. & HEIJNE, G. V. 1998. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science*, 7, 1029-1038.
- WATERHAM, H. R., DIGAN, M. E., KOUTZ, P. J., LAIR, S. V. & CREGG, J. M. 1997. Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. *Gene*, 186, 37-44.
- WATERS, M. G., EVANS, E. A. & BLOBEL, G. 1988. Prepro-alpha-factor has a cleavable signal sequence. *Journal of Biological Chemistry*, 263, 6209-6214.
- WEGNER, G. H. 1990. Emerging applications of the methylotrophic yeasts. *FEMS Microbiology Letters*, 87, 279-283.
- WEISS, H., HAASE, W., MICHEL, H. & REILANDER, H. 1998. Comparative biochemical and pharmacological characterization of the mouse 5HT_{5A} 5-hydroxytryptamine receptor and the human β ₂-adrenergic receptor produced in the methylotrophic yeast *Pichia pastoris*. *Biochem. J*, 330, 1137-1147.

- WEIß, H. M. & GRISSHAMMER, R. 2002. Purification and characterization of the human adenosine A2a receptor functionally expressed in Escherichia coli. *European Journal of Biochemistry*, 269, 82-92.
- WERNER-WASHBURNE, M., BRAUN, E. L., CRAWFORD, M. E. & PECK, V. M. 1996. Stationary phase in *Saccharomyces cerevisiae*. *Molecular microbiology*, 19, 1159-1166.
- WIENER, M. C. & WHITE, S. H. 1992. Structure of a fluid dioleoylphosphatidylcholine bilayer determined by joint refinement of x-ray and neutron diffraction data. III. Complete structure. *Biophysical journal*, 61, 434-447.
- XU, F., WU, H., KATRITCH, V., HAN, G. W., JACOBSON, K. A., GAO, Z. G., CHEREZOV, V. & STEVENS, R. C. 2011. Structure of an agonist-bound human A2A adenosine receptor. *Science*, 332, 322-7.
- YANG, Z., DAY, Y.-J., TOUFEKTSIAN, M.-C., XU, Y., RAMOS, S. I., MARSHALL, M. A., FRENCH, B. A. & LINDEN, J. 2006. Myocardial Infarct-Sparing Effect of Adenosine A2A Receptor Activation Is due to Its Action on CD4+ T Lymphocytes. *Circulation*, 114, 2056-2064.
- YELISEEV, A. A., WONG, K. K., SOUBIAS, O. & GAWRISCH, K. 2005. Expression of human peripheral cannabinoid receptor for structural studies. *Protein science*, 14, 2638-2653.
- ZHANG, R., KIM, T. K., QIAO, Z. H., CAI, J., PIERCE, W. M., JR. & SONG, Z. H. 2007. Biochemical and mass spectrometric characterization of the human CB2 cannabinoid receptor expressed in *Pichia pastoris*--importance of correct processing of the N-terminus. *Protein Expr Purif*, 55, 225-35.
- ZHANG, W., BEVINS, M. A., PLANTZ, B. A., SMITH, L. A. & MEAGHER, M. M. 2000. Modeling *Pichia pastoris* growth on methanol and optimizing the production of a recombinant protein, the heavy-chain fragment C of botulinum neurotoxin, serotype A. *Papers in Biotechnology*, 18.
- ZHANG, Z., MOO-YOUNG, M. & CHISTI, Y. 1996. Plasmid stability in recombinant *Saccharomyces cerevisiae*. *Biotechnology advances*, 14, 401-435.
- ZINSER, E. & DAUM, G. 1995. Isolation and biochemical characterization of organelles from the yeast, *Saccharomyces cerevisiae*. *Yeast*, 11, 493-536.

8. Appendices

A1. Classification of GPCRs

Since the GPCRs comprise of a ‘super-family’ of membrane proteins, there have been many efforts to classify them into classes and groups on the basis of sequences, mutation data and ligand binding data (Schiøth and Fredriksson, 2005). In spite of this, there still is no international standardised system to classify them. The two most common systems that are currently used to classify GPCRs are the A–F system developed in 1994 by Attwood and Findlay and also Kolakowski (Kolakowski Jr, 1994, Attwood and Findlay, 1994) and the GRAFS (**G**lutamate, **R**hodopsin, **A**dhesion, **F**rizzled/Taste2, **S**ecretin) system developed by Fredriksson and colleagues in 2003 (Fredriksson et al., 2003). The A-F system includes:

- Class A = Rhodopsin-like
- Class B = Secretin receptor family
- Class C = Metabotropic glutamate
- Class D = Fungal mating pheromone receptors
- Class E = Cyclic AMP receptors
- Class F = Frizzled/Smoothened

This system developed by Attwood, Findlay and Kolakowski (Attwood and Findlay, 1994, Kolakowski Jr, 1994), grouped GPCRs into six families for both vertebrates and invertebrates based on >20% amino acid homology within the transmembrane helix domains. However, the shortcomings of this type of classification include no categorisation for mammalian GPCRs such as adhesion and vomeronasal 1 and taste 2 receptors. This is main reason why the GRAFS system is becoming more popular for studying the taxonomy of GPCRs.

The **G**lutamate (metabotropic) receptor family are typically characterised by a “venus fly trap” amino acid N – terminus which is extremely large (~ 600 amino acids in length) and also a very short intracellular loop. The family includes eight metabotropic glutamate receptors, two γ aminobutyric acid (GABA) receptors, a single calcium sensing receptor (CaSR) and taste receptor type 1. There is a recognized disulphide bridge between TM3 and ECL2, and no other defining signatures or motifs for this family.of GPCRs (Schiøth and Fredriksson, 2005).

The **R**hodopsin receptor family are the largest family of GPCRs in vertebrates. The GRAFS system further classifies this family into four smaller groups: α , β , γ and δ . The α group include amine binding GPCRs, several peptide binding and prostaglandin receptors. The β group include receptors that bind peptides as long as the ligands are known. The γ group consists of peptide binding receptors such as chemokine receptors and ones that may bind to neuropeptides such as somatostatins and opioids. The final group, δ , these receptors are olfactory, purine and glycoprotein receptors (Schiøth and Fredriksson, 2005). This family has highly conserved

residues such as NSxxNPxxY motif in TM7, the DRY motif or D(E)-R-Y(F) at the border between TM3 and IL2, (N/S)LxxxD in TM2, the CWXP motif in TM6, GN in TM1 and finally HX in the cytoplasmic helix. There is also the presence of a palmitoylated cysteine in the carboxyl tail (Fredriksson et al., 2003).

The Adhesion family are a small group of GPCRs and in general tend to have long N-termini containing a high amount of serine and threonine residues that can function as O- and N-glycosylation sites. They are thought to form rigid structures on the surface of cells due to the mucin-derived domains present (Schiøth and Fredriksson, 2005).

The Frizzled/Taste2 family include a relatively recent group of GPCRs. Frizzled receptors are thought to be involved in cell fate and proliferation by mediating signals from secreted glycoproteins. They are about 200 amino acids residues in length with a long N-termini and conserved cysteines. The Taste2 receptors are thought to be 'bitter taste' receptors. They share several common features among the human consensus sequences with the frizzled receptors and hence are in the same group despite their quite different downstream physiological function (Schiøth and Fredriksson, 2005).

The Secretin family has structural similarities with the adhesion family, and in the A-F system, they are classed the same. The GRAFS separates them due to major differences in the N-termini. The N-termini are very long, between 60 and 80 amino acids with cysteine bridges (Schiøth and Fredriksson, 2005).

A2. MT-hA_{2A}R DNA and amino acid sequence alignment

BssHII EcoRI
AscI BstBI
 GGCGCGCCGAATTCGAAACGATGAGAATGAAGGTTTTGATCGTTTTGTTGGCTATCTTCG
 1 -----+-----+-----+-----+-----+-----+-----+-----+
 CCGCGCGGCTTAAGCTTTGCTACTCTTACTTCCAAAACCTAGCAAAAACAACCGATAGAAGC
 E T M R M K V L I V L L A I F A
 1 3 5 7 9 11 13 15

CTGCTTTGCCATTGGCTTTGGCTTTGGTTCCAAGAGGTTCTCATCACCATCACCATCATC
 61 -----+-----+-----+-----+-----+-----+-----+-----+
 GACGAAACGGTAACCGAAACCGAAACCAAGGTTCTCCAAGAGTAGTGGTAGTGGTAGTAG
 A L P L A L A L V P R G S H H H H H H H
 17 19 21 23 25 27 29 31 33 35

ACCACCATCACGAGAACTTGTAAGGTTCTGGTCCGGTCTGGATCTGGTTCCA
 121 -----+-----+-----+-----+-----+-----+-----+-----+
 TGGTGGTAGTGTCTTGAACATGAAGGTTCCAAGACCAAGGCAAGACCTAGACCAAGGT
 H H H E N L Y F Q G S G S G S G S G S M
 37 39 41 43 45 47 49 51 53 55

TGCCAATTATGGGTTCTCCGTTTACATCACTGTTGAGTTGGCTATCGCTGTTTTGGCTA
 181 -----+-----+-----+-----+-----+-----+-----+-----+
 ACGGTTAATACCCAAGGAGGCAAATGTAGTGACAACCTCAACCGATAGCGACAAAACCGAT
 P I M G S S V Y I T V E L A I A V L A I
 57 59 61 63 65 67 69 71 73 75

TCTTGGGTAACGTTTTGGTTGTTGGGCTGTTTGGTTGAACTCCAACCTGCAGAACGTTA
 241 -----+-----+-----+-----+-----+-----+-----+-----+
 AGAACCCATTGCAAAACCAACAACCCGACAAACCAACTTGAGGTTGAACGCTCTGCAAT
 L G N V L V C W A V W L N S N L Q N V T
 77 79 81 83 85 87 89 91 93 95

CAAACTACTTCGTTGTTTCCCTTGGCTGCTGCTGACATTGCTGTTGGAGTTTTGGCTATTC
 301 -----+-----+-----+-----+-----+-----+-----+-----+
 GTTTGTGAAGCAACAAAGGAACCGACGACACTGTAACGACAACCTCAAAACCGATAAG
 N Y F V V S L A A A D I A V G V L A I P
 97 99 101 103 105 107 109 111 113 115

CATTGCTATCACTATCTCCACTGGTTTCTGTGCTGCTTGTCACGGTTGTTTGTTCATTG
 361 -----+-----+-----+-----+-----+-----+-----+-----+
 GTAAGCGATAGTGATAGAGGTGACCAAAGACACGACGAACAGTGCCAACAAACAAGTAAC
 F A I T I S T G F C A A C H G C L F I A
 117 119 121 123 125 127 129 131 133 135

CTTGTTTCGTTTTGGTTTTGACTCAGTCCTCTATCTTCTCCTTGTGGCTATGCTATCG
 421 -----+-----+-----+-----+-----+-----+-----+-----+
 GAACAAAGCAAAACCAAAACTGAGTCAGGAGATAGAAGAGGAACAACCGATAACGATAGC
 C F V L V L T Q S S I F S L L A I A I D
 137 139 141 143 145 147 149 151 153 155

ACAGATATATCGCTATCAGAATCCCATGAGATACAACGGTTTGGTTACTGGTACTAGAG
 481 -----+-----+-----+-----+-----+-----+-----+-----+
 TGTCTATATAGCGATAGTCTTAGGGTAACTCTATGTTGCCAAACCAATGACCATGATCTC
 R Y I A I R I P L R Y N G L V T G T R A
 157 159 161 163 165 167 169 171 173 175

CTAAGGTTATTATCGCTATCTGTTGGGTTTTGTCCCTTCGCTATCGGTTTGACTCCAATGT
 541 -----+-----+-----+-----+-----+-----+-----+-----+
 GATTCCCATATAAGCGATAGACAACCCAAAACAGGAAGCGATAGCCAACCTGAGGTTACA
 K G I I A I C W V L S F A I G L T P M L
 177 179 181 183 185 187 189 191 193 195

TGGGTTGGAACAACTGTGGTCAGCCAAAAGAGGTAAGCAACACTCTCAAGGTTGTGGTG
 601 -----+-----+-----+-----+-----+-----+-----+-----+
 ACCCAACCTTGTGACACCAGTCGGTTTTCTCCATTCGTTGTGAGAGTTCCAACACCAC
 G W N N C G Q P K E G K Q H S Q G C G E
 197 199 201 203 205 207 209 211 213 215

661 AGGGTCAGGTTGCTTGTGTTTGTTCGAGGACGTTGTTCCAATGAACTACATGGTTTACTTCA
-----+-----+-----+-----+-----+-----+-----+
TCCCAGTCCAACGAACAAACAAGCTCCTGCAACAAGTTACTTGATGTACCAAATGAAGT
 G Q V A C L F E D V V P M N Y M V Y F N
 217 219 221 223 225 227 229 231 233 235

721 ACTTTTTGCTTGTGTTTGGTTCCTTTGTGTTGATGTTGGGTGTTACTTGAGAATCT
-----+-----+-----+-----+-----+-----+-----+
TGAAAAAGCGAACACAAAACCAAGGAAACAACAACACTACAACCCACAAATGAACTCTTAGA
 F F A C V L V P L L L M L G V Y L R I F
 237 239 241 243 245 247 249 251 253 255

BbsI

781 TCTTGGCTGCTAGAACAGTGAAGCAGATGGAATCCCAGCCATTGCCAGGTGAAAGAG
-----+-----+-----+-----+-----+-----+-----+
AGAACCAGCAGATCTTCTGTCAACTTCGTCTACCTTAGGGTCGGTAACGGTCCACTTTCTC
 L A A R R Q L K Q M E S Q P L P G E R A
 257 259 261 263 265 267 269 271 273 275

841 CTAGATCCACATTGCAGAAAGAGGTTTCACGCTGCTAAGTCCTTGGCTATCATCGTTGGTT
-----+-----+-----+-----+-----+-----+-----+
GATCTAGGTGTAACGTCTTTCTCCAAGTGCACGATTCAGGAACCGATAGTAGCAACCAA
 R S T L Q K E V H A A K S L A I I V G L
 277 279 281 283 285 287 289 291 293 295

901 TGTTGCTTTGTGTTGGTTGCCATTGCACATCATCAACTGTTTTACTTTCTTCTGTCTG
-----+-----+-----+-----+-----+-----+-----+
ACAAGCGAAACACAACCAACGGTAACGTGTAGTAGTTGACAAAATGAAAGAAGACAGGAC
 F A L C W L P L H I I N C F T F F C P D
 297 299 301 303 305 307 309 311 313 315

961 ACTGTTCCCACGCTCCATTGTGGTTGATGTACTTGGCTATCGTTTTGTCCCACACTAACT
-----+-----+-----+-----+-----+-----+-----+
TGACAAGGGTGCAGGTAACCAACTACATGAACCGATAGCAAAAACAGGGTGTGATTGA
 C S H A P L W L M Y L A I V L S H T N S
 317 319 321 323 325 327 329 331 333 335

HincII

1021 CCGTTGTTAACCCATTTCATCTACGCTTACAGAATCAGAGAGTTCAGACAGACTTTTCAGAA
-----+-----+-----+-----+-----+-----+-----+
GGCAACAATTGGGTAAGTAGATGCGAATGTCTTAGTCTCTCAAGTCTGTCTGAAAGTCTT
 V V N P F I Y A Y R I R E F R Q T F R K
 337 339 341 343 345 347 349 351 353 355

1081 AGATCATCAGATCCCACGTTTTGAGACAGCAAGAGCCATTCAAGGCTGACTACAAGGATG
-----+-----+-----+-----+-----+-----+-----+
TCTAGTAGTCTAGGGTGCAAACTCTGTGCTTCTCGGTAAGTTCGGACTGATGTTCCCTAC
 I I R S H V L R Q Q E P F K A D Y K D D
 357 359 361 363 365 367 369 371 373 375

1141 ATGACGACAAGGGTCCGGATCAGGTTCTGGATCAGGATCAGGTCAATTCCGGTGGTGGTA
-----+-----+-----+-----+-----+-----+-----+
TACTGCTGTTCCCAAGGCCTAGTCCAAGACCTAGTCCAGTCCAGTTAAGCCACCACCAT
 D D K G S G S G S G S G S G Q F G G G T
 377 379 381 383 385 387 389 391 393 395

PvuII

PflMI

1201 CTGGTGGTCTCCAGCTCCAGCTGCTGGTGGTCTGGTGGTAAAGCTGGTGAAGGTG
-----+-----+-----+-----+-----+-----+-----+
GACCACCACGAGGTCGAGGTCGACGACCACCACGACCACGACCATTTCGACCCTCCAC
 G G A P A P A A G G A G A G K A G E G E
 397 399 401 403 405 407 409 411 413 415

PflMI *BglII*

1261 AAATTCAGCTCCATTGGCTGGTACTGTTTCCAAGATCTTGGTTAAGGAAGGTGACACTG
-----+-----+-----+-----+-----+-----+
TTAAGGTCGAGGTAACCGACCATGACAAAGGTTCTAGAACCAATTCCTTCCACTGTGAC
I P A P L A G T V S K I L V K E G D T V
417 419 421 423 425 427 429 431 433 435

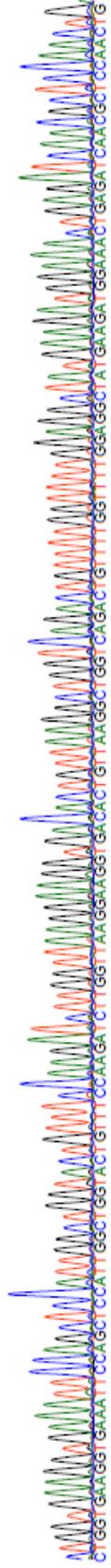
1321 TTAAGGCTGGTCAGACTGTTTTGGTTTTGGAGGCTATGAAGATGGAAACTGAGATCAACG
-----+-----+-----+-----+-----+-----+
AATTCGACAGTCTGACAAAACCAAACCTCCGATACTTCTACCTTTGACTCTAGTTGC
K A G Q T V L V L E A M K M E T E I N A
437 439 441 443 445 447 449 451 453 455

1381 CTCCAAGTACGGTAAGGTTGAGAAGTTTTGGTCAAAGAAAGAGATGCTGTTGAGGGAG
-----+-----+-----+-----+-----+-----+
GAGTTGACTGCCATTCCAACCTTCCAAAACAGTTTCTTTCTCTACGACAAGTCCCTC
P T D G K V E K V L V K E R D A V Q G G
457 459 461 463 465 467 469 471 473 475

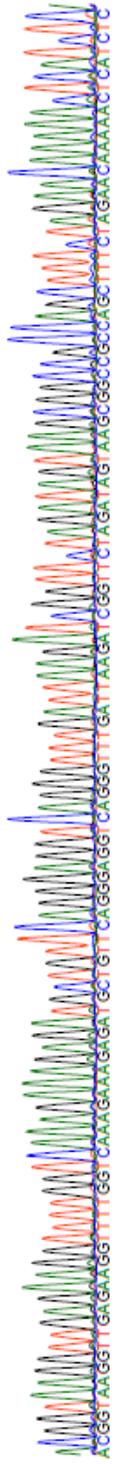
NotI
XbaI *EagI* *PacI*

1441 GTCAGGTTTGATTAAGATCGGTTCTAGATAGTAAGCGCCGCTTAATTAA
-----+-----+-----+-----+-----+-----+
CAGTCCCAAATAATTCTAGCCAAGATCTATCATTCGCCGGCGAATTAATT
Q G L I K I G S R * *
477 479 481 483 485 487

1250 1260 1270 1280 1290 1300 1310 1320 1330 1340 1350 1360 1370 1380
CTGGTGAAGGTTGAAATTCAGGCTCCATGGCTGGTACTGTTTCCAAAGATCTTGGTTAAGGAGGCTGTTAAAGGCTGGTCAACTGTTTGGTTTTGGAGGCTATGAAGATGGAAACTGAGATCAACGGCTCCAACCTG



1390 1400 1410 1420 1430 1440 1450 1460 1470 1480 1490
ACGGTAAGGTTGAGAAAGTTTTGGTCAAAGAAAGAGATGCTGTTCAAGGAGGTCAGGGTTTGAATTAAGATCGGTTCTAGATAGTAAAGCGGCGCCGCAAGCTTTCTAGAACAAAACTCATCTC



A4. Linear regression analysis for d-optimal DoE (Minitab® output data)

General Linear Model: BMS1 binding versus Temp, pH, dO, DMSO

Factor	Type	Levels	Values
Temp	fixed	3	22, 24, 26
pH	fixed	3	5.0, 5.5, 6.0
dO	fixed	3	30, 40, 50
DMSO	fixed	2	Present, Absent

Analysis of Variance for BMS1 binding, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Temp	2	540.2	28.1	14.1	0.14	0.872
pH	2	1477.0	1465.8	732.9	7.15	0.002
dO	2	784.3	677.0	338.5	3.30	0.044
DMSO	1	96.8	67.6	67.6	0.66	0.420
Temp*DMSO	2	17.2	15.1	7.5	0.07	0.929
pH*dO	4	1204.6	1168.9	292.2	2.85	0.032
pH*DMSO	2	88.1	68.8	34.4	0.34	0.716
dO*DMSO	2	69.7	69.7	34.9	0.34	0.713
Error	54	5538.2	5538.2	102.6		
Total	71	9816.0				

S = 10.1271 R-Sq = 43.58% R-Sq(adj) = 25.82%

Term	Coef	SE Coef	T	P
Constant	16.250	1.307	12.44	0.000
Temp				
22	0.933	4.318	0.22	0.830
24	-1.916	4.318	-0.44	0.659
pH				
5.0	-4.412	2.050	-2.15	0.036
5.5	-2.229	2.349	-0.95	0.347
dO				
30	-2.979	2.050	-1.45	0.152
40	4.520	1.853	2.44	0.018
DMSO				
Present	-1.020	1.256	-0.81	0.420
Temp*DMSO				
22 Present	-1.062	2.926	-0.36	0.718
24 Present	0.578	1.864	0.31	0.758
pH*dO				
5.0 30	3.880	4.256	0.91	0.366
5.0 40	-6.466	3.228	-2.00	0.050
5.5 30	1.470	3.007	0.49	0.627
5.5 40	-7.054	3.469	-2.03	0.047
pH*DMSO				
5.0 Present	0.125	2.333	0.05	0.957
5.5 Present	1.383	2.671	0.52	0.607
dO*DMSO				
30 Present	0.018	2.333	0.01	0.994
40 Present	-1.504	1.860	-0.81	0.422

Unusual Observations for BMS1 binding

Obs	binding	Fit	SE Fit	Residual	St Resid
13	65.1400	42.4705	5.4832	22.6695	2.66 R
37	24.2100	42.4705	5.4832	-18.2605	-2.14 R

R denotes an observation with a large standardized residual.

General Linear Model: WT binding versus Temp, pH, dO, DMSO

Factor	Type	Levels	Values
Temp	fixed	3	22, 24, 26
pH	fixed	3	5.0, 5.5, 6.0
dO	fixed	3	30, 40, 50
DMSO	fixed	2	Present, Absent

Analysis of Variance for WT binding, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Temp	2	23706	69002	34501	14.00	0.000
pH	2	35748	66068	33034	13.40	0.000
dO	2	17938	13732	6866	2.79	0.071
DMSO	1	12744	12484	12484	5.07	0.028
Temp*DMSO	2	44983	49351	24675	10.01	0.000
pH*dO	4	23547	31295	7824	3.17	0.020
pH*DMSO	2	59891	70268	35134	14.26	0.000
dO*DMSO	2	18633	18633	9316	3.78	0.029
Error	54	133078	133078	2464		
Total	71	370267				

S = 49.6428 R-Sq = 64.06% R-Sq(adj) = 52.74%

Term	Coef	SE Coef	T	P
Constant	32.550	6.405	5.08	0.000
Temp				
22	93.68	21.17	4.43	0.000
24	-44.12	21.17	-2.08	0.042
pH				
5.0	10.17	10.05	1.01	0.316
5.5	-51.90	11.52	-4.51	0.000
dO				
30	-4.68	10.05	-0.47	0.643
40	-19.007	9.082	-2.09	0.041
DMSO				
Present	-13.855	6.156	-2.25	0.028
Temp*DMSO				
22 Present	-64.17	14.34	-4.47	0.000
24 Present	24.456	9.140	2.68	0.010
pH*dO				
5.0 30	15.02	20.86	0.72	0.475
5.0 40	19.93	15.83	1.26	0.213
5.5 30	-30.54	14.74	-2.07	0.043
5.5 40	-25.20	17.00	-1.48	0.144
pH*DMSO				
5.0 Present	-28.55	11.44	-2.50	0.016
5.5 Present	66.05	13.09	5.04	0.000
dO*DMSO				
30 Present	14.54	11.44	1.27	0.209
40 Present	19.402	9.117	2.13	0.038

Unusual Observations for WT binding

Obs	WT binding	Fit	SE Fit	Residual	St Resid
7	421.300	279.780	26.604	141.520	3.38 R
31	462.960	279.780	26.604	183.180	4.37 R
55	77.510	279.780	26.604	-202.270	-4.83 R

R denotes an observation with a large standardized residual.

General Linear Model: TM6 Binding versus Temp, pH, dO, DMSO

Factor	Type	Levels	Values
Temp	fixed	3	22, 24, 26
pH	fixed	3	5.0, 5.5, 6.0
dO	fixed	3	30, 40, 50
DMSO	fixed	2	Present, Absent

Analysis of Variance for TM6 Binding, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Temp	2	200.51	16.93	8.46	0.23	0.795
pH	2	322.38	93.53	46.77	1.27	0.288
dO	2	357.18	631.05	315.53	8.60	0.001
DMSO	1	55.28	49.67	49.67	1.35	0.250
Temp*DMSO	2	275.99	339.09	169.55	4.62	0.014
pH*dO	4	452.83	341.64	85.41	2.33	0.068
pH*DMSO	2	453.37	172.95	86.47	2.36	0.105
dO*DMSO	2	179.50	179.50	89.75	2.45	0.096
Error	54	1982.26	1982.26	36.71		
Total	71	4279.30				

S = 6.05875 R-Sq = 53.68% R-Sq(adj) = 39.10%

Term	Coef	SE Coef	T	P
Constant	6.9435	0.7817	8.88	0.000
Temp				
22	1.723	2.583	0.67	0.508
24	-1.619	2.583	-0.63	0.533
pH				
5.0	0.140	1.226	0.11	0.910
5.5	1.591	1.405	1.13	0.263
dO				
30	-3.828	1.226	-3.12	0.003
40	-1.941	1.108	-1.75	0.086
DMSO				
Present	0.8740	0.7513	1.16	0.250
Temp*DMSO				
22 Present	-4.643	1.751	-2.65	0.010
24 Present	3.056	1.115	2.74	0.008
pH*dO				
5.0 30	3.476	2.546	1.37	0.178
5.0 40	0.440	1.931	0.23	0.821
5.5 30	-3.897	1.799	-2.17	0.035
5.5 40	-3.477	2.075	-1.68	0.100
pH*DMSO				
5.0 Present	-2.409	1.396	-1.73	0.090
5.5 Present	3.452	1.598	2.16	0.035
dO*DMSO				
30 Present	-1.109	1.396	-0.79	0.431
40 Present	-2.069	1.113	-1.86	0.068

Unusual Observations for TM6 Binding

Obs	TM6 Binding	Fit	SE Fit	Residual	St Resid
6	60.6500	30.6173	3.2470	30.0327	5.87 R

R denotes an observation with a large standardized residual.

A5. Response surface regression analysis for RSM DoE (Minitab® output data)

Response Surface Regression: WT binding versus Temp, pH, dO, DMSO1

The following terms cannot be estimated, and were removed.

DMSO1*DMSO1

The analysis was done using coded units.

Estimated Regression Coefficients for WT binding

Term	Coef	SE Coef	T	P
Constant	-47.039	24.446	-1.924	0.063
Temp	-24.282	11.583	-2.096	0.044
pH	38.203	10.137	3.769	0.001
dO	-9.841	12.207	-0.806	0.426
DMSO1	3.557	9.621	0.370	0.714
Temp*Temp	14.641	17.706	0.827	0.414
pH*pH	57.073	21.012	2.716	0.010
dO*dO	46.972	17.652	2.661	0.012
Temp*pH	-48.646	14.327	-3.395	0.002
Temp*dO	51.063	16.233	3.146	0.003
Temp*DMSO1	21.934	15.234	1.440	0.159
pH*dO	-34.651	13.004	-2.665	0.012
pH*DMSO1	-33.166	11.262	-2.945	0.006
dO*DMSO1	3.757	12.307	0.305	0.762

S = 53.5882 PRESS = 207603
R-Sq = 73.29% R-Sq(pred) = 43.20% R-Sq(adj) = 63.07%

Analysis of Variance for WT binding

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	13	267884	267884	20606.5	7.18	0.000
Linear	4	85983	51639	12909.9	4.50	0.005
Temp	1	19259	12621	12620.6	4.39	0.044
pH	1	35281	40788	40788.3	14.20	0.001
dO	1	12583	1866	1866.4	0.65	0.426
DMSO1	1	18860	393	392.5	0.14	0.714
Square	3	32897	38252	12750.7	4.44	0.010
Temp*Temp	1	7566	1964	1963.6	0.68	0.414
pH*pH	1	17886	21187	21187.5	7.38	0.010
dO*dO	1	7445	20334	20333.6	7.08	0.012
Interaction	6	149004	149004	24834.0	8.65	0.000
Temp*pH	1	32019	33107	33106.9	11.53	0.002
Temp*dO	1	54971	28416	28415.6	9.90	0.003
Temp*DMSO1	1	14899	5953	5953.4	2.07	0.159
pH*dO	1	20916	20390	20390.0	7.10	0.012
pH*DMSO1	1	25933	24904	24904.3	8.67	0.006
dO*DMSO1	1	268	268	267.5	0.09	0.762
Residual Error	34	97638	97638	2871.7		
Lack-of-Fit	8	95599	95599	11949.9	152.43	0.000
Pure Error	26	2038	2038	78.4		
Total	47	365522				

Unusual Observations for WT binding

Obs	StdOrder	WT binding	Fit	SE Fit	Residual	St Resid
7	7	421.300	333.633	32.779	87.667	2.07 R
17	17	2.410	121.894	27.182	-119.484	-2.59 R
31	31	462.960	333.633	32.779	129.327	3.05 R
41	41	10.230	121.894	27.182	-111.664	-2.42 R

R denotes an observation with a large standardized residual.

Estimated Regression Coefficients for WT binding using data in uncoded units

Term	Coef
Constant	4134.36
Temp	-22.4126
pH	-990.099
dO	-61.7212
DMSO1	90.1532
Temp*Temp	3.66037
pH*pH	228.292
dO*dO	0.469716
Temp*pH	-48.6460
Temp*dO	2.55317
Temp*DMSO1	10.9670
pH*dO	-6.93021
pH*DMSO1	-66.3327
dO*DMSO1	0.375654

TEMP*TEMP REMOVED ONLY

Response Surface Regression: WT binding versus Temp, pH, dO, DMSO1

The following terms cannot be estimated, and were removed.

DMSO1*DMSO1

The analysis was done using coded units.

Estimated Regression Coefficients for WT binding

Term	Coef	SE Coef	T	P
Constant	-36.111	20.473	-1.764	0.086
Temp	-23.371	11.478	-2.036	0.049
pH	39.056	10.038	3.891	0.000
dO	-9.605	12.149	-0.791	0.435
DMSO1	3.509	9.577	0.366	0.716
pH*pH	56.942	20.916	2.722	0.010
dO*dO	45.002	17.412	2.585	0.014
Temp*pH	-49.926	14.179	-3.521	0.001
Temp*dO	49.963	16.105	3.102	0.004
Temp*DMSO1	21.661	15.161	1.429	0.162
pH*dO	-36.773	12.690	-2.898	0.006
pH*DMSO1	-34.365	11.118	-3.091	0.004
dO*DMSO1	5.616	12.045	0.466	0.644

S = 53.3456 PRESS = 200786
R-Sq = 72.75% R-Sq(pred) = 45.07% R-Sq(adj) = 63.41%

Analysis of Variance for WT binding

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	12	265920	265920	22160.0	7.79	0.000
Linear	4	85983	52711	13177.7	4.63	0.004
Temp	1	19259	11798	11798.3	4.15	0.049
pH	1	35281	43077	43077.3	15.14	0.000
dO	1	12583	1779	1778.7	0.63	0.435
DMSO1	1	18860	382	382.0	0.13	0.716
Square	2	24790	36288	18144.2	6.38	0.004
pH*pH	1	19287	21091	21091.3	7.41	0.010

dO*dO	1	5503	19010	19010.3	6.68	0.014
Interaction	6	155148	155148	25858.0	9.09	0.000
Temp*pH	1	33235	35284	35283.7	12.40	0.001
Temp*dO	1	50320	27389	27388.6	9.62	0.004
Temp*DMSO1	1	15084	5809	5809.1	2.04	0.162
pH*dO	1	26804	23895	23895.1	8.40	0.006
pH*DMSO1	1	29087	27188	27187.8	9.55	0.004
dO*DMSO1	1	619	619	618.7	0.22	0.644
Residual Error	35	99601	99601	2845.7		
Lack-of-Fit	9	97563	97563	10840.3	138.28	0.000
Pure Error	26	2038	2038	78.4		
Total	47	365522				

Unusual Observations for WT binding

Obs	StdOrder	WT binding	Fit	SE Fit	Residual	St Resid
7	7	421.300	332.662	32.610	88.638	2.10 R
17	17	2.410	115.599	25.977	-113.189	-2.43 R
31	31	462.960	332.662	32.610	130.298	3.09 R
41	41	10.230	115.599	25.977	-105.369	-2.26 R

R denotes an observation with a large standardized residual.

Estimated Regression Coefficients for WT binding using data in uncoded units

Term	Coef
Constant	1653.14
Temp	162.980
pH	-934.920
dO	-56.4675
DMSO1	99.1231
pH*pH	227.767
dO*dO	0.450021
Temp*pH	-49.9259
Temp*dO	2.49817
Temp*DMSO1	10.8307
pH*dO	-7.35469
pH*DMSO1	-68.7304
dO*DMSO1	0.561647

A6. Patent application for PMAS work showing claims (application number GB 1312343.5, authors, Bawa Z, Bill R.M., Campbell D, Mahomed A and Tighe B.J.)



Page removed for copyright restrictions.

**A7. Oral presentation at Aston University Post-graduate research day (June 2012).
Awarded 1st prize.**



Improving membrane protein production in yeast -

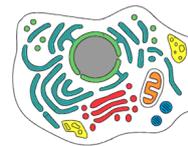
the effect of induction phase growth rates on the
functional yield of adenosine 2a receptor produced in
P. pastoris

by

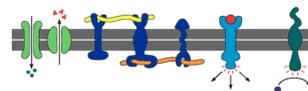
Zharain Bawa

Membrane proteins play a disproportionately large role in health and disease

They play pivotal roles in all cellular processes
They mediate a wide variety of processes
e.g. cell-cell interactions



They are ubiquitous
~30% of all proteins in all cells are membrane proteins



They are already key drug targets
The majority of prescription pharmaceuticals on the market
are targeted to membrane proteins, particularly
G protein-coupled receptors (GPCRs)



Human adenosine 2a receptor (hA_{2a}R)

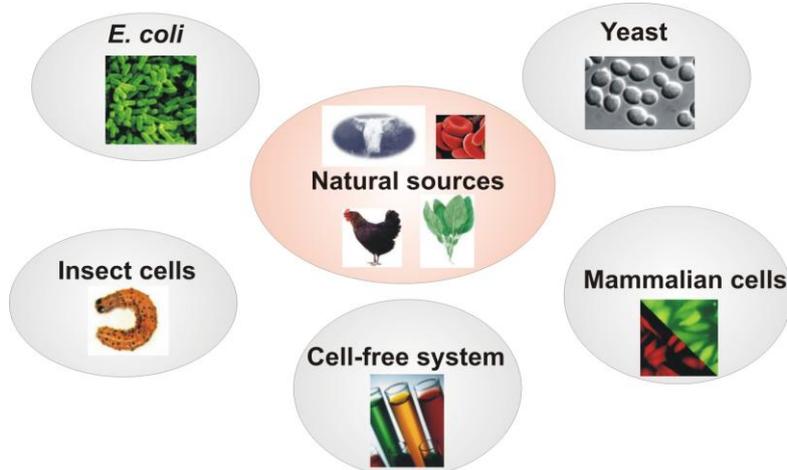


Membrane proteins: Drug discovery challenges

- ▶ To enable further drug discovery in a rational manner
 - ▶ Structure
 - ▶ Function
 - ▶ Biochemistry
- ▶ This is proving to be challenging because:
 - ▶ Membrane proteins rely on the membrane to retain their structure and function
 - ▶ Low abundance

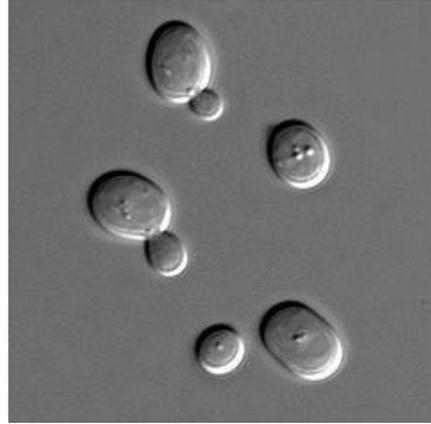


Hosts for membrane protein production



Yeast – *Pichia pastoris*

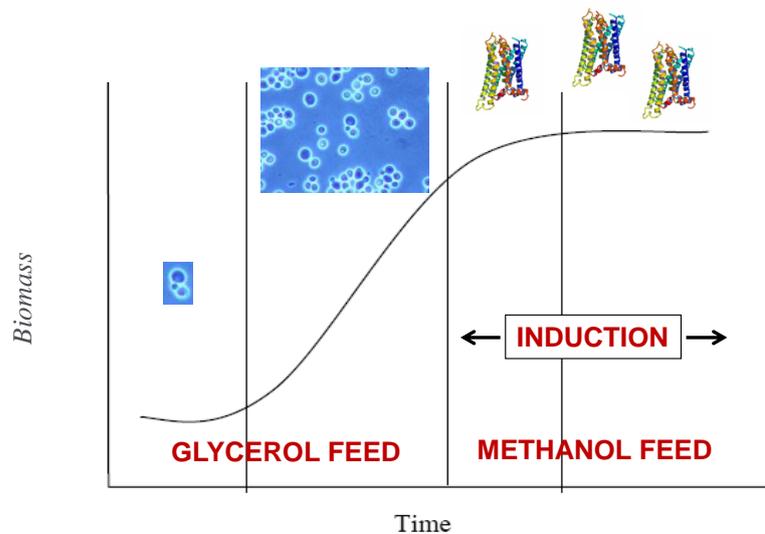
- ▶ Eukaryote: similar to mammalian cells
- ▶ Rapid cell growth, high yields and cheap
- ▶ Grow on methanol
- ▶ Protein structures resolved which have contributed to the drug discovery pipeline



The Biochemist 2010 32:5 pg15

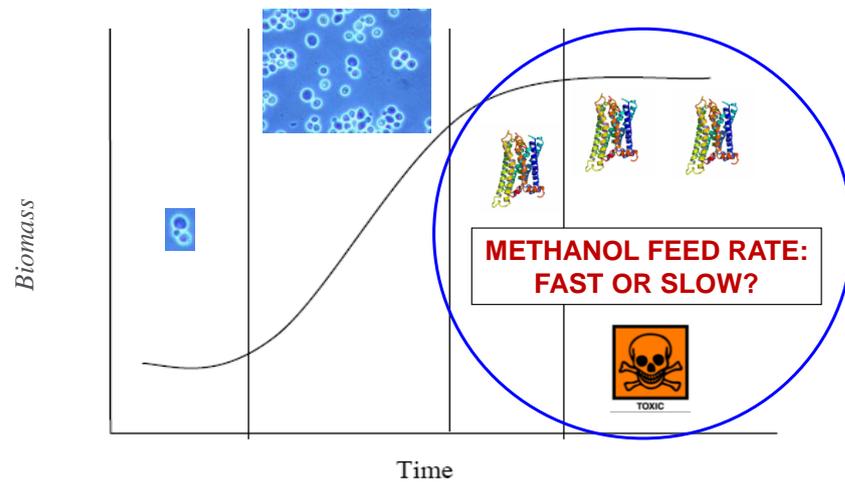
Pichia pastoris and making GPCRs

- ▶ Genetically engineered to over-produce the protein we want by simply changing the type of feed (carbon source) → induction



Aim

- ▶ To explore the influence of the methanol feeding rate on the production of hA_{2a}R in *P. pastoris*

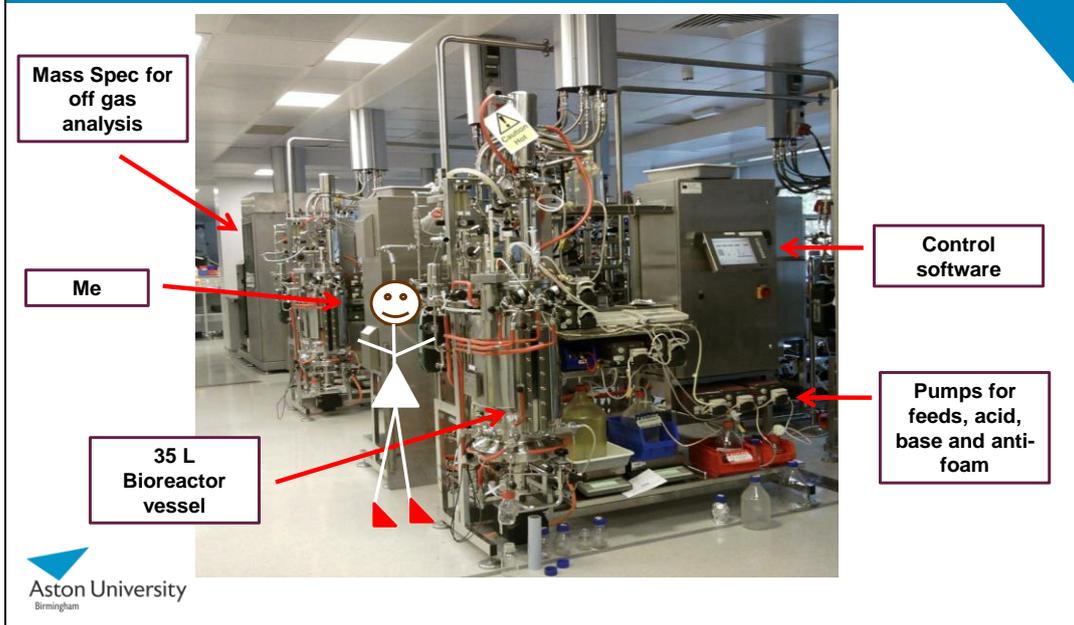


How? By using Bioreactors

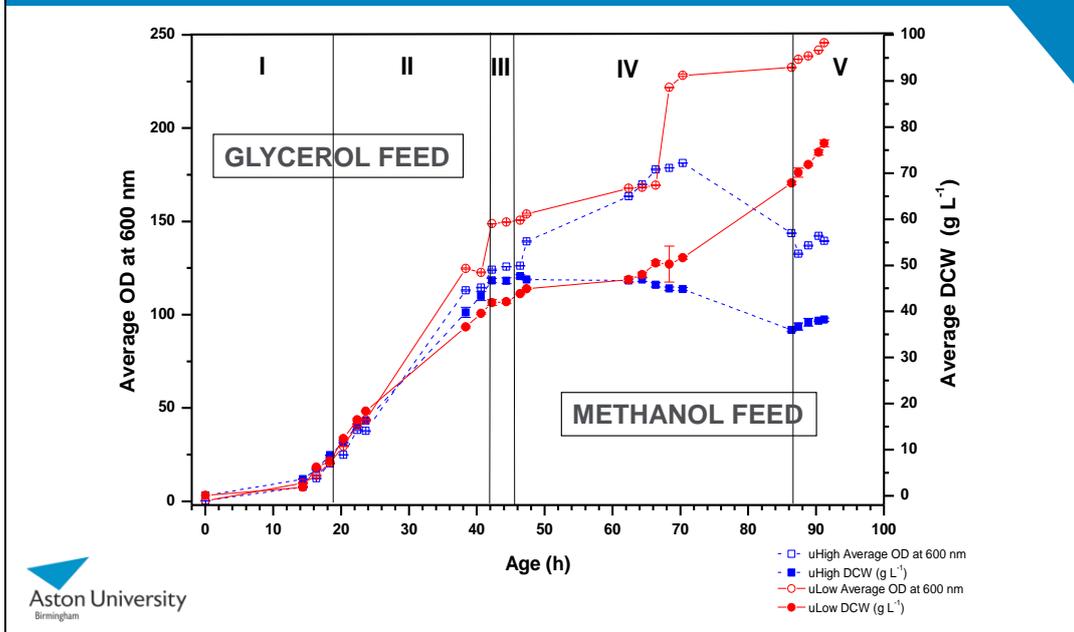
- ▶ Bioreactors used to control process conditions such as temperature, pH, aeration and addition of carbon source – not possible in shake flasks



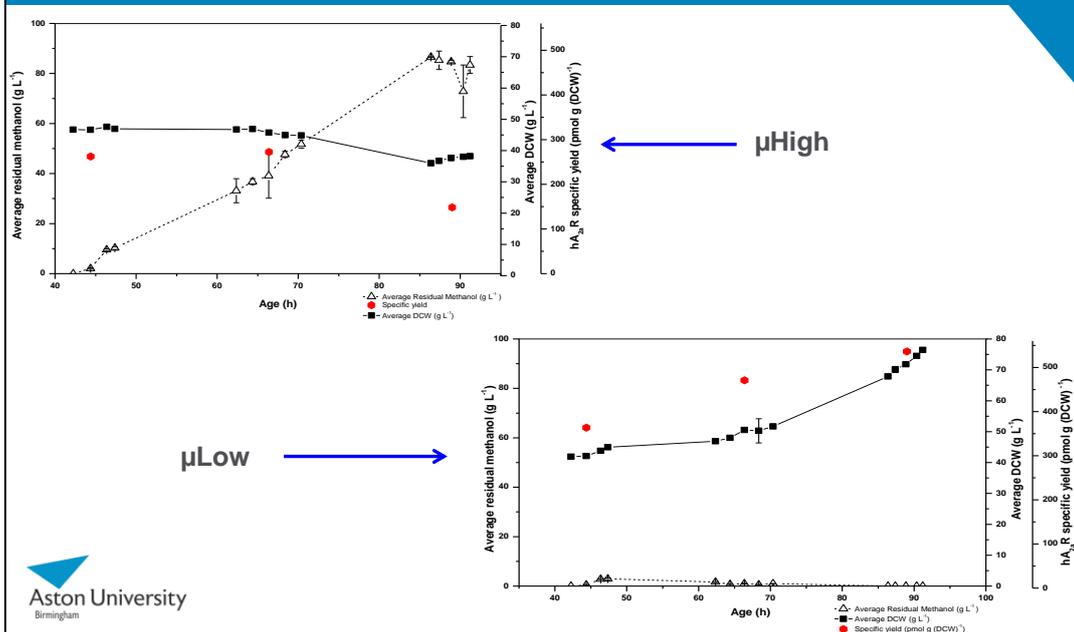
35 L Bioreactor at AstraZeneca Facility



Cell biomass increases with lower methanol feed rate (μ Low). Final biomass of μ Low is double the final biomass of μ High



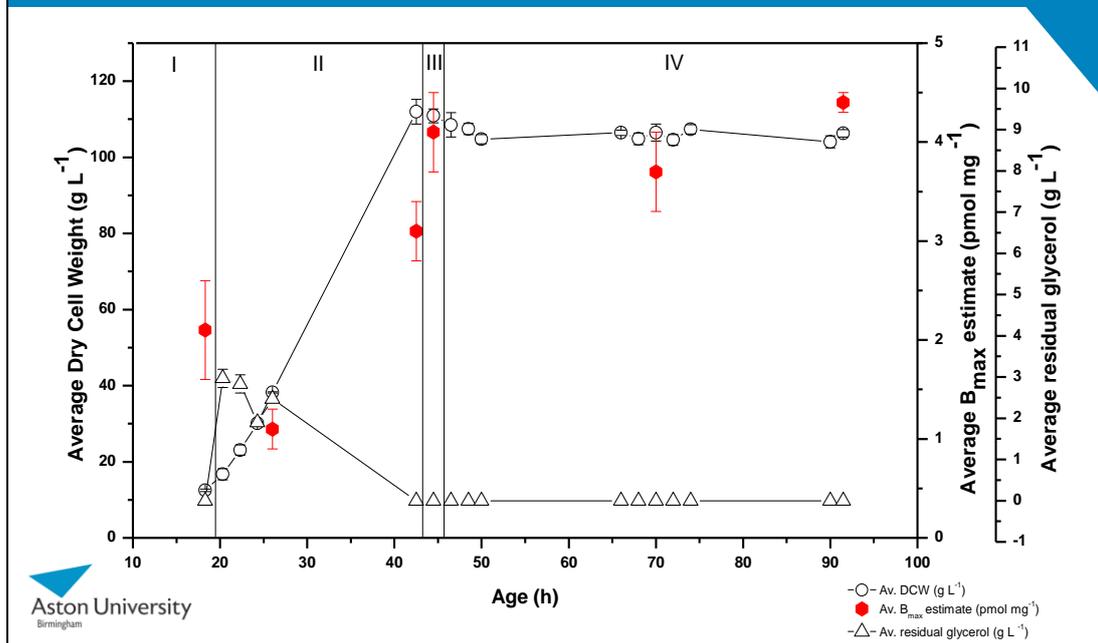
A low methanol induction feed rate has an overall positive effect on cell biomass and hA_{2a}R production



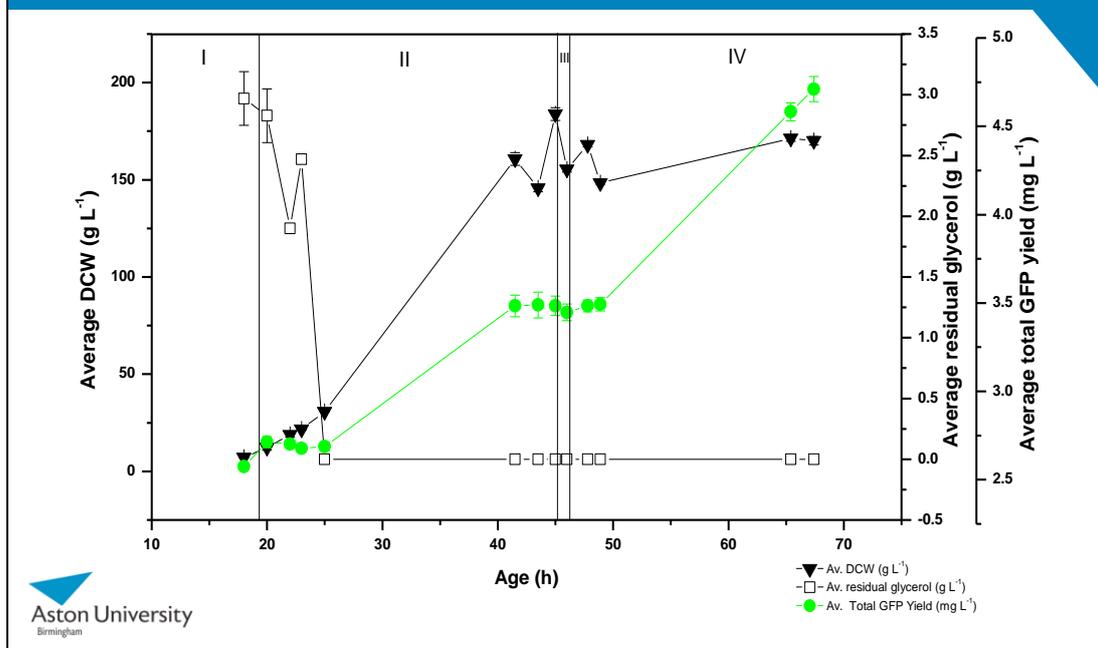
The presence of glycerol does not repress production of hA_{2a}R

Cultivation phase	Age (h)	B _{max} estimate (pmol mg ⁻¹)		Residual glycerol (g L ⁻¹)	
		μ_{low}	μ_{high}	μ_{low}	μ_{high}
IIIA	42.2	4.4 (0.1)	4.4 (0.2)	0 (0)	2.9 (0.1)
IIIB	44.4	7.1 (0.6) **	4.0 (0.0)	1.1 (0.2)	4.2 (0.2)
IV	66.4	4.5 (0.2) *	3.8 (0.1)	0 (0)	1.1 (0.2)
V	88.9	5.0 (0.2) **	3.1 (0.1)	0 (0)	0 (0)

Baseline hA_{2a}R production is present during glycerol phase in another cultivation at Aston



This is also true for another protein, Green Fluorescent Protein (GFP)



Discussion Part 1

- ▶ μ_{Low} cultivation gave overall high biomass and high $hA_{2a}R$ yields
- ▶ μ_{Low} showed low amounts of residual methanol in cultivation ($< 3 \text{ g L}^{-1}$) from feeding the cells slowly
- ▶ μ_{High} showed increasing amounts of residual methanol ($\sim 80 \text{ g L}^{-1}$) from feeding the cells quickly

Discussion Part 2

- ▶ In literature, soluble proteins tend to increase during low feeding rates and biomass only increases during high feeding rates
- ▶ Different to what we found: low feeding rate gave high $hA_{2a}R$ AND biomass
- ▶ Suggests other factors influence production of membrane proteins
- ▶ Maybe more important to increase biomass? More cells = more membranes = more membrane proteins to be inserted

Discussion Part 3

- ▶ hA_{2a}R production was detected when glycerol present in cultivation
- ▶ Unexpected as glycerol represses protein production
- ▶ This was the same for GFP as well
- ▶ Unusual and potentially controversial finding

Conclusion

- ▶ Applying a slow methanol feed strategy to yeast (*P. pastoris*) is critical for maximum GPCR yields

Acknowledgements

Aston University

Roslyn Bill

**Matthew Conner
Charlotte Bland
Sarah Routledge
Stephanie Cartwright
Marvin Dilworth
Michelle Clare**

**David Poyner
James Barwell**

Lab 347



AstraZeneca Ltd.

Markus Ganzlin



A8. Poster presentations

Biochemical Society, Cambridge, United Kingdom, January 2011

Recombinant Protein Production, Vienna, Austria, February, 2011

British Pharmacological Society, Leicester, United Kingdom, April, 2012

A9. Publications arising from this thesis



Page removed for copyright restrictions.