Recombinant Protein Production in *Escherichia coli*: Optimisation of Improved Protocols

By

Christopher Wyre

A thesis submitted to the

University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemical Engineering

School of Chemical Engineering

College of Engineering and Physical Sciences

January 2014

Abstract

Recombinant protein production (RPP) in *E. coli* is a cornerstone of modern bioprocessing, especially for biopharmaceutical production. This study presents work aimed at developing analytical techniques and production protocols for RPP in *E. coli*, particularly with reference to industrial applications.

Flow cytometry (FCM) was used routinely to monitor cell physiology and RPP itself through the production of CheY::GFP a fluorescent model protein. Further applications of FCM for monitoring RPP processes were also developed. A protocol was developed to identify CheY::GFP inclusion bodies produced under high-stress conditions using the amyloidophilic fluorescent dye Congo red. FCM analysis was also applied to RPP process stages that are currently little studied, agar plate and early-stage liquid cultures. FCM screening of transformants expressing CheY::GFP on agar plates identified abnormal, likely mutated cells. Further analysis identified 3 populations of varying fluorescence intensity and the progressive transfer of cells from the high fluorescence population to one of intermediate fluorescence along with a progressive loss of culturability. This was shown to be the result of amyloid inclusion body formation by Congo red staining. FCM analysis of agar-plate and early-stage liquid culture is therefore proposed as a useful (and currently under–exploited) analytical step for RPP processes.

RPP conditions that minimise physiological stress by reducing growth temperature and inducer concentration can increase product yields, solubility and biomass yields. In this study stress-minimised production conditions were applied to industrially-derived RPP fermentation protocols. The original complex-medium based fermentation protocol used for stress-minimised RPP showed substantial limitations for industrial use. Application of stress-minimised conditions to a semi-defined medium-based protocol using both early and late-phase induction and glucose or glycerol as carbon source was shown to be successful with high yields of biomass, total CheY::GFP and soluble CheY::GFP (up to 77.5 $g \cdot L^{-1}$, 15.7 $g \cdot L^{-1}$ and 7.1 $g \cdot L^{-1}$, respectively). The protocols developed in this study improved biomass generation, product formation and reproducibility over both the original stress-minimised and unmodified industrially-derived protocols. It is therefore concluded that stress-minimisation is of potential industrial use.

Adaptation of a defined growth medium fermentation for stress-minimisation showed limited success due to nitrogen and phosphate limitation in the initial growth medium. Further analysis of nitrogen and phosphorus content showed that the complex medium was likely phosphate limited and that the semi-defined medium was likely nitrogen limited. It was concluded that with further adaptation a stress-minimised defined-medium protocol will be possible.

This work is dedicated to my grandparents.

I still can't believe you won't see this

Acknowledgments

First I would like to thank Dr Tim Overton for his excellent supervision, support and guidance throughout this project and for allowing me this opportunity for which I am very grateful. Thanks also to Prof Colin Thomas for his valuable comments on this thesis.

For their help and support both in the lab and out my thanks go to Isaac Vizcaino-Caston, Amir Anvarian and Louise Hackett.

Also I would like to thank all my colleagues from Biochemical Engineering especially, Reza Jalalirad, Evan Hsu, Neeraj Jumbu and Trong Duangkanok.

For their excellent technical assistance I would like to thank Elaine Mitchell, David French, Ronnie Baglin, Christine Fletcher and the late Hazel Jennings.

From the School of Biosciences I would like to thank Prof Jeff Cole for his insights and assistance, and to members of his research group past and present for their help and the provision of strains, plasmids etc., especially Yanina Sevastsyanovich, Sara Alfasi and Lenny Zaffaroni.

Financially I would like to acknowledge the support of the BBSRC who funded this project and to the SGM for kindly providing me with a student travel grant.

Finally, I must thank my family without whose constant support I could not have done this.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

A-β	Amyloid-β peptide
AFR	Aerobic fluorescence recovery
APS	Ammonium persulphate
*av*GFP	*Aequorea victoria* GFP
BCA	Bicinchoninic acid
BFP	Blue fluorescent protein
BHK	Baby hamster kidney
BOX	Bis-oxonol
cAMP	Cyclic adenosine monophosphate
CAT	Chloramphenicol acetyltransferase
CCP	Cytochrome *c* peroxidase
CDC	Carbon dioxide evolution
CFU	Colony forming units
cGMP	Current good manufacturing practice
CHO	Chinese hamster ovary
CR	Congo red
CRP	cAMP receptor protein
CSTR	Constant stirred tank reactor
CV	Coefficient of variance
CW	Cell wall
ddH$_2$O	Double-distilled H$_2$O
DCW	Dry cell weight
DMSO	Dimethylsulphoxide
DoE	Design of experiments
DOT	Dissolved oxygen tension
DSB	Disulphide bond
DSP	Downstream processing
EGFP	Enhanced GFP
ER	Endoplasmic reticulum
ERAD	ER-associated protein degradation
Fab	Fragment antigen-binding
FACS	Fluorescence-activated cell sorting
FALS	Forward angle light scatter
FCM	Flow cytometry
FDA	United States Food & Drugs Administration
FITC	Fluorescein isothiocyanate
FM	Fluorescence microscopy
FMN	Flavin mononucleotide
FMN-BP	Flavin mononucleotide binding protein
FP	Fluorescent protein
FRET	Förster resonance energy transfer
FSC	Forward scatter
GDH	Glutamate dehydrogenase
GFP	Green fluorescent protein
GOGAT	Glutamine oxoglutarate aminotransferase
GRAS	Generally regarded as safe
GST	Glutathione-s-transferase
HCDC	High cell density culture

HSP	Heat shock protein
IB	Inclusion body
IM	Inner membrane
imGFP	Immature GFP
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lennox broth
LPS	Lipopolysaccharide
MBP	Maltose binding protein
MTB	Modified terrific broth
MS	Mass spectrometry
NA	Nutrient agar
OD$_{xxx\ nm}$	Optical density at XXX nm
OM	Outer membrane
ORT	Operator/repressor titration
OXC	Oxygen consumed
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
PFCF	Plasmid-free cell fraction
PI	Propidium Iodide
PMMA	Poly(methylmethacrylate)
PMT	Photomultiplier tube
PP	Polypropylene
PPG	Polypropylene glycol
ppGpp	5-3-guanosine pyrophosphate
PPM	Parts per million
PS	Polystyrene
PTM	Post-translational modification
PWA	Pulse-width analysis
RALS	Right angle light scatter
ROS	Reactive oxygen species
RP	Recombinant protein
RPP	Recombinant protein production
RPM	Revolutions per minute
RQ	Respiratory quotient
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SSC	Side scatter
TA	Toxin/antitoxin
Tat	Twin arginine transportation
TCA	Tricarboxylic acid
TEMED	Tetramethylethylenediamine
TGS	Tris, glycine, SDS
Th-S	Thioflavin-S
Th-T	Thioflavin-T
Tris	Tris(hydroxymethyl)aminomethane
UPR	Unfolded protein response
VBNC	Viable but non-culturable cells
VVM	Volume per volume per minute
wtGFP	Wild-type GFP

Chapter 1: Introduction & Literature Review

This chapter contains a review of relevant scientific literature, covering the following areas: a general summary of recombinant protein production (RPP) and its application in *Escherichia coli*, the physiological stresses RPP imposes and the methods to ameliorate them, industrially-relevant methods of *E. coli* culture and the uses of flow cytometry (FCM) and of autofluorescent proteins in microbial biotechnology. Finally, the aims of this project are described.

1.1. Uses & importance of RPP

The earliest reported instances of the production of a protein from a recombinant gene were from the late 1970s (Itakura *et al.*, 1977; Goeddel *et al.*, 1979), leading on from recent advances in recombinant DNA technology. In both cases small medically-relevant polypeptide hormones (somatostatin and insulin) were produced in *E. coli* as plasmid-encoded fusions to β-galactosidase. Since then RPP has become a highly useful aspect of modern biotechnology, finding applications in medicine, industry and research, being worth billions of pounds globally every year.

1.1.1. Pharmaceutical, biotechnological & research uses

One of the best-known uses of RPP is the production of protein-based pharmaceuticals, such as insulin, human growth factor and antibody-based drugs like Herceptin (Walsh, 2010a). Industrial uses include recombinant enzymes for chemical and food processing such as bovine chymosin (rennet), a milk-curdling enzyme used in cheese production that has been synthesised in a number of microbes including *E. coli* and is approved for human consumption by the United States Food and Drug Administration (FDA) (Flamm, 1991; Olempska-Beer *et al.*, 2006). RPP is also of great importance for modern biological research in the production of commonly used laboratory enzymes such as the restriction

endonucleases and thermostable DNA polymerases vital for genetic manipulation; the majority of restriction endonucleases available from New England Biolabs are recombinant (New England Biolabs, 2013/2014). Smaller-scale RPP is also directly used in research to synthesise proteins of interest for structural or functional analysis, recent examples include: Rouyi *et al.* (2014), Świeżawska *et al.* (2014), Overman *et al.* (2014) and Um *et al.* (2013).

1.1.2. Systems, uses & limitations.

Numerous expression systems exist for RPP, from microbes such as bacteria and yeast, to higher eukaryotes such as filamentous fungi, plant, insect and mammalian cells, and even cell-free systems. The wide range of expression systems available is necessitated by the range of proteins that have been targeted for production. The properties of the recombinant protein (RP) product and its intended use will affect the choice of host system; as such, there is no single 'ideal' system. Below follows a summary of common expression systems with example products from each in Table 1.1 and a summary of the advantages and disadvantages of each in Table 1.2.

1.1.2.1. Bacterial systems

In the interests of efficiency, it is advantageous to utilise the simplest system possible for the product in question. In this regard bacteria hold many advantages: they are simple organisms and can grow rapidly to high cell densities, with high product yields in relatively simple bioreactors using inexpensive growth media. Microbial hosts can therefore be advantageous for RPP processes.

Bacteria do however possess certain limitations, primarily involving their ability to produce more complex, eukaryotic proteins. First, eukaryotic proteins are often subject to significantly higher levels of post-translational modification (PTM) than in

Table 1.1: Selected commercialised and industrially-relevant RPP products from common expression systems

Expression system			Proteins/Products		EU/FDA Approved	References
Microbial systems	Bacteria	*Escherichia coli*	Thrombolytics	tPA (Ecokinase etc.)	Yes	Schmidt, 2004; Walsh, 2010a; Flamm, 1991
			Hormones	Insulin (Humulin etc.)	Yes	
				HGH (Protropin etc.)	Yes	
			Interferons	Betaferon etc.	Yes	
			Vaccines	*Borellia burgdorferi* (Lymarix)	Yes	
			Antibodies (Fab)	Cimzia	Yes	
			Enzymes	Chymosin	Yes	
		Bacillus subtilis	Hormones	Proinsulin	No	Westers *et al.*, 2004; Terpe, 2006
			Enzymes	Lipase A	No	
			Other	Streptavidin	No	
		Bacillus megaterium	Enzymes	Dextransucrase	No	Terpe, 2006
			Other	Toxin A (*Clostridium difficile)*	No	
	Yeast	*Saccharomyces cerevisiae*	Anticoagulants	Hirudin (Revasc etc.)	Yes	Schmidt, 2004; Walsh, 2010a; Martínez *et al.*, 2012
			Hormones	Insulin (Novolog etc.)	Yes	
				HGH (Valtropin)	Yes	
			Growth factors	PDGF (Regranex)	Yes	
			Vaccines	Hepatitis B (Recombivax etc.)	Yes	
				HPV (Gardisil)	Yes	
		Pichia pastoris	Blood factors	Kalbitor	Yes	Martínez *et al.*, 2012
			Hormones	Insulin precursor	No	
			Others	Human serum albumin	No	
		Hansenula polymorpha	Vaccines	Hepatitis B	No	Martínez *et al.*, 2012
			Interferons	Interferon α2b	No	
	Filamentous Fungi	*Aspergillus niger*	Interleukins	Human interleukin 6	No	Martínez *et al.*, 2012
Higher eukaryotes	Insect cells	Baculovirus	Vaccines	HPV (Cervarix)	Yes	Walsh, 2010; Martínez *et al.*, 2012
			Others	Human proapoliproprotein AI	No	
	Mammalian cells	Chinese hamster ovarian (CHO) cells	Blood factors	Factor VIII (Bioclate)	Yes	Schmidt, 2004; Walsh, 2010a
			Thrombolytics	tPA (Activase etc.)	Yes	
			Hormones	FSH (Gonal F etc.)	Yes	
			Growth factors	EPO (Neorecormin etc.)	Yes	
			Interferons	Interferon β1a (Avonex etc.)	Yes	
			Antibodies	Herceptin etc.	Yes	
		Baby hamster kidney (BHK) cells	Blood factors	Factor VIII (Kogenate etc.)	Yes	Schmidt, 2004; Walsh, 2010a
				Factor VIIa (Novoseven)	Yes	
		Human cell lines	Interleukins	Oprelvekin	Yes	Schmidt, 2004; Walsh, 2010a
			Anti-thrombotics	Protein C (Xigris)	Yes	
			Enzymes	VPRIV	Yes	
				Elaprase	Yes	

Table 1.2: Advantages and disadvantages of common industrially-relevant expression systems

Information taken from Palomares *et al.* (2004), Schmidt (2004), Terpe (2006), Demain & Vaishnav (2009), Ferrer-Miralles *et al.* (2009), Adrio & Demain (2010) and Assenberg *et al.* (2013).

Host	Advantages	Disadvantages
Bacteria – general[1]	Simple Inexpensive High biomass yields High product yields Rapid growth & production	Low levels of PTMs Difficulty expressing larger proteins Poor production of proteins with high DSB content Non-mammalian codon bias
E. coli	Well characterised, highly studied & used Extensive range of tools and systems available Easily manipulated	Poor secretion capacity Tendency for RP aggregation (IBs) Production of pyrogenic LPS endotoxin
Bacillus sp.	Efficient secretion GRAS status Generally well characterised	High protease activity Sporulation
Yeast – general[2]	Increased PTM Efficient secretion Simple Inexpensive High biomass yields High product yields Rapid growth & production	Non-mammalian PTM
S. cerevisiae	Well characterised, highly studied & used GRAS status	High protease activity in secretory vesicles Hyper glycosylation (mannosylation) can result in immune response
Methylotrophic yeasts (*P. pastoris* & *H. polymorpha*)	Efficient secretion Does not hyperglycosylate (mannosylate) as *S. cerevisiae*	Methanol feed flammable & toxic to humans Non-GRAS
Filamentous Fungi	Efficient secretion More complex PTM some GRAS	High protease activity
Insect cell culture (baculovirus)	More complex PTM Efficient secretion Cheaper than mammalian cell culture	Higher cost than microbial culture Slower Complex procedure N-glycosylation is non-sialated
Mammalian cell culture	PTMs should be almost identical to native protein	Expensive Complex fermentation Potential virus & prion contamination

[1] – Applies to all bacteria listed
[2] – Applies to all yeast listed

prokaryotes, especially glycosylation, but also including carboxylation, hydroxylation, amidation, sulphation and disulphide bond (DSB) formation (Walsh, 2010b). Often PTMs are essential for function or, for proteins of pharmaceutical use, may be necessary to avoid immunological responses. Bacteria do not produce most PTMs. For example, a protein that is glycosylated when made in eukaryotic cells will be unglycosylated when produced in *E. coli*. Synthesis of a protein that is natively glycosylated with partial or no glycosylation could significantly limit its functionality.

Bacteria show difficulties in synthesising larger RPs, although production of a 210 kDa protein in *E. coli* has been reported (Doekel *et al.*, 2002; Palomares *et al.*, 2004), typically bacteria are used to produce proteins <30 kDa (Demain & Vaishnav, 2009).

Proteins with high DSB content are also more difficult to produce in bacteria as the cytoplasm, where the bulk of RPP systems direct the nascent polypeptides, has too reducing an environment to favour DSB formation (Terpe, 2006; Demain & Vaishnav, 2009; Assenberg *et al.*, 2013).

An additional limitation in the production of exogenous proteins in bacteria results from differences in codon usage during translation, often termed codon bias. The genetic code is degenerate; the number of trinucleotide codons available is greater than the number of amino acids coded for (61 codons for 20 amino acids (plus 3 stop codons)), therefore many amino acids are coded for by multiple codons and hence multiple tRNA molecules. The relative abundance of different tRNAs for a particular amino acid can vary greatly between species and results in species-specific preferences (reviewed by Terpe, 2006) that will be reflected in the codon usage of a gene from that species (Dong *et al.*, 1995). Rapid production of a protein from an organism with a significantly different codon bias can exhaust the supply of a less-abundant tRNA. This can have severe effects on translation, resulting in ribosomal stalling, translational frame-shifts, amino acid

misincorporation, and premature termination. It may also affect mRNA stability (Goldman *et al.*, 1995; Calderone *et al.*, 1996; Kane, 1995; Kurland & Gallant, 1996). Differences in codon bias can be remedied by supplementing genes for rare tRNAs or modification of the recombinant gene to match host organism preferences, by site-directed mutagenesis or more recently by chemical DNA synthesis and *in silico* design (Burgess-Brown *et al.*, 2008).

1.1.2.2. E. coli

The most common bacterial host for RPP is *E. coli*, the reasons for this appear to be principally historical. *E. coli* is one of the most studied model organisms in the biological sciences therefore much is known about its genetics, physiology, growth and metabolism and *E. coli* was integral to the development of RPP, therefore a wide variety of expression systems and strains tailored for RPP are available (Section 1.2.2). In addition *E. coli* is also easily genetically manipulated allowing ready production of novel strains tailored to the task in hand (Demain & Vaishnav, 2009; Assenberg *et al.*, 2013).

E. coli exemplifies many of the general advantages to bacterial RPP: *E. coli* can be grown to high cell densities (>100 g·L^{-1} dry cell weight (DCW) using inexpensive media (see Chapter 1.4). *E. coli* can accumulate RP to a very high percentage of its total protein content, i.e. up to ~50% of soluble protein (Sung *et al.*, 1987; Sevastsyanovich *et al.*, 2009) and up to 50% of total cellular protein when produced in inclusion bodies (IBs) (Jevševar *et al.*, 2005). For an approximate comparison; the most abundant *E. coli* protein in the PaxDB database of protein abundance, the translational elongation factor EF-Tu (Tuf), has an average abundance of 15497 PPM, equivalent to ~1.5% of protein molecules in the cell (Wang *et al.*, 2012). Growth of *E. coli* is very rapid; under optimal conditions, doubling time can be as little as 20 minutes.

Although *E. coli* will not natively glycosylate RPs there have advances in this regard, primarily involving metabolic engineering and incorporation of the glycosylation machinery from *Campylobacter jejuni* (Lizak *et al.*, 2011; Pandhal *et al.*, 2011, 2012; Valderrama-Rincon *et al.*, 2012).

There are disadvantages to the use of *E. coli* as a host for RPP, it is relatively poor at protein secretion (see Section 1.2.1.1) and rapid RPP often causes aggregation of the RP as IBs (see Section 1.2.1.2). In addition, *E. coli* natively produces pyrogenic lipopolysaccharide endotoxins that must be removed in additional processing stages (Petsch & Anspach, 2000) or requires the use of strains deficient in endotoxin production such as the commercialised CleanColi™ BL21 (DE3) strain (Lucigen Corp.) (Lucigen Corp., 2013).

1.1.2.3. Other Bacteria

In addition to *E. coli,* other bacteria have been used as hosts for RPP particularly the Gram-positive *Bacillus subtilis* and *Bacillus megaterium*. As bacteria they share many of the advantages of *E. coli* and in addition are generally regarded as safe (GRAS) organisms, do not produce endotoxins and are more efficient at secretion. However, high secreted protease activity along with the tendency of *Bacilli* to sporulate under stressful conditions can result in lower productivity and as yet, there are no approved RP pharmaceuticals produced in *Bacillus* species (Palomares *et al.*, 2004; Westers *et al.*, 2004; Terpe, 2006; Walsh, 2010a)*.*

1.1.2.4. Yeast

Many species of yeast have been used as RPP-hosts, particularly *Saccharomyces cerevisiae* and the methylotrophic yeasts *Pichia pastoris* and *Hansenula polymorpha* (Walsh, 2010a; Walsh, 2010b; Darby *et al.*, 2012; Nielsen, 2013). Growth of yeast to high cell densities and high RP yields is again relatively simple and inexpensive. In addition,

yeast are able to secrete and glycosylate proteins. Glycosylation in yeast is however significantly different to that of mammalian cells and may adversely affect product function (Walsh, 2010b). *S. cerevisiae* particularly, is known to hypermannosylate proteins, which can cause an immune response (Palomares *et al.*,2004; Walsh 2010b; Darby *et al.*, 2012).

1.1.2.5. Higher eukaryotic systems

Due to the failings of microbial systems in producing more complex proteins it is sometimes necessary to use species closer to their source in order to ensure adequate RP folding and processing, despite increasing complexity and cost. Common types of higher eukaryotes used for RPP include filamentous fungi, plant, insect, mammalian and human cell lines. Again, notable examples of proteins produced in these systems are given in Table 1.1 and summaries of the advantages and disadvantages given in Table 1.2.

1.1.2.6. Cell-free systems

There are also RPP systems that occur entirely *in vitro*, with no whole, viable cells involved, using either cell-free lysate or a reconstituted expression system. Similar to whole-cell systems, *E. coli*-derived lysates tend to provide higher yields of protein and are among the most commonly used, but eukaryote-derived systems such as those from wheat-germ and rabbit reticulocytes can be used if significant PTM is required. The major advantages of cell-free RPP over cell-based systems are that it can be performed at lower volumes (<1 mL) aiding high-throughput studies, it is less sensitive to toxic proteins (proteins whose production, even when uninduced, can cause cell death, growth defects and decreases in productivity (Saïda *et al.*, 2006)), DSP is simplified as cell lysis is not required and reaction conditions can be easily modified to the target protein (Katzen *et al.*, 2005). A major limitation of cell-free RPP is that it has not yet

been applied industrially, although recently Zawada *et al.* (2011) reported production of a granulocyte-macrophage colony-stimulating factor (rhGM-CSF) with *E. coli* cell extract at the 100 L scale, suggesting possible future industrial applications.

1.2. RPP in *E. coli*

E. coli is a Gram-negative rod-shaped γ-proteobacterium that is a natural coloniser and strain-dependent pathogen of the mammalian intestine. First isolated by Theodor Escherich in the late 19th century and later named after him, it is one of the most widely studied model organisms in the biological sciences. With the extent of knowledge available, ease of growth and easily manipulated genetics it is logical that *E. coli* is also one of the most common hosts for RPP (Overton, 2014).

E. coli is inextricably linked to the history of RPP: the earliest instances of RPP used *E. coli*, the first FDA-approved RP pharmaceutical was produced in *E. coli* (Humulin® (Eli Lilly), recombinant human insulin) (Johnson, 1983) and it remains a preferred host. The utility of *E. coli* for a wide variety of RPP applications is demonstrated by the number of high-value pharmaceuticals (~50% of all approved RP pharmaceuticals; Walsh, 2010a), industrially relevant enzymes and proteins for structural determination (up to 80% of structures in the protein data bank; Sørensen & Mortensen, 2005) produced in *E. coli*.

1.2.1. Methods of production

There are many of strategies for RPP in *E. coli*, relating to the subcellular localisation of the RP and the state in which it is produced, summarised in Figure 1.1. Each strategy has advantages and limitations as summarised in Table 1.3.
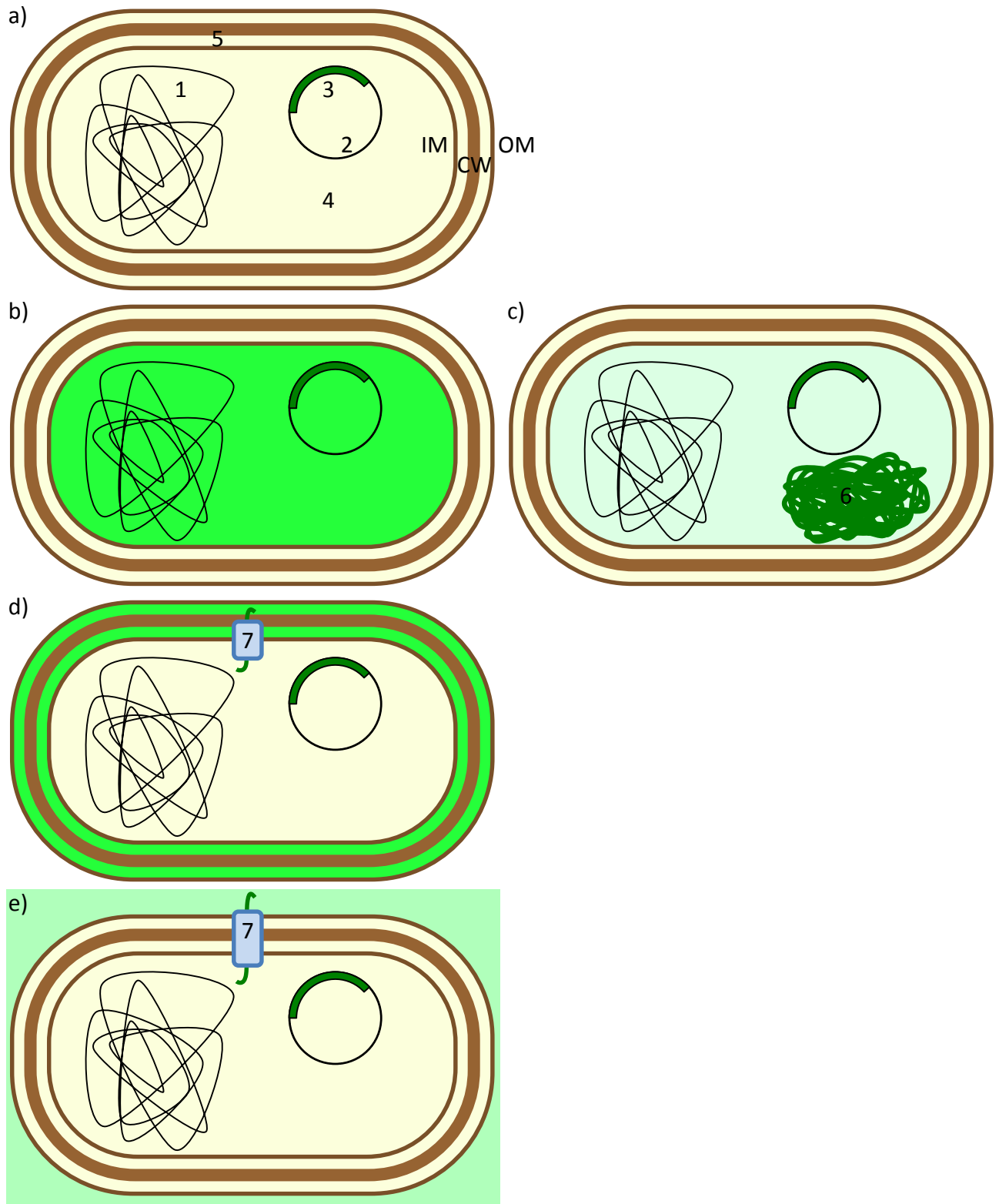
Figure 1.1: Sub-cellular localisation & production methods for RPP in *E. coli*.
Presence of the RP product is denoted by green colouration. a) uninduced cell. b) cytoplasmic, soluble. c) cytoplasmic, insoluble. d) periplasmic. e) secreted. (Key: 1 – chromosome, 2 – expression vector, 3 – gene encoding RP, 4 - cytoplasm, 5 – periplasm, 6 – inclusion body, 7 – protein export machinery, IM – inner membrane, OM – outer membrane, CW – cell wall).

Table 1.3: Advantages and disadvantages of subcellular localisation and conformational state strategies for RPP in *E. coli*

Location/State	Advantages	Disadvantages
Cytoplasm	Simplest expression system Generally high RP yields	More complex DSP (requirement for cell breakage) Reducing environment disfavours DSB formation High protease activity
Periplasm	Simpler DSP Oxidising environment favours DSB formation Lower protease activity N-terminus authenticity	More complex expression systems Transportation systems introduce additional bottlenecks
Extracellular	Simpler DSP Improved folding Lower protease activity N-terminus authenticity	*E. coli* is a poor secretor More complex expression systems Transportation systems introduce additional bottlenecks
Soluble	Correctly folded Functional 'Default' state	Many proteins cannot be produced solubly in *E. coli* Toxic products functional
IBs	Substantial purification by centrifugation Protease resistant Toxic products inactive High cellular RP yields	Requires effective refolding strategy Refolding not always successful Stressful on cells

1.2.1.1. Subcellular localisation

While protein synthesis occurs solely in the cytoplasm *E. coli* can accumulate RP in 3 subcellular locations; the cytoplasm, periplasm and extracellular milieu, accumulation in the latter 2 often being interchangeably referred to as secretion.

Cytoplasmic accumulation is the most common strategy for RPP in *E. coli* as it requires the simplest expression system and generally allows highest RP yields (Makrides, 1996). The cytoplasm however, has many disadvantages for RP accumulation; it has a reducing environment disfavouring the formation of DSBs (Choi & Lee, 2004; Overton, 2014) (although there are advances in this area, see Section 1.2.3.1) and relatively high protease activity (Makrides, 1996; Overton, 2014). Furthermore cytoplasmic accumulation is problematic for DSP, release of proteins from the cytoplasm requires disruption of the entire cell, which requires substantial energy input, can generate heat and produces a large volume of a complex, relatively impure mixture of RP, other cellular proteins, nucleic acids, endotoxin and other cell fragments (Overton, 2014).

Accumulation in the periplasm requires the RP to be translocated across the cytoplasmic membrane. The principal mechanism of translocation for RPP is through the SecYEG complex that exports unfolded proteins, although it has also been demonstrated that RPs can also be exported in a folded state by the twin-arginine translocation (Tat) system (Georgiou & Segatori, 2005; Matos *et al.*, 2012). Periplasmic localisation holds many benefits over cytoplasmic; the periplasmic environment is more oxidising and hence DSB formation is favoured (Makrides, 1996; Mergulhão *et al.*, 2005), protease activity is lower in the periplasm and hence likely increases RP stability (Makrides, 1996; Mergulhão *et al.*, 2005; Overton, 2014), and cleavage of signal peptides increases the likelihood of the RP displaying its native N-terminus (N-terminus authenticity) (Makrides, 1996; Mergulhão *et al.*, 2005). The principal benefit of periplasmic

localisation is that purification of RP from selective disruption of the periplasm is substantially simpler than from the cytoplasm (Makrides, 1996; Mergulhão *et al.*, 2005). The periplasm contains fewer proteins (40-50, equivalent to <5% of cellular protein) and should not contain any nucleic acids, reducing contaminants in the cell extract (Makrides, 1996; Mergulhão *et al.*, 2005). As the volume of the periplasm is substantially lower than that of the cytoplasm (20-40% of total cell volume) (Stock *et al.*, 1977) the subsequent volume of cell extract to be processed is also greatly reduced. Finally, disruption of the periplasm alone can be achieved using gentler methods for example by osmotic shock treatment (French *et al.*, 1996). Periplasmic accumulation does however require a more complex expression system that introduces additional bottlenecks in the process, i.e. the capacity of the translocation apparatus and cellular RP yields are generally lower (Makrides, 1996).

Finally, RP can be transported out of the cell entirely to accumulate in the extracellular milieu. This is a common strategy in hosts such as CHO cells, but it has been less effective when applied to *E. coli* because secretion in wild type *E. coli* is limited, typically only occurring for proteins involved in pathogenesis (Ni & Chen, 2009; Overton, 2014). RP secretion can be split into 2 categories: leakage of periplasmic contents and targeted transport of RP through the secretory apparatus. In the former, proteins are periplasmically targeted under conditions where the outer membrane (OM) is selectively permeablised either by coexpression of genes known to permeablise the OM such as bacteriocin release protein and *kil* or addition of chemical agents such as Triton X-100 and glycine (Makrides, 1996; Choi & Lee, 2004). An alternate strategy involves periplasmic targeting in L-form cells that lack a cell wall and periplasm, however L-form cells show growth defects and are less robust than their cell-wall containing counterparts and are therefore generally unsuited to large-scale production (Gumpert &

Hoischen, 1998; Mergulhão *et al.*, 2005). True secretion through existing apparatus has been achieved by fusion to OM proteins such as OmpF or natively secreted proteins such as haemolysin (Choi & Lee, 2004; Mergulhão *et al.*, 2005). Fusion of RPs to IgG-binding domains of *S. aureus* protein A (or synthetic derivatives thereof) have been demonstrated to direct the fusion product to the extracellular space in addition to aiding RP folding and affinity purification (also see Section 1.3.6) (Abrahmsén *et al.,* 1986; Moks *et al.*, 1987). Recently, Sevastsyanovich *et al.* (2012) reported the development of a novel RP secretion module using the type V or autotransporter system. Accumulation of RP into the extracellular milieu is an attractive option, particularly with regards to DSP. Purification of RP directly from fermentation broth does not require any form of cell disruption and whole cells can be removed simply, for example by centrifugation. This removes a harsh processing stage that may adversely affect the product and, as *E. coli* does not secrete many proteins extracellularly, greatly decreases contaminants (Makrides, 1996). In addition, protease activity is much lower extracellularly, protein folding is often improved and similarly to periplasmic localization, cleavage of signal peptides can increase N-terminus authenticity (Makrides, 1996; Choi & Lee, 2004; Mergulhão *et al.*, 2005). Much like periplasmic accumulation, yields of extracellular RP tend to be lower than those achieved cytoplasmically, the expressions systems required are more complex and introduce additional bottlenecks in processing (Makrides, 1996). RP secretion holds many advantages for industrial RPP, particularly in simplifying DSP as this can represent substantial savings for a process (purification is estimated to account for 50-90% of the total costs for a bioprocess (Cunha & Aires-Barros, 2002)). However, the complexity of the systems involved in secretion, variability in efficacy and extremely high yields possible with cytoplasmic accumulation result in many processes still utilising cytoplasmic accumulation. It is therefore likely that further research into

more robust secretion systems will likely see more widespread adoption of secretory RPP systems.

1.2.1.2. Conformational state

RPs can be made in two conformational states, either in their soluble folded form or as insoluble aggregates called inclusion bodies.

Under normal growth conditions *E. coli* will produce native globular proteins so that they fold into their stable, soluble, form, and can therefore be regarded as the 'default' production option. Production of RP in its soluble form is advantageous as it will most likely exhibit native structure and function. However, many proteins, particularly those of eukaryotic origin cannot be produced solubly in *E. coli* and some RPs when expressed in active form may be toxic to the cells (Makrides, 1996; Saïda *et al.*, 2006).

Conventional RPP protocols often exceed the cells' capacity for producing correctly folded protein, in these circumstances RPs aggregate into IBs. IBs are large, relatively dense, proteolysis-resistant aggregates that often exhibit considerable amyloid character i.e. large amounts of cross-beta sheet structure (de Groot *et al.*, 2009) and have been regarded as being inactive relative to their constituent protein in its native state (Singh & Panda, 2005) (though there exists evidence that they do exhibit some activity, see Section 1.6.6). Although RPs are generally required in their soluble, active state as a final product, production as IBs is frequently utilised as it can provide advantages to the process. IBs are relatively pure, dense particles therefore it is possible to achieve significant RP purification from crude cell lysate in a single centrifugation step; this process is used to produce many products such as insulin (Lilie *et al.*, 1998; Schmidt *et al.*, 1999; Vallejo & Rinas, 2004; Sevastsyanovich *et al.*, 2010). IBs are also resistant to proteolysis and present a method to produce toxic proteins in an inactive form (Makrides, 1996). In addition cellular yields of RP in IBs are often extremely high,

up to 50% of total cellular protein (Jevševar *et al.*, 2005). The principle issue with IBs is that they require substantial processing to refold the constituent RPs into their biologically active forms. Refolding of proteins from IBs involves two steps: first, solubilisation of the IB preparation in buffer typically containing a chaotropic agent such as urea or guanidium chloride and a reducing agent such as β-mercaptoethanol, then removal of the solubilisation agents by dialysis, allowing gradual refolding. While refolding can be highly effective and is frequently used industrially it is an additional processing stage that introduces additional costs, it is not universally effective and therefore the quality of refolded protein can be low. Yields of biologically active refolded RP can be as low as 15-25% of total protein (Lilie *et al.*, 1998; Singh & Panda, 2005). In addition, RP misfolding and IB formation is associated with substantial physiological stresses that may be detrimental to cell productivity (see Section 1.3).

The choice of conformational state for RPP depends primarily on the RP in question and its intended use, for a complex eukaryotic protein that is unstable and potentially toxic when produced in bacteria, IBs are the logical option as this would be the only method to ensure sufficient RP yields, offsetting the additional costs involved with refolding. However for an RP that can be stably expressed in soluble form, soluble production *ab initio* may be a preferable strategy.

1.2.2. Common RPP systems in *E. coli*

Many systems have been developed for RPP in *E. coli*, comprising different mechanisms for induction, host strains and additional features to aid production.

1.2.2.1. Induction

RPP is often known to result in a rapid loss of productivity and cell growth either due to RPP-related stress or simply redirection of metabolites towards RPP (the metabolic

burden). As such, regulated production of RP is generally favoured (Makrides, 1996).

One of the principal steps in the regulation of gene expression in *E. coli* is at the stage of

transcription initiation, similarly it is at this stage that RPP is principally regulated. In its

simplest form the recombinant gene of interest is placed under the control of a

promoter that will become active only under specific controlled conditions. On exposure

of the cells to these conditions (termed induction) RNA polymerase can bind the

promoter, initiating transcription and hence production of the RP. While it is

theoretically possible to use any promoter from *E. coli* and any synthetic derivative

thereof, only a limited number have proved to be effective and hence are routinely used.

Terpe (2006) provides a useful list of features that a promoter must possess to be useful

for RPP: low basal transcription (i.e. tight regulation); readily transferable to other

strains; simple and inexpensive to induce; and induction should be independent of

extraneous stimuli. In addition it is also possible to modulate the rate of RPP using a

dose-dependent promoter such as the *araC*-P_{BAD} system (Guzman *et al.,* 1995).

1.2.2.2. lac promoter-based systems

Regulation of the *lac*ZYA operon, encoding the machinery for utilisation of the

disaccharide lactose, was the first gene regulatory circuit elucidated and as such is the

archetype for gene expression in *E. coli*. In wild type *E. coli* the *lac* operon is under the

control of the *lac* repressor LacI, via the *lac* promoter (P_{lac}). On binding the lactose

isomer, allolactose, LacI undergoes a conformational change that prevents DNA binding,

allowing RNA polymerase access to the promoter and hence initiating transcription

(summarised in Figure 1.2). In a synthetic setting deliberate addition of lactose to a

culture will induce recombinant gene expression. This has been aided by the utilisation

of metabolically-inactive synthetic inducer analogues such as isopropyl β-D-1-

thiogalactopyranoside (IPTG) the structure of which in comparison to lactose and
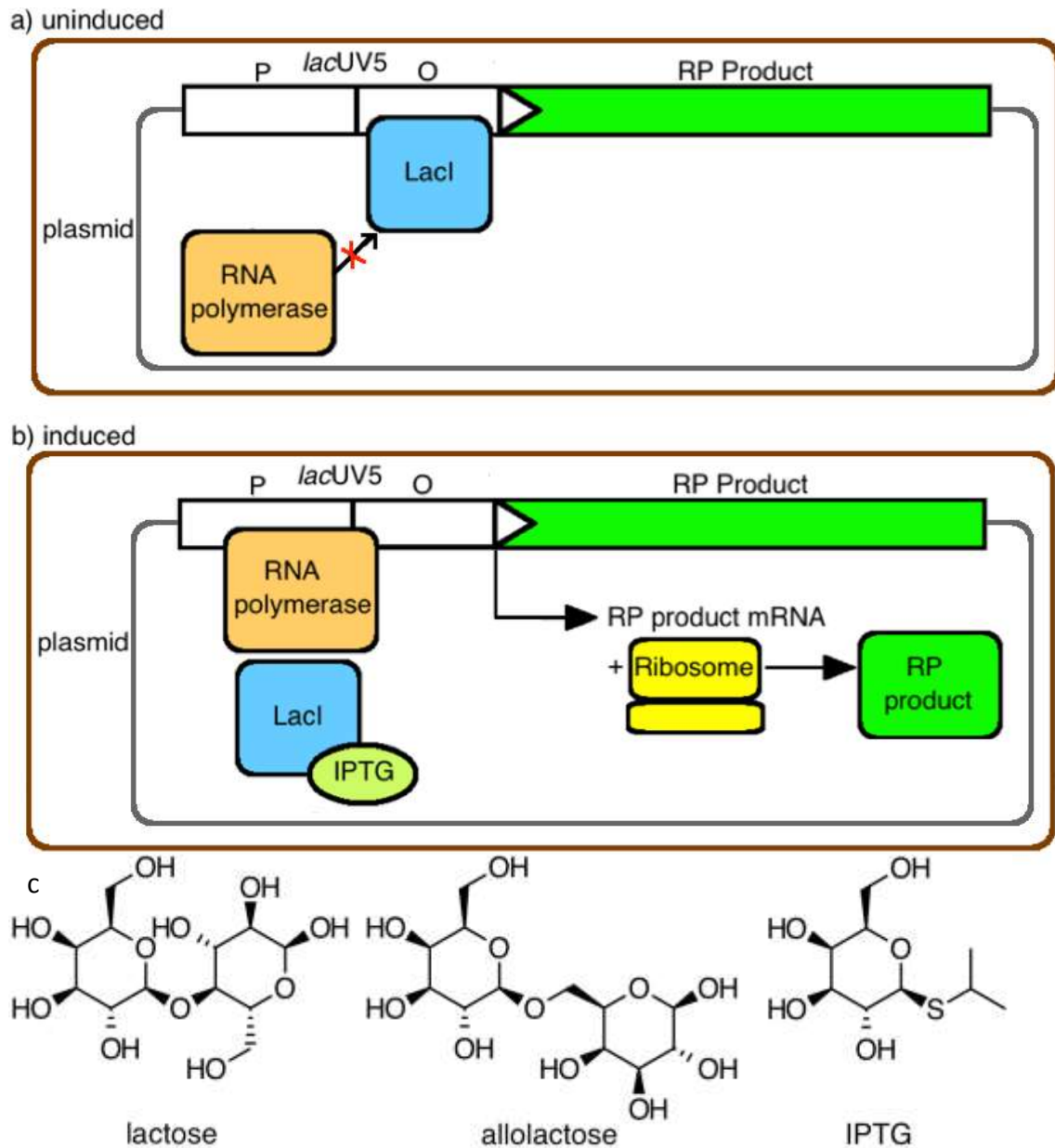
Figure 1.2: Mechanism of induction for LacI-based RPP systems
a) Uninduced cells: in the absence of inducer, LacI is bound to the *lac* operator, blocking access of RNA polymerase, therefore preventing expression of the RP product. b) Induced cells: LacI binds IPTG and cannot bind DNA, allowing access of RNA polymerase to the T7 polymerase gene, resulting in synthesis of the RP product. c) Chemical structures of LacI effector molecules. Adapted from Sweet (2003).

allolactose is given in Figure 1.2. As IPTG is metabolically inactive it is not subject to decreases in concentration due to catabolism of lactose (Sweet, 2003). Given the status of P$_{lac}$ as a model expression system it is unsurprising that some of the earliest examples of RPP used transcriptional *lac*-fusions and IPTG to induce RPP (Goeddel *et al.*, 1979). There are 3 limitations of wild-type P$_{lac}$ for RPP: it is subject to catabolite repression and as such requires activation by cyclic AMP (cAMP) and cAMP receptor protein (CRP) for full activity and therefore will not reach full induction during growth on glucose; it is a relatively weak promoter; and finally, it has significant basal or 'leaky' activity in the absence of inducer (Sweet, 2003; Terpe, 2006). A number of P$_{lac}$ variations have been developed to address these limitations. To reduce reliance on CRP/cAMP a mutant variant of P$_{lac}$, called *lac*UV5 can be used that allows significant transcription without CRP (Eron & Block, 1971). To increase promoter activity, hybrid promoters were developed by replacing the *lac*UV5 -35 region with that of the *trp* operon (encoding genes for tryptophan biosynthesis). Two promoters were developed, *tac* and *trc,* with the sole difference being the region between the -10 and -35 elements is 16 and 17 bp respectively (de Boer *et al.*, 1983; Brosius *et al.*, 1985). *tac* was shown to have 5-10 fold higher activity than *lac* (Amann *et al.*, 1983) and *trc* was shown to retain approximately 90% activity compared to *tac* (Brosius *et al.*, 1985). To reduce basal activity it is possible to use the mutant LacI variant, LacI^q. LacI^q binds the *lac* operator more tightly, reducing basal transcription 10-fold (Calos, 1978).

1.2.2.3. Other chemically-induced systems

Of the simple chemically-induced systems *lac*-based ones remain the most developed, but others are available. These include promoters derived from other carbohydrate utilisation operons such as the L-arabinose-induced P$_{BAD}$ (Guzman *et al.*, 1995) and L-rhamnose-induced *rha*P$_{BAD}$ promoters (Haldimann *et al.*, 1998) and others such as the

*tet*A promoter that is induced by anhydrotetracycline, a less active form of the antibiotic tetracycline (Skerra, 1994).

1.2.2.4. Bacteriophage λ-based systems

In addition to *E. coli*-derived, chemically-induced systems, it is also possible to use regulatory systems derived from bacteriophage and physical signals for induction. For example, the system reported by Elvin *et al.* (1990) that used the strong promoter p_L from bacteriophage λ in conjunction with a mutant form of the λ repressor protein, λ *cI*857. λ *cI*857 is temperature dependent, at increased temperatures it loses the ability to repress and therefore RPP can be induced by an upshift of temperature from 30°C to 42°C (Elvin *et al.*, 1990; Terpe, 2006).

1.2.2.5. Bacteriophage T7 RNA polymerase-based systems

One of the most common RPP systems in use and the one most pertinent to this study is a combination of the *lac*-based, chemically-induced and bacteriophage-derived systems; the T7 polymerase system (Studier & Moffat, 1986; Sweet, 2003; Sørensen & Mortensen, 2005; Terpe, 2006). In the T7 polymerase system the recombinant gene, supplied on a plasmid, is under the control of a promoter that is recognised solely by the T7 bacteriophage RNA polymerase. Expression of the recombinant gene occurs upon the expression of the T7 polymerase, which is chromosomally located (Figure 1.3). The reason for chromosomal location of the T7 polymerase gene was that in initial experiments where the gene was provided on a plasmid along with the recombinant gene, basal expression of the polymerase was sufficient to severely affect plasmid maintenance. Studier & Moffat (1986) proposed two induction systems, first providing the T7 polymerase only on induction by infection with bacteriophage λ and second with T7 polymerase already chromosomally integrated through a prophage of the λ-derivative DE3 under the control of *lac*UV5 and IPTG induction. The former was
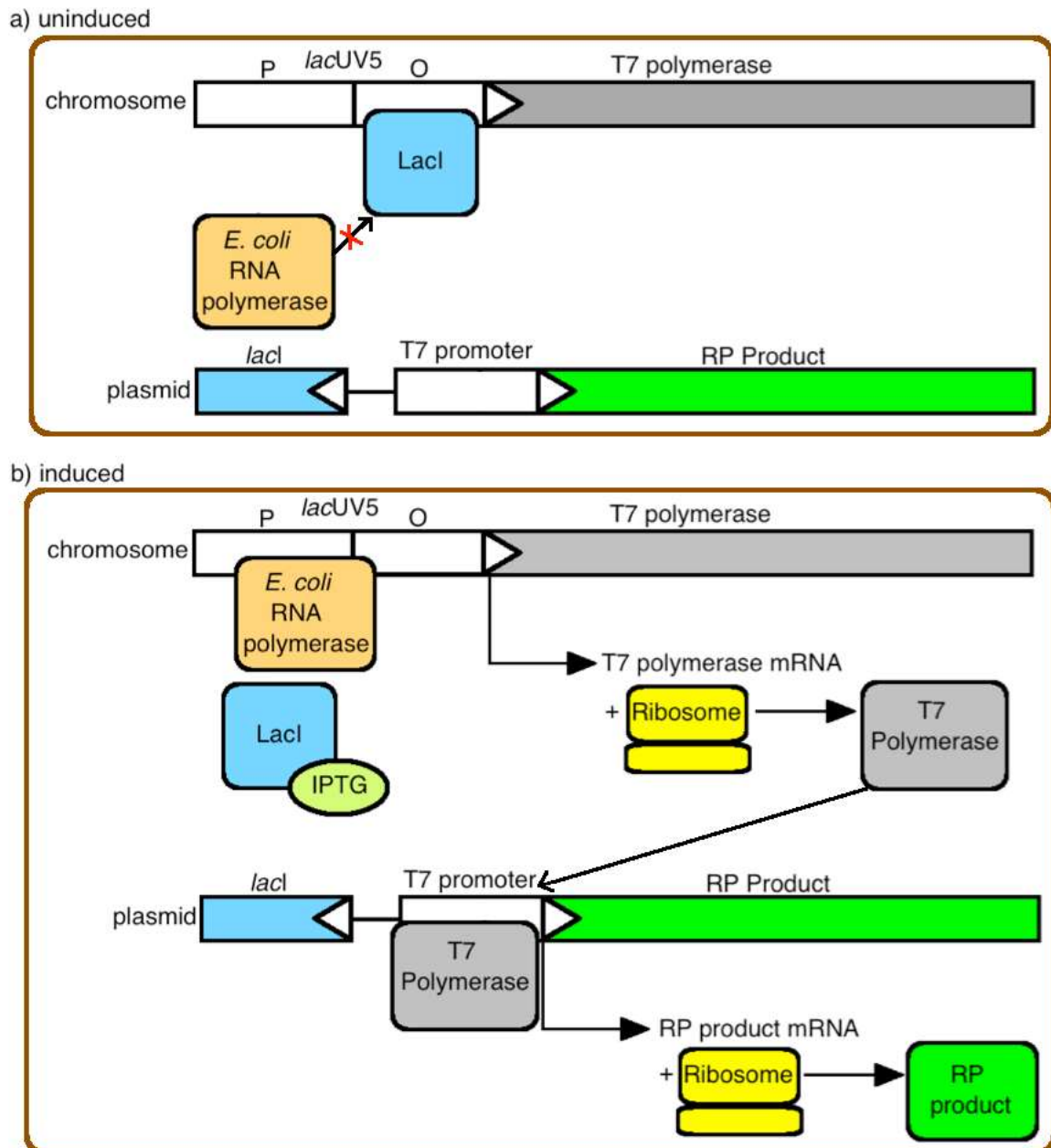
Figure 1.3: Mechanism of induction for the T7 polymerase RPP system.
a) Uninduced cells: in the absence of inducer, LacI is bound to the *lac* operator, blocking access of *E. coli* RNA polymerase, therefore preventing expression of the T7 RNA polymerase. b) Induced cells: LacI binds IPTG and cannot bind DNA, allowing access of *E. coli* RNA polymerase to the T7 polymerase gene, resulting in synthesis of T7 RNA polymerase. T7 RNA polymerase then allows synthesis of the RP product. Adapted from Sweet (2003).

suggested for production of toxic products where even basal activity from a *lac* promoter is sufficient to cause growth defects, however it is the latter that has become the more popular method. Following the development of the T7 polymerase system numerous DE3 strains have been commercialised along with many T7 expression vectors, most notably the pET series.

1.2.3. Strains

Numerous strains of *E. coli* exist that may be used for RPP or have been specifically developed as a host for RPP, depending on the expression system used, selected examples are summarised in Table 1.4. Simple plasmid-based expression systems may be used in many strains providing they contain compatible genotypes, whether they are specifically developed for RPP or not such as *E. coli* W3110 a wild-type K12-derivative strain commonly used for RPP (Huang *et al.,* 2012). T7-polymerase-based systems however, require strains specifically created for RPP by integration of the DE3 prophage. While many DE3 strains are available the most common are BL21 (DE3) and its derivatives.

1.2.3.1. E. coli BL21 and derivatives

E. coli BL21 (DE3) was one of the original strains developed by Studier & Moffat (1986) by integration of the DE3 prophage into BL21, a derivative of B834 (Wood, 1966). An advantage of BL21 and derivatives for RPP is that they are deficient in the proteases *lon* and *ompT* (Sørensen & Mortensen, 2005; Terpe, 2006). In addition to BL21 a number of derivatives have been developed with improvements in various aspects of RPP. BL21*, as used in this study, is a derivative of BL21 that contains the *rne*131 mutation. *rne* is an essential gene encoding the enzyme RNase E. *rne*131 produces a truncated enzyme that results in increased RNA stability (Lopez *et al.*, 1999) and is claimed to therefore

Table 1.4: Selected *E. coli* strains commonly used for RPP
Adapted from Terpe (2006).

Strain	Derivation	Features
W3110	K12	Wild-type laboratory strain
Origami	K12	*trxB* & *gor* mutant for improved DSB formation
HMS174	K12	DE3 prophage for T7 polymerase expression system *recA* mutant for increased plasmid stability
BL21	B834	DE3 prophage for T7 polymerase expression system *lon* & *ompT* protease mutant
BL21*	BL21	*rne*131 mutation for increased mRNA stability
BLR	BL21	*recA* mutant for increased plasmid stability
Origami B	BL21	*trxB* & *gor* mutant for improved DSB formation
Rosetta	BL21	pRARE to supplement tRNAs for rare codons
Rosetta-gami	BL21	pRARE to supplement tRNAs for rare codons *trxB* & *gor* mutant for improved DSB formation
C41	BL21	Mutant strain for improved production of membrane proteins
C43	BL21	Double mutant strain for improved production of membrane proteins
pLysS, pLysE	Plasmids	Encode T7 lysozyme to reduce basal RP expression

increase yields of recombinant protein (Life Technologies Corporation, 2010). To reduce basal or 'leaky' expression of toxic proteins in T7 polymerase systems the pLysS and pLysE plasmids may be used encoding the T7 lysozyme that degrades the T7 polymerase and hence decreases uninduced RP expression (Sweet, 2003; Terpe, 2006; Merck KGaA, 2013a). To aid formation of cytoplasmic disulphide bonds Origami™ (DE3) (Novagen) can be used; a *trxB* and *gor* mutant, BL21 (DE3)-derivative strain with a more oxidising cytoplasm (Terpe, 2006; Merck KGaA, 2013b). To aid production from non-codon optimised genes Rosetta™ 2(DE3) (Novagen) can be used; a BL21(DE3) derivative containing pRARE2, a plasmid encoding tRNAs for 7 rare codons (Terpe, 2006; Merck KGaA, 2013c). Finally, Rosetta-gami™ B(DE3) possesses features from both the Rosetta and Origami strains (Terpe, 2006; Merck KGaA, 2013d). *E. coli* C41 and C43 are BL21-derivative strains that were originally isolated by Miroux & Walker (1996) that had acquired mutations allowing improved production of a number of globular and membrane proteins.

A major advantage of BL21-derivatives in industrial bioprocesses is that BL21 has been demonstrated to produce less acetate in high cell density culture than K-12-derived strains such as JM109. This reduces culture acidification and increases the efficiency of conversion from glucose to biomass (see Section 1.4.3.2). This was shown to be due to increased flux through the glyoxalate shunt and acetyl-CoA synthetase (Noronha *et al.*, 2000; Phue & Shiloach, 2004).

One consideration ought to be made when using BL21 and derivatives for industrial cultivation, BL21(DE3) has been shown to be FNR deficient due to the acquisition of a nonsense mutation at codon 141 in an ancestral strain (Studier *et al.*, 2009). FNR is a global gene regulator associated with the shift to anaerobiosis. Lacking FNR, BL21(DE3)

therefore displays metabolic impairments including the inability to respire anaerobically (Pinske *et al.*, 2011).

1.2.3.2. Other (DE3) strains

BL21 (DE3)-derivative strains are not the only strains that use the T7 polymerase system. In addition to BL21(DE3) Studier & Moffat (1986) also inserted DE3 into HMS174, a K-12 derivative strain. HMS174 is a *recA* mutant and as such may have increased vector stability. HMS174(DE3) is commercially available and there are also K-12 derivative versions of the previously mentioned Rosetta™, Origami™ and Rosetta-gami™ strains (Terpe, 2006).

1.3. Physiological stress during RPP & the improved protocol

RPP generally requires an organism to produce a single protein to a significant proportion of its total protein content at a rate far in excess of that during normal growth. *E. coli* has not evolved to do this and adverse effect on the cells may be expected. There have been many studies regarding the considerable stresses RPP imposes on bacteria (Gill *et al.*, 2000; Jürgen *et al.*, 2000; Haddadin & Harcum, 2005; Harcum & Haddadin, 2006; Dürrschmid *et al.*, 2008). RPP has been shown to induce an overlapping stress response; for example, the transcriptional response of *E. coli* to production of chloramphenicol acetyltransferase (CAT) with a high-stress RPP protocol (37°C, 5 mM IPTG) was found by Haddadin & Harcum (2005) to include elements of a general stress response regulated by RpoS, the heat-shock response regulated by RpoH and the nutrient limitation-detecting stringent response, causing detrimental effects to cell growth, plasmid retention and quality of the RP product. Decreases in expression of genes involved in oxidative phosphorylation and the machinery for transcription and

translation led the authors to conclude that the response to RPP induction down-regulates genes that should be up-regulated to continue production.

1.3.1. Mis-folded protein and the heat shock response

During RPP the rate of protein production often exceeds the capacity of the post-translational folding and modification systems including the chaperones and transporters that are essential for some proteins to reach their soluble, stable, native states (Baneyx & Mujacic, 2004). This results in an accumulation of incorrectly folded protein, often as IBs. The presence of mis-folded protein is one of the key signals for induction of the RpoH (heat shock sigma factor) regulon, through an unfolded protein titration model, summarised in Figure 1.4 (Yura & Nakahigashi, 1999; Guisbert *et al.*, 2008). RpoH is synthesised constitutively, however under normal growth conditions RpoH becomes associated with the ubiquitous chaperones DnaK, DnaJ, GrpE and GroELS and is directed towards proteolysis (primarily by the ATP-dependent metalloprotease FtsH) (Yura & Nakahigashi, 1999; Guisbert *et al.*, 2008). There is therefore, insufficient RpoH available to induce expression of its regulon significantly. Mis-folded protein caused by either RPP or denaturation (e.g. due to heat) will bind to the chaperones and hence titrate them away from binding RpoH. Free RpoH is then available to bind RNA polymerase and induce expression of its regulon comprising of >20 proteins (Gamer *et al.*, 1996; Yura & Nakahigashi, 1999; Arsène *et al.*, 2000; Guisbert *et al.*, 2004, 2008). Effects of RpoH induction include up-regulation of a number of 'heat shock proteins' (HSPs) principally the molecular chaperones DnaK, DnaJ, GrpE and GroELS, the ATP-dependent proteases Lon and ClpAB and the inclusion body-associated proteins IbpA and IbpB, allowing mis-folded proteins to be either refolded by the chaperone systems or degraded by the proteases. Regulation at the σ-factor level as opposed to
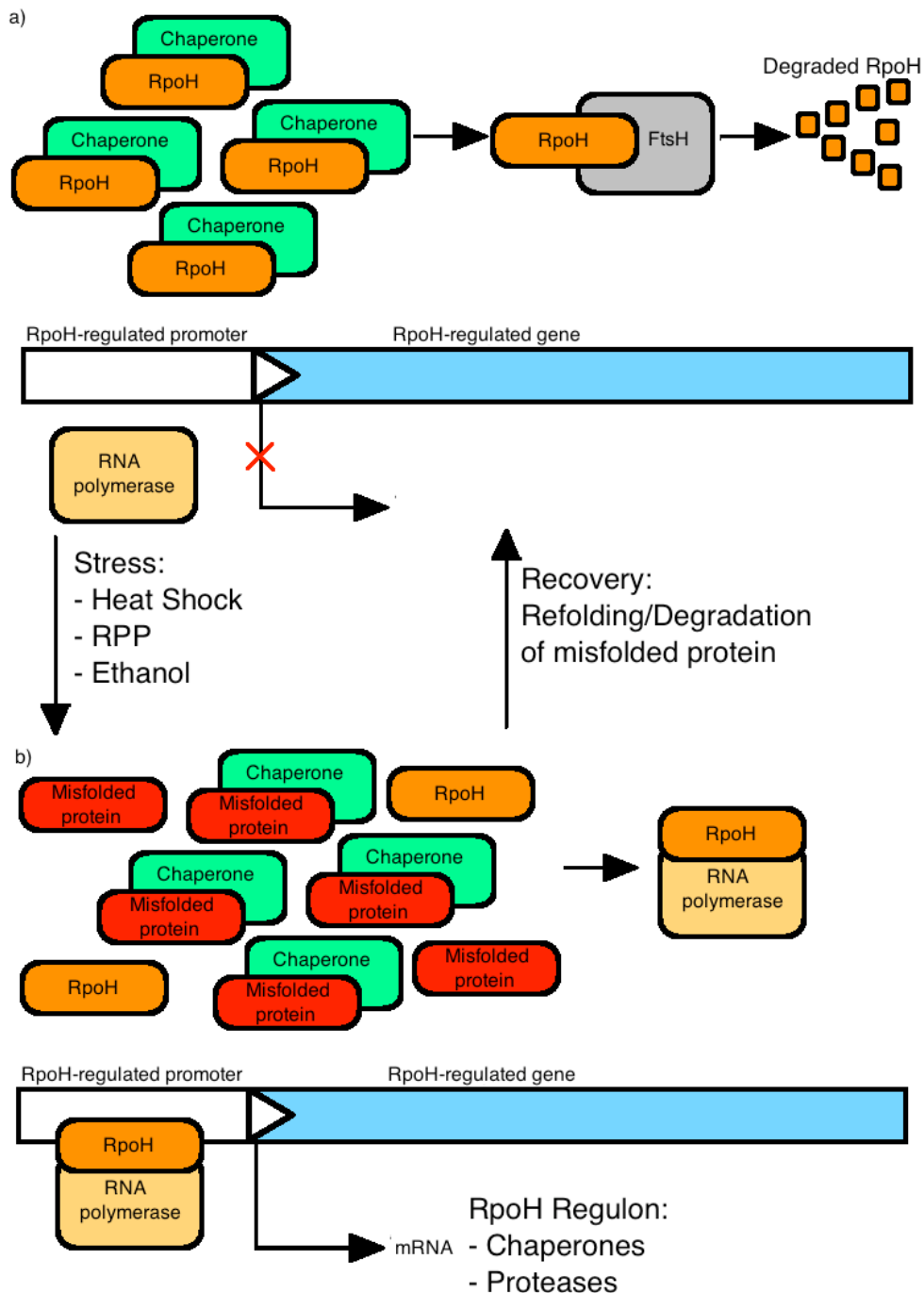
Figure 1.4: Mechanism of RpoH induction through the misfolded protein titration model.

a) Under normal growth conditions: RpoH is sequestered by chaperones and targeted to FtsH for proteolysis and hence unable to bind RNA polymerase and initiate transcription. b) Under heat shock: Increased amounts of mis-folded protein titrate chaperones away from RpoH, resulting in free RpoH that then is able to bind RNA polymerase and direct transcription of HSP operons (adapted from Yura & Nakahigashi (1999) and Guisbert *et al.* (2008)).

transcription factor allows the transition to a stress-responsive state at a higher hierarchical level, essentially reprogramming RNA polymerase away from 'housekeeping' genes. This is reflected in the fact that at 46°C, 15-20% of cellular protein is comprised of the two major chaperone systems DnaK and GroE (Yura & Nakahigashi, 1999; Arsène *et al.*, 2000). Similarly, mis-folded protein in the periplasm will induce expression of the envelope stress response via RpoE (Yura & Nakahigashi, 1999; Ades, 2004; Alba & Gross, 2004).

1.3.2. Nutrient depletion and the stringent response

RPP imposes significant metabolic challenges to cells, not only through synthesis of the product itself, but also through maintenance of a high copy number plasmid expression vector with associated antibiotic resistance genes that must be expressed (See Chapter 1.3.4.) and also in the case of the *E. coli* BL21/pET and similar systems through synthesis of the T7 RNA polymerase. It is only to be expected that diverting cellular resources from growth towards RPP and supporting processes will result in some degree of growth arrest. Bentley *et al.* (1990) observed that RPP caused a decrease in growth rate with increasing RP content, Dong *et al.* (1995) and Kurland & Dong (1996) then observed that RPP triggered a breakdown of rRNA, and hence of ribosomes, causing a decrease in protein synthesis, followed by a loss of viability. It is not even necessary to produce the RP to cause growth defects. Soriano *et al.* (1999) observed that induction (37°C, 1.68 mM IPTG) of a T7-based expression vector with no RP gene still limited growth. The authors concluded that this was a specific effect of T7-type plasmids as this effect was not observed with another, non-T7 type plasmid.

Synthesis of all the necessary components for RPP represents a significant diversion of cellular resources that would otherwise be available for normal cellular processes,

termed a 'metabolic burden' (Bentley & Kompala, 1990). If the metabolic burden of RPP becomes sufficient to deplete cellular resources (amino acids, ATP and other metabolites) it can activate a state known as the stringent response. The stringent response detects a state of nutrient depletion that has become limiting for growth and facilitates a shift of gene expression from one suited for rapid growth to stationary phase (Chatterji & Ojha, 2001). Depletion of amino acids results in an increase in uncharged tRNAs and stalled ribosomes. Triggered by the presence of stalled ribosomes, RelA synthesises the alarmone 5',3' guanosine pyrophosphate (ppGpp). ppGpp is then the signal to induce the stringent response. Effects of the stringent response include a decrease in the amount of stable RNAs produced, a decrease in the number of ribosomes and an up-regulation of genes involved in amino acid biosynthesis. In non-RPP growth conditions the effects of these changes would be to prepare cells for survival in stationary phase. ppGpp has in fact been found to positively regulate *rpoS* and hence entry into stationary phase (Gentry *et al.*, 1993; Chatterji & Ojha, 2001).

1.3.3. RpoS and the general stress response

When stressed, *E. coli* can induce not only specific stress responses such as that for heat shock, but also a general stress response mediated by RpoS that provides some resistance to further and additional stresses often termed cross-protection and is involved in the transition to stationary phase on nutrient limitation (Battesti *et al.*, 2011). Gill *et al.* (2000) showed upregulation of rpoS in response to RPP and Haddadin & Harcum (2005) showed RPP induction to have an overlapping transcriptional response to that of the shift to stationary phase.

1.3.4. Effects of RPP-induced stress

The combination of these responses (heat shock, stringent and general stress responses) has 3 effects. First, the productivity of the cells and quality of the product decreases, changes in gene expression ultimately reduce the cells capacity for producing the RP and protein that is made is either aggregated into IBs or degraded by proteases (Villaverde & Carrió, 2003; Haddadin & Harcum, 2005; Sevastsyanovich *et al.*, 2009). Second, stress-responsive gene regulatory elements titrate RNA polymerase away from σ^{70}-regulated housekeeping and growth related genes, severely arresting the growth rate and culturability of RP-producing bacteria (Andersson *et al.*, 1996; Sundström *et al.*, 2004; Haddadin & Harcum, 2005). This growth arrest also is related to a third detrimental phenomenon observed during RPP; loss of the plasmid expression vector. To ensure retention, plasmids usually also contain a gene encoding antibiotic resistance. Supplementation of the growth medium with the relevant antibiotic creates a selective pressure for retention of the plasmid as plasmid loss will result in a cell becoming antibiotic sensitive and hence unable to grow. Despite antibiotic based selection, plasmid-free cells often eventually outgrow their protein-producing counterparts (Sevastsyanovich *et al.*, 2009), presumably as the selective pressure conferred by the antibiotic resistance genes becomes insufficient to counterbalance the severe disadvantage posed by RPP-induced stress. This effect is probably related to a second issue with antibiotic-based selection in that some of the most common antibiotics used are the β-lactams such as ampicillin. β-lactam resistance relies on enzymatic degradation of the antibiotics by β-lactamase (Bla) and may, after sufficient time, reduce concentrations below the minimum inhibitory concentration (Baneyx, 1999). This situation is exacerbated by the fact that the Bla can accumulate in the extracellular environment as a result of its leakage from lysed cells (Baneyx, 1999). There are a

number of ways that this can be avoided; the simplest is by using an alternate antibiotic, either a more stable β-lactam antibiotic variant such as carbenicillin (Butler *et al.*, 1970) or by using an alternate resistance marker that may exhibit better stability. Other options remove the need for antibiotics altogether through either chromosomal insertion of the expression system, or so called 'plasmid addiction systems' where loss of the plasmid directly causes cell death. There are three categories of plasmid addiction systems; toxin/antitoxin (TA)-based systems, metabolism-based systems and operator/repressor titration (ORT) systems (Kroll *et al.*, 2010). In TA-based systems the plasmid encodes a stable toxin and unstable antitoxin, plasmid loss prevents further synthesis of both components but the antitoxin degrades more rapidly than the toxin resulting in free toxin within the cell and therefore cell death. There are 3 types of TA-systems used, i.e. antisense RNA-regulated, protein-regulated and restriction modification-based; these are described in detail by Kroll *et al.* (2010). In metabolism-based systems an essential, chromosomal gene is knocked out and provided in *trans* on a plasmid, such as *ispH* an essential gene in isoprenoid biosynthesis (Kroll *et al.*, 2009). Plasmid loss removes the essential gene, impairing growth of plasmid-free cells. Finally, in operator/repressor titration systems, an essential chromosomal gene is placed under the control of the *lac* promoter and the plasmid contains multiple copies of the *lac* operator. When the plasmid is present lacI is titrated away by the multiple *lac* operator sites allowing essential gene expression however, when the plasmid is lost lacI binds and represses essential gene expression, causing cell death. ORT-systems were originally developed by Williams *et al.* (1998) using chromosomally located *kan* encoding kanamycin resistance, allowing growth on kanamycin-containing media. Cranenburgh *et al.,* (2001) then developed a system using *dapD* an essential gene involved in lysine biosynthesis and peptidoglycan crosslinking that did not require

antibiotics. A limitation of the ORT-systems described here for RPP is that the use of LacI to repress the essential gene precludes the use of *lac*-based RPP systems as induction would also induce essential gene expression independent of the plasmid.

1.3.5. RPP-related stress in other host species

RPP does not, of course, result in substantial physiological stress in *E. coli* alone. Similarly to *E. coli*, production of an aggregation prone protein in *B. subtilis* was shown to form IBs and proteome analysis showed increased synthesis of HSPs including chaperone & proteases (Jürgen *et al.*, 2000; Westers *et al.*, 2004)

RPP-related stress can also affect eukaryotes e.g. yeasts and mammalian cell culture hosts in addition to prokaryotes. Typically the nascent RP polypeptide is translocated from the cytoplasm into the endoplasmic reticulum (ER) where it folds and is processed by the PTM machinery before secretion. High rates of RPP can exceed the capacity of the folding and PTM systems in the ER and cause accumulation of misfolded RP (Schröder & Kaufman, 2005; Hussain *et al.*, 2014). As a result, detection of misfolded protein typically occurs in the ER (Mattanovich *et al.,* 2004). Detection of misfolded protein in the ER ultimately results in the expression of a number of genes known as the unfolded protein response (UPR). The UPR is a wide acting response (transcriptomic studies in yeast showed >350 genes (>5% of the total gene complement) were UPR-regulated (Patil & Walter, 2001;)). The UPR can increase the protein processing capability of the ER by upregulating for example ER-resident chaperones, decrease the workload of the ER by limiting transcription and translation or by degrading misfolded protein by ER-associated protein degradation (ERAD) whereby misfolded protein is translocated back into the cytoplasm, ubiquitinylated and targeted for proteolysis (Patil & Walter, 2001; Mattanovich *et al.*, 2004; Gasser *et al.*, 2008; Hussain *et al.*, 2014). The UPR is also

thought to affect the entirety of the secretory pathway, beyond that of the ER (Patil & Walter, 2001;). If the effects of the UPR are insufficient to counteract that of the misfolded protein itself this is known in higher eukaryotes to be a trigger for apoptosis (Mattanovich *et al.*, 2004; Hussain *et al.*, 2014). Although there are substantial differences in the specifc mechanisms of RPP related stress between *E. coli* and eukaryotic systems there are superficial similarities; the principle trigger detected is that of misfolded protein and in both cases the response include upregulation of chaperones and proteolysis. It is also possible that high rates of RPP could result in an exhaustion of metabolites as with the *E. coli* stringent response but the lower synthesis rates typical of eukaryotic RPP make this unlikely (Mattanovich *et al.*, 2004).

1.3.6. Previous attempts to ameliorate RPP-induced stress

Protein aggregation/degradation, growth arrest and plasmid loss are serious problems for the efficacy of RPP-related processes and there have been a number of attempts to ameliorate these. Induction of the heat-shock response appears to be the cell's mechanism by which it adapts to and copes with the presence of RPP-generated mis-folded protein. If this is the case then deliberately stimulating the heat shock response, or artificially simulating it could be beneficial to RPP processes. Heat-shock immediately prior to induction or using heat shock to induce protein production (for example with λ-cI857, a temperature sensitive mutant of the bacteriophage λ repressor protein) induces chaperone synthesis that can increase the cells capacity for protein folding (Schmidt *et al.*, 1999). However, heat shock will also induce the synthesis of potentially detrimental proteases. Others have had some success co-expressing chaperones from a second plasmid (de Marco *et al.*, 2007; reviewed by Martínez-Alonso *et al.*, 2010). This, however, requires maintenance of a second plasmid with the requirement for further

antibiotic selection and imposes an additional metabolic burden in synthesis of both the plasmid and chaperones. In addition, its efficacy is by no means consistent. For instance, depending on the RP and chaperones used, chaperone co-expression can result in a decrease in productivity. It is also possible to follow induction with a decrease in temperature (e.g. 37°C to 25°C), the lower temperature potentially favouring RP folding (Sevastsyanovich *et al.,* 2009; Alfasi, 2010).

An additional strategy for increasing RPP productivity is to modify the growth media to favour protein solubility and folding by the addition of chemical agents that promote protein folding (reviewed by Fahnert, 2012). This can be as simple as using rich growth media (Moore *et al.*, 1993) or supplementing a limiting metabolite such as an amino acid (Ramírez & Bentley, 1995). It is also possible to supplement chemicals known to modify protein production and folding; Kusano *et al.* (1999) reported sub-lethal doses of the protein synthesis inhibiting and cold-shock inducing antibiotic chloramphenicol enhanced production of a number of eukaryotic cytochrome P450s and production of another P450 was enhanced by addition of ethanol that is known to induce heat shock and hence increase chaperone production. In addition, the original protocol for producing the CheY::GFP model protein used in this study involved the addition of 20 g·L^{-1} ethanol at 6 h post-induction to enhance RP solubility (Jones, 2007).

The methods discussed above for alleviating RPP-related stress have focused on adapting process and culture conditions, it is also possible to adapt the target protein and host organism; a protein that correctly folds more readily is less likely to cause stress to the host and the benefits of producing a host more able to withstand stress is obvious. In this regard there has been some success in fusing the protein of interest to another protein with useful properties. Fusion to a protein known to fold more robustly can increase solubility over that of the native protein (reviewed by Walls & Loughran,

2011), common fusion partners include maltose binding protein (MBP), glutathione-s-transferase (GST), thioredoxin (TRX) and NusA. MBP is the most common solubilisation partner and has been shown to be more effective than either GST or thioredoxin (Kapust & Waugh, 1999). N-terminal fusions have been shown to be more effective for *in vivo* folding, suggesting that initial correct folding of the fusion partner positively influences folding of the RP product (Sachdev & Chirgwin, 1998a). MBP and TRX fusions have been demonstrated to improve refolding of proteins *in vitro* (Sachdev & Chirgwin, 1998b). Samuelsson *et al.* (1994) also demonstrated that refolded human insulin-like growth factor I (IGF-I) showed higher solubility and stability when fused to ZZ, 2 copies of a synthetic IgG binding domain derived from *Staphylococcus aureus* protein A. An additional benefit to some solubility-enhancing fusions is that the fusion partner may enable affinity purification, such as MBP, GST, TRX and ZZ, potentially simplifying DSP (Moks *et al.*, 1987; Walls & Loughran, 2011). While solubility-enhancing fusions are effective they are not universally so, the protein may still mis-fold. The fusion partner often must be first cleaved from the RP for it to be used, creating additional processing stages and the fusion partner will represent a substantial proportion of the total RP produced, therefore decreasing final yields of the cleaved RP product.

Isolation of host organisms that have acquired mutations allowing more efficient production and RPs whose coding sequence has acquired mutations that improve folding have also been reported, such as the work reported by Miroux & Walker (1996) who isolated the strains C41(DE3) and C43(DE3) that showed improved production of a number of membrane and globular proteins over unmodified BL21(DE3). As fluorescent protein-RP fusions have been particularly effective in this regard this is discussed in Section 1.6.5.

Many of these methods for ameliorating stress have analogues among other host species. For example, overproduction of chaperones during RPP in *B. subtilis* has been shown to increase RP solubility and yields (Westers *et al.* 2004), although similar to *E. coli* the effects of chaperone co-production in yeast appear to be mixed (Mattanovich *et al.*, 2004). It appears that decreasing the growth temperature during RPP may be beneficial in yeast (Mattanovich *et al.*, 2004; Gasser *et al.*, 2008) and potentially also in mammalian cell culture (Al-Fageeh & Marchant, 2006). Finally, the selection of high-producing cell lines is a common stage in process development for eukaryotic RPP, it has been possible to use the response of the UPR to guide cell line selection for mammalian cell culture, selecting cell lines with higher tolerance to RPP-related stress (Kober *et al.*, 2012; Du *et al.*, 2013; Hussain *et al.*, 2014).

1.3.7. Minimising RPP-induced stress: The 'improved' protocol

A new strategy has been proposed that, rather than mitigating the effects of the stress responses inherent to a 'standard' RPP protocol (i.e. those described in the manufacturers instructions for *E. coli* BL21* (Life Technologies Corp., 2010), growth at 37°C pre-induction and 25°C post-induction, induced with 0.5mM IPTG), aims to minimise these stresses from the outset (Sevastsyanovich *et al.*, 2009; Alfasi, 2010). Growth conditions were set to decrease physiological stress and the inducer concentration was optimised not on the amount of RP generated or the rate of RP formation, but on the level of protein function observed at the end of the experiment. The temperature was maintained at 25°C throughout the experiment, to favour correct protein folding (Vera *et al.*, 2007). In addition this would also decrease growth and hence protein synthesis rates. A constant temperature was used as opposed to decreasing temperature from 37°C on induction to limit any potential stress involved in

a rapid change of temperature. Induction of cold shock is known to result in a temporary block in translation that is thought, at least in part, to be sensed directly by the ribosome (VanBogelen & Niedhardt, 1990; Yamanaka, 1999). This protocol was originally developed using a 42 kDa model protein comprising *E. coli* CheY (a regulator protein involved in chemotaxis (Hess *et al.*, 1988)) with a C-terminal green fluorescent protein (GFP) fusion (incorporating the S65T red-shift and F64L folding mutations) expressed from the pET20b vector derivative pET20bhc-CheY::GFP (Waldo *et al.*, 1999; Jones *et al.*, 2004; Jones, 2007). CheY is a good model protein for testing the ability to form soluble protein as it does not readily fold into its soluble form under 'standard' RPP conditions. The GFP-fusion allows a simple method to estimate the quantity of soluble protein present via fluorimetry and also allows for monitoring of the process by FCM.

The hypothesis that RPP under conditions aimed to minimise physiological stress can be beneficial for production of CheY::GFP was found to be true. Under stress-minimising conditions (25°C, 8 μM IPTG) total amount of biomass, RP solubility, the percentage of cellular protein that was the recombinant product and therefore specific yield of protein were significantly higher than those produced under 'standard' conditions despite the total amount of T7 polymerase present being too low to be detected by Western analysis. The modified stress-minimising conditions were therefore termed an 'improved' protocol. This protocol was then applied to laboratory-scale batch and fed batch fermentation and it was shown that product yield (in terms of g product per L culture broth) in batch growth increased under improved conditions and that prolonging expression in fed-batch growth under improved further increased product yield over batch (Sevastsyanovich *et al.*, 2009).

The physiological basis for this effect appears to be reduction of the growth arrest in protein producing bacteria described earlier. In the standard protocol culturability,

plasmid retention and the proportion of productive cells as determined by green fluorescence via FCM dramatically decreased immediately post-induction with viable cell numbers only recovering following a lag period and the majority having lost the plasmid. In contrast, during the improved protocol both productive cell counts and plasmid retention remained above 90% and viable cell numbers increased throughout. Despite protein synthesis occurring at a much lower rate, the increased cell viability allowed for continued growth and hence protein synthesis continued for much longer, with the net effect of increased product yield (Sevastsyanovich *et al.*, 2009). This conclusion is supported by a study by Wagner *et al.* (2008) into the 'Walker' mutant strains of *E. coli* BL21; C41(DE3) and C43(DE3) (Miroux & Walker, 1996). It was observed that the reason the mutant strains were able to accumulate soluble protein to a much greater extent than the wild-type was due to the IPTG-inducible promoter controlling the T7 polymerase acquiring mutations that decreased its activity and hence reduced the amount of T7 RNA polymerase produced on induction.

Application of the improved protocol was also shown to be effective for production of not only the soluble, cytoplasmic model protein CheY::GFP. Similarly improved RP yields were obtained for three additional proteins: cytochrome c_2 from *Neisseria gonorrhoeae* that requires periplasmic localisation and significant post-translational modification in the attachment of haem, cytochrome *c* peroxidase (CCP) from *N. gonorrhoeae* (Turner *et al.*, 2003) and D-GFP another non-*E. coli* protein (GSK) (Sevastsyanovich *et al.*, 2009).

The success of the improved protocol in increasing RPP productivity in *E. coli* represents a new paradigm for the design of future processes. When it is identified that a process causes physiological stress, it may be recommended not just to counteract the effects, but also to reduce or remove the stress itself.

1.4. Fermentation & industrial cell culture

The scale of RPP processes can vary greatly, from laboratory uses that may only require milligram to gram batches of the target protein to industrial uses such as the production of pharmaceuticals or industrially relevant proteins that will often require vast quantities of the product regularly, often many kg per year. As such, RPP culture volumes may vary from the millilitre to 100,000s of litres scale. In order to cater to such varieties in scale there exist a wide range of options available for growth of RPP cultures, including culture vessel design, growth strategy and growth media composition.

1.4.1. Culture vessel design

Obviously the design of a culture vessel will alter greatly based on the volume of culture it is intended for and on any specific requirements of the organism being cultured, from a simple shake-flask through to many thousand litre industrial bioreactors. A summary of many variations on these is given below, however for further detail Doran (2013) is recommended.

1.4.1.1. Shake-flask cultures

The simplest culture vessel used for RPP is the shake flask. They are relatively inexpensive, requiring little specialised equipment and due to their low volume they are suitable for screening studies. However, the low maximum volumes preclude large-scale production, and sample taking and making additions cannot be readily automated and require the removal of the culture vessel from the incubator, potentially disturbing growth. This means that there can be no real-time monitoring. Although it is interesting to note adhesive sensors have been developed recently that can be attached to the interior of a flask and can be used to monitor culture parameters such as pH and

dissolved oxygen tension (DOT) (the partial pressure of O_2 dissolved in the culture broth) (Wittman *et al.*, 2003; Schneider *et al.*, 2010; Doran, 2013)

1.4.1.2. Bioreactor/fermenter cultures

Bioreactors or fermenters are vessels specifically designed for cell culture, the most common type for *E. coli* being the stirred-tank reactor. Stirred tank reactors generally consist of a vessel made of glass in the case of smaller-scale laboratory models or photobioreactors and stainless steel for larger scale models. The vessels themselves are stationary, agitation is provided by a central rotating shaft with impellers attached. Baffles are often also included to provide additional turbulence for mixing and to prevent vortexing. Samples are removed through a submerged tube. The culture is aerated by sparging filtered air into the vessel and if necessary this can be supplemented with oxygen. Temperature is sensed by a submerged probe and controlled automatically either by a water jacket around the vessel or a combination of heating mats and cold fingers. Additional control mechanisms are often included; most bioreactors will contain ports to attach a number of analytical probes, the most frequent being those for DOT, pH as these are useful parameters to asses the physiological condition of the culture, and the presence of foam in order to prevent blockage of air filters and culture leakage. These probes are then connected to an external control unit where the outputs can be used to drive control systems and correct observed readings to user-defined set values. The agitation speed and aeration rate can be controlled to regulate DOT, sterile concentrated acid and base solutions can be automatically added to control pH and foaming can be prevented by the automated addition of an antifoam solution. Outputs from these systems (temperature, pH, DOT, agitation speed and addition rates) are often logged into a computer system for later analysis. Vessels will also include ports for the aseptic addition of material, such as for inoculation, or a solution of chemical inducer

such as IPTG (Doran, 2013). Bioreactors hold many advantages over simpler techniques, increased aeration allows the culture volume to be scaled to the many 1000s of litres required for an industrial process. The use of probes to detect, monitor and automatically regulate culture parameters allows the operator to be more informed about the process and for the process itself to be more reliable and reproducible. The main disadvantages of bioreactor culture are that it requires specialised equipment, due to the larger volumes it is inherently more expensive per run and hence it is more difficult to multiplex experiments especially for screening studies. Also, the complexity of the equipment requires not only significantly more time for vessel set-up and shut-down, but operators must be trained to a higher level.

1.4.1.3. Additional features available for bioreactors

As with any technology, there are many additional and variant features available for bioreactors, those more frequently used include: the analysis of exhaust gas components; often by gas-mass spectrometry (MS); probes for parameters such as glucose concentration (Phelps *et al.*, 1995) and fluorescence (Randers-Eichorn *et al.*, 1997; Jones *et al.*, 2004); additional control systems, such as controlling feeding rate by pH or DOT; and if connected to a network the ability to monitor and operate the vessel remotely. A more recent variation, disposable vessels, are finding popularity in industry as cleaning costs and requisite down-time associated with re-usable vessels is reduced, especially for mammalian cell culture (reviewed by Eibl *et al.*, 2010). A final innovation in many ways contradicts previously stated points, a scaled-down microtitre plate fermentation system has been developed called the Biolector (M2P Labs, Germany). Based on a microtitre plate reader, it uses modified microplates allowing parallel monitoring of up to 96 individual cultures, detecting OD, pH, O_2, CO_2 and fluorescence. It is possible to grow in both batch and fed batch mode and was demonstrated to be

scalable to laboratory scale fermentation (1.4 L) (Kensy *et al.*, 2009a, 2009b; Scheidle *et al.*, 2010).

1.4.2. Cultivation techniques

1.4.2.1. Batch growth

Batch growth is the simplest cultivation technique available and involves a single step process where the growth medium is inoculated and the culture allowed to grow. Batch cultures can be grown at all scales from shake-flask to production-scale vessels but they rely only on the medium components initially provided and as such the maximum biomass that can be produced is limited.

1.4.2.2. Fed-batch growth

Industrial processes generally require large amounts of biomass to be produced and it would be more efficient if this is can be done in as small a culture volume as possible, reducing either the capacity (and hence, space and cost) of the bioreactors used or the number of runs necessary and the volume of waste water produced (Choi *et al.*, 2006). It is possible to produce concentrations of biomass vastly in excess of those possible in batch processes with the addition of concentrated, sterile nutrients following depletion of the initial (batch-phase) medium. Most fed-batch growth strategies operate under carbon-limiting conditions, feeding additional carbon source, but this can be supplemented by additional metabolites. The main use of fed-batch cultures are in bioreactors where feed can be added automatically, requiring little operator intervention.

1.4.2.3. Continuous culture

A final culture method is the continuous stirred-tank reactor (CSTR) or chemostat. A culture is grown and maintained in steady-state growth at a fixed volume by

simultaneous feeding and removal of culture. While CSTRs hold some advantages such as the removal of potentially limiting waste products, they are not often used for RPP processes, the long run-times increasing risk of failure and likelihood of increased genetic variability. Industrially CSTRs are used in the growth of bakers yeast, brewing and waste processing (Doran, 2013), but their main application is in research for studying the behaviour of cultures in a steady-state such as Sunya *et al.* (2012, 2013); who determined the response of *E. coli* to glucose pulses in steady-state chemostat cultures.

1.4.3. Growth media

A final consideration for industrial cell culture is the composition of growth media. This can be divided into two areas, the growth medium itself and the carbon source.

1.4.3.1. Growth media

Microbial growth media are classified on the extent to which they include defined chemical components or biologically-derived complex components such as yeast extract (powdered autolysate of yeast), tryptone (tryptic hydrolysate of casein) and peptone (enzymatic autolysate of milk, meat or soy protein).

Complex media consist of predominantly complex components. The staple laboratory medium Lysogeny Broth (LB) (also known as Luria, Luria-Bertani or Lennox broth), as a complex medium, contains only tryptone, yeast extract and NaCl. All nitrogen, vitamins and trace elements derive from tryptone and yeast extract. Sometimes it may be necessary to supplement complex media for example as a buffer or to counteract a deficiency in a trace metabolite. In contrast, defined media consist solely of simple, non-biologically derived components such as salts. Semi-defined media contain aspects of both.

Complex media are commonly used in a laboratory setting due to their simplicity of production and the generally good growth achievable, but they have two major limitations for industrial application. First, being biologically derived, complex media components are inherently variable in content between manufacturer and product batch. Industrial processes require reproducibility and therefore any variability introduced by using a different batch of media components would be undesirable. Second, many complex components such as tryptone and peptone are animal-derived and their use will introduce additional regulatory concerns for pharmaceutical production.

1.4.3.2. Carbon source

Carbon is a necessary element for biomass and while complex media will contain sufficient carbon to sustain some growth, for efficient biomass yields it is necessary (for non-photosynthetic cultures) to supplement carbon. The two most common carbon sources for industrial growth of *E. coli* are glucose and glycerol, although others are available.

Glucose appears to be the most commonly used carbon source in the culture of *E. coli* (Choi *et al.*, 2006). It is relatively inexpensive and is consumed preferentially to other carbon sources in wild type cells (Lee, 1996; Overton, 2014). Glucose does possess some disadvantages, for example, under oxygen limited conditions, it is subject to mixed-acid fermentation producing in varying quantities acetate, lactate, formate, succinate, ethanol and hydrogen. When glucose concentrations, even in highly aerated cultures, exceed the cell's requirements (>5 g·L^{-1} (Lee, 1996)) overflow metabolism is triggered, producing acetate. The overall effect of these phenomena is culture acidification and reduction in the efficiency of biomass conversion, leading to perturbation of growth and productivity (Lee, 1996). Additional considerations with the use of glucose are that it can react with

amine groups in Maillard reactions and therefore must be autoclaved separately to the growth medium, and that as a solid it must generally be dissolved for feeding, diluting the vessel's contents. This was addressed by Matsui *et al.* (1989) who devised an apparatus that fed powdered glucose, hence limiting dilution.

Glycerol is the second most commonly used carbon source in *E. coli* culture and holds many advantages over glucose. It is not as readily converted to acetate and therefore higher concentrations may be used, allowing longer batch growth and less strenuous monitoring of feeding (although acetate accumulation has been observed using glycerol in fed-batch cultures (Korz *et al.*, 1995)). It does not undergo Maillard reactions and may be autoclaved with the growth medium, removing an addition step, decreasing the risk of contamination. As a liquid it can be fed directly into the culture, with minimal dilution. However, this is complicated by its viscosity and therefore it is most commonly used in solution, albeit at higher concentrations than is possible with glucose. Glycerol is however more expensive than glucose and growth rates are typically lower (Lee, 1996).

Other carbon sources sometimes used include lactose and fructose. Lactose is an interesting case in RPP cultures as it can act not only as carbon source but also can be interconverted into its isomer allolactose, the natural effector molecule of LacI and hence induce RPP in LacI-mediated systems (Figure 1.2). This was used by Humphreys *et al.* (2002) who, to induce RPP in a fed-batch fermentation, changed the carbon source from glycerol to lactose. The use of fructose was argued for by Aristidou *et al.* (1999) as a potential rival to both glucose and glycerol as fructose, like glycerol, does not cause growth-limiting levels of acetate production and it is only marginally more expensive than glucose (and hence cheaper than glycerol). However, as a solid, it retains the related limitations of glucose and in addition can convert to glucose on autoclaving, requiring filter-sterilisation and therefore additional cost.

1.4.4. Strategies for fed-batch RPP in *E. coli*

It is economically advantageous to grow cells in an industrial bioprocess to high densities, generally through fed-batch fermentation. There have been numerous studies aimed at optimising high cell density culture (HCDC) in *E. coli* both solely for biomass and also in conjunction with RPP, summarised below.

1.4.4.1. HCDC techniques

In an ideal HCDC fermentation the only limitation on growth would be the theoretical maximum cell density in liquid suspension. This is estimated at between 200-400 g·L^{-1} DCW (Reisenberg *et al.*, 1991; Märkl *et al.*, 1993). As Mori *et al.* (1979) demonstrated that over 200 g·L^{-1} DCW viscosity sharply increases with increasing cell density a theoretical maximum of approximately 200 g·L^{-1} for a stirred tank reactor is logical (Lee, 1996). To corroborate this, cell concentrations in excess of 100 g·L^{-1} are frequently reported (Lee, 1996; Choi *et al.*, 2006). There does not appear to be a single 'ideal' strategy for HCDC, successful instances are reported with numerous differences although certain commonalities do occur. Few successful HCDC strategies use complex media, the majority opting for semi-defined or defined (Lee, 1996; Choi *et al.*, 2006), presumably this is because of the variability and lack of control inherent in complex components. The majority of HCDC strategies use glucose as carbon source although a significant minority use glycerol (Lee, 1996; Choi *et al.*, 2006). The most variable aspect of HCDC appears to be feeding strategy. Lee (1996) summarises the options into methods without feedback control, methods with indirect control and methods with direct control. In methods without feedback control the feed rate is determined independently of sensor outputs, including constant feeding, increased feeding and exponential feeding for a fixed specific growth rate (µ). In methods with indirect feedback control the feed rate is controlled by sensor outputs indicative of substrate

concentration such as the DOT (a DO-stat), pH (a pH-stat), carbon dioxide evolution and cell concentration such that the feed rate maintains a constant value of the detected parameter and therefore likely, also a constant substrate concentration. In direct feedback control the feed rate is determined directly from the substrate concentration via an on-line probe allowing a constant substrate concentration.

The major limitation of HCDC is that, especially in the later stages of culture, mixing and aeration efficiency decreases, causing transient anaerobiosis (Garcia *et al.*, 2009). This was partly addressed by Soini *et al.* (2008) who demonstrated that common HCDC media are designed solely for aerobic growth and supplementation with trace elements typically only necessary for anaerobic growth such as selenium, nickel and molybdenum can be beneficial to HCDC.

1.4.4.2. HCDC for RPP

While it has been observed that yields of recombinant protein per unit biomass can be higher in shake-flask cultures (Jeong & Lee, 1999; Choi *et al.*, 2006) the vast limitations of the format for industrial processes result in economic benefits for HCDC of potentially lower-producing cells due to the quantity of biomass that can be produced. HCDC protocols for RPP are generally similar to those for other purposes such as for biomass accumulation, the primary difference being RPP induction. For shake-flask RPP of a soluble, non-toxic protein induction is recommended at mid-log phase (optical density (OD_{650}) ~0.4) harvesting 2-3 h later (Life Technologies Corporation, 2010) typically reaching a final OD_{650} of ~10. A single HCDC-RPP run will typically last for 20-48 h or possibly longer (Lee, 1996), this therefore raises the question of when should RPP be induced in HCDC. It has been demonstrated that RPP can dramatically affect growth rates therefore late induction would limit the period cells are subject to growth-limiting conditions, allowing more rapid biomass accumulation and would therefore result in

shorter run-times and decrease overall process costs. RPP can also reduce viability and plasmid retention, which may result in a limited productivity window, late induction therefore allowing maximal cell counts. On the other hand sufficient time must be allowed for accumulation of an economically viable quantity of RP. An effective induction strategy for HCDC-RPP should balance these two competing factors and it appears that most studies opt for separating the growth and production phases with induction at relatively high biomass concentrations (Choi *et al.*, 2006). The dichotomy of protein production and biomass yield was highlighted by Want *et al.* (2009) who analysed production of an antibody Fab (fragment antigen binding) fragment using HCDC inducing at three different points: at high biomass during the fed-batch phase (OD$_{600}$ ~50); and at two points during the transition from batch to fed-batch growth (OD$_{600}$ ~21). Earlier induction resulted in decreased biomass formation, but a concomitant increase in Fab production; the earliest induction point showed a ~50% decrease in peak biomass (DCW per L culture broth) and 4-fold increase in Fab production (concentration of Fab in culture broth) when compared to the latest (16 g·L^{-1} versus 25 g·L^{-1} biomass and 1 µg·mL^{-1} versus 4.1 µg·mL^{-1} Fab, respectively). The HCDC-RPP fermentation protocol used by Sevastsyanovich *et al.* (2009) when reporting the improved protocol induced RPP very early, in mid-logarithmic growth (OD$_{650}$ ~0.5). It is possible that the lower RP synthesis rates achieved during the improved protocol favour increased production times for sufficient product formation.

1.4.5. Fermentation development

Many factors may influence the outcome of a fermentation for example; temperature, pH, DOT, medium, feeding regime etc., all of which must be examined to optimise the fermentation protocol. For an RPP process this extends to include, for example; host

strain, inducer, inducer concentration, induction point and harvest time. Traditionally fermentation development involved a trial and error approach examining a single or limited number of features, this is not only time consuming but also does not account for combinatorial effects (Papneophytou & Kontopidis, 2014). This has however began to be superseded by Design of Experiments (DoE), a more systematic methodology allowing analysis of multiple parameters within a single series of experiments. A number of variables at an appropriate range of values are combined in a factorial or fractional factorial manner to produce a reduced number of experiments from which the effects of individual parameters can be calculated statistically (Bora *et al.*, 2012; Papneophytou & Kontopidis, 2014). The reduced number of experiments required represents a saving of both time and money during development. DoE methodologies have been successfully applied to RPP-fermentation process development in *E. coli* (Swalley *et al.*, 2006; Islam *et al.*, 2007; Coutard *et al.*, 2008) and with the development of microscale fermentation equipment such as the Biolector (Kensy *et al.*, 2009a, 2009b; Scheidle *et al.*, 2010) (Section 1.4.1.3) its use will continue to increase.

1.5. Flow cytometry & its applications in biotechnology & bioprocessing

1.5.1. Heterogeneity and the importance of single cell analysis

Flow cytometry (FCM) is a single-cell analysis method, in that it detects and analyses individual cells as opposed to bulk measurement methods that provide values averaged over a population. While the most common parameters detected by FCM are similar to existing bulk measurement techniques such as OD and fluorimetry, the utility of FCM is in its ability to detect culture heterogeneity. Even in cultures of relatively 'simple' organisms such as microbes there exists significant levels of heterogeneity within members of a population (reviewed by Davey & Kell, 1996; Díaz *et al.*, 2010). Davey &

Kell (1996) identified three major causes of culture heterogeneity. First, the position of the cells in the cell cycle, as cultures are rarely in synchronous growth. Heterogeneity is also caused by physiological differences as a result of the cellular environment, of particular relevance for sub-optimally mixed bioreactors and also as a result of stochasticity in gene expression, thought to allow a range of phenotypes in order to more readily respond to a change in environmental conditions (Thattai & van Oudenaarden, 2004). Cultures are also eventually likely to show genetic variability as DNA replication in *E. coli* is not perfect and every division event holds a small possibility of a mutation arising. Davey & Kell (1996) estimate that after 1 generation 99.5% of cells remain wild-type. Considering that a large-scale fed-batch fermentation may have many litres of high cell density culture (OD >100), all derived from a single initial transformant, the scope for mutational variability is substantial (assuming OD 1 $\sim 10^9$ cells·mL^{-1}, 1 L of an OD 100 culture will contain approximately 10^{14} cells, equivalent to at least 47 generations and, assuming no selective pressures, would be expected to result in approximately 79% wild-type cells). This is attested to in the success of studies involving forced-evolution and isolation of mutant strains (Waldo *et al.*, 1999; Alfasi *et al.*, 2011). In addition to these causes, RPP is likely to introduce heterogeneity due to cell viability, plasmid loss and IB formation.

Reproducibility and reliability are of great importance for successful bioprocesses, particularly at industrial scale. As bulk measurement techniques provide only an average value it is not possible to detect changes in the distribution of a parameter. For example, a single bulk fluorescence value from fluorescent protein (FP)-producing cells may correspond to a single population of modest fluorescence or a mixture of high producers and non-producers and any gradation between. Being able to detect such a

situation is of great utility in bioprocess monitoring and herein lies the benefit of single cell analysis.

1.5.2. FCM

FCM was originally developed for the US Army during the Second World War by Gucker *et al.* (1947). Light scatter and a sheath of filtered air was used to measure smoke particles for testing gas masks and it was also found to be effective in detecting airborne bacterial spores (Gucker & O'Konski, 1949; Ferry *et al.*, 1949; Shapiro, 2005). Further advances occurred in the 1950s with Wallace H. Coulter developing what would become the Coulter Counter a device which measured the electrical impedance of cells in liquid suspension as they passed through a narrow aperture (Davey & Kell, 1996). Since then FCM has become a powerful technique for the analysis of biological samples at the single cell level and has been shown to have great potential utility in biotechnology and bioprocessing (Hewitt & Nebe-Von-Caron, 2001; Reiseberg *et al.*, 2001; Mattanovich & Borth, 2006; Want *et al.*, 2009; Díaz *et al.*, 2010; Broger *et al.*, 2011; Gatza *et al.*, 2012). As the name suggests flow cytometry involves the measurement of cells as they flow past a detector. The following section will provide an overview of FCM, for further information on the subject the reader is recommended the extensive work by Shapiro (2005), in addition there are many resources regarding FCM available on the website of the Purdue University Cytometry Laboratories http://www.pucl.purdue.edu.

1.5.3. Components and function of a flow cytometer

Modern flow cytometers typically operate utilising light-scattering principles as in the Gucker particle counter and in liquid suspension. Schematic representations of key components of the cytometers used in this study are summarised in Figure 1.5. Flow
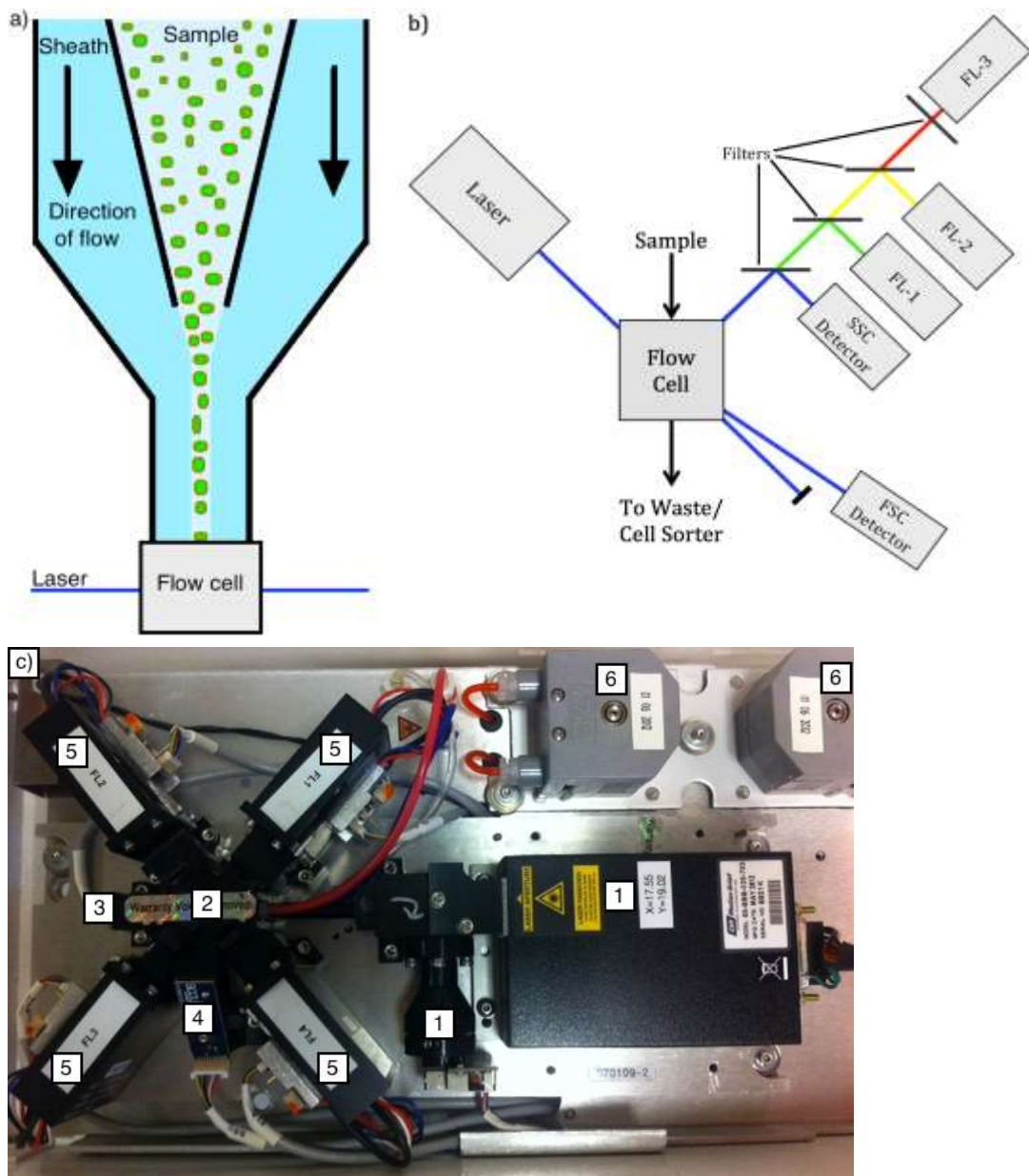
Figure 1.5: Operation of a typical flow cytometer.
a) Sample insertion and hydrodynamic focusing: sample leaves nozzle into stream of sheath fluid, achieving laminar flow and aligning cells in single file before passing to the flow cell. b) Optical system of a typical flow cytometer: Cells interact with laser at the flow cell, forward scatter is detected at <10° from incidental light with incidental light masked, light at 90° passes through a series of filters allowing detection of side scatter and fluorescence. c) Optical & fluidics systems from the BD-Accuri C6: optical detectors are spaced around flow cell (Key: 1 – lasers, 2 – flow cell, 3 – FSC-detector, 4 – SSC-detector, 5 – fluorescence detectors, 6 – peristaltic pumps).

cytometers can generally be divided into three main sections: fluidics, optics and electronics/computational.

1.5.3.1. Fluidics

The fluidics system of a flow cytometer is involved in the delivery of cells into the detection (optics) system. Cells suspensions are passed into the centre of a channel of pressurised sheath fluid such that laminar flow is achieved and the sample is hydrodynamically focused, with the result that particles pass in single-file through the flow cell (Figure 1.5a). Once in the flow cell, particles pass through the detector and then flow either into waste collection or in more specialised cytometers a cell sorter or imaging system.

1.5.3.2. Optics

The detection systems in most flow cytometers are light-based involving a laser and optical sensors. Once cells reach the flow cell they pass through a laser beam and the interaction of the cell and its contents with the laser is measured. Two interactions are generally detected: light scattering and fluorescence emission. In order to detect fluorescence the operational wavelength of the laser will be set to the excitation wavelength of common fluorophores e.g. ~ 488 nm for enhanced GFP (EGFP) and fluorescein isothiocyanate (FITC) and will include filters and detectors at relevant emission wavelengths either in series with the side scatter (SSC) detector (Figure 1.5b) or spaced around the flow cell (Figure 1.5c).

1.5.3.3. Electronics & computational

The final common component in a flow cytometer is the electronic or computational system. Scattered or emitted light is first converted into an electrical signal by photomultiplier tubes or photodiodes then converted into numeric data by the included computer. Should a peak in the detector output fall inside user-defined controls to

eliminate intrinsic sample noise (termed a threshold or discriminator) it is identified as an event and recorded. Data from each event is recorded from all output parameters (light scattering, fluorescence, signal length, pulse width, timestamp) so that it is possible to perform multivariate analysis between all parameters recorded. The computer system will also control other aspects of the cytometers operation such as control of the fluidics. An additional use of computing is for gating data, the highlighting and isolation of a particular sub-population for further analysis.

1.5.3.4. Additional features

Many cytometers also include additional features, most commonly fluorescence activated cell sorting (FACS). FACS was originally developed by Fulwyler (1965) and involves the selection of a cellular population based on the data gathered as the cells passed through the laser and its isolation (sorting) from the remainder of the sample for further analysis. Having passed through the laser the sample flow is divided by agitation into droplets that due to the dilution of cells used will be statistically likely to contain individual cells which can then be 'sorted' based upon user-defined properties. If a cell falls within the sort conditions its droplet is electrostatically charged and directed by electrically charged deflection plates into collection tubes; cells outside of the sort conditions remain uncharged and fall into the waste (Davey & Kell, 1996). Cell sorting is of great utility when particular cells are required for further use/analysis such as for the isolation of high-producing mutant strains as described in Section 1.6.5.

Another feature that has been incorporated into some flow cytometers is cellular imaging, where a photomicrographic image of a cell is also recorded as it passes through the machine. This is available on cytometers such as the Cytosense and its variants produced by CytoBuoy for analysis of marine phytoplankton and the Amnis FlowSight

(Millipore) (examples of research using these machines have been reported by Rutten *et al.* (2005) and Grimwade *et al.* (2012), respectively).

1.5.4. Flow cytometric data & its applications

1.5.4.1. Light scattering

The primary data output in FCM is the amount of light scattered by a cell, (Figure 1.6). Forward scatter (FSC) or forward angle light scatter (FALS) is detected at angles typically <10° from that of the incidental laser light. It has been shown that FSC is primarily dependent on cell size, volume or biomass and has been used to predict these (Koch *et al.*, 1996). Many studies however have shown this relationship is not always directly proportional and can be affected by many other factors including the shape, structure, refractive index and chemical composition of the cell and also the design of the instrument itself (Wållberg *et al.*, 2005). In some cases it was found that mathematical modelling could be used to predict biomass from FSC (Vives-Rego *et al.*, 2000). SSC or right angle light scatter (RALS) is detected at 90° to the incidental light and is primarily indicative of intracellular complexity or granularity (Davey & Kell, 1996).

The primary use of light scatter in FCM is to determine the morphology of cells (Hewitt & Nebe-Von-Caron, 2001; Shapiro, 2005). Cells are initially identified from background noise (thresholding) either based on FSC alone or in combination with other parameters such SSC and fluorescence (Shapiro, 2005). It is also possible to differentiate cell types based upon scatter if there is sufficient morphological difference, useful for detecting contamination (Hewitt & Nebe-Von-Caron, 2001). It is relatively simple to differentiate between, for example, mammalian, yeast and bacterial cells based upon size alone however it is also possible to differentiate between different species of bacteria. *E. coli*
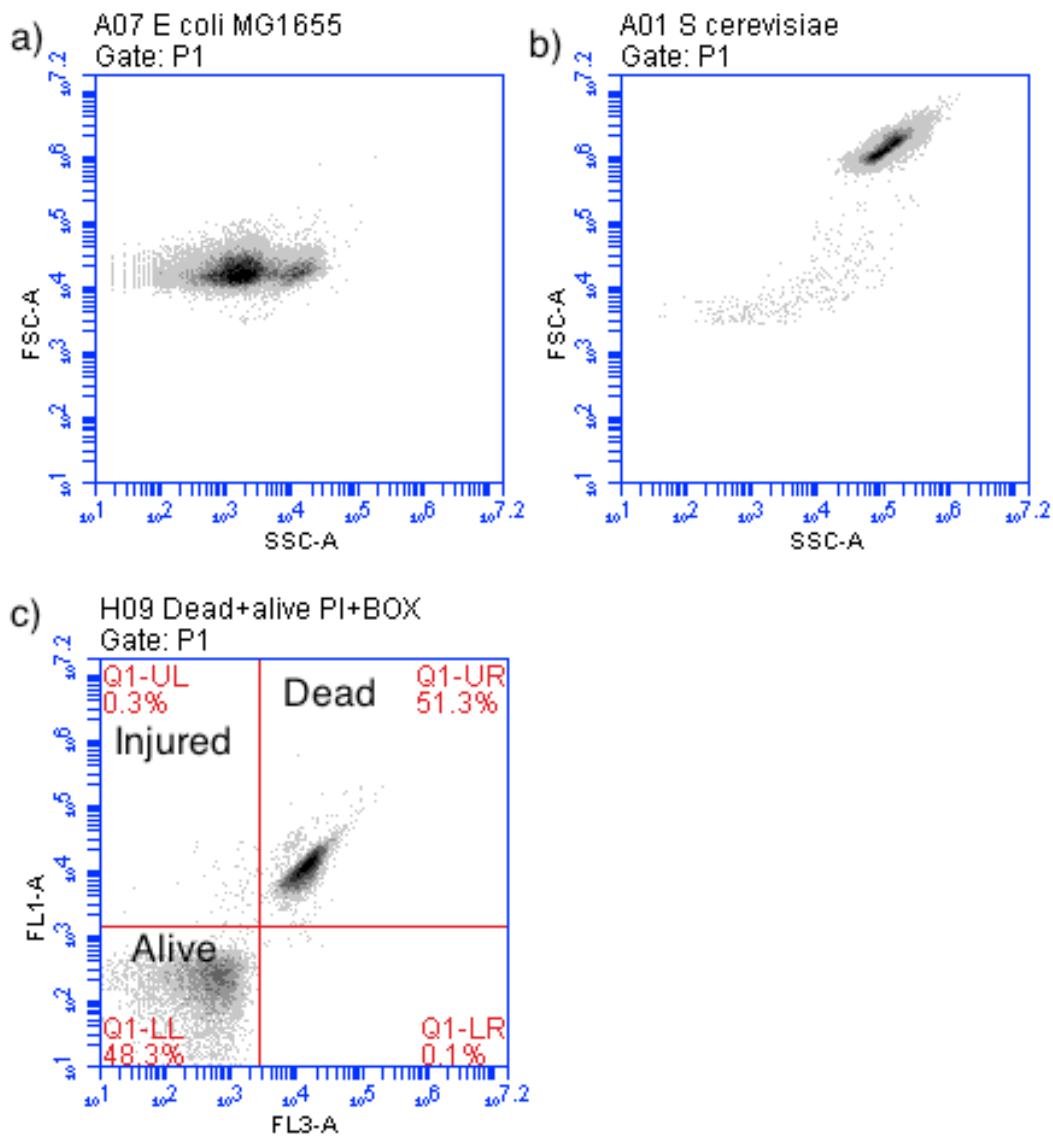
Figure 1.6: Sample FCM data

a) *E. coli* MG1655 FSC/SSC plot showing distinctive bimodal SSC distribution of rod-shaped cells. b) *S. cerevisiae* FSC/SSC plot showing higher FSC and SSC signals than *E. coli* caused by increased cell size and complexity. c) FL1 (green fluorescence)/FL3 (red fluorescence) quadrant plot for a mixture of living and ethanol-killed *E. coli* BL21 stained with propidium iodide (PI) (red) and bis-oxonol (DiBAC$_4$(3)) (green), quadrants are labelled based on viability of cells within (Alive – PI$^-$ BOX$^-$, Injured (cells with depolarised membrane potential) PI$^-$ BOX$^+$, Dead PI$^+$ BOX$^+$). All data generated using a BD-Accuri C6 flow cytometer.

as a rod-shaped bacterium displays a bimodal distribution in SSC but not FSC (Figure 1.6a). While this may appear on initial inspection to be two separate sub-populations Hewitt & Nebe-Von-Caron (2001) showed that following sorting and re-analysis identical distributions were obtained from both populations and was attributed to the rod-shaped morphology of *E. coli*, the two populations being due to cells passing through the laser in different orientations relative to their length and width. This is supported by the observation that coccoid species such as *Rhodococcus* produce unimodal distributions.

A final use of light scatter of particular relevance to this study involves the detection of IBs. Lewis *et al.* (2004) reported that *E. coli* producing a model mammalian protein AP50 displayed an increase in both FSC and SSC when the protein formed IBs and suggested that this effect may be exploited as a diagnostic technique. Hedhammar *et al.* (2005) used EGFP fusions to demonstrate a positive correlation between the standard deviation of FSC signals and the formation of IBs (along with green fluorescence to indicate the amount of soluble protein formed). Wållberg *et al.* (2005) however, concluded that light scattering alone was not sufficient to identify promegapoietin IBs in *E. coli* and found it necessary to use a fluorescent antibody. The authors concluded that this may be due to process-specific effects or more likely due to differences in FCM optical set-up. It could therefore be concluded that while light-scattering may provide information on IB formation it appears to be case-dependent.

1.5.4.2. Fluorescence

While light scatter can provide much information, it is the inclusion of fluorescence detection that has made FCM as powerful an analytical technique as it now is. The advantages of fluorescence detection are first that there are a large number of fluorescent species that may be used and detected, from the native autofluorescence of

cellular components to fluorescent proteins and a wide range of fluorescent dyes. There is also a wide spectrum of fluorescence that can be detected, allowing multiplexed experiments analysing multiple aspects of cellular function on different fluorescence channels (Shapiro, 2005).

1.5.4.3. Autofluorescence & its uses

All cells exhibit autofluorescence to varying extents, which can be exploited for analysis (Shapiro, 2005). One of the most widely used applications in microbial FCM is in the analysis of marine phytoplankton due to the presence of fluorescent pigments such as chlorophyll and phycoethrythrin (Vives-Rego *et al.*, 2000).

1.5.4.4. Fluorescent dyes & their uses

A major advantage of FCM is the range of fluorescent dyes that are available for use and its ability to detect multiple coloured dyes in a single experiment. Due to the sheer number of dyes available and myriad uses thereof it would be impossible to summarise them fully here and as such only a limited number are discussed. Some of the most common dyes of use in microbial biotechnology are summarised in Table 1.5 (reviewed Veal *et al.*, 2000; Díaz *et al.*, 2010) and for further information on the variety of fluorescent dyes available the reader is directed towards the Molecular Probes™ Handbook (Johnson & Spence, 2010).

Probably the most common use of FCM and fluorescent dyes is for determining cellular viability, often using a dual staining system of propidium iodide (PI) and bis-oxonol (BOX). PI is a red-fluorescent DNA binding dye to which the cell membrane is impermeable. When bound to DNA PI shows a 20-30 fold enhancement in fluorescence, the excitation maximum shifts ~30-40 nm towards the red and the emission ~15 nm towards the blue (excitation 530 nm, emission 625 nm when bound to DNA, an additional stronger excitation peak exists at ~300 nm however this is less commonly

Table 1.5: Summary of fluorescent dyes commonly used in microbial FCM

Adapted from Díaz *et al.* (2010), additional information from Johnson & Spence (2010) and Esparagó *et al.* (2012)

Type of dye/fluorophore or feature detected	Dye / Fluorophore		Excitation/emission maxima (nm)
Nucleic acids	DNA, RNA	Ethidium bromide	510/595
		Propidium iodide	300,536/623
		Cyanines — TO-PRO	515/533 (TO-PRO 1) 462/661 (TO-PRO 3)
		Cyanines — TOTO	514/533 (TOTO 1) 642/660 (TOTO 3)
		Cyanines — SYTO	485-508/498-527
		Cyanines — SYBR	467/520 (SYBR Green)
	DNA	DAPI	350/470
		Hoechst 33342	350/461
		DRAQ5	647/700
Lipids		Nile Red	551/636
		BODIPY	503/512
Fluorogenic substrates		ChemChrome (CY, CB, CV6)	488/520
		FDA, CFDA, CFDA/SE, CFDA-AM	492/519
		Calcein-AM	494/519
		CTC	450/630
		FUN-1	480/580
Intracellular pH		BCECF-AM	482/520
		SNARF-A	510-580/587-635
Membrane potential		Rh123	507/529
		$DiOC_6(3)$	484/501
		$DiOC_2(3)$	482/497
		$DiBAC_4(3)$	488/525
		JC-1 (mitochondrial)	498, 593/525, 585
Cytoplasmic Ca^{2+}		Indo-1	361, 330/405, 480
Antibodies or oligonucleotides		PE, PE-Cy5 conjugates	490/575, 690
		Alexa Fluor 488	495/519
		Oregon Green	496/524
Autofluorescent proteins		GFP & derivatives	Various (488/510 for S65T)
IBs/amyloid		Thioflavin-S	375/455

utilised). As the cell membrane is impermeable to PI, it will only be able to enter a cell and bind DNA when the cell membrane is compromised and as such is used as a stain for dead cells where dead cells fluoresce red (Johnson & Spence, 2010). BOX refers to a family of anionic, lipophilic dyes that are able to cross the cell membrane, however due to their charge are excluded by cells with an active membrane potential and hence will only cause a cell to become fluorescent when either the cell is membrane compromised (dead) or membrane depolarised (termed 'injured' cells). The most common variant used is $DiBAC_4(3)$ that fluoresces green (excitation 493 nm, emission 515 nm) and as such is commonly used with PI as a dual-staining system as illustrated in Figure 1.6c (Johnson & Spence, 2010). Other variants of BOX exist with different fluorescence maxima to accommodate alternate secondary stains. PI/BOX staining has been used to successfully monitor RPP in *E. coli* fermentations many times (Hewitt & Nebe-Von-Caron, 2001; Lewis *et al.*, 2004; Want *et al.*, 2009). An alternative viability staining method described by Lehtinen *et al.* (2004) uses GFP and PI to determine dead cells where loss of GFP fluorescence also indicates cell death. Flow cytometric viability determination holds many advantages over the more traditional techniques, namely counting colonies grown on agar plates. Plate counting is a very limited technique in that, first it requires a minimum of approximately 16 h growth to generate a result and second it provides only a single parameter for defining cellular 'life', that of culturability on agar. FCM-based techniques for enumerating viable cells are comparatively rapid, generally requiring only a few minutes incubation with the dyes and they provide a much more detailed indication of cell viability. What defines the terms alive and dead at a cellular level is not absolute (Davey & Kell, 1996). It is obvious that a cell with an intact membrane, actively respiring and able to undergo cell division can be called alive and this is the basis of colony forming units (CFU) counts. Cells with a ruptured membrane,

as indicated by fluorescence of membrane-excluded DNA-binding dyes, are typically defined as dead (although Votyakova *et al.* (1994) demonstrated that *Micrococcus luteus* cells stained with PO-PRO-3, a membrane-excluded DNA-binding dye usually used as a dead cell stain, were able to be resuscitated). Cells may also exist between these two states. Many terms have been given to these intermediate viability states, however one of particular relevance is that of the viable but non-culturable (VBNC) phenotype. VBNC cells cannot form colonies on agar plates (and are therefore undetected by conventional enumeration) but retain metabolic activity and may still contribute to the culture (Davey & Kell, 1996; Nebe-Von-Caron *et al.*, 2000; Díaz *et al.*, 2010). It is also possible with many flow cytometers to count cells, either with the use of ratiometric beads or directly by measurement of sample volume. A simple comparison of FCM-derived absolute cell counts with CFU counts can easily determine the proportion of the total cells that are unable to grow on agar.

In addition to viability staining an interesting use of fluorescent dyes for RPP has recently been reported by Esparagó *et al.* (2012) who expressed amyloid-β (A-β) peptide in *E. coli* and were able to positively identify cells containing A-β IBs using the amyloidophilic fluorescent dye thioflavin-s (Th-S).

1.5.4.5. FPs & their uses

Fluorescent proteins have found many applications within microbial FCM and more generally within microbial biotechnology as a whole (Vizcaino-Caston *et al.*, 2012), these applications are discussed in greater detail below in Section 1.6.

1.5.5. The benefits of FCM

As has been demonstrated single-cell measurements and specifically FCM provide in a single technique multiparametric data that can include information regarding cell count,

morphology, viability, contamination, RP product formation and general culture heterogeneity, significantly more data than can be generated by any single conventional analytical technique. With a trained operator data can be generated within 5-10 minutes of a sample being removed from a culture vessel, unlike many conventional techniques that allow analysis many hours later. The flexibility and rapidity of FCM is ideal for a bioprocess, allowing the operator considerably more data on which to make a control decision. These benefits only stand to increase: Novel technologies such as automated sampling and preparation systems can allow regular samples to be taken from the culture vessel, prepared and analysed without the presence of an operator allowing real-time culture monitoring. An example of this is reported by Gatza *et al.* (2012) where batch fermentations of GFP-expressing *E. coli* in either LB or M9 minimal media were automatically monitored throughout the length of the fermentation (8 h LB, 30 h M9) using a BD-Accuri C6 FCM coupled with an MSP M5000 FlowCytoPrep automated sampler. Finally, it is conceivable that improvements to process control software could see direct integration of automated FCM data into bioreactor feedback and control mechanisms.

1.6. GFP & its applications in biotechnology & bioprocessing

1.6.1. The history of FPs

The green fluorescent protein (GFP) of *Aequorea victoria* (*av*GFP), a species of deep-sea jellyfish, was first discovered and characterised by Shimomura *et al.* (1962) as a companion protein to the Ca^{2+}-dependent blue-emitting chemiluminescent aequorin. The blue light emitted by aequorin excites the GFP chromophore with the result that *A. victoria* appears green not blue (Morin & Hastings, 1971; Shimomura *et al.*, 1974). It was not until the 1990s however that key breakthroughs would allow its exploitation within

research. First, the gene sequence of GFP was cloned and sequenced by Prasher *et al.* (1992) then recombinant GFP was cloned into, and produced in exogenous hosts by Chalfie (1995) (*E. coli* & *Caenorhabitidis elegans,* also showing the first use of GFP as a gene reporter and for protein localisation *in vivo*) and Inouye & Tsuji (1994) (*E. coli*). Finally, screening of mutant GFPs yielded variants with altered and improved spectral properties, including a blue fluorescent species, that greatly extended the application of *av*FPs (Heim *et al.*, 1994, 1995). The use of autofluorescent proteins has since become a mainstay of research in the biological sciences (March *et al.*, 2003; Su, 2005; Vizcaino-Caston *et al.*, 2012). The extent to which FPs have been incorporated in research was reflected in the decision in 2008 to award the Nobel Prize for Chemistry to Shimomura, Chalfie and Tsien "for the discovery and development of the green fluorescent protein, GFP" (Zimmer, 2009).

1.6.2. How FPs work

There are 2 main forms of FPs in use: those derived from *A. victoria* (the archetypal GFP & derivatives): and those derived from coral of the genus *Discosoma* such as dsRed and the mFruit series of proteins (Shaner *et al.*, 2004; Shaner, 2013). In this study the focus will be primarily on the *Aequorea*-derived FPs, specifically *av*GFP.

Wild-type GFP (wtGFP) is a soluble, 238 amino acid, 27 kDa protein comprised of an 11-strand β-barrel surrounding a central α-helix to which the chromophore is attached. The chromophore is a 4-(*p*-hydroxybenzylidene)imidazolidin-5-one formed directly out of three amino acids, in wtGFP S65-Y66-G67, in a multiple stage oxygen-dependent reaction, summarised in Figure 1.7a.

wtGFP has excitation and emission maxima of 395-397 nm and 504 nm respectively; however, mutations to amino acids surrounding the chromophore or the chromophore
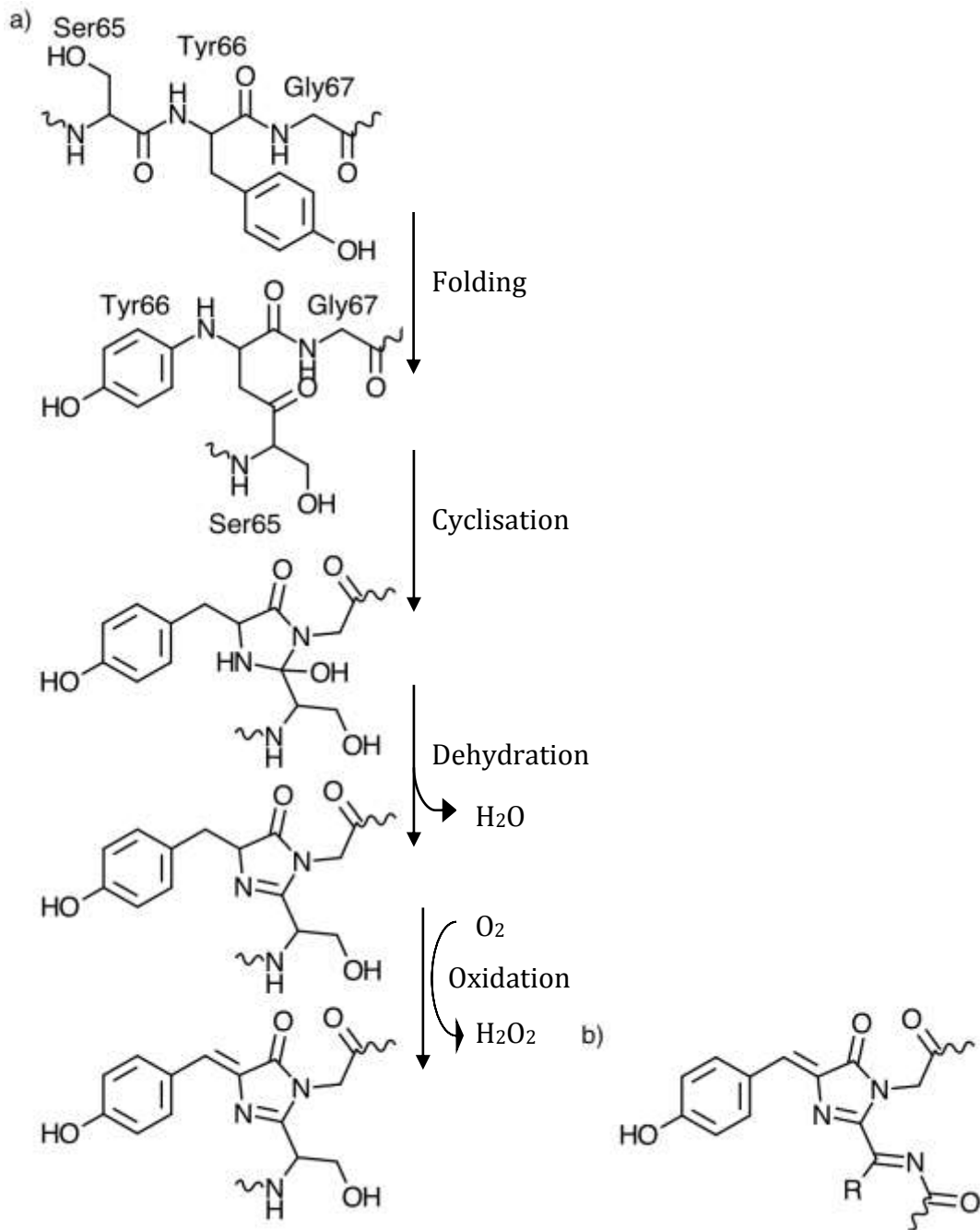
Figure 1.7: Maturation & structure of the avGFP and dsRed chromophores.
a) Proposed mechanism for maturation of the avGFP chromophore including oxidation by molecular oxygen and production of hydrogen peroxide. b) Structure of the dsRed chromophore for comparison to that of avGFP (wavy lines denote continuation of polypeptide chain) (Adapted from Tsien (1998) & Sample *et al.* (2009)).

itself can modify the spectral properties. One of the most common mutations is S65T found in enhanced GFP (EGFP) that shifts the excitation and emission maxima to the red (488 nm and 507-509 nm respectively). Other mutations affecting spectral properties include yellow, cyan and blue fluorescent proteins. Further mutations have produced other beneficial characteristics, such as the F64L mutation that improves folding in *E. coli* (Cormack *et al.*, 1996) and Andersen *et al.* (1998) who modified GFPmut3* with C-terminal short peptide additions to reduce the half-life and hence limit persistence of GFP that could reduce sensitivity when measuring small changes over time for example in gene reporter experiments. The *Discosoma*-derived FPs typically fluoresce at longer wavelengths than *av*GFP and as *av*GFP has been mutated to fluoresce at shorter wavelengths the two complement each other for a full colour palette. The chromophore structures of both categories of proteins are also highly similar (see Figure 1.7b) (Craggs, 2009).

1.6.3. Advantages & limitations of FPs

FPs hold many advantages over more traditional reporter proteins: they exhibit a broad host range, having been successfully synthesised in organisms from bacteria (Chalfie, 1995; Inouye & Tsuji, 1994) through to mammalian cells (Marshall *et al.*, 1995) and even in full multicellular organisms (Okabe *et al.*, 1997), being relatively simple proteins that do not require extensive PTM. Their use does not require the addition of a chemical substrate or cofactor such as the archetypal reporter protein β-galactosidase (Chalfie, 1995). Detection methods are typically non-destructive, allowing cells to be monitored *in vivo.* Fluorescence detection can also be exquisitely sensitive, allowing not only bulk but also single cell measurements via fluorescence microscopy (FM) and FCM and even at the subcellular level for example for organelle localisation (de Giorgi *et al.*, 1996). The

wide colour palette available allows multiple factors to be investigated simultaneously, or for spatial studies involving Förster resonance energy transfer (FRET) (Day & Davidson, 2009).

FPs, however hold some limitations; chromophore maturation is oxygen-dependent therefore precluding the use of *av*FPs in anaerobic conditions without potentially destructive sample processing. A further feature of *av*FPs under anaerobic conditions is that when mature, aerobically-produced GFP is subject to anoxia, fluorescence maxima can shift dramatically such that the protein fluoresces in the red spectrum (excitation, 543 nm; emission, 570-630 nm) (Elowitz *et al.*, 1997).

In addition to the classical FPs and derivatives there are further autofluorescent proteins now available with properties targeted towards the limitations of those already in use such as the flavin mononucleotide binding protein (FMN-BP) described by Drepper *et al.* (2007). The chromophore of FMN-BP, unlike the amino acid-derived chromophore of GFP is a bound flavin mononucleotide. While it does therefore require a cofactor to fluoresce, FMN is ubiquitous in the cell and therefore it can be assumed for convenience that the protein fluoresces naturally. The primary advantage of FMN-BP is that fluorescence is oxygen-independent and therefore can be produced and measured in anaerobic environments and the spectra does not change when proteins are transferred from aerobic to anaerobic environments.

1.6.4. General applications of GFP in microbial biotechnology

The main focus of this work is the application of FPs specifically to RPP in *E. coli*, but FPs have also provided many important advances for microbial biotechnology in general and brief examples are given below. For further reading the reader is directed to the recent review by Vizcaino-Caston *et al.* (2012).

The inclusion of an FP-tag in a microorganism can allow for whole-cell localisation and monitoring especially when combined with single-cell analyses such as FCM. Fluorescence-based whole cell monitoring has applications in many aspects of microbial biotechnology including food production, environmental monitoring and in any bioprocess that requires co-culture. An example of this comes from Miao *et al.* (2009) who developed *E. coli* SCC1, a derivative of *E. coli* MG1655 with GFPmut3* inserted on the chromosome under the control of a constitutive promoter, as the GFP gene is chromosomally encoded this eliminated the inherent problems involved with loss of a plasmid-encoded system. The GFP-tag enabled differentiation by FCM between *E. coli* SCC1 and non *E. coli* strains and the introduction of RFP reporter plasmids showed differential gene expression patterns when *E. coli* SCC1 was co-cultured with *Klebsiella pneumoniae* and *Enterococcus faecalis*.

FPs are ideal reporters for gene expression studies and there exist large libraries of reporter plasmids such as the pUA66 library developed by Zaslaver *et al.* (2006) consisting of about 2000 *E. coli* MG1655 promoters transcriptionally fused to GFPmut2. pUA66 reporters have been used to monitor the expression of genes for the chaperones GroE and DnaK during RPP in *E. coli* (Vizcaino-Caston, 2012). GFP reporters have also been used to monitor conditions in microbial fermentations for example, Garcia *et al.* (2009) who used an oxygen-responsive reporter pNar-GFPuv comprising the promoter for the anaerobically-expressed nitrate reductase (*nar*G) transcriptionally fused to GFPuv, to demonstrate that mixing inefficiencies in aerobic bioreactors caused localised anaerobiosis.

There have also been many advances in using modified FP molecules themselves as a direct sensor for environmental conditions such as intracellular pH (Wilks & Slonczewski, 2007) and redox state (Delic *et al.*, 2010). A direct sensor has advantages

over genetically-encoded reporters as there would be less potential for interference from other gene regulatory elements; this area is reviewed in more detail by Frommer *et al.* (2009), Wang *et al.* (2009a) and Vizcaino-Caston *et al.* (2012).

1.6.5. Applications of FPs for RPP in *E. coli*

Autofluorescent proteins have also shown considerable impact as reporters of RPP and protein folding, primarily as RP-FP fusion products. This is based upon the observation that GFP must fold correctly before the chromophore is able to form and the protein becomes fluorescent. The assumption therefore is that a correctly-folded, soluble GFP-fusion protein will be fluorescent whereas a mis-folded, insoluble protein will be non-fluorescent and therefore fluorescence can be to estimate RPP yields (Waldo *et al.*, 1999; Vizcaino-Caston *et al.*, 2012).

This concept was first explored by Waldo *et al.* (1999) who, using a panel of 20 cytoplasmic proteins from *Pyrobaculum aerophilum,* a hyperthermophilic archaeon, and their equivalent C-terminal GFP fusions, demonstrated that the fluorescence of *E. coli* producing GFP fusions showed a strong positive correlation to the solubility of the proteins when produced under the same conditions without the GFP fusion. Assuming therefore, that fluorescence of an FP fusion protein can be an effective indicator of solubility, this correlation was then used to isolate mutated forms of predominantly insoluble proteins that showed improved folding and solubility while retaining protein function (the ability of Ferritin to oxidise Fe(II)) without the need for a specialised screen for protein activity. A recent, complementary use of GFP-fusion proteins as a screen for the presence of advantageous mutations is reported by Alfasi *et al.* (2011) who, using the protein CheY::GFP isolated mutant forms of the host strain *E. coli* BL21* displaying improved production characteristics. Following standard protocol RPP (37°C,

0.5 mM IPTG) a small amount of cells retained green fluorescence, suggesting the ability to maintain production under physiologically stressful conditions. It is possible that these cells had acquired spontaneous mutations enabling higher productivity. Two methods to identify these potentially high-producing strains were used, first by identification of GFP+ colonies on agar plates following standard protocol RPP and second by FACS, sorting GFP-positive cells following multiple rounds of RPP.

In addition to isolation of mutants a common use of FP-fusions in RPP is to screen for favourable production conditions, increased fluorescence being assumed to be indicative of more productive conditions. Jones *et al.* (2004), while reporting the production of an on-line fluorescence probe that would allow real-time monitoring of FP-fusion protein production in a bioreactor (a development on the fluorescence probe reported by Randers-Eichorn *et al.* (1997)), demonstrated the utility of RP-FP fusions in identifying conditions for improved production and optimal harvesting time. By monitoring fluorescence probe readings in 20 L batch fermentations with 5 FP-RP fusions they were able to demonstrate that for one of the 5 proteins a standard 24 h post-induction harvesting time was not optimal as fluorescence was still accumulating and harvesting at 48 h was optimal. Further, they were able to develop the growth media used; RP solubility was lower when cells were grown in Modified Terrific Broth (MTB) than in LB and adding pH regulation to LB increased production further (two-fold over LB alone and four-fold over MTB). Comparison of CO_2 evolution over time for fermentations producing GFP alone, a model RP and that protein fused to GFP showed that the pattern of CO_2 production for the GFP fusion showed greater similarity to that of the RP alone than that of GFP. This suggests that metabolism during the production of the GFP-fusion protein is affected to a greater extent by the protein of interest than the GFP moiety. As previously mentioned Sevastsyanovich *et al.* (2009) used an RP-GFP

fusion CheY::GFP with fluorimetry and FCM to validate their hypothesis that production conditions designed to eliminate metabolic stress inherent to standard RPP protocols can improve productivity and RP quality.

The ease at which fluorescence measurements can be made make FP-RP fusions highly useful for testing novel expression systems and technology. Siller *et al.* (2010) proposed a novel technique to produce soluble eukaryotic proteins in *E. coli* using decreased translation speeds. Based on the observation that protein synthesis is typically much slower in eukaryotes than *E. coli* (3-8 amino acids per second (aa·s^{-1}) versus 10-20 aa·s^{-1}, respectively) it was hypothesised that decreasing translation rates in *E. coli* could improve RP solubility during the synthesis of eukaryotic proteins. This was tested using a streptomycin pseudo-dependent strain of *E. coli* that requires streptomycin for near wild-type translation rates and in the absence of streptomycin exhibits translation speeds equivalent to eukaryotic ribosomes (~5 aa·s^{-1}). Three RPs were tested, GFP alone and two GFP-fusion proteins; MPB-GFP and GFP-enolase. This approach was shown to be effective with GFP production under slower translation conditions exhibiting a ~2-fold increase in fluorescence over fast translation conditions and a ~3-fold increase for the fusion proteins, although in all cases the increase in solubility was lower than that of fluorescence.

Modern biotechnology is now heavily reliant on high-throughput studies to streamline process development. With rapid detection of fluorescence being a standard feature in most 96-well plate readers, RP-FP fusions are ideal for high-throughput studies. This was shown by Coutard *et al.* (2008) who, using an incomplete-factorial approach (testing 12 combinations of expression conditions out of a possible 36 from 4 *E. coli* strains, 3 growth media and 3 growth temperatures, see Section 1.4.5) and C-terminal GFP fusions, determined optimal production conditions for ten proteins, with "at least

the same sensitivity and specificity" as more expensive Western analysis. Combining both high-throughput screening and novel technology, the microtitre-plate based scale-down fermentation system 'Biolector' was tested in batch and fed-batch growth with both *E. coli* and the methylotrophic yeast *Hansenula polymorpha* producing GFP. Growth kinetics and RP production (measured by DCW and GFP-fluorescence respectively) for microplate cultures were compared to equivalent 1.4 L fermentations (a 7000-fold increase in scale) and in all cases data from microplate cultures closely matched those from fermentations; as such it was concluded that scale-up was successful (Kensy *et al.*, 2009a, 2009b; Scheidle *et al.*, 2010).

1.6.6. Disadvantages of using GFP as a solubility/production reporter

A large number of studies have successfully used FP-RP fluorescence as a predictor of protein folding and solubility, demonstrating its utility. However, numerous studies now suggest the direct relationship assumed between FP-RP fluorescence, solubility and correct protein folding may be more complicated and that caution may be advised when interpreting these data. First, it has been shown that technically insoluble proteins in an IB may still retain functionality. This was demonstrated by García-Fruitós *et al.* (2005) who produced IBs in *E. coli* from fusions of aggregation-prone proteins to GFP and blue fluorescent protein (BFP) that still retained considerable fluorescence. Further work in this area has since shown that fluorescence in FP-IBs is localised towards the centre of the aggregate (García-Fruitós *et al.*, 2007a) and while cells grown at 16°C produce fewer FP-IBs than cells grown at temperatures more likely to result in misfolded RP (30°C & 37°C) those that did form were significantly more fluorescent. The presence of biological activity in IBs, whether fluorescence or enzymatic, is unlikely to be as a result of active site or chromophore amino acids being directly incorporated in the aggregate and more

likely as a result of folded or partially-folded protein becoming 'trapped' within the aggregate. Supporting this, it has also been shown that soluble, active recombinant protein (including GFP) can be extracted from IBs when incubated under mild, non-denaturing conditions (Jevševar *et al.*, 2005; Peternel *et al.*, 2008). It is interesting to note that as IBs have been demonstrated to retain some function their direct use as biologically-active nanoparticles has been suggested, both in biotechnology as an immobilised biocatalytic particle (Sans *et al.*, 2012) and in medicine as a method for intracellular therapeutic protein delivery (Villaverde *et al.*, 2012). In addition to insolubility not necessarily being indicative of completely incorrect folding and loss of protein function, solubility is not necessarily indicative of correct folding and functionality. It was shown by Martínez-Alonso *et al.* (2008) that GFP fused to an aggregation-prone protein formed microaggregates (~0.2 μm) and protein fibrils that remained in the soluble fraction (supernatant following centrifugation of cell lysate formed by sonication). Siller *et al.* (2010) then showed that slowed translation rates resulted in increases in both fluorescence and solubility (as determined by SDS-PAGE analysis of protein fractions obtained from chemical lysis and centrifugation of cells), however the improvement in fluorescence was found to be higher than that of solubility. This discrepancy was found to be as a result of the presence, under normal production conditions, of non-fluorescent aggregates of the FP-fusion protein that remained in the soluble fraction.

A second possible issue with the use of FP-RP fusions as protein folding/solubility reporters is that if the fusion is used to develop specific methods for later production of the native fusion partner, should the FP moiety affect solubility in any way this may cause erroneous conclusions. Many studies have been referenced above claiming a close relationship between native RPs and their equivalent FP fusions, however some suggest

that GFP may act as a solubilisation partner. Jones *et al.* (2004) stated that out of a panel of 5 test proteins the solubility of one was enhanced when expressed as its GFP-fusion. Japrung *et al.* (2005) reported that a *Plasmodium falciparum* dihydrofolate reductase-GFP fusion purified from *E. coli* only aggregated *in vitro* at concentrations above 30 mg.mL^{-1} whereas the native enzyme aggregated at 2 mg.mL^{-1}. Coutard *et al.* (2008) demonstrated that fusion to GFP resulted in moderate increases in solubility for a number of proteins (up to 8.5% for a glycosyl transferase from *Mycobacterium tuberculosis*). Although interestingly, they also reported that two fusion proteins showed decreased solubility (up to ~11.5% for an *M. tuberculosis* esterase).

While these studies show that using FP-RP fusions as reporters of protein folding and solubility may not be as straightforward as initially thought, the sheer volume of successful studies suggest that there is no reason at this time to abandon the technique.

1.6.7. FPs as reporters of plasmid loss

A final use of RP-FP fusions in RPP is as a means of monitoring retention of the plasmid expression vector. It is established that RPP-related stress can cause significant levels of plasmid loss and hence be detrimental to productivity (Baneyx, 1999; Sevastsyanovich *et al.*, 2009). Standard techniques for enumerating plasmid loss involve plating cultures for CFU then replica plating onto antibiotic agar, requiring 2 rounds of bacterial growth and as such may take up to 2 days from initial sampling time. Such a lengthy time delay means that it is impossible to use these data for process control. If the expression vector encodes an FP, either as an RP-FP fusion or separately this would mean that plasmid-containing cells should be fluorescent and non-fluorescent cells should be plasmid-free. If combined with single-cell analytical techniques such as FCM, this can be used to estimate plasmid retention without lengthy growth stages and therefore in a timescale

that can be used to drive process control decisions. This has been used for RPP in both *E. coli* and yeast: Patkar *et al.* (2002) analysed a variety of expression strategies and systems for the production of an elastic, protein-based polymer $(GVGIP)_{260}GVGVP$ fused to GFP in *E. coli* and used FCM for rapid evaluation of expression levels and the plasmid free cell fraction (PFCF). Similarly Ishii *et al.* (2010) evaluated a number of expression vectors and strains of *Saccharomyces cerevisiae* producing GFP with FCM, not only did they analyse expression levels and the levels of GFP$^-$ cells but also showed that the proportion of GFP$^-$ cells accurately reflected plasmid retention rates as determined by plate-counting. There are however, limitations to these techniques: First, classical plasmid retention analysis is a growth-based technique whereas FCM will analyse all cells present whether alive and culturable, dead or displaying a VBNC phenotype. If plasmid retention is not equally distributed between the culturable and non-culturable populations then there may be discrepancies between the two figures. Second, GFP is generally a relatively stable molecule (Andersen *et al.* (1998) conservatively estimated the *in vivo* half life of unmodified GFPmut3* in *E. coli* and *Pseudomonas putida* to be >1 day) and therefore recently plasmid-free cells will likely retain some fluorescence and hence overestimate the retention rates. For a culture in logarithmic growth this should not present a significant issue as rapid cell division would quickly dilute residual fluorescence below detectable levels, however stationary phase or slow-growing cells may be more problematic. The latter limitation was addressed by Bahl *et al.* (2004) who inserted into *E. coli* a chromosomally-encoded GFP regulated by a LacI-controlled promoter, which in conjunction with plasmid-encoded LacI, only fluoresced when the plasmid was lost. While effective, this system requires modification of both the host strain and plasmid, to be used generally in RPP it would be necessary to produce a novel expression system, especially as LacI would interfere with any system using DE3.

1.7. Project aims & objectives

Following an assessment of related literature the following areas were identified as being of particular interest for further study:

- FCM is a powerful technique for the analysis of cell cultures in bioprocesses including RPP, allowing rapid and even real time analysis of multiple physiological parameters that can be used to streamline development and influence process control (Section 1.5). However, its use as a process analytical technique is by no means universal, it is therefore concluded that research aimed at extending the potential uses of FCM in the context of industrial RPP in *E. coli* would increase its uptake as a standard technique.

- The stress-minimising 'improved' RPP protocol by Sevastsyanovich *et al.* (2009) represents a significant change in the theory behind RPP in *E. coli* and has the potential for wider application in both laboratory and industrial contexts (Section 1.3.7), but the fermentation protocol it is applied to has substantial limitations for industrial use, particularly as complex growth media are rarely used industrially (Section 1.4.3.1). It is therefore concluded that the optimisation of 'improved' RPP for industrial HCDC would be beneficial.

From these areas of study the following aims project aims were developed:

- To investigate and develop additional methods by which FCM can be used for the monitoring and analysis of RPP cultures in *E. coli* (Chapter 3).

- To optimise protocols for RPP in *E. coli* using the physiological stress-minimising improved protocol of Sevastsyanovich *et al.* (2009), specifically regarding its application to industrially-relevant fermentation conditions (Chapter 4).

Chapter 2: Materials & Methods

2.1. Materials

Details of all chemicals; commercially-produced reagents, kits and bacterial strains; plasticware/consumables and principal pieces of equipment are given in Tables 2.1-4. Unless otherwise stated all solutions were made using double distilled water (ddH$_2$O).

2.2. Buffers & solutions

2.2.1. Antibiotics & other chemical additions

Stock solutions of culture additions such as antibiotics and RPP inducers were made to 1000x working concentration and added to culture media at 1 µL·mL^{-1}. Carbenicillin (Melford), a more stable variant of ampicillin (Butler *et al.*, 1970) was dissolved in ddH$_2$O to make a 100 mg·mL^{-1} stock solution and filter-sterilised through a 0.22 µm disposable filter unit (Millipore). If not used immediately this solution was stored at -20°C.

Stock solutions of IPTG (Melford) were required at 3 concentrations dependent on RPP protocol; 500 mM, 100 mM and 8 mM. For the 500 and 100 mM stock solutions an appropriate amount of powder was dissolved in ddH$_2$O to produce the correct concentration directly, then filter sterilised as above. For the 8 mM solution, due to the inherent inaccuracies in weighing small amounts, an 80 mM (10x) stock solution was made and filter-sterilised, when required this was diluted with sterile ddH$_2$O to produce the correct concentration. If not used immediately, all solutions were protected from light by wrapping in aluminium foil and stored at 4°C.

E. coli sulphur-free salts solution was added to some fermentations (as indicated) to supplement trace elements, the solution consisted of, in 1 L: 82 g MgCl$_2$·7H$_2$O, 10 g MnCl$_2$·4H$_2$O, 4 g FeCl$_2$·6H$_2$O, 1 g CaCl$_2$·6H$_2$O, and 20 mL concentrated HCl.

Table 2.1: Chemicals used during this study

Item	Manufacturer
Acetic acid (glacial)	Fisher Scientific UK Ltd, Loughborough, UK
Ammonium persulphate ($\geq$ 98%) (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
Asparagine (>98.5%) (BDH)	VWR International Ltd, Lutterworth, UK
β-mercaptoethanol (for electrophoresis) (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
Bromophenol blue (sulphone form) (BDH)	VWR International Ltd, Lutterworth, UK
$CaCl_2 \cdot 6H_2O$ ($\geq$ 98%) (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
Carbenicillin disodium salt	Melford Laboratories Ltd, Ipswich, UK
Citric acid (>99.7%) (BDH AnalaR)	VWR International Ltd, Lutterworth, UK
Coomassie Brilliant blue R (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
$CoSO_4 \cdot 7H_2O$ (>99%) (Sigma ReagentPlus™)	Sigma-Aldrich Co., St Louis, MO, USA
$CuSO_4 \cdot 5H_2O$ (Technical) (Fisher)	Fisher Scientific UK Ltd, Loughborough, UK
Dimethylsulphoxide	Fisher Scientific UK Ltd, Loughborough, UK
Ethanol	Fisher Scientific UK Ltd, Loughborough, UK
$FeCl_2 \cdot 6H_2O$ (Reagent Grade) (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
$FeSO_4 \cdot 7H_2O$ (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
Glucose (BDH)	VWR International Ltd, Lutterworth, UK
Glycerol	Fisher Scientific UK Ltd, Loughborough, UK
H_3BO_3 (Electronic Grade) (Fisons)	Sanofi S.A., Paris, France
H_3PO_4 (concentrated)	Fisher Scientific UK Ltd, Loughborough, UK
HCl (concentrated)	Fisher Scientific UK Ltd, Loughborough, UK
IPTG (>99%)	Melford Laboratories Ltd, Ipswich, UK
K_2HPO_4 (>99%) (BDH AnalaR)	VWR International Ltd, Lutterworth, UK
KH_2PO_4 (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
Methanol	Fisher Scientific UK Ltd, Loughborough, UK
$MgCl_2 \cdot 6H_2O$ (Sigma ReagentPlus™)	Sigma-Aldrich Co., St Louis, MO, USA
$MgSO_4 \cdot 7H_2O$ (Sigma ReagentPlus™)	Sigma-Aldrich Co., St Louis, MO, USA
$MnCl_2 \cdot 4H_2O$ (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
$MnSO_4 \cdot H_2O$ (>98.5%) (BDH AnalaR)	VWR International Ltd, Lutterworth, UK
$Na_2MoO_4 \cdot 2H_2O$ (Analytical Grade) (Fisons)	Sanofi S.A., Paris, France
NaCl	Sigma-Aldrich Co., St Louis, MO, USA
$(NH_4)_2SO_4$ (CertiFied AR) (Fisher)	Fisher Scientific UK Ltd, Loughborough, UK
Polyethylene glycol 6000 (BDH)	VWR International Ltd, Lutterworth, UK
Polypropylene glycol	Sigma-Aldrich Co., St Louis, MO, USA
Propidium Iodide (Invitrogen)	Life Technologies Ltd (Invitrogen), Paisley, UK
Sodium dodecyl sulphate ($\sim$95%) (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
Serine (>99%) (Sigma ReagentPlus™)	Sigma-Aldrich Co., St Louis, MO, USA
Tetramethylethylenediamine	Melford Laboratories Ltd, Ipswich, UK
Threonine (>98%) (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
Tris HCl (Trizma®) (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
Tryptone (Fisher Bioreagents)	Fisher Scientific UK Ltd, Loughborough, UK
Tryptone	Sigma-Aldrich Co., St Louis, MO, USA
Tryptone-peptone (Bacto)	Becton, Dickinson & Company, Oxford, UK
Yeast extract	Oxoid Ltd, Basingstoke, UK
Yeast extract	Duhefa Biochemie, Haarlem, Netherlands
Yeast extract (Bacto)	Becton, Dickinson & Company, Oxford, UK
$ZnSO_4 \cdot 7H_2O$ (>99.5%) (BDH AnalaR)	VWR International Ltd, Lutterworth, UK

Table 2.2: Pre-prepared reagents and kits used during this study

Item	Manufacturer
10x TGS buffer	Bio-Rad Laboratories Ltd, Hemel Hempsted, UK
20% (w/v) SDS	Bio-Rad Laboratories Ltd, Hemel Hempsted, UK
BugBuster®	Merck KGaA (Novagen), Darmstadt, Germany
E. coli BL21*	Life Technologies Ltd (Invitrogen), Paisley, UK
E. coli JM109	Promega Corp., Madison, WI, USA
Flow-Check Fluorospheres	Beckman Coulter (UK) Ltd, High Wycomb, UK
LB Broth (powdered) (Sigma)	Sigma-Aldrich Co., St Louis, MO, USA
Nutrient Agar	Oxoid Ltd, Basingstoke, UK
PBS Tablets	Oxoid Ltd, Basingstoke, UK
Protogel	National Diagnostics, Atlanta, GA, USA
QIAprep Spin Miniprep kit	QIAGEN Ltd, Manchester, UK
Pierce™ BCA Protein Assay Kit	Pierce Biotechnology, Rockford, IL, USA
Silicone antifoam	Dow Corning Corp, Midland, MI, USA

Table 2.3: Consumables used during this study

Item	Manufacturer
Bijou, 7 mL, polystyrene	Appleton Woods Ltd, Birmingham, UK
Cellulose gel drying film	Promega Corp., Madison, WI, USA
Centrifuge tubes (Falcon), 50 mL, polypropylene, sterile	Becton, Dickinson & Company, Oxford, UK
Cuvettes, semi-micro, polystyrene	Fisher Scientific UK Ltd, Loughborough, UK
Fluorescence cuvettes, 3 mL polymethylmethacrylate	Kartell Labware Division, Noviglio, MI, Italy
Microcentrifuge tube, 1.5 mL, polypropylene	Sarstedt Ltd, Leicester, UK
Microcentrifuge tube, 2 mL, polypropylene	Sarstedt Ltd, Leicester, UK
Needles, 21 gauge regular, Microlance®	Becton, Dickinson & Company, Oxford, UK
Nescofilm sealing film	Nippon Shoji Kaisha Ltd, Osaka, Japan
Petri dishes, 90 mm diameter, polystyrene, sterile	Sarstedt Ltd, Leicester, UK
Pipette tips, 1 mL, polypropylene	Sarstedt Ltd, Leicester, UK
Pipette tips, 200 µL, polypropylene	Sarstedt Ltd, Leicester, UK
Pipette tips, 5 mL, polypropylene, Eppendorf fit	VWR International Ltd, Lutterworth, UK
Syringe filters, 0.22µm PES membrane, sterile, Millex	Merck Millipore, Billerica, MA, USA
Syringe, 1 mL, Plastipak	Becton, Dickinson & Company, Oxford, UK
Syringe, 20 mL, Plastipak	Becton, Dickinson & Company, Oxford, UK
Syringe, 5 mL, Plastipak	Becton, Dickinson & Company, Oxford, UK
Syringe, 60 mL, Plastipak	Becton, Dickinson & Company, Oxford, UK

Table 2.4: Principal equipment used during this study

Item	Manufacturer
Midisart® 2000 0.2 μm PTFE air filters	Sartorius AG, Goettingen, Germany
Sarto-Capsule air filters	Sartorius AG, Goettingen, Germany
BD-Accuri C6 Flow cytometer	Becton, Dickinson & Company, Oxford, UK
Canoscan 9000F scanner	Canon Inc., Tokyo, Japan
Coulter EPICS-Elite Flow cytometer	Beckman Coulter (UK) Ltd, High Wycomb, UK
D150 Oxyprobe DOT probe	Broadley-James Corp., Irvine, CA, USA
F-695 Fermprobe pH probe	Broadley-James Corp., Irvine, CA, USA
Fermac 310/60 bioreactor	Electrolab Biotech Ltd, Tewkesbury, UK
Omni-Page mini SDS-PAGE apparatus	Cleaver Scientific Ltd, Rugby, UK
Perkin Elmer Luminescence Spectrometer LS 50B	PerkinElmer Inc., Waltham, MA, USA
PrimaDB process gas mass spectrometer	Thermo Fisher Scientific, Waltham, MA, USA
PT100 temperature sensor	Electrolab Biotech Ltd, Tewkesbury, UK
Uvikon Spectrophotometer 922	Kontron AG, Eching, Germany

2.2.2. Phosphate buffered saline (PBS)

Where appropriate, cultures and cell suspensions were diluted in PBS to maintain an appropriate osmolarity and pH. This was made according to Dulbeccos formulation 'A' (NaCl 8.0 $g \cdot L^{-1}$, KCl 0.2 $g \cdot L^{-1}$, Na_2HPO_4 1.15 $g \cdot L^{-1}$, KH_2PO_4 0.2 $g \cdot L^{-1}$, pH 7.3) from commercially produced tablets (Oxoid) according to the manufacturer's instructions.

2.2.3. Buffers & solutions for protein analysis (SDS-PAGE & BCA Assay)

The compositions of solutions made for the production of SDS-PAGE gels were as follows:

- 2x stock running gel buffer – 0.75 M Tris(hydroxymethyl)aminomethane (Tris)/HCl, pH 8.3 (pH adjusted by addition of HCl).

- 10x stock stacking gel buffer – 1.25 M Tris/HCl, pH 6.8 (pH adjusted by addition of HCl).

- Ammonium persulphate (APS) – An 80 $mg \cdot mL^{-1}$ solution of APS was made by weighing 80 mg of APS in a 1.5 mL microcentrifuge tube and dissolving in 1 mL ddH_2O.

- Sample buffer – 2 g SDS, 20 mL glycerol and 5 mg bromophenol blue (sulphone form) were dissolved in 1x stacking buffer (made by diluting 10 mL of 10x stock stacking buffer in ddH_2O to a final volume of 100 mL) to a final volume of 92 mL. This was divided into 1 mL aliquots in 1.5 mL microcentrifuge tubes and stored at ambient temperature. Immediately prior to use 87 μL of β-mercaptoethanol was added to the tube.

- Running buffer – 1x Tris-glycine-SDS (TGS) running buffer (25 mM Tris, 192 mM glycine, 0.1% w/v SDS, pH 8.6) was made from a commercially available 10x stock solution (BioRad), diluting in ddH_2O.

- BCA assay lysis/solubilisation buffer – 2% (w/v) SDS and 1 mM EDTA were dissolved in 1x SDS-PAGE stacking buffer (125 mM Tris/HCl, pH 6.8).

2.2.4. Solutions for the development of Coomassie blue-stained protein following SDS-PAGE

The composition of solutions used to develop Coomassie blue-stained SDS-PAGE gels were as follows:

- Coomassie blue stain – 2 $g \cdot L^{-1}$ Coomassie Brilliant blue R in 50% (v/v) methanol, 10% (v/v) glacial acetic acid and 40% (v/v) ddH$_2$O.

- Fast destain – 40% (v/v) methanol, 10% (v/v) glacial acetic acid and 50% (v/v) ddH$_2$O.

- Slow destain – 10% (v/v) methanol, 10% (v/v) glacial acetic acid and 80% (v/v) ddH$_2$O.

- Shrink solution – 48% (v/v) methanol, 2% (v/v) glycerol and 50% (v/v) ddH$_2$O.

2.2.5. Dyes for flow cytometry

A 200 $\mu g \cdot mL^{-1}$ stock solution of PI (Sigma) was made by dissolving 4 mg of PI powder in 20 mL of ddH$_2$O. This was divided into 1 mL aliquots in PP microcentrifuge tubes and stored at 4°C for up to 6 months.

Stock solutions of Congo Red (CR) (BDH Chemicals) were produced in two ways. For initial experiments a saturated stock solution was made by dissolving 100 mg of dye and 100 mg of NaCl in 5 mL 80% (v/v) ethanol, filtered through a 0.22 μm filter unit to remove undissolved solids and used only on the day produced as the solution rapidly precipitated during storage. For subsequent experiments a 2 $mg \cdot mL^{-1}$ solution was

made by dissolving 2 mg of dye in 1 mL of dimethylsulphoxide (DMSO) (Fisher) and stored at 4°C.

2.3. Bacterial strains & plasmids & basic growth media

2.3.1. Bacterial strains

E. coli JM109 (*end*A1 *rec*A1 *gyr*A96 *thi hsd*R17 (rk– mk+) *rel*A1 *sup*E44 Δ(*lac-pro*AB), [F′ *tra*D36 *pro*AB *laq*IqZΔM15]) (Promega) was used to produce stocks of the expression vector. *E. coli* BL21* (DE3) (F$^-$ *ompT hsd*S$_B$ (r$_B^-$m$_B^-$) *gal dcm rne131* λ(DE3)) (Invitrogen) was used for all subsequent cultures and RPP. For work involving anaerobic production and aerobic recovery of GFP the strain *E. coli* SCC1 was used. *E. coli* SCC1 is a derivative of wild-type *E. coli* MG1655 (F$^-$ λ$^-$ *ilvG*$^-$ *rfb*-50 *rph*-1) containing a chromosomal insertion of P$_{A1/04/03}$-gfpmut3* allowing constitutive expression of GFP (Miao *et al.*, 2009).

2.3.2. Plasmids

Expression vector pET20bhc-CheY::GFP encoding the model RPP product CheY::GFP (Jones *et al.*, 2004, Jones, 2007) was used for all RPP experiments, a plasmid map is given in Figure 2.1.

2.3.3. Solid growth media

To make solid growth media 11.2 g of nutrient agar (NA) (Oxoid) was dissolved in 400 ml of ddH$_2$O and sterilised by autoclaving for 15 minutes at 121°C (1 bar). If selective agar was required the sterile molten agar was allowed to cool to 60°C to limit thermal degradation of the antibiotic, then carbenicillin was added to a final concentration of 100 μg·mL^{-1}. Molten agar was poured into sterile 90 mm diameter polystyrene (PS) petri dishes (Sarstedt) at approximately 20 mL per plate, allowed to set then dried to remove
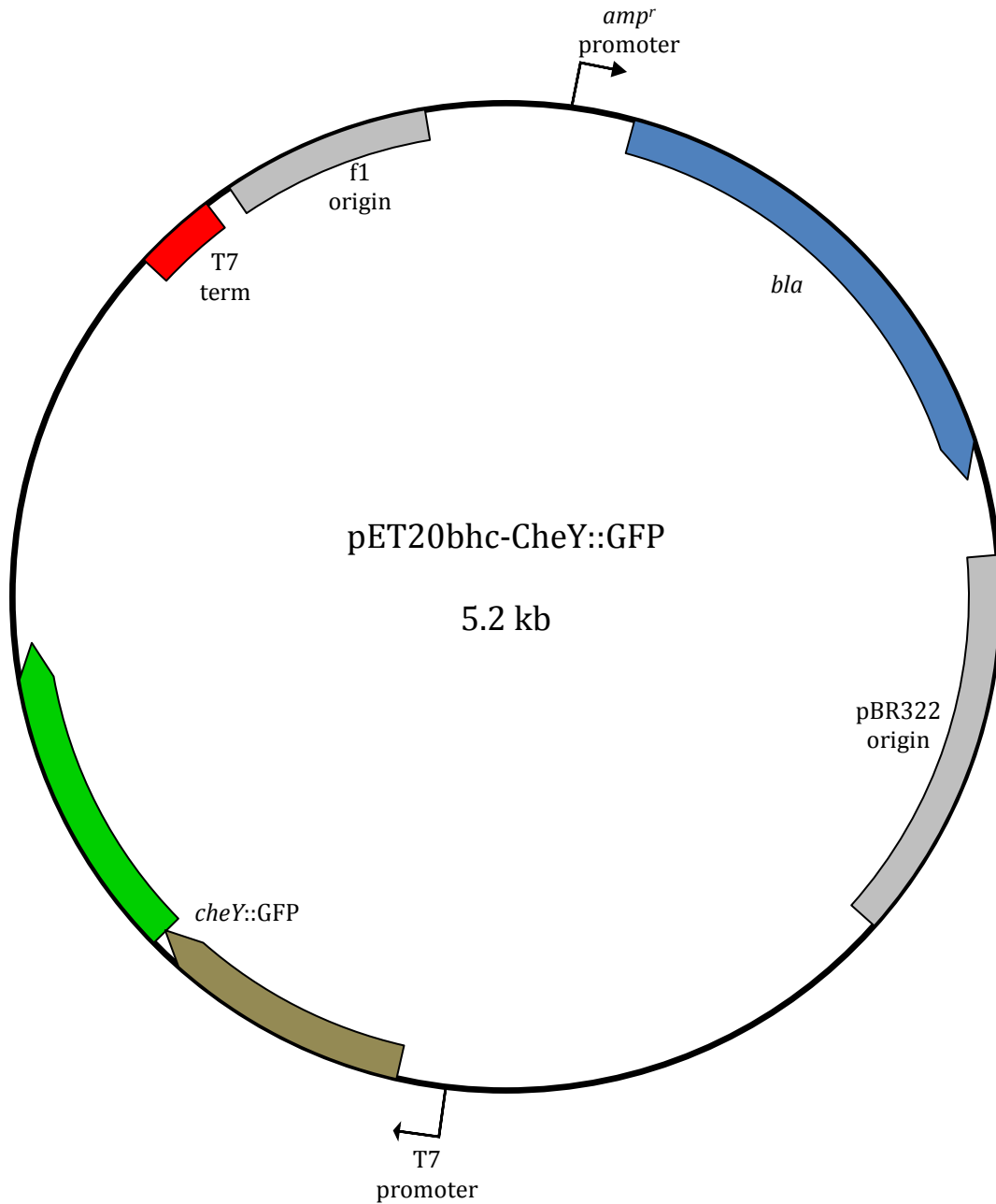
Figure 2.1: Plasmid map for expression vector pET20bhc-CheY::GFP

Key: Grey – origin of replication; Blue – gene encoding β-lactamase (antibiotic resistance marker); Brown – gene encoding CheY; Green – gene encoding GFP; Red – transcriptional terminator; Bent arrow – transcriptional promoter. Not to scale.

excess moisture. Plates were stored at 4°C and those containing antibiotic were used within 2 months.

2.3.4. Liquid growth media

For work involving *E. coli* SCC1, cells were grown in LB with the following composition: 20 g·L^{-1} tryptone-peptone (BD Bacto), 10 g·L^{-1} yeast extract (BD Bacto) and 10 g·L^{-1} NaCl (Sigma). This was dissolved in ddH$_2$O and sterilised by autoclave for 15 minutes.

Unless otherwise stated *E. coli* BL21* and JM109 were grown in LB with the following composition: 10 g·L^{-1} tryptone-peptone, 5 g·L^{-1} yeast extract and 5 g·L^{-1} NaCl. This was prepared as above.

If glucose was required in the medium, a 40% (w/v) solution in ddH$_2$O was autoclaved separately and added post-sterilisation to the correct final concentration at ambient temperature to prevent Maillard reactions between the glucose and amine groups in LB. If antibiotics were required these were also added post-sterilisation to prevent thermal inactivation.

2.4. Recombinant DNA techniques

2.4.1. Production and transformation of competent *E. coli*

MgCl$_2$-competent cells (both *E. coli* JM109 and BL21* (DE3)) were prepared according to the protocol by Nishimura *et al.* (1990). 50 mL of growth medium A (LB (1% tryptone, 0.5% yeast extract, 0.5% NaCl) supplemented with 10 mM MgSO$_4$·7H$_2$O, 0.2% (w/v) glucose) was inoculated with 0.5 mL of an overnight culture (2 mL LB, 37°C, 150 RPM, 16 h) and grown (37°C, 200 RPM) to mid-logarithmic phase (OD$_{650}$ ~1.0). Cultures were decanted into sterile, disposable 50 mL centrifuge tubes (BD-Falcon), cooled on ice for 10 minutes then pelleted by centrifugation (1500 x g, 4°C, 10 minutes) and resuspended in 0.5 mL of pre-cooled medium A. To this cell suspension, 2.5 mL of storage medium B

(LB supplemented with 36% (w/v) glycerol, 12% (w/v) polyethylene glycol (PEG) 6000, 12 mM $MgSO_4 \cdot 7H_2O$, filter-sterilised) was added and mixed without vortexing then divided into 100 μL aliquots in sterile 1.5 mL PP microcentrifuge tubes and stored at -80°C.

Competent cells were transformed according to a modified protocol based on that by Nishimura *et al.* (1990). Aliquots of cells were thawed on ice, 2 μL of miniprep plasmid DNA was added then incubated on ice for 15-30 minutes. Cells were heat-shocked at 42°C for 1 minute, chilled on ice for 1-2 minutes, 900 μL of LB was added then the cells were allowed to recover and express antibiotic resistance for 1 h at 37°C. Following recovery of the cells in LB, cells were pelleted by centrifugation (16,873 g, 5 minutes, ambient temperature), resuspended in 100 μL LB and plated onto selective NA. Plates were then incubated to allow for cell growth (25°C, 48 h). Single colonies from the transformation plate were then restreaked onto fresh antibiotic agar and grown under the same conditions to create a master plate of cells derived from a single colony. Unless otherwise stated, plates were stored at 4°C and discarded after 2 weeks. To investigate the effect of oxygen limitation during storage some plates were sealed between use with Nescofilm, a gas-resistant sealing film. If it was required to analyse transformants prior to growth in liquid culture, a small amount of cells were removed from the plate with a sterile 200 μL pipette tip and resuspended in 1-2 mL sterile PBS, this could then be used both for FCM, CFU and SDS-PAGE analysis.

2.4.2. Production & purification of plasmid stocks

To produce stocks of plasmid *E. coli* JM109 was transformed as above and single colonies from the transformation plate restreaked onto fresh selective agar. From these plates 2 mL cultures of LB supplemented with 100 μg·mL^{-1} carbenicillin were inoculated

and grown for 16 h at 37°C with agitation. These cultures were harvested by centrifugation in sterile 2 mL PP microcentrifuge tubes (16,873 g, 5 minutes, ambient temperature) and plasmid DNA was extracted using a QIAprep Spin Miniprep kit (QIAGEN) according to manufacturers instructions, eluting into 50 µL sterile ddH$_2$O.

2.5. Bacterial growth & RPP – Shake flasks

2.5.1. E. coli SCC1

2.5.1.1. Growth Conditions

E. coli SCC1 was grown in 100 mL LB cultures supplemented with 0.4% (w/v) glucose, from a 2% (v/v) inoculum. The inoculum was grown from a single colony of bacteria in LB at 37°C and 200 RPM. Aerobic cultures were grown in 500 mL conical flasks and incubated at 37°C and 200 RPM, oxygen-limited cultures were obtained by growth in 100 mL conical flasks at 37°C without agitation.

2.5.1.2. Aerobic fluorescence recovery

For aerobic fluorescence recovery (AFR) of GFP (Zhang *et al.*, 2005), both aerobic and oxygen-limited cultures were grown for 3 h then 15 mL of culture was removed and cells harvested by centrifugation (10 minutes, 2556 x g, 4°C) in sterile PP disposable centrifuge tubes (BD-Falcon). Cell pellets were then re-suspended in 15 mL PBS and incubated at 37°C, 200 RPM in 50 mL conical flasks for 2 h.

2.5.2. E. coli BL21*

2.5.2.1. Inoculum set-up

All inocula for RPP experiments were prepared in such a way to be analogous to those used for fermentations i.e. 35 mL of LB supplemented with 100 µg·mL^{-1} carbenicillin in a 250 mL conical flask was inoculated with a sweep of cells taken from a master plate of

transformants and incubated with agitation (150 RPM). Temperature and duration of incubation varied with experiments.

2.5.2.2. Growth conditions & RPP

For shake-flask RPP cultures 100 mL of LB supplemented with 100 µg·mL^{-1} carbenicillin and 0.5 % (w/v) glucose in a 500 mL conical flask was inoculated at 2% (v/v) from an overnight culture prepared as described above and incubated with agitation (200 RPM). 3 protocols were used for shake-flask RPP during this study, summarised in Table 2.5.

2.6. Bacterial growth & RPP – Fed-batch fermentation

2.6.1. Vessel & fermentation equipment

For all fed-batch fermentations a Fermac 310/60 bioreactor (Electrolab) was used with a 5 L cylindrical vessel. The vessel was equipped with 4 baffles and an agitator with 2 six-bladed Rushton turbines. Aeration was achieved by sparging air from below the lower impeller at a rate of 3 L·minute^{-1} (2 vvm initial culture volume) through a reusable, autoclavable 0.22 µm filter (Sartorius). Off-gas was passed through a condensing column at approximately 2°C to prevent moisture loss and then through a larger volume reusable, autoclavable 0.22 µm filter (Sartorius). To prevent blockage of the exhaust filter by foamed culture two 1 L catch pots were placed between the condenser and filter, each containing approximately 1 mL of polypropylene glycol (PPG) as antifoam (Sigma). When filled with medium the vessel was sterilised by autoclaving at 121°C for 20 minutes.

2.6.2. Media & growth conditions

A number of different fermentation protocols were used during this study, including differences in growth media, feed composition and feed rate, these are summarised in

Table 2.5: Growth temperature and induction conditions for shake-flask RPP cultures

Protocol	S1	S2	SA
Source	Sevastyanovich *et al.*, 2009 ('standard conditions')	This study	Sevastyanovich *et al.*, 2009 ('improved conditions')
Induction point	$OD_{650} \sim 0.5$ (mid logarithmic phase)	$OD_{650} \sim 0.5$ (mid logarithmic phase)	$OD_{650} \sim 0.5$ (mid logarithmic phase)
Temperature pre-induction	37°C	37°C	25°C
Temperature post-induction	25°C	37°C	25°C
[IPTG]	0.5 mM	0.5 mM	8 µM

Tables 2.6 and 2.7. All fermentations began with an initial volume of 1.5 L medium supplemented with 100 μg·mL^{-1} carbenicillin, adjusted to pH 6.3 and were inoculated at 2% (v/v) from an overnight culture. Inocula were grown from a sweep of cells in 35 mL of LB supplemented with 100 μg·mL^{-1} carbenicillin in a 250 mL conical flask, at either 30°C and 150 RPM for 12-14 h or 25°C and 150 RPM for 18-21 h. Prior to addition to the vessel 5 mL of inoculum was removed and used for screening and analysis. Fermentations were terminated after it was apparent that the culture had ceased to grow, indicated by OD$_{650}$ and/or DOT/off-gas analysis.

2.6.3. Process monitoring and control

Unless otherwise stated data logging and process control were via the on-board control unit or Electrolab fermentation management software on a connected PC.

2.6.3.1. Temperature sensing & control

Temperature sensing was via an external PT100 Sensor (Electrolab). To control temperature a wrap around heating mat and cold finger were used. Temperature regulation was periodically calibrated by inserting an independent thermometer through an open port when the vessel was set up containing water. It was observed that during the later stages of fed-batch growth that cooling with a single cold finger was insufficient to maintain the set point, therefore once an increase in temperature was observed the chilled water supply for the condenser was rerouted to a second cold finger.

2.6.3.2. DOT sensing & control

DOT sensing was via a reusable, autoclavable D150 Oxyprobe (Broadley James). DOT was calibrated both prior to assembly with H$_2$O in the vessel and after autoclaving prior to inoculation by alternately sparging air and nitrogen. DOT was maintained above a set

Table 2.6: Media composition for basic fermentations protocols

Protocol	A (LB-based)	B (Semi-defined)	C (Chemically-defined SM6E)
Source	Sevastyanovich *et al.*, 2009; S. Alfasi – Personal communication	Want *et al.*, 2009	Humphreys *et al.*, 2002; C. Hsu – Personal communication
Components	10 g·L^{-1} Tryptone 5 g·L^{-1} Yeast extract 5 g·L^{-1} NaCl 1 mL *E. coli* sulphur free salts 1 mL·L^{-1} Silicone antifoam (Corning)	14 g·L^{-1} $(NH_4)_2SO_4$ 35 g·L^{-1} Glycerol 20 g·L^{-1} Yeast extract 2 g·L^{-1} KH_2PO_4 16.5 g·L^{-1} K_2HPO_4 7.5 g·L^{-1} Citric acid. 1.5 mL·L^{-1} Conc. H_3PO_4 0.66 mL·L^{-1} polypropylene glycol (PPG) (as antifoam)	5.2 g·L^{-1} $(NH_4)_2SO_4$ 3.86 g·L^{-1} NaH_2PO_4·H_2O 4.03 g·L^{-1} KCl 1.04 g·L^{-1} $MgSO_4$·$7H_2O$ 0.25 g·L^{-1} $CaCl_2$·$2H_2O$ 10 mL·L^{-1} SM6E Trace elements solution 4.16 g·L^{-1} Citric acid monohydrate 31.11 g·L^{-1} Glycerol 0.66 mL·L^{-1} PPG
Post-sterilisation additions	12.5 mL·L^{-1} 40% (w/v) Glucose 1 mL·L^{-1} 100 mg·mL^{-1} Carbenicillin	34 mL·L^{-1} Trace metal solution 10 mL·L^{-1} 1 M $MgSO_4$·$7H_2O$ 2 mL·L^{-1} 1 M $CaCl_2$·$2H_2O$ 1 mL·L^{-1} 100 mg·mL^{-1} Carbenicillin	1 ml·L^{-1} 100 mg·mL^{-1} Carbenicillin
Further additions	At 5 h post-induction: 1 mM Serine 1 mM Threonine 1 mM Asparagine		
Trace elements composition		3.36 g·L^{-1} $FeSO_4$·$7H_2O$ 0.84 g·L^{-1} $ZnSO_4$·$7H_2O$ 0.15 g·L^{-1} $MnSO_4$·H_2O 0.25 g·L^{-1} Na_2MoO_4·$2H_2O$ 0.12 g·L^{-1} $CuSO_4$·$5H_2O$ 0.36 g·L^{-1} H_3BO_3 48 mL·L^{-1} Conc. H_3PO_4	104 g·L^{-1} Citric acid monohydrate 5.22 g·L^{-1} $CaCl_2$·$2H_2O$ 2.06 g·L^{-1} $ZnSO_4$·$7H_2O$ 2.72 g·L^{-1} $MnSO_4$·$4H_2O$ 0.81 g·L^{-1} $CuSO_4$·$5H_2O$ 0.42 ·L^{-1} $CoSO_4$·$7H_2O$ 10.06 g·L^{-1} $FeCl_3$·$7H_2O$ 0.03 g·L^{-1} H_3BO_3 0.02 g·L^{-1} $NaMoO_4$·$2H_2O$
Feed composition	100 g·L^{-1} Tryptone 50 g·L^{-1} Yeast extract 200 g·L^{-1} Glucose 10 mM Serine 10 mM Threonine 10 mM Asparagine 1 ml 100 mg.l^{-1} Carbenicillin stock 1 mL·L^{-1} 8mM IPTG stock 1 mL·L^{-1} *E. coli* sulphur-free salts 0.1% (v/v) silicone antifoam (Corning)	714 g·L^{-1} Glycerol 30 ml.l^{-1} 1 M $MgSO_4$·$7H_2O$ 1 mL·L^{-1} 100 mg.l^{-1} Carbenicillin stock 1 mL·L^{-1} IPTG stock (100mM/8mM as necessary)	714 g·L^{-1} Glycerol 30 ml.l^{-1} 1 M $MgSO_4$·$7H_2O$ 1 mL·L^{-1} 100 mg.l^{-1} Carbenicillin stock 1 mL·L^{-1} 8mM IPTG stock

Table 2.7 Fermentation protocols used in this study

Protocol	A	B	B1	B2	B3	C1	C2	C3
Growth medium	Medium A	Medium B	Medium B	Medium B	Medium B	Medium C	Medium C	Medium C
Carbon source in vessel	5 g·L⁻¹ glucose	35 g·L⁻¹ glycerol	35 g·L⁻¹ glycerol	35 g·L⁻¹ glycerol	5 g·L⁻¹ glucose	31.11 g·L⁻¹ glycerol	31.11 g·L⁻¹ glycerol	31.11 g·L⁻¹ glycerol
Feed	1 L Feed A	0.5 L Feed B	0.5 L Feed B	0.5 L Feed B	0.5 L 400 g·L⁻¹ glucose 7.4 g·L⁻¹ MgSO₄·7H₂O	0.5 L Feed B	0.5 L Feed B	0.5 L Feed B
Feed rate	Stepped linear: 7 hours post-induction 13.69 mL⁻¹·h⁻¹ 32.5 h post-induction 21.13 mL⁻¹·h⁻¹ 48.5 h post-induction 27.17 mL⁻¹·h⁻¹ 54.5 h post-induction 38.0 mL⁻¹·h⁻¹	67.5 mL·h⁻¹	67.5 mL·h⁻¹	67.5 mL·h⁻¹	Exponential at µ = 0.2 up to 67.5 mL·h⁻¹	67.5 mL·h⁻¹	67.5 mL·h⁻¹	67.5 mL·h⁻¹
Temperature	25°C	37°C	25°C	25°C	25°C	25°C	25°C	25°C
[IPTG]	8 µM	100 µM	8 µM	8 µM	8 µM	8 µM	8 µM	8 µM
Induction point	Mid-logarithmic phase (OD$_{650}$ ~0.5)	With feeding (OD$_{650}$ ~40-50)	With feeding (OD$_{650}$ ~40-50)	Mid-logarithmic phase (OD$_{650}$ ~0.5)	Mid-logarithmic phase (OD$_{650}$ ~0.5)	OD$_{650}$ ~80-90	OD$_{650}$ ~80-90	OD$_{650}$ ~80-90

point of 30% by increasing agitation to a maximum of 1000 RPM from a minimum of 200-500 RPM.

2.6.3.3. pH sensing & control

pH sensing was via a reusable, autoclavable F-695 Fermprobe (Broadley James). Prior to assembly a 2-point calibration was done using pH 4 and 7 buffers then prior to inoculation a sample of media was removed and tested on an external pH meter for a final single-point calibration. pH was controlled at a set point of 6.3 ± 0.1 with the automated addition of sterile 10% (v/v) NH_3 and 5% (v/v) HCl.

2.6.3.4. Off-gas analysis

Off-gas was automatically collected for compositional analysis using a PrimaDB process gas mass spectrometer (MS) (Thermo) and compared to atmospheric air in order to calculate the proportion of O_2 consumed (OXC), CO_2 evolved (CDC) and from these the respiratory quotient (RQ), these data were logged automatically by GasWorks v1.0 (Thermo). This was done not only to provide additional information regarding the cells metabolism but also the live data plotting facility incorporated into the software allowed monitoring of changes in the respiratory activity of the culture during the stages of the fermentation where DOT was depleted and therefore less sensitive to small changes.

2.6.4. Aerobic fluorescence recovery

For AFR of fed-batch fermentation cultures a sample was taken from the fermenter when the DOT was approaching 0% despite full agitation and aeration and 1 mL aliquots were pelleted by centrifugation (16,873 x g, 5 minutes, ambient temperature). Pelleted cells were resuspended in 10 mL sterile PBS, transferred to a sterile 50 mL conical flask and incubated with agitation (25°C, 150 RPM).

2.7. Cell analysis techniques

2.7.1. Optical density

Optical density was measured in disposable semi-micro polystyrene (PS) cuvettes (Fisher) at 650 nm using a Uvikon Spectrophotometer 922 (Kontron Instruments), diluting with PBS where necessary to obtain a reading of 0.1-0.6, zeroing the machine with ddH$_2$O.

2.7.2. Fluorimetry

Fluorescence measurements were determined in disposable 3 ml polymethylmethacrylate (PMMA) fluorescence cuvettes (Kartell) using a Perkin Elmer Luminescence Spectrometer LS 50B, exciting at 488 nm, detecting at 510 nm with 10 nm slit width, diluting with PBS where necessary. As it was observed that LB has an inherent fluorescence under the conditions used, a sample of medium was taken from each flask prior to inoculation, the fluorescence determined, then subtracted from the readings for the growing culture.

2.7.3. Colony forming units (CFU) & plasmid retention

To determine culturability and plasmid retention, cultures were serially diluted in sterile PBS, plated on to non-selective NA plates to allow 50-500 colonies per plate and incubated at 25°C for 48 h. Growth at 25°C as opposed to 37°C allowed colonies to develop a clear GFP[+] phenotype from 'leaky' recombinant gene expression. Expression vector retention was determined in two ways, first the proportion of GFP positive colonies were identified by eye, then colonies were replica-plated on to NA and NA-carbenicillin plates.

2.7.4. Dry cell weight

To determine DCW from fermentations 4 x 2 mL aliquots of sample were pelleted by centrifugation (16,873 g, 15 minutes) in 2 mL pre-weighed and labelled PP microcentrifuge tubes stored under desiccating conditions, the supernatant was removed by syringe and discarded, tubes were then dried at 100°C for a minimum of 24 h before weighing.

2.7.5. Flow cytometry

Two flow cytometers were used during this study, initially an EPICS Elite (Coulter) then subsequently a BD-Accuri C6 (BD).

2.7.5.1. Sample preparation & staining for cell physiology

For both cytometers samples were diluted in 0.22 μm filtered PBS to an appropriate data rate (Table 2.3) and data were recorded. Cell physiology was then tested by staining with fluorescent dyes, PI for dead cells and CR for the presence of amyloid-like protein deposits. To the unstained samples, having been diluted to give an appropriate data rate, dyes were added to final concentrations previously determined to give optimal results and data was recorded again. Live/dead differentiation using PI was calibrated using a mixture of live and ethanol-killed cells, 2 x 1 mL aliquots of cells were pelleted by centrifugation (16,873 x g, 10 minutes, ambient temperature), one was resuspended in PBS, the other in ethanol. These cell suspensions were diluted in filtered PBS to an appropriate data rate, mixed in approximately equal volumes, stained with PI, analysed by FCM and gates drawn to differentiate the 2 populations (examples for both cytometers are given in Figure 2.2). These gates were then applied to experimental samples to estimate viability.
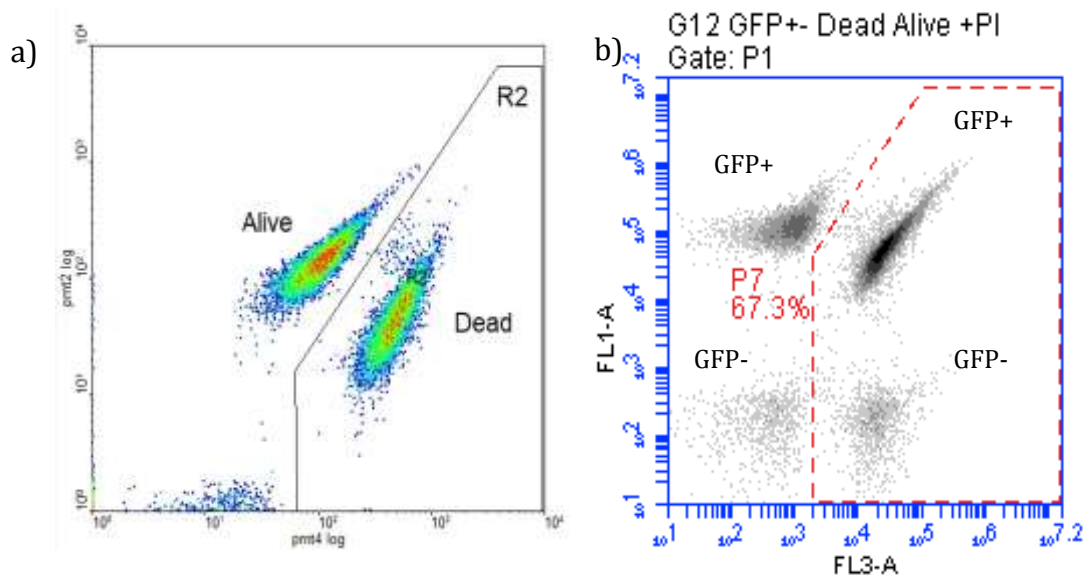
Figure 2.2: Propidium iodide staining for FCM

Control data for FCM-derived viability staining with PI. a) & b) show typical gating for identification of dead (PI+) cells for both cytometers used during this study.

a) Coulter EPICS Elite-derived control PMT2 (green fluorescence) versus PMT4 (red fluorescence) plot of a mix of live and ethanol-killed *E. coli* BL21*pETCheY::GFP stained with PI, dead cells fall within the defined gate R2. b) BD-Accuri C6-derived control FL1-A (green fluorescence) versus FL3-A (red fluorescence) plot of a mix of live and ethanol-killed *E. coli* BL21*pETCheY::GFP and untransformed cells stained with PI, dead cells fall within the defined gate P7.

2.7.5.2. Gating of fluorescent populations

It was observed that when there were populations of varying average fluorescent intensity these populations would often overlap when visualised as a histogram, due to a correlation between cell size/volume and total cellular fluorescence i.e. a small but highly fluorescent cell would give a similar fluorescence reading to a larger less fluorescent cell. This phenomenon prevented accurate gating and hence enumeration of populations, but an intensity/dot plot of forward scatter versus green fluorescence exploited the relationship between FSC and cell size and hence allowed more accurate gating (Figure 2.3), an additional benefit of this was that when a population was small enough to not be visible on a histogram it could often still be visualised on a bivariate plot.

2.7.5.3. EPICS Elite

The EPICS-Elite uses a 15 mW air-cooled Argon ion laser at 488 nm for sample excitation and for this study fluorescence was detected by photomultiplier tubes (PMTs) at 525 nm (PMT2, GFP) and 630 nm (PMT4, PI). The flow cell and optics were manually aligned whenever the machine was used using Flow-Check Fluorospheres (Beckman-Coulter). PMT voltages and fluidics settings were variable and typical values are given in Table 2.8. To eliminate background noise a variable discriminator was used on PMT1 (FSC) and SSC then 20,000 events were collected. To prevent both concurrent events and a disproportionate percentage of noise in the sample the data rate was maintained between 1-2000 where possible. The staining protocol for PI was derived from previous studies using this equipment; stock dye solution was added to cell suspensions to a final concentration of 7.5 $\mu g \cdot mL^{-1}$, incubated at ambient temperature for 1 minute then analysed.
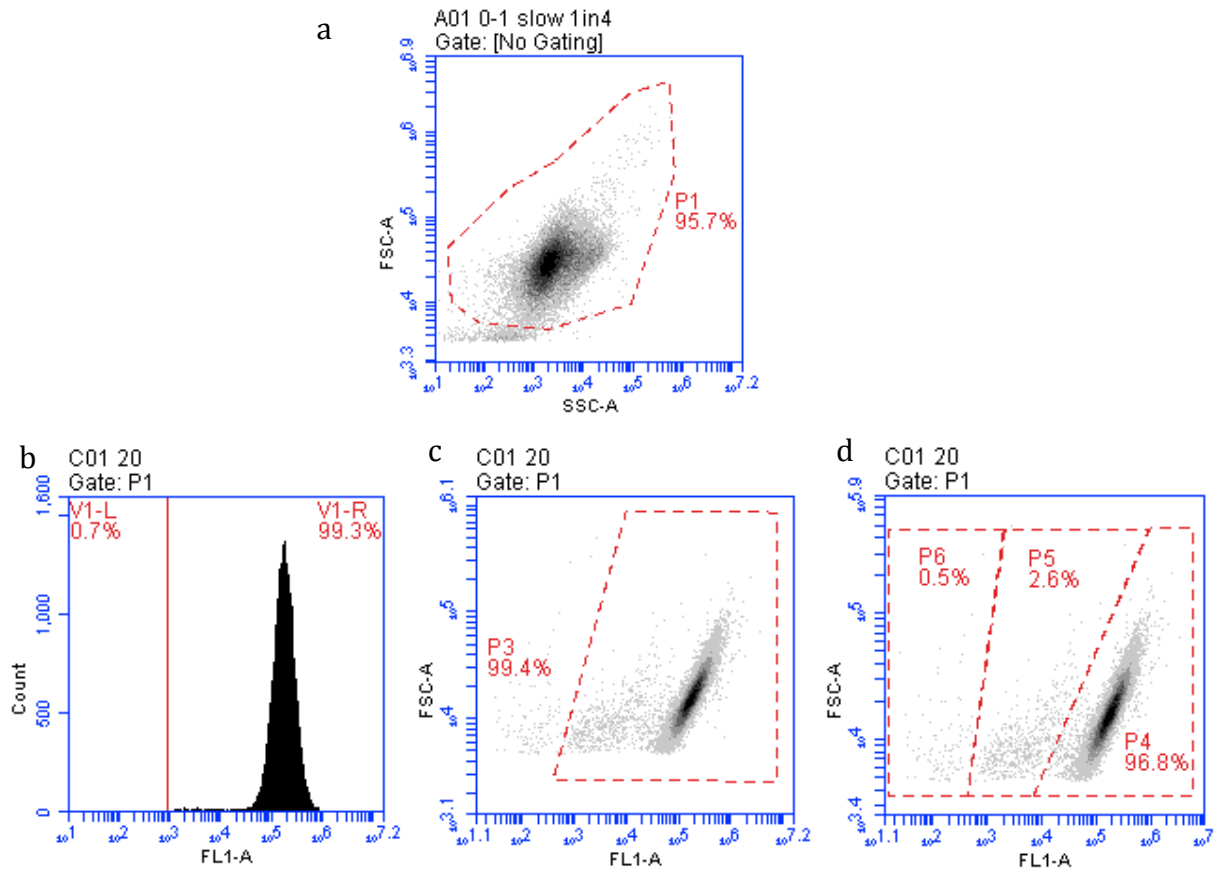
Figure 2.3: Typical gating of FCM data

a) FSC-A versus SSC-A plot of *E. coli* BL21*pETCheY::GFP showing typical gating to eliminate intrinsic noise (P1) a small amount of noise can be seen below the gate.

Typical gating of FCM data to enumerate GFP+ cells and fluorescent sub-populations: b) Histogram of mean FL1-A values for *E. coli* BL21*pETCheY::GFP (sample data) showing typical gating for GFP+/GFP- cells V1-R = GFP+, V1-L = GFP-. c) FSC-A versus FL1-A plot of same data as (b), gated such that GFP+ cells are within gate P3, additional fluorescent sub-populations can be seen that were not distinguishable in (b), unless otherwise stated GFP+ data in this work is gated in this manner. c) FSC-A versus FL1-A plot of same data as (b), gated to enumerate fluorescent sub-populations.

Table 2.8: Operating parameters for FCM data acquisition

Coulter EPICS Elite

Optical settings – PMT values used			
FS	**PMT1 (SSC)**	**PMT2 (GFP)**	**PMT4 (PI)**
150	600	650	1200

Fluidics & data collection settings	
Sheath pressure	5 psi
Sample pressure	Variable
Data rate	1000-2000 events·s^{-1}
Events collected	25000

BD-Accuri C6

Optical settings	Fixed, manufacturers settings. No artificial voltage manipulation or colour compensation applied
Fluidics settings[1]	Default slow fluidics setting: Flow rate – 14 µL·mL^{-1} Core size – 10 µm
Data rate[1]	1000-2500 events·s^{-1}
Events collected[1]	20000 (in gate P1)

[1] – Unless otherwise stated

2.7.5.4. C6

The C6 uses a 488 nm solid-state laser for sample excitation and for this study fluorescence was detected using a 533/30 BP filter for GFP (FL1-A) and 670 nm LP filter for PI (FL3-A). The flow cell alignment and fluorescence detection on the C6 is fixed therefore no alterations were made to these. To eliminate background noise data was thresholded on FSC-H (typical values 7-20,000), and 20,000 events were recorded in a gated region on FSC-A/SSC-A plots corresponding to the scatter profile of *E. coli* (P1, typical gating given in Figure 2.3). For routine analysis the pre-set fluidics setting 'Slow' was used to maximise the passage time of cells through the laser due to the relatively small size of *E. coli*. Where accurate volume measurements were required 'Medium' fluidics were used as manufacturers instructions state 'Slow' fluidics do not give accurate volume measurements. According to manufacturer's instructions samples were diluted to give a data rate of 1,000-2,500 where possible. The BD-Accuri C6 was a new acquisition during the course of this study, as such effective dye concentrations were calibrated before experimental use. PI staining was determined to show adequate differentiation between dead and alive cells at a final concentration of 4 $\mu g\cdot mL^{-1}$, therefore stock dye solution was added to this concentration (10 μL dye per 1 mL sample), incubated at ambient temperature for 3 minutes then analysed. As CR staining has not been used previously to stain for IBs in *E. coli* a range of dye concentrations and incubation times were tested, these are detailed in Section 3.2.

2.8. Protein analysis

2.8.1. SDS-PAGE

Proteins extracted from various preparations of culture samples were separated by mass using denaturing SDS-PAGE, using an Omni-PAGE mini apparatus (Cleaver

Scientific). The Omni-PAGE apparatus allows 2 gels to be run simultaneously; therefore all following amounts are for 2 gels. A 15% (w/v) acrylamide resolving gel was made by mixing 7.5 mL of protogel, 7.5 mL of 2x resolving gel buffer and 75 µL of 20% SDS. Polymerisation was initiated by the addition of 150 µL 80 mg·mL^{-1} APS solution and 7.5 µL Tetramethylethylenediamine (TEMED), the mixture was mixed briefly then immediately poured between 2 100x100 mm glass plates to 0.5 cm below the level of the combs used for producing the sample wells. To exclude oxygen from the polymerisation and to prevent evaporation of water a layer of 0.1 % (w/v) SDS was pipetted over the gel, this was then left to polymerise for approximately 1 h. Once the resolving gel had polymerised the layer of 0.1% SDS was removed and a 6% acrylamide stacking gel was added. To make the stacking gel 2 mL of protogel, 1 mL of 10x stacking gel buffer, 7 mL of ddH$_2$O and 50 µL 20% SDS were mixed and polymerisation was initiated by adding 100 µL APS and 5 µL TEMED. After mixing, a 1 mL layer of stacking gel was pipetted over the resolving gel and immediately discarded to wash any remaining 0.1% SDS away, then stacking gel was poured over to the top of the plates and a comb was inserted at an angle to prevent bubbles forming under the wells. After allowing approximately 1 h for the gel to polymerise the combs were gently removed and the wells were washed with ddH$_2$O to remove any unpolymerised acrylamide. Gels were loaded with samples as described below and run in 1x TGS buffer at 110 V for approximately 2 h at ambient temperature.

2.8.2. Sample preparation & BugBuster®

Samples for both total protein and solubility analysis were taken from cultures in the same way, so that all samples were standardised for biomass. A volume of culture in mL equal to 0.9/OD$_{650}$ was pelleted by centrifugation (16,873 g, 15 minutes) in sterile 1.5

mL or 2 mL (where appropriate) PP microcentrifuge tubes, supernatant removed by pipetting and unless used immediately stored at -20°C. For total protein analysis pelleted cells were thawed (if necessary), resuspended in 60 μL of sample buffer, boiled for 10 minutes and 5 μL was loaded on the gel (assuming OD_{650} 1 = 0.4 g·L^{-1} DCW this equates to a loading of approximately 0.03 mg biomass). To separate the soluble and insoluble protein fraction the chemical lysis agent BugBuster® (Novagen) was used. Cell pellets were thawed and resuspended in a volume of BugBuster® equal to that of sample buffer for total protein samples (60 μL), incubated at room temperature for 10 minutes then the insoluble fraction was pelleted by centrifugation (16,873 x g, 20 minutes, ambient temperature). Following centrifugation the soluble fraction (supernatant) was removed by pipetting into a fresh, sterile microcentrifuge tube. The insoluble fraction (pellet) was washed to remove any residual soluble protein by resuspension in 180 μL sterile PBS then pelleted by centrifugation (16,873 x g, 10 minutes, ambient temperature) and the supernatant removed. Both soluble and insoluble samples were then resuspended in 60 μL of sample buffer and boiled for 10 minutes. 5 μL of insoluble samples were loaded on the gel however as the soluble samples had a final volume of 120 μL, a 2x dilution of the original extract, 10 μL was loaded so that the amount loaded for total protein, soluble and insoluble corresponded to the same initial biomass.

2.8.3. Coomassie Blue Staining & Drying

Proteins were visualised by staining with a 0.2% (w/v) solution of Coomassie Brilliant Blue R. Gels were removed from the casting/running apparatus, rinsed in ddH$_2$O and incubated with the stain at ambient temperature for 30 minutes with gentle agitation. To remove non-specifically bound dye; gels were washed in fast destain twice for 30 minutes then three times in slow destain for 1 h with gentle agitation. Stained gels were

then removed from the destain solution, placed in shrink solution with agitation for 1 h then mounted between 2 sheets of cellulose gel drying film (Promega), stretched over a frame and left to dry under ambient conditions for 1-2 days.

2.8.4. Bicinchoninic acid assay

Protein concentrations from end-point fermentation samples were determined by the bicinchoninic acid (BCA) assay using a commercially-produced kit (Pierce™ BCA Protein Assay Kit) in 96-well plates according to manufacturers instructions and analysed at 560_{nm} using a GloMax®-Multi microplate reader (Promega).

2.8.4.1. Sample preparation

2 mL fermentation samples were taken at termination, cells harvested by centrifugation (16,873 g, 15 minutes) in sterile 2 mL microcentrifuge tubes, the supernatant removed by syringe and stored at -20°C until required. When required, cell pellets were thawed and resuspended in 10 mL PBS then divided into 200 μL aliquots and cells were pelleted by centrifugation (16,873 g, 10 minutes) in sterile 1.5 mL microcentrifuge tubes. Supernatants were removed by pipette and retained for analysis to determine concentration of protein released on freeze-thaw of initial cell pellet.

To determine total cellular protein content cell pellets were resuspended in 500 μL solubilisation/lysis buffer, boiled (10 minutes, 100°C) diluted as necessary in PBS and assayed.

2.9. Computational analysis

2.9.1. Flow cytometry

Due to differences in age and manufacturer; data from the 2 cytometers could not be analysed using the same software package.

2.9.1.1. WinMDI

All data gathered using the EPICS-Elite flow cytometer was analysed using the Windows Multiple Document Interface for Flow Cytometry (Win MDI) software package (http://facs.scripps.edu/software.html).

2.9.1.2. Cflow

All data gathered using the BD-Accuri C6 cytometer was initially and unless otherwise stated analysed by the software package provided, Cflow (BD).

2.9.2. SDS-PAGE

Dried SDS-PAGE gels were scanned with a Canoscan 9000F (Canon) at 600 dpi and bands were quantified by densitometry with the software package ImageJ (Abàamoff *et al.*, 2004; Schneider *et al.*, 2012) (Figure 2.4). Images were first subjected to background subtraction (default settings) followed by lane definition. Histograms of colour intensity were produced from defined lanes, peaks corresponding to CheY::GFP were defined and the areas underneath calculated. To estimate the percentage of soluble CheY::GFP, CheY::GFP peak areas for soluble and insoluble samples were compared. To estimate the percentage of total protein that was CheY::GFP, either the intensity of the CheY::GFP peak was calculated as a percentage of the intensity of the total lane or the intensities of the CheY::GFP peaks of both the soluble and insoluble samples were calculated as a percentage of the intensities of both samples combined.
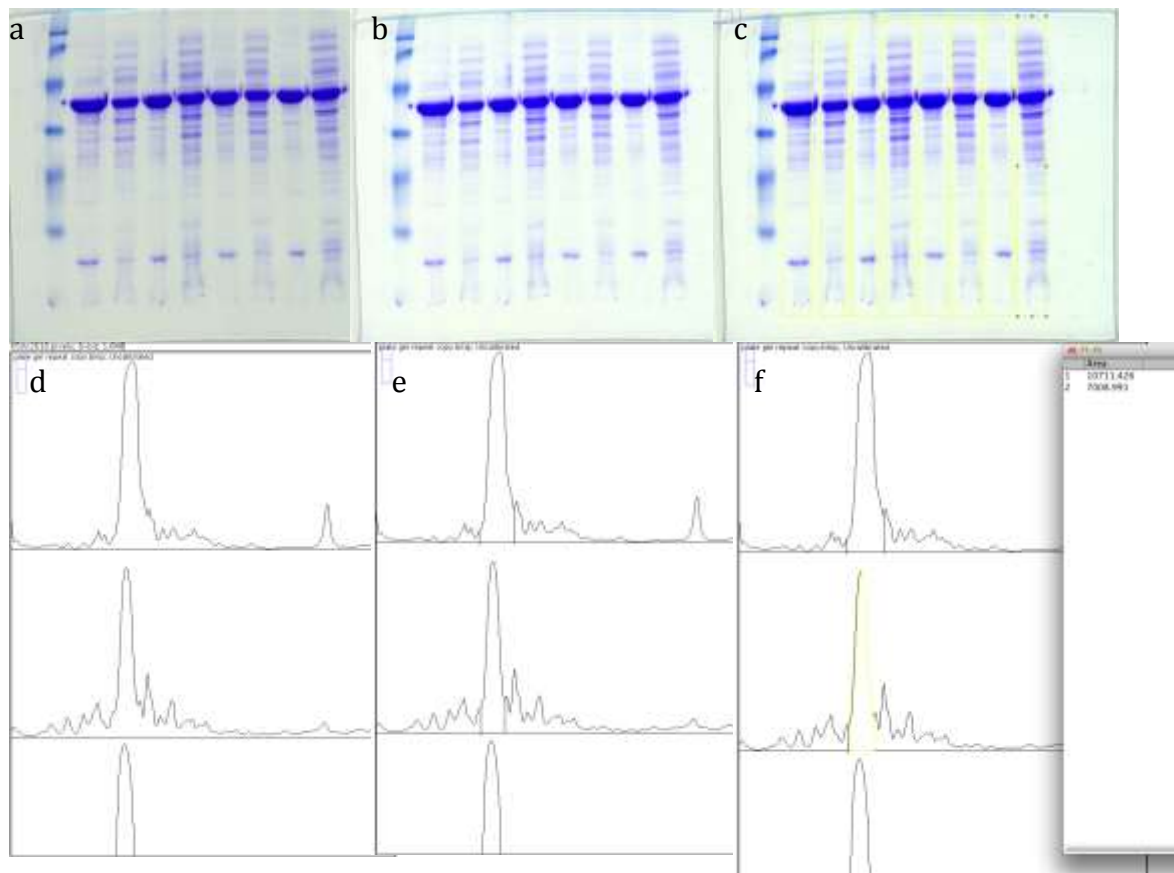
Figure 2.4: Densitometric analysis of SDS-PAGE gels
Analysis of a sample SDS-PAGE gel by ImageJ. a) Gel was scanned at 600 dpi and opened in ImageJ. b) Background was removed by default settings. c) Lanes were defined. d) Lane histograms were plotted. e) CheY::GFP peaks were defined. f) CheY::GFP peaks were selected, area under the peak automatically calculated and used for analysis (for solubility analysis, area of insoluble and soluble CheY::GFP peaks were compared, for total production analysis, area under CheY::GFP peaks were compared to area under remainder of the histogram).

2.9.2.1. Protein yield analysis

Initial estimates of CheY::GFP yields were predicted in a similar way to Sevastsyanovich *et al.* (2009) from biomass values and SDS-PAGE data assuming protein comprises 60% of *E. coli* dry cell mass (based on 50-61% estimates from Valgepea *et al.* (2013) and 70% from Sevastsyanovich *et al.* (2009)). This method however does not account for any variation in protein content that may have occurred therefore yields were also calculated from BCA assay-derived total cellular protein concentrations.

2.10. Preliminary experiments

In order to ensure consistent growth of microbial cultures and accuracy of analytical methods a series of preliminary experiments were made.

2.10.1. Batch variation of complex growth medium components

During the initial stages of this study tryptone from Fisher Scientific and yeast extract from Oxoid were used for media production, but on replacement of supplies of both a decrease in the final OD_{650} for overnight cultures was observed (from approximately 7 to 5). Despite replacement of the yeast extract from the vendor it was not possible to recover growth to previous levels. In order to ensure reproducible cell growth a brief study was made into batch variation of tryptone and yeast extract. 35 mL overnight cultures (LB, 30°C, 150 RPM) were grown using a number of sources and batches of both tryptone (BD Bacto, Oxoid, Sigma and Fisher bioreagents) and yeast extract (BD Bacto, Oxoid x2 and Duchefa Biochemie). For ease of comparison 2 series of cultures were made; the first containing each tryptone sample along with BD yeast extract, the second containing BD tryptone with each yeast extract sample as those from BD were available in the largest amount. In addition a batch of pre-mixed LB of the same composition was

also used (Sigma). OD$_{650}$ measurements were taken for each combination following 14 h growth (Figure 2.5).

OD$_{650}$ data showed a range of cell densities generated. For tryptone the range was from 5.87 (Oxoid) to 8.41 (Sigma). For yeast extract the range was from 4.73 for Oxoid 1 to 7.14 for BD, both Oxoid samples gave similar results (4.81 for Oxoid 2). LB using BD tryptone gave the second highest OD$_{650}$ of all tryptone samples and BD yeast extract gave the highest (7.14), similar to that of initial work. Both BD tryptone and yeast extract were available in single batches sufficient for the remainder of this study and thus were used.

2.10.2. Production of GFP under oxygen-limited conditions

This work relied heavily on the estimation of CheY::GFP accumulation by fluorescence, either by fluorimetry or FCM. However, the *av*GFP fluorophore requires O$_2$ to produce its mature, fluorescent form (Figure 1.7) (Hansen *et al.* (2001) demonstrated that GFP produced in *Streptococcus gordonii* ceased to fluoresce between 0.1-0.025 ppm) and GFP produced in anaerobic environments can accumulate in a folded but non-fluorescent, immature form (imGFP) (Zhang *et al.*, 2005). Bioreactor cultures, even under aerobic conditions, can be subject to inefficiencies in mixing and aeration, resulting in microoxic and anoxic microenvironments sufficient to result in cells upregulating genes for anaerobic growth, e.g. the *nar* genes for nitrate reduction (Garcia *et al.*, 2009). This could only be exacerbated in HCDC where cellular oxygen demands can exceed the aeration capacity of the bioreactor, as frequently observed during this study. It is therefore possible that during later stages of fermentation where the DOT is severely limited (~0%) immature, non-fluorescent CheY::GFP could accumulate and hence cause underestimation of RPP productivity.
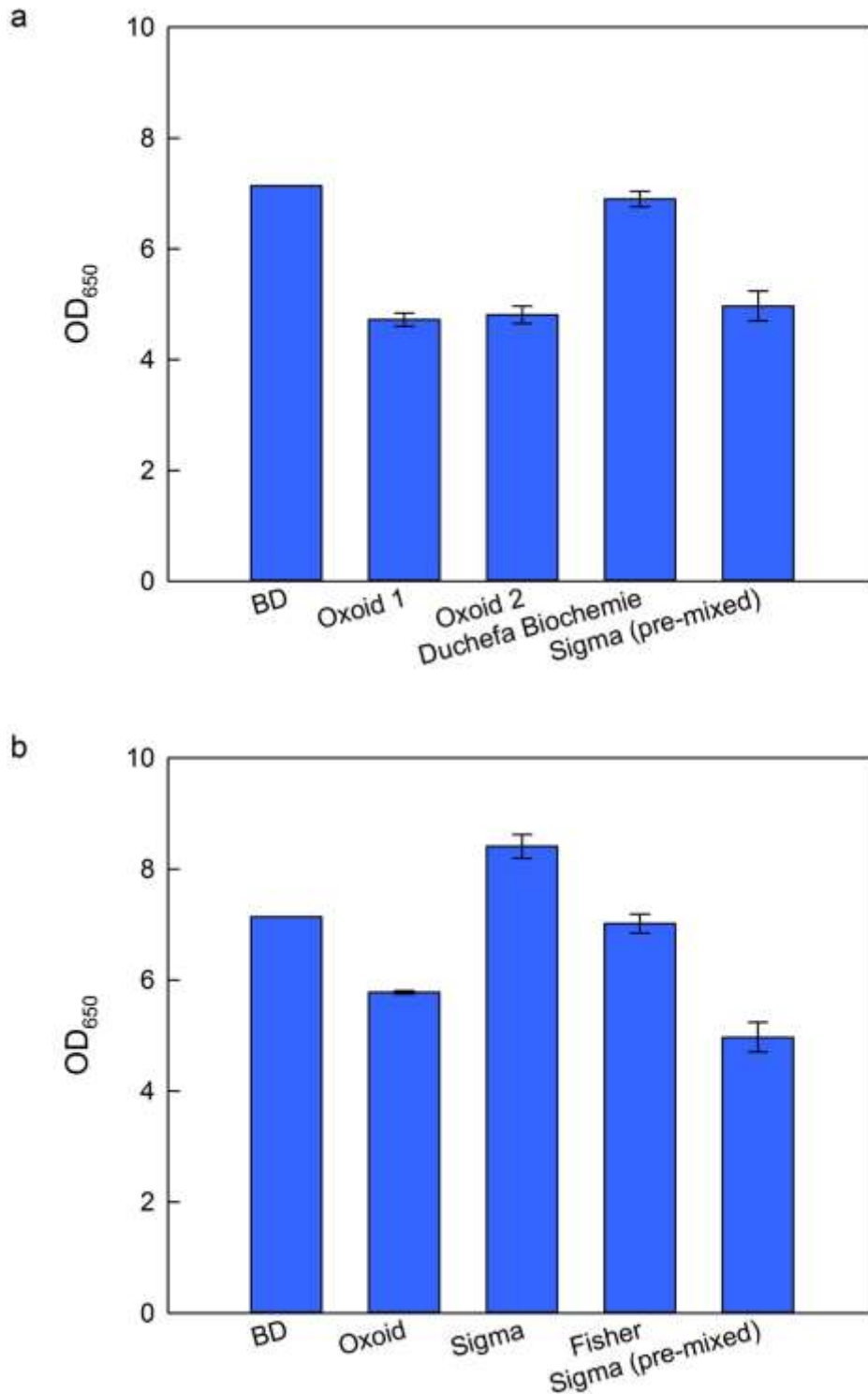
Figure 2.5: OD$_{650}$ data for overnight cultures of *E. coli* BL21* pET CheY::GFP using different sources of tryptone and yeast extract
a) Varying sources of tryptone with BD bacto yeast extract. b) Varying sources of yeast extract with BD bacto tryptone. Data are mean values derived from 2 replicate cultures, error bars are ± 1x standard deviation.

To determine whether imGFP can accumulate in oxygen limited cultures of *E. coli* an aerobic fluorescence recovery (AFR) protocol was used; anaerobically grown GFP-expressing cells were resuspended in PBS to limit further cell growth and protein production, then incubated under aerobic conditions. If imGFP was present, exposure to oxygen during AFR would allow maturation of the chromophore and hence cause an increase in fluorescence measurements. AFR was initially proposed by Zhang *et al.* (2005), to observe GFP-producing *Enterobacter aerogenes* grown under the stringently anaerobic conditions required for biohydrogen production. AFR was used for this study as it would not only identify imGFP accumulation but would also be a useful sample processing stage if significant amounts of imGFP were shown to form.

AFR was initially tested using *E. coli* SCC1, a K-12-derivative strain that expresses GFPmut3* constitutively. The reason for using SCC1 is that GFPmut3*, being a less-stable variant of GFP limits accumulation of fluorescent protein that might mask small changes in fluorescence and constitutive expression eliminates variability resulting from RPP induction, also, *E. coli* BL21* as a B-derivative strain is unable to be grow anaerobically (Studier *et al.* 2009; Pinske *et al.* 2011).

E. coli SCC1 was grown in 100 mL shake flask cultures (LB, 0.5% glucose) under both aerobic and oxygen-limited conditions then harvested for AFR after 3 h growth. Cells were harvested at 3 h based on data from preliminary growth curves. After 3 h both aerobic and oxygen-limited cultures were in exponential growth phase (Figure 2.6a) and lower cell densities for the latter suggested that cells were in fact oxygen-limited. Also, peak specific fluorescence (fluorescence per unit OD_{650}) was observed for both aerobic and oxygen-limited cultures after 3 h growth (Figure 2.6b). At harvesting it was observed that the specific fluorescence of aerobic cultures was lower than that obtained
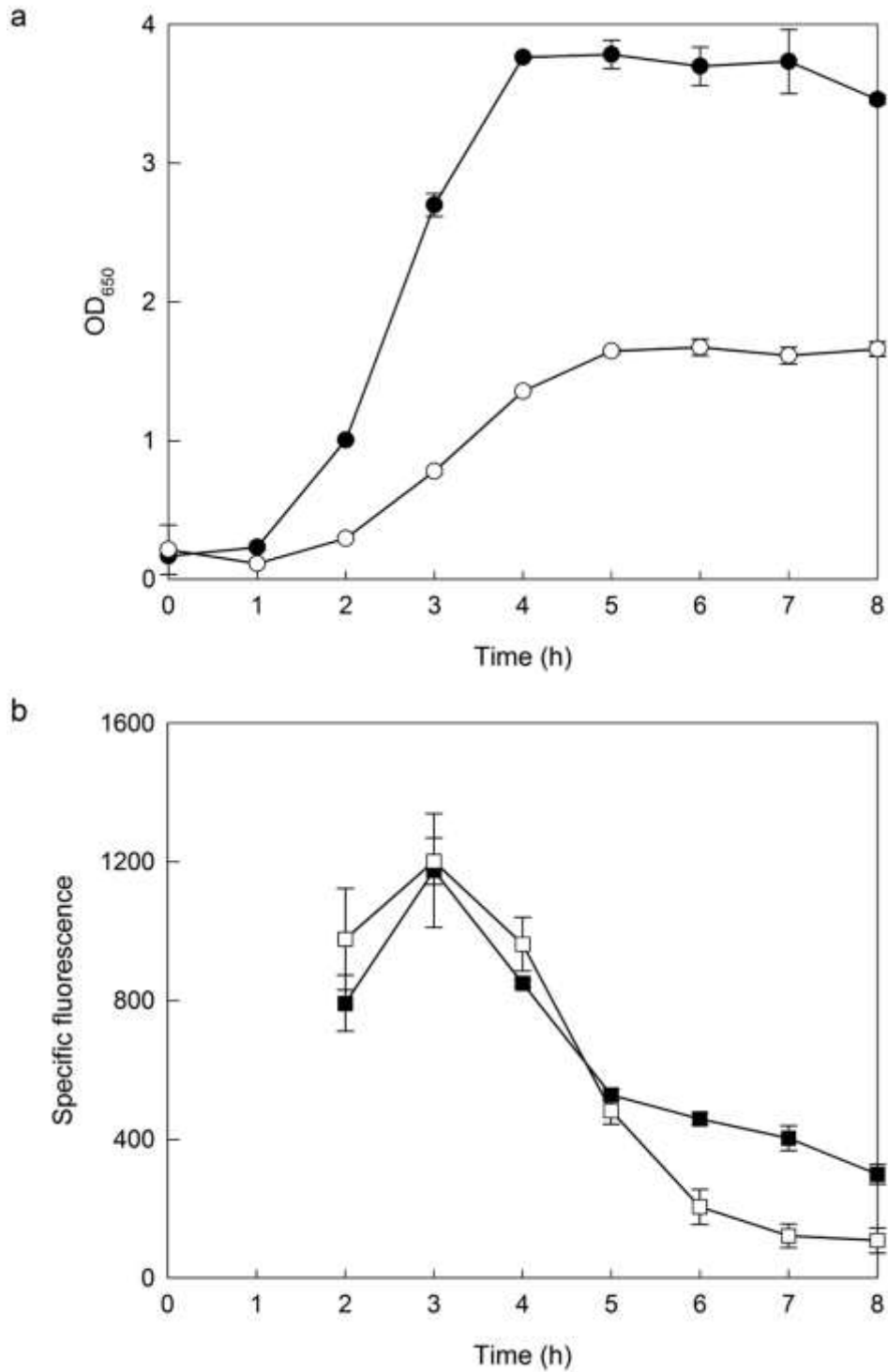
Figure 2.6: Growth and fluorescence profiles for *E. coli* SCC1 over 8 h under aerobic and oxygen-limited conditions

a) OD_{650} data for aerobic ($\bigcirc$) and oxygen-limited ($\bullet$) cultures. b) Relative fluorescence (fluorescence per unit OD_{650}) data for aerobic ($\square$) and oxygen-limited ($\blacksquare$) cultures. Data are mean values from 2 replicate cultures from a single representative experiment, error bars are ± 1 standard deviation.

during preliminary growth curves, this was interpreted as possibly being due to differences in culture handling i.e. cultures used for solely growth analysis would have been removed from the incubator more often than those grown for AFR, possibly retarding growth. Cells were resuspended in PBS and incubated under AFR conditions (37°C, 200 RPM) for 2 h, sampling for OD_{650} and fluorescence every 30 minutes; in addition samples of the original culture at the point of harvesting were also analysed. After 2 h AFR specific fluorescence values for anaerobic cultures were lower than that of the starting culture. (Figure 2.7a), presumably due to an increase in OD_{650} observed (Figure 2.7b). Unexpectedly, aerobically-grown cells showed an increase in specific fluorescence during AFR and in comparison to the starting culture, presumably as a result of a decrease in OD_{650}. As AFR did not cause an increase in specific fluorescence for oxygen-limited cultures, it was concluded that under the conditions used there either was insufficient imGFP to warrant routine use of AFR or that sample handling introduced sufficient oxygen for maturation of any imGFP produced.

2.10.2.2. Oxygen limitation in HCDC-RPP fermentation

To further validate the conclusion that the AFR protocol was not necessary for routine use during this study, specifically when applied to oxygen-limited, HCDC fermentations, the AFR protocol was tested on cells grown under these conditions. Cells were taken from a fermentation following growth to high cell density ($OD_{650 \, nm} \sim 250$) when the DOT was severely depleted ($\sim 5.0\%$), divided into replicates and subjected to AFR. In order to limit potential aggregation of CheY:GFP, at 37°C and hence decrease in fluorescence, the AFR protocol was modified: The incubation temperature was decreased to 25°C and as this may have caused a decrease in the rate of the maturation reaction the incubation time was increased to 4 h. Samples were then analysed by FCM. FCM analysis showed that in the original culture and during the 4 h AFR 97-98% of cells
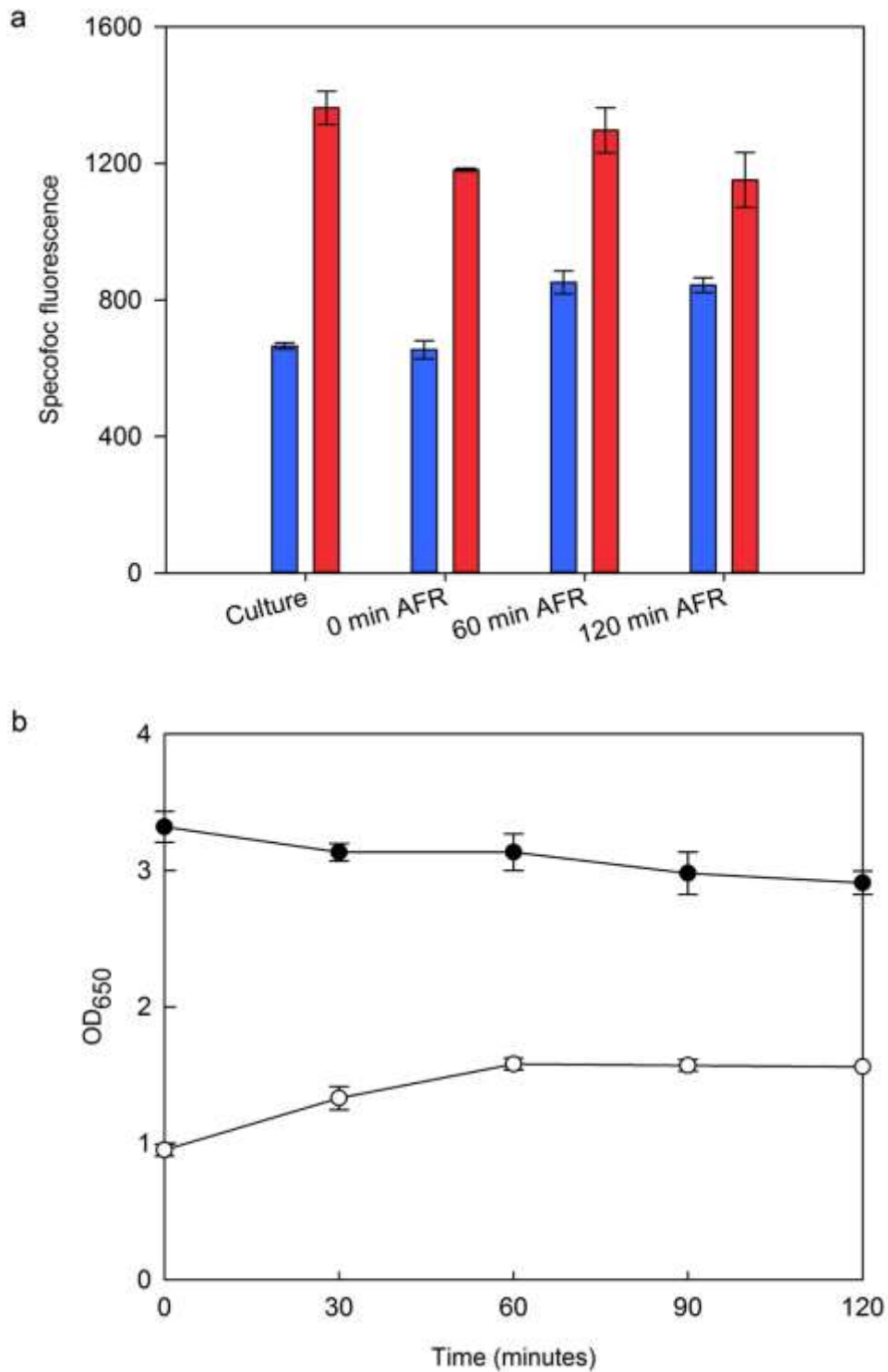
Figure 2.7: Data from AFR treatment of aerobic and oxygen-limited *E. coli* Scc1
a) Comparison of relative fluorescence from AFR to the original culture for aerobic (blue) and oxygen-limited (red) samples. b) OD_{650} over 2 hours AFR for aerobic ($\bigcirc$) and oxygen-limited ($\bullet$) cultures. Data are mean values from 2 replicate cultures from a single representative experiment, error bars are ± 1 standard deviation.

were GFP[+] and there was no overall trend in FSC (data not shown). Although a small (2.97% overall) increase was observed in green fluorescence during AFR and in comparison to the original culture (Figure 2.8a) and there was an increase in the proportion of dead (PI[+]) cells after 4 h (Figure 2.8b) neither trend was statistically significant. It was therefore concluded that the additional time and expense in sample processing for AFR outweighs the benefits of its inclusion.

2.10.2.3. Conclusion

The application of AFR to the analysis of fluorescence in oxygen-limited *E. coli* cultures was shown to be unnecessary for both GFPmut3* and CheY::GFP grown under shake flask and HCDC conditions. The limited concentration of oxygen available in the instances tested appears to have been sufficient to allow maturation of the GFP produced. Alternately it is possible that small quantities of imGFP present in the culture on sampling may have been exposed to sufficient oxygen during sample handling to allow for maturation. As the rate constant for the oxidation stage of GFP maturation (S65T variant) is estimated at 19-83 minutes (Heim *et al.* 1995, Reid & Flynn 1997, Tsien 1998) there may have been sufficient time for this to occur. During preparation for both fluorimetry and flow cytometry samples were diluted in buffer that was not subject to degassing and hence will have contained dissolved oxygen and samples were agitated/vortexed for mixing. It is possible that this provided sufficient oxygen for maturation. The use of an autosampler connected to the culture apparatus may possibly minimise the amount of agitation as a result of sample handling and therefore demonstrate whether this is the case. It is also possible that in future comparisons of fluorescence data to predicted fluorescence based on SDS-PAGE derived yields of GFP may also assist in assessing this phenomenon. This method would be of particular use if imGFP shows decreased stability relative to mature GFP and hence misfold.
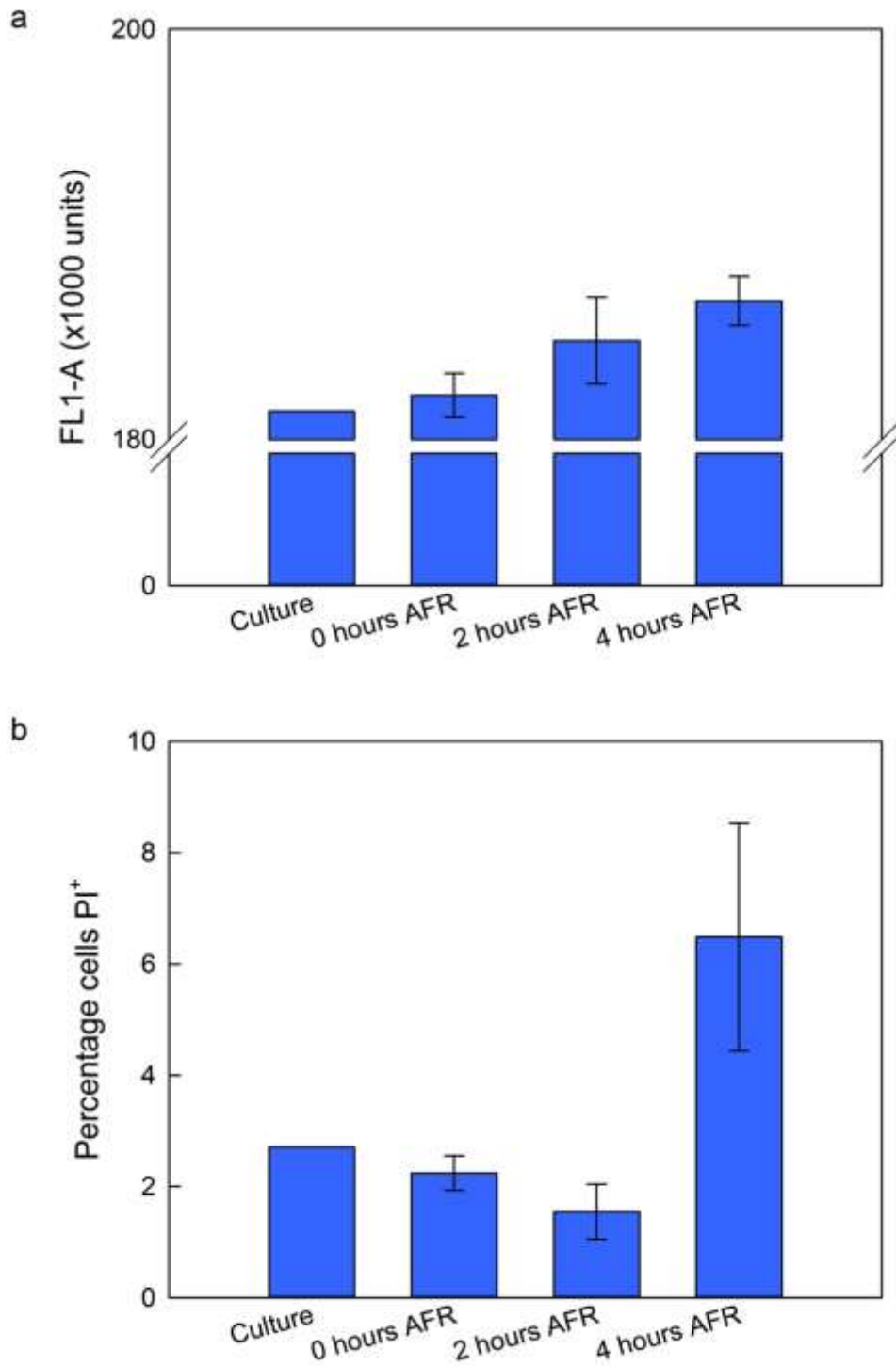
Figure 2.8: FCM data from modified AFR treatment of oxygen-limited HCDC *E. coli* BL21* pETcheY::GFP

a) Comparison of mean FL1-A (green fluorescence) from 4 hours AFR to the original culture. b) Percentage dead cells (PI+) in original culture and during 4 hours AFR. Data shown are single values for culture samples and mean values from 2 replica flasks for AFR samples, error bars are ± 1x standard deviation.

Chapter 3: Results 1 – Novel Applications of Flow Cytometry for

Bioprocess Monitoring and Control

3.1. Introduction

FCM is of great utility in RPP bioprocess monitoring using established techniques such as viability staining and the use of recombinant protein-FP fusions, but this is by no means the full scope of applications FCM could have. The first aim of this work was to "To investigate and develop additional methods by which FCM can be used for the monitoring and analysis of RPP cultures in *E. coli*", this chapter therefore describes a series of experiments that develop a novel method for RPP monitoring by FCM and apply FCM analysis to aspects of RPP and fermentation not previously reported in *E. coli*.

3.2. Use of the amyloidophilic dye Congo red as a stain for the flow cytometric detection of inclusion bodies

3.2.1. Introduction

There are many advantages to the production of RPs in their native, soluble form, as opposed to in IBs, therefore a method to detect the formation of IBs during a bioprocess would be beneficial. This has been attempted using bulk measurements, such as the work reported by Jin *et al.* (1994) who showed that inclusion body presence can be predicted based on the OD_{600}/OD_{420} ratio of *E. coli* cell debris during purification. There have also been other attempts using FCM. The simplest methods use light scattering, based on the observation that IBs are highly refractive particles and therefore their presence should result in an altered light scatter profile. While there have been successes in this regard (Lewis *et al.*, 2004; Hedhammar *et al.*, 2005), as light scattering can be affected by numerous other cellular parameters and by cytometer design, success appears to be case-dependent (Wållberg *et al.,*. 2005).

The use of FP fusions as solubility indicators have been in many cases successfully applied to FCM but these do not directly detect IBs, only a decrease in fluorescence

indicative of IB formation and may be complicated by the fact that FP-IBs can retain some fluorescence (García-Fruitós *et al.*, 2005). Ramdzan *et al.* (2012) addressed these concerns for RPP in mammalian cells using pulse width analysis (PWA). Soluble FPs will be diffuse throughout the cytoplasm and therefore the width of fluorescent pulses detected will be equivalent to that of the whole cell whereas an inclusion body will be spatially constricted and the resultant fluorescent pulse will be narrower than that of the cell. This technique is unlikely to be as effective in bacteria due to the much smaller cell size.

The area that shows the greatest promise in flow cytometric IB detection is that of fluorescent dyes; they are generally influenced by fewer factors than light scatter and hence can be more specific, and can be applied to any RP without the use of an FP-fusion.

An obvious choice of dye for detecting IBs are fluorescent amyloidophilic dyes such as Thioflavin-S (Th-S), Thioflavin-T (Th-T) and Congo red (CR), as IBs are known to possess considerable amyloid character (Wang *et al.*, 2008, de Groot *et al.*, 2009) and these dyes are used for the histological detection of amyloid. There have been a limited number of studies applying these dyes to the detection of *E. coli* RP-IBs with and without FCM. Esparagó *et al.* (2012) positively identified IBs formed from the amyloid β peptide (A-β) in *E. coli* using Th-S by FCM. Upadhyay *et al.* (2012) demonstrated binding and fluorescence of CR to isolated *E. coli* RP-IBs and Wall & Solomon (1999) and Wall (2002) demonstrated FCM detection of CR-stained amyloid fibrils from mammalian samples. These studies leave ample opportunity for further investigation as CR has not been applied to *in vivo* detection of IBs and while thioflavin has been used *in vivo*, A-β has little utility as a model RPP product. A series of experiments were therefore devised to

determine the efficacy of CR in staining IBs formed by the model RPP product CheY::GFP *in vivo* and their detection by FCM.

3.2.2. Results & discussion

3.2.2.1. Initial trial

For an initial limited trial to determine if it was possible to stain IB-containing *E. coli* with CR and to detect these cells by FCM, a 100 mL shake flask culture of *E. coli* BL21* pETCheY::GFP was grown according to protocol S1 (37°C pre-induction, 25°C post-induction, 0.5 mM IPTG) (Table 2.5) and sampled at 24 h post-induction. Samples were analysed by FCM first unstained then stained with CR by the addition of 20 $\mu L \cdot mL^{-1}$ of a stock dye solution (100 mg CR, 100 mg NaCl in 5 mL 80% (v/v) ethanol, 0.22 μm filtered), incubated (> 10 minutes, ambient conditions) and reanalysed. A comparison of the stained and unstained cells (Figure 3.1) showed a slight increase in red fluorescence (FL3-A) after staining for all cells except the population with highest green fluorescence (FL1-A) and the presence of an additional population (4.6%) with further increased FL3-A. This additional population was assumed to correspond to cells containing CR-bound IBs. The smaller increase in FL3-A in the majority of cells was assumed to correspond to non-specific binding for example to curli, amyloid-containing fibres found on the exterior of many *Enterobacteraciae* and the usual target for CR staining in *E. coli* (Barnhart & Chapman, 2006). It was therefore concluded that the test was successful. Having established that cells grown under conditions favouring IB formation display a population with increased red fluorescence when stained with CR, it was decided to investigate further: First to develop a more optimal staining protocol and second to determine the efficacy of the optimised protocol as a potential process diagnostic methodology.

119
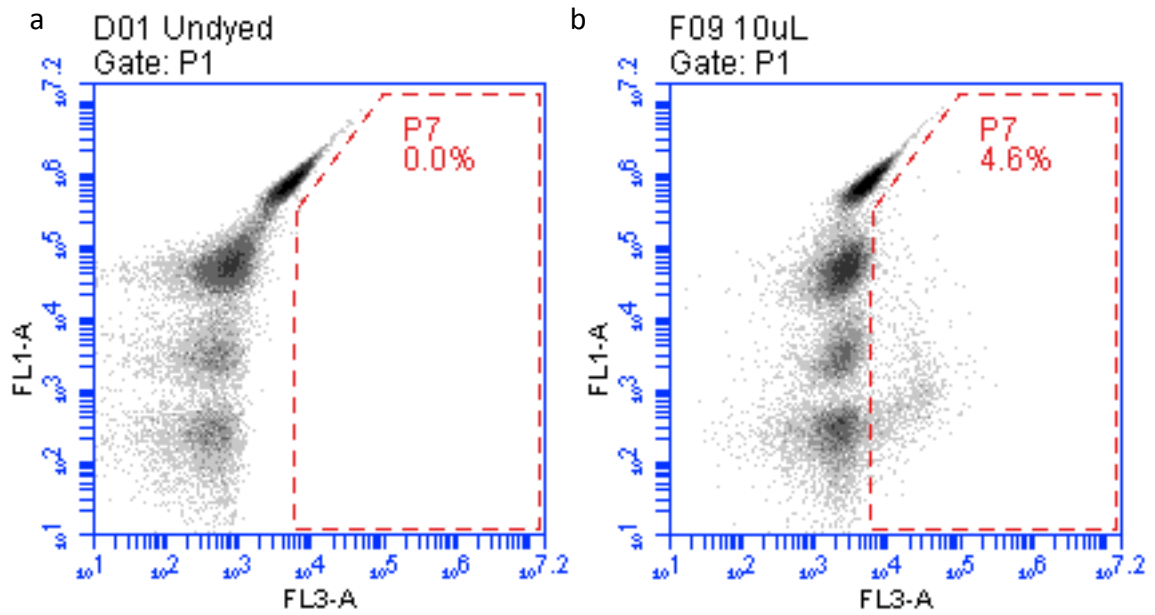
Figure 3.1: FCM data of CR (in 80% EtOH/NaCl) stained *E. coli* BL21* pETCheY::GFP FL1-A (green fluorescence) versus FL3-A (red fluorescence) plots of cells taken from a single culture 24 h post-induction following protocol S1 production conditions. a) unstained cells. b) CR-stained cells. Gate P7 corresponds to CR+ cells.

In order to further test the staining protocol 100 mL shake flask cultures were grown under 3 conditions with varying propensity for RPP related stress and IB formation: protocols SA (mild/improved conditions) (25°C, 8 µM IPTG), S1 (standard conditions) (37°C pre-induction, 25°C post-induction, 100 mM IPTG) and S2 (increased stress conditions) (37°C, 100 mM IPTG) (Table 2.5). As in Section 3.2.2.1 samples were taken from cultures at 24 h post induction and analysed by FCM. Samples were stained with 20 µL·mL^{-1} of CR stock solution and incubated for 30 minutes as this was determined to be sufficient, duplicate samples were also stained with PI.

Following staining all protocols clearly showed a CR$^+$ population of cells and the proportions of CR$^+$ cells appeared consistent with the expected amount of IBs in each protocol (Figure 3.2). Protocol SA should have produced the lowest amount of aggregated protein, S1 an intermediate amount and S2 the largest amount. SA contained the fewest CR$^+$ cells (1.0%), S1 contained an intermediate amount (3.2%) and S2 the most (12.2%). It was also observed that the percentage of CR$^+$ cells was in all cases lower than the percentage of PI$^+$ cells, while this may have been an accurate reflection of culture state (i.e. CR$^+$ cells were also generally PI$^+$) it may also indicate issues with permeability of the cell membrane to CR. PI only enters cells with a compromised membrane and it is likely that not all PI$^+$ cells would contain IBs therefore if CR was also excluded by intact membranes CR$^+$ cells would form a subset of PI$^+$ cells. To determine whether this was the case it was decided to modify the staining protocol to improve permeability of the cell membrane to CR. Attempts to permeablise the cells with both ethanol (25%) and DMSO (50%) proved inconclusive, the percentage of CR$^+$ cells did not increase substantially and cell morphology, as indicated by light scatter, was altered. CR was therefore dissolved in DMSO, an amphiphillic solvent known for its ability to

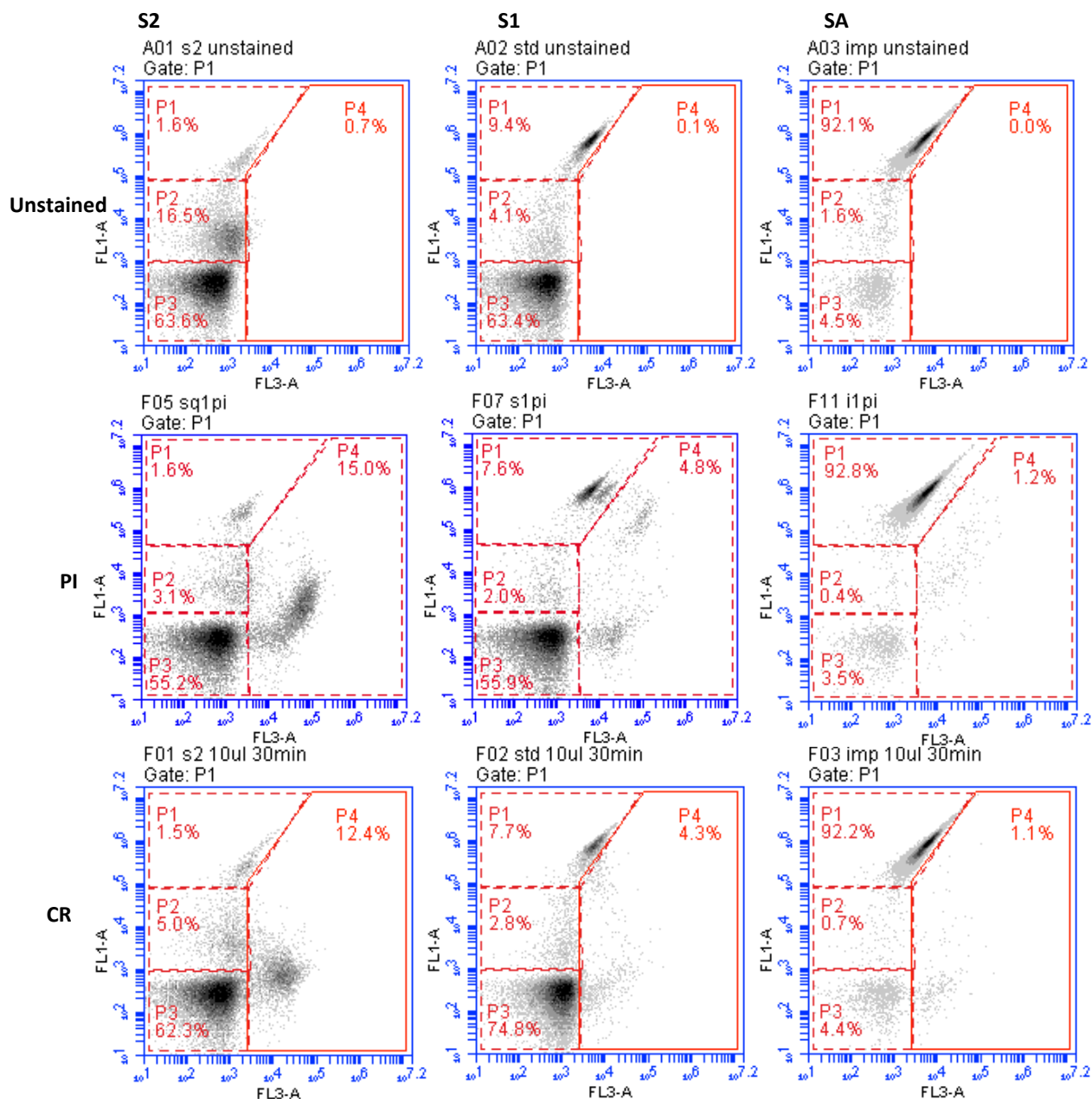**Figure 3.2: FCM data of CR (in 80% EtOH/NaCl) stained _E. coli_ BL21* pETCheY::GFP**
FL1-A (green fluorescence) versus FL3-A (red fluorescence) plots of cells taken from single cultures grown to 24 h post-induction under protocols S2 (increased stress conditions), S1 (standard conditions) and SA (improved conditions). Gates P1,2 & 3 corresponds to populations P1,2 & 3 respectively, gate P4 corresponds to PI[+] or CR[+] cells.

permeate cell membranes and to increase membrane permeability to solutes (Yu & Quinn, 1994). Advantageously CR appeared to be more soluble in DMSO than in NaCl/80% ethanol and the DMSO-based stock solution was more stable as evaporation of ethanol would frequently cause precipitation.

To ensure the CR+ cells observed were due to production of CheY::GFP and not for example due to production of curli, 2 cultures of untransformed *E. coli* BL21* were grown to 24 h post-inoculation and analysed by FCM with and without staining (20 μL·mL^{-1} of 2 mg·mL^{-1} CR (in DMSO), 30 minutes ambient conditions). Unstained samples showed 0% of cells within the CR+ gate and stained only 0.5% (Figure 3.3), fewer than observed for the lowest RPP sample (1.0% for improved conditions). As the percentage of CR+ cells corresponded to the predicted amounts of IBs for each set of production conditions and that only minimal levels of CR+ cells were observed in plasmid⁻ cells it can be reasonably concluded that CR appeared to stain for IBs and thus warranted further study.

3.2.2.3. Test of staining protocol

To test fully the potential of flow cytometric CR staining for IB detection a series of shake flask experiments were set up (100 mL, LB, 0.5% glucose) and grown according to protocols SA, S1 and S2. Cultures were analysed throughout by FCM, staining with PI and CR and by SDS-PAGE.

The proportion of CR+ cells remained below 5% throughout for SA cells and also for S1 (Figure 3.4a). For S2 cells the proportion of CR+ cells remained low during the early stages of culture, only showing substantial amounts at 24 h post-induction. CheY::GFP solubility at 24 h post-induction appeared in agreement with the predicted stress levels of each protocol, SA produced the highest proportion of soluble protein, S2 the least and S1 an intermediate amount, whereas the proportion of CR+ cells only increased for S2
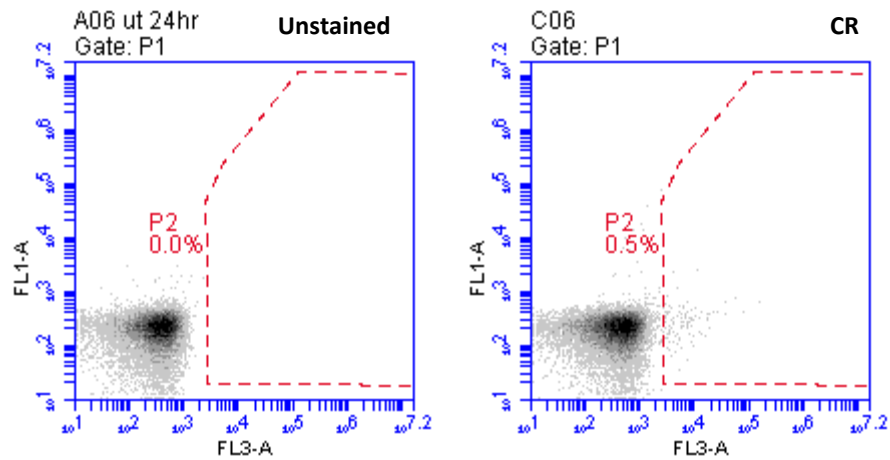
Figure 3.3: FCM data of CR (2 mg·mL^{-1} in DMSO) stained untransformed *E. coli* BL21*

FL1-A (green fluorescence) versus FL3-A (red fluorescence) plots of cells taken from a single culture (of 2 replicates) 24 h post-inoculation following protocol S1 production conditions. Gate P2 corresponds to CR$^+$ cells.
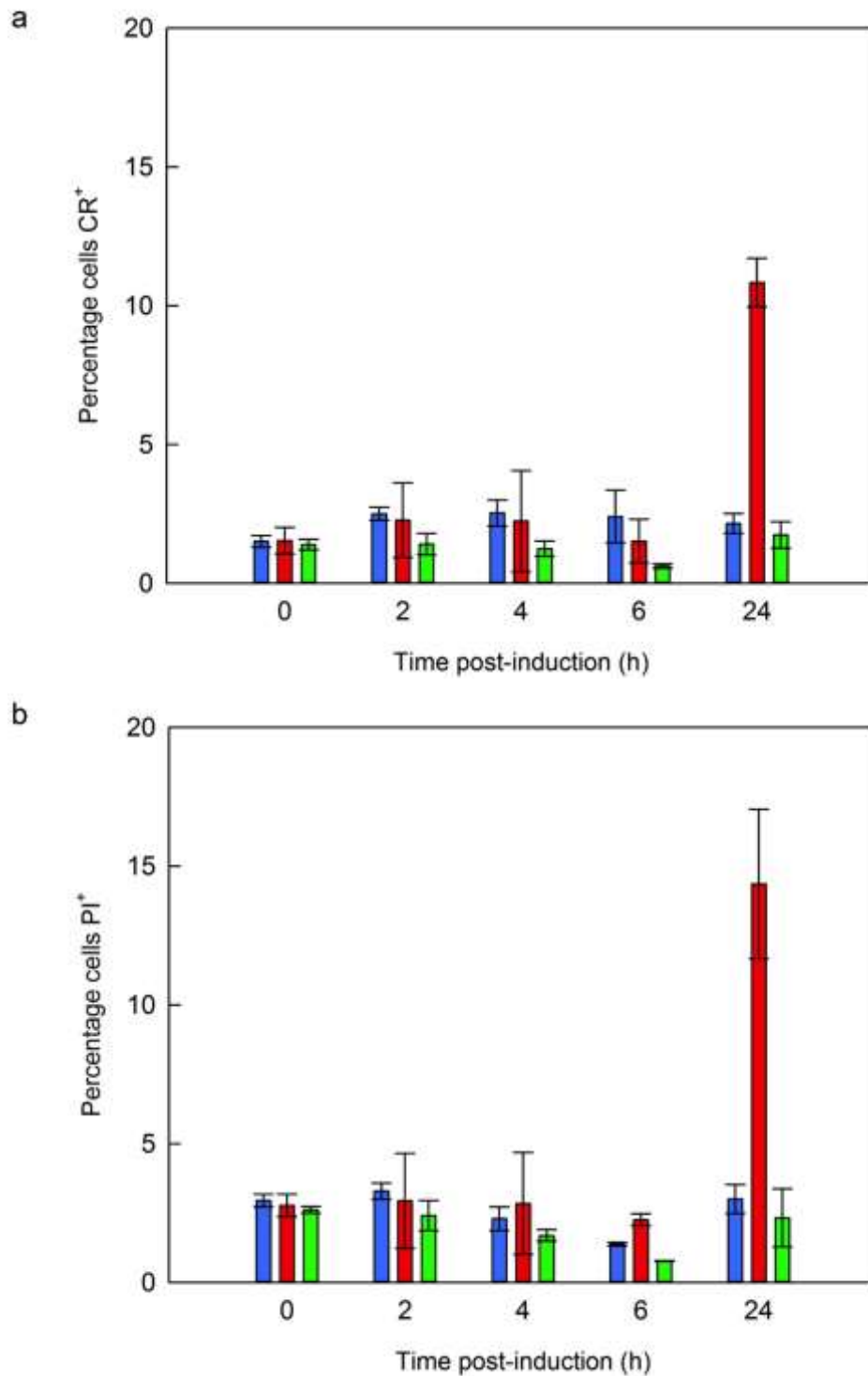
Figure 3.4: FCM data of CR and PI-stained shake-flask *E. coli* BL21* pETCheY::GFP RPP cultures

a) Percentage of CR+ cells for cultures grown under protocols S1 (blue), S2 (red) and SA (green). b) As (a) but for PI+ cells. Data are mean values of 4 cultures from 2 replicate experiments, error bars are ± 1x standard deviation.

with SA and S1 showing similar amounts (Figure 3.5). These data suggest that the relationship between formation of CheY::GFP IBs and the amount of amyloid character exhibited is not as simple as originally assumed and that IBs formed under lower temperatures possess negligible amounts of amyloid whereas those produced under higher temperatures posses substantial amounts. S1 and S2 differ only by the production temperature (25°C to 37°C, respectively) and therefore this is the likely cause for the difference in proportion of CR[+] cells observed. A temperature-dependent relationship whereby increased temperature favours increased amyloid content is logical. IBs can be divided into classical and non-classical forms; classical IBs are typically less biologically active, more resistant to proteolysis and denaturation and contain higher amyloid content (Jevševar *et al.*, 2005; Peternel *et al.*, 2008; Peternel & Komel, 2010). Increasing growth temperature is known to not only decrease biological activity of RP-IBs (García-Fruitós *et al.*, 2007a) but also increase resistance to proteolysis and denaturation (de Groot & Ventura, 2006) i.e. increase classical characteristics. It is therefore highly likely that increasing growth temperature would also increase amyloid content.

The proportion of CR[+] cells remained lower than that of PI[+] cells in most cases as previously observed (Figure 3.4a.b), except for 2 points; 4 and 6 h post-induction for S1 cells, with 2.52 and 2.4% CR[+] to 2.30 & 1.38% PI[+] respectively. These data points suggest that CR is able to enter non membrane-compromised cells and the observation that CR[+] cells frequently appeared to be a sub-set of PI[+] cells is likely accurate. This therefore suggests that CheY::GFP IBs with high amyloid character may be cytotoxic. The presence of IBs in cells is known to cause growth defects (Lindner *et al.*, 2008), amyloid deposits may be cytotoxic in the absence of the molecular chaperones DnaK or GroEL (González-Montalbán *et al.*, 2005) and amyloid peptides may have antimicrobial properties (Kagan
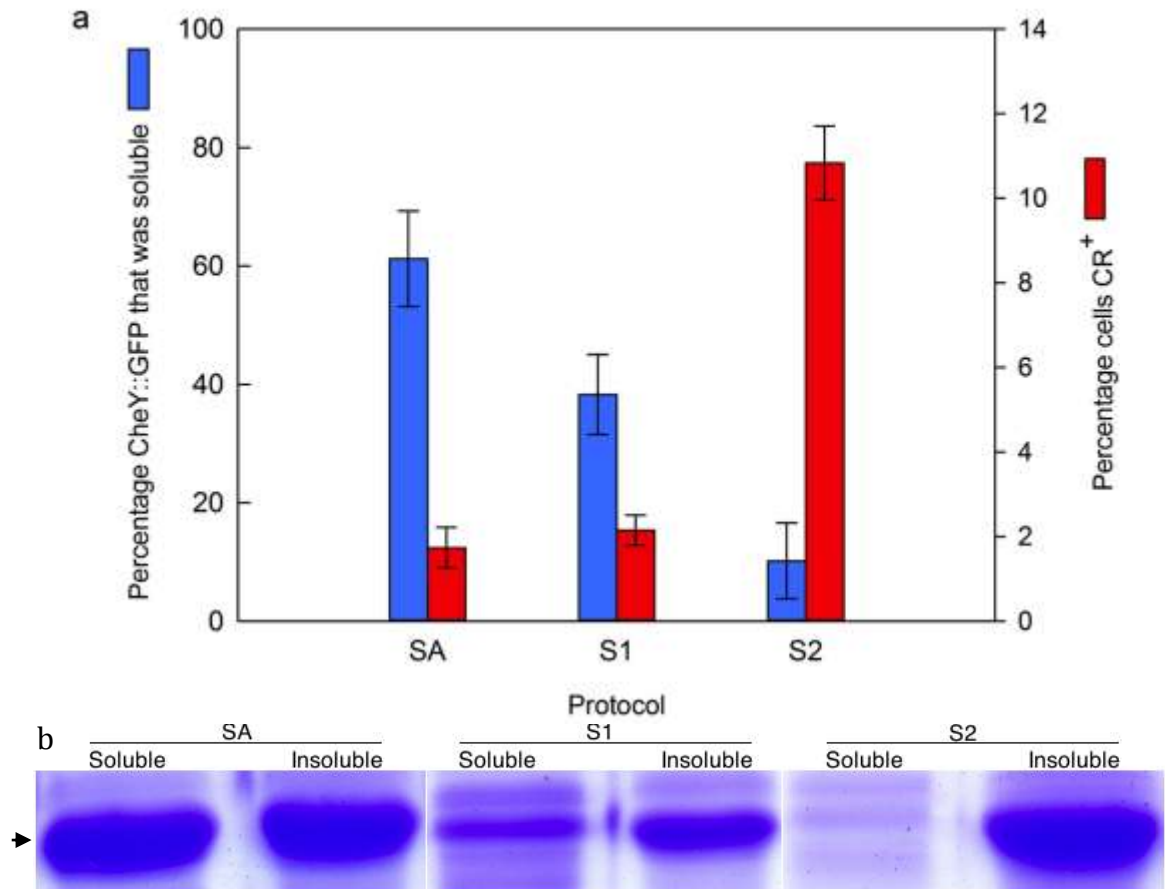
Figure 3.5: FCM and SDS-PAGE data of shake-flask _E. coli_ BL21* pETCheY::GFP RPP cultures at 24 h post-induction

a) Percentage of CR+ cells at 24 h post-induction as determined by FCM (blue), percentage of CheY::GFP that was soluble at 24 h post-induction as determined by BugBuster® fractionation followed by SDS-PAGE (red). Data are mean values of 4 cultures from 2 replicate experiments, error bars are ± 1x standard deviation. b) Representative SDS-PAGE samples separated into soluble and insoluble fractions using BugBuster® from data in (a), arrow indicates approximately 42 kDa.

et al., 2012). There is however evidence to the contrary, that amyloid[+] IBs produced in *E. coli* without toxic effects are toxic to mammalian cells (González-Montalbán *et al.*, 2007).

3.2.3. Conclusion

It has been demonstrated that the amyloidophilic fluorescent dye Congo red positively stains *E. coli* cells producing the model RPP product CheY::GFP. It was further demonstrated that the proportion of CR[+] cells is higher for production conditions favouring the formation of IBs and hence it is concluded that CR positively stains CheY::GFP IBs produced in *E. coli*. While it was expected that the proportion of CR[+] cells would correlate with CheY::GFP insolubility a direct relationship was not observed. It was concluded that amyloid content of CheY::GFP IBs was affected primarily by growth temperature as the percentage of CR[+] cells increased most when CheY::GFP was produced at 37°C.

It is therefore concluded that CR can be used to stain for IBs in *E. coli* under specific conditions and likely in a case-dependent manner as amyloid formation is heavily influenced by amino acid sequence. It is proposed that a possible application of CR staining may be for monitoring the intentional production of IBs either for refolding or the production of bioactive nanoparticles. 'Classical' highly-amyloidogenic IBs are known to be more resistant to denaturation and solubilisation (de Groot & Ventura, 2006; Upadhyay *et al.*, 2012) and hence limitation of amyloid content (as indicated by CR[+] cells) during production of IBs for refolding would be favourable. The use of IBs as biologically active nanoparticles is an area of active research (e.g. Sans *et al.*, 2012; Villaverde *et al.*, 2012) and would again likely require limiting 'classical' high-amyloid character in the IBs produced.

3.3. Flow cytometric screening of *E. coli* transformants for RPP

3.3.1. Introduction

Most RPP systems in *E. coli* provide the gene encoding the RP product on a plasmid (Baneyx, 1999). It is therefore necessary to first transform the host strain. Transformations frequently generate large numbers of transformants of which only a limited amount are cultured further, either to directly inoculate cultures or to establish a cell bank. During process development it is likely transformants will be screened to ensure that the cells used for production are optimal and have not, for example, acquired detrimental mutations. In the interests of efficiency it is advantageous for screening to occur at the earliest opportunity. FCM is ideally suited to these needs, as only a small number of cells are required and can be taken directly from agar plate cultures of transformants, without additional growth stages. This section presents a case-study that highlights the use of FCM to identify an aberrant transformant at the earliest possible stage, from agar plate culture.

3.3.2. Results & discussion

E. coli BL21* was transformed with pETCheY::GFP and grown on NA supplemented with carbenicillin. 6 transformants grew and all were restreaked onto fresh NA-carbenicillin plates. Following growth, all transformants were screened by FCM. 5 transformants showed distributions of light scatter and green fluorescence that were highly similar not only to each other but also to those previously observed (termed T1 cells). One however, showed atypical distributions (termed T2 cells) (Figure 3.6). Fluorescence distribution for both T1 and T2 transformants showed 2-3 populations, a high fluorescence population and a lower fluorescence population that could be further subdivided into an intermediate and a low fluorescent population, presumed to be plasmid free cells and
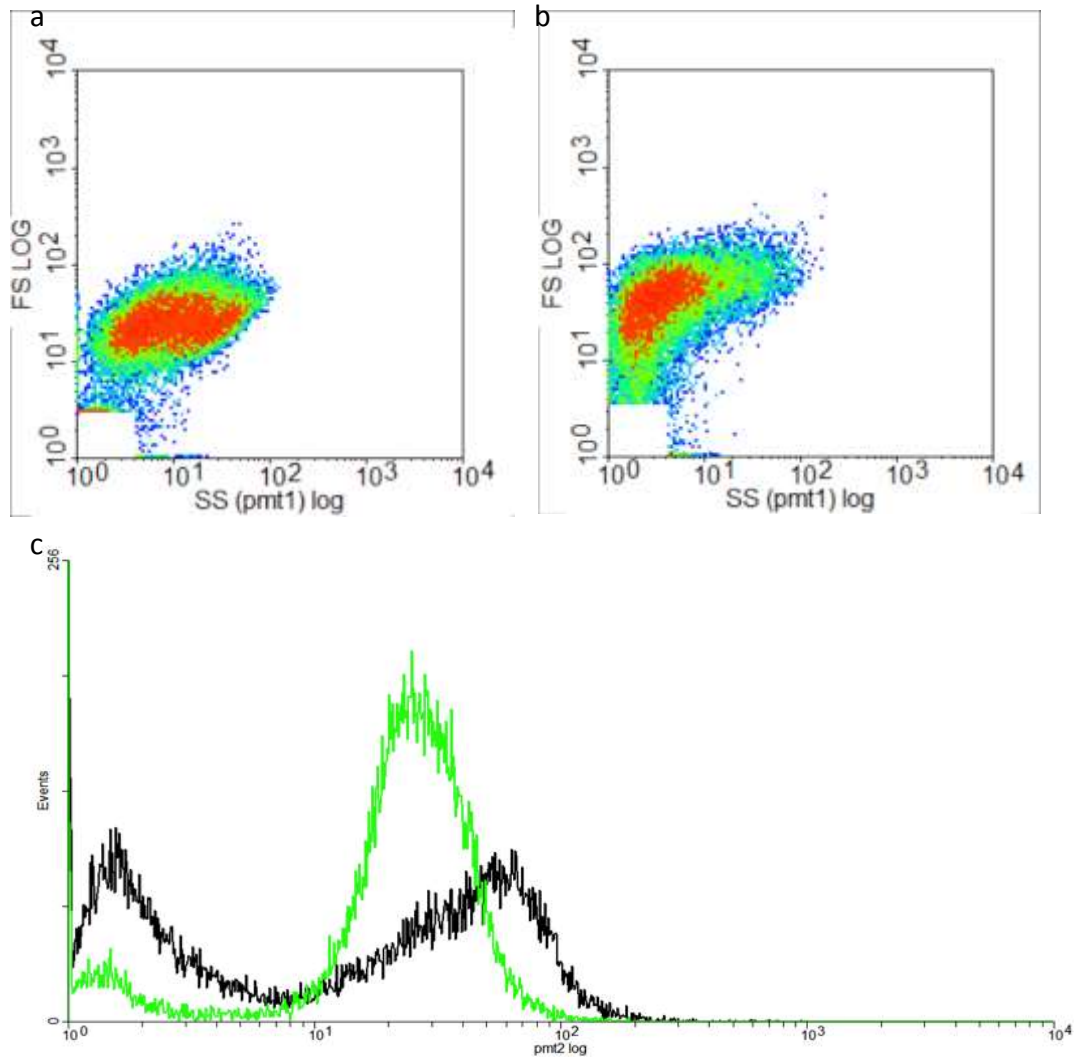
Figure 3.6: FCM data of T1 and T2 cells
a) Forward scatter (FS LOG) versus side scatter (SS pmt1 LOG) density plots of a representative T1 transformant. b) As (a) for the T2 transformant. c) Histogram of green fluorescence (pmt 2 log) for a representative T1 (green) and T2 transformant (black).

sample noise. T1 cells were predominantly in the higher fluorescence population and T2 cells showed more cells in the lower fluorescence population. Interestingly, the higher fluorescence population of T2 cells, while being a lower proportion overall had a higher fluorescence than that of the T1 cells. The most likely explanation for the differing light scatter and fluorescence distributions was that the T2 cells had mutated. Furthermore, as the higher fluorescence population of the T2 cells showed an increase in fluorescence over that of the T1 cells it was hypothesised that the mutations acquired may be beneficial for RPP. As Alfasi *et al.* (2011) were able to isolate improved RPP host strains using GFP-fusions and FCM, it was decided to investigate T2 cells further. In order to test this, overnight cultures (35 mL LB-carbenicillin, 14 h, 30°C) were grown from all transformants at intervals (1, 3, 4 & 8 weeks) while the plates were stored at 4°C. FCM analysis of cultures at 1 weeks storage showed that cultures grown from T2 cells showed lower fluorescence than for T1 cells (Figure 3.7a). Comparing OD_{650} data for overnight cultures grown from T1 and T2 cells showed that T2 cells were subject to growth defects of increasing severity; by 8 weeks of storage it was not possible to cultivate T2 cells (Figure 3.7b). From these data it was concluded that while it was likely that the T2 cells had acquired a mutation, the resultant growth defects made these cells unsuitable for RPP.

An additional phenomenon was also observed during the course of this experiment, that the distribution of fluorescence for T1 cells changed during the 8 weeks of monitoring. In the initial stages of storage the majority of cells were in a high fluorescence population but as the cells were stored a population of intermediate fluorescence developed and it appeared that cells transitioned from the high fluorescence population to the intermediate (Figure 3.8a,b). Transition to the intermediate fluorescence population did not appear to affect subsequent culture as the majority of cells following
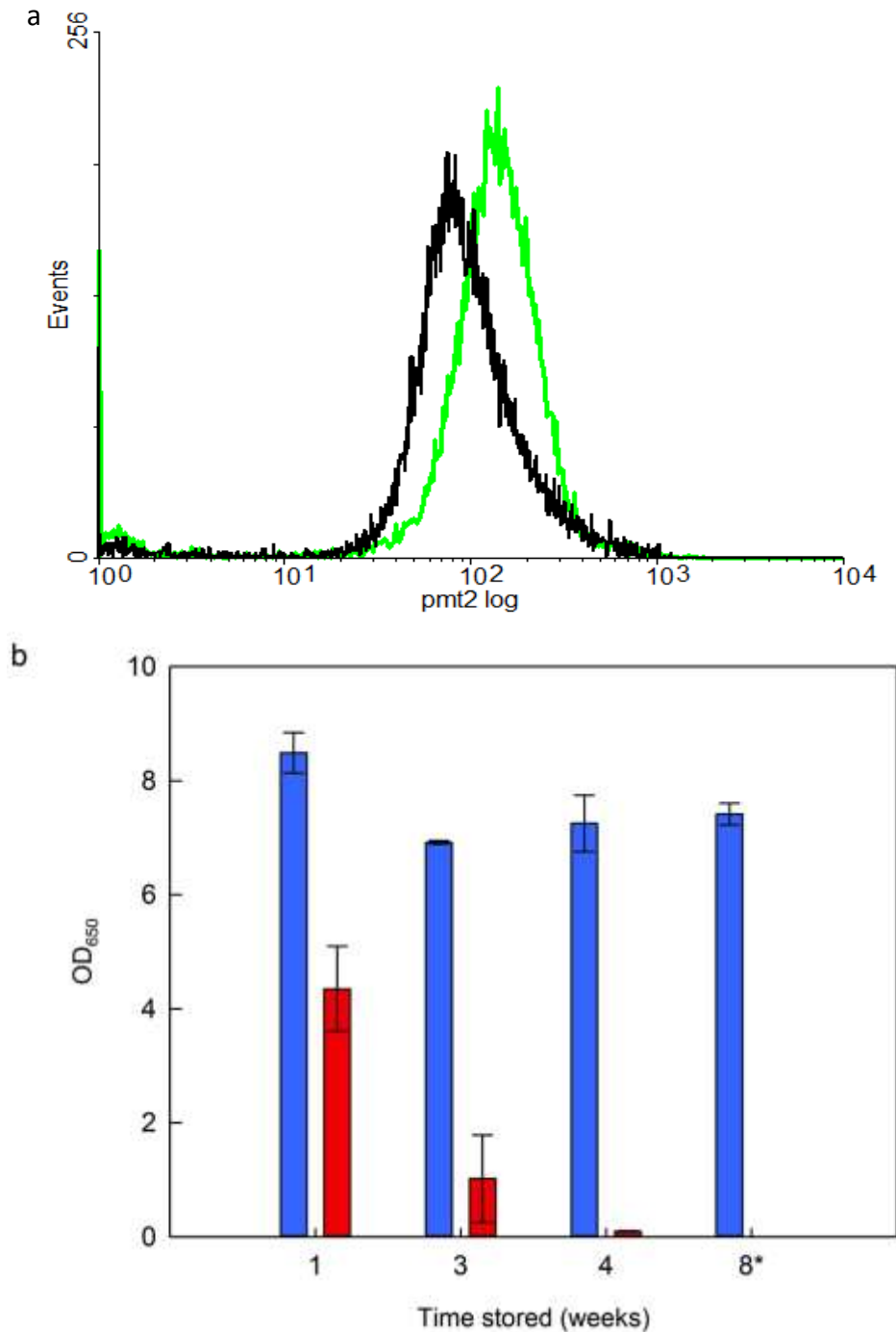
Figure 3.7: Analysis of overnight cultures from T1 and T2 cells

a) Green fluorescence (pmt 2) histogram comparing representative overnight cultures grown from T1 (green) and T2 (black) transformants after 1 weeks storage at 4°C. b) OD_{650} data of representative overnight cultures grown from T1 (blue) and T2 (red) transformants during 8 weeks storage at 4°C, data are mean values derived from 2 replicate cultures, error bars are ± 1x standard deviation. * - denotes timepoint at which it was not possible to culture the atypical (T2) transformant.
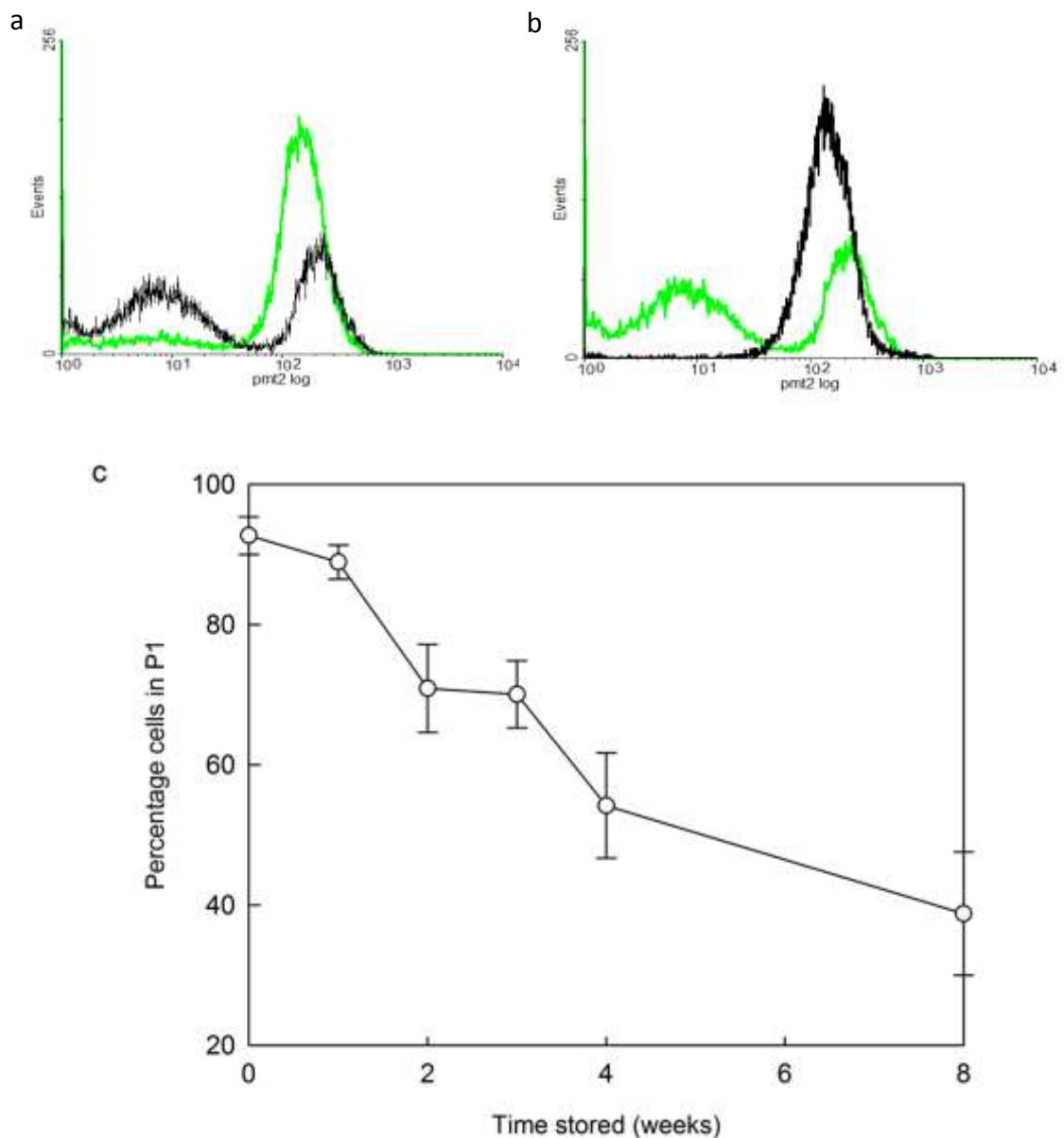
Figure 3.8: Analysis of fluorescence population shift for typical transformants during storage

a) Fluorescence intensity histograms for a representative T1 transformant at 1 week (green) and 8 weeks (black) storage. b) Fluorescence intensity histograms for a representative typical (T1) transformant at 8 weeks storage (green) and after overnight growth (black). c) Percentage of cells in the high fluorescence population for 5 T1 transformants during storage, data are mean values of 5 samples, error bars are ± 1x standard deviation.

overnight growth remained in the high fluorescence population despite being inoculated with large amounts of intermediate fluorescent cells (Figure 3.8c). The formation of intermediate fluorescence cells on agar plates was unexpected and therefore it was decided to investigate further.

3.3.3. Conclusion

FCM screening of RPP transformants was demonstrated to be effective in identifying atypical (T2) transformants simply, quickly and at an early experimental stage. Further study showed T2 cells to be inappropriate for further use. In the context of a bioprocess this technique would allow identification and discarding of aberrant transformants almost immediately during process development, representing a potential saving of time. Although in this case the primary identifier of atypicality was an altered green fluorescence distribution, T2 cells also showed altered light scatter suggesting the potential of transfer to non-fluorescent products.

3.4. Flow cytometric monitoring of aging in agar plate cultures of RP producing *E. coli*

3.4.1. Introduction

In Section 3.3.2 it was observed that during storage at 4°C, cells on agar plates formed 2 discrete populations of varying green fluorescence intensity and that as cells aged they gradually transitioned from the highest fluorescence population to the lower. This phenomenon was unexpected and therefore it was decided to investigate further in order to determine the specific cause of this phenomenon and any effects it might have on subsequent RPP processes.

3.4.2. Results & discussion

3.4.2.1. Monitoring the aging of agar-plate cultures

E. coli BL21* was transformed with pETCheY::GFP and grown on NA-carbenicillin (25°C, 48 h) then restreaked on to fresh NA-carbenicillin plates, grown under the same conditions then stored at 4°C for the remainder of the experiment. At regular intervals during storage, cells were removed from the plates, resuspended in PBS and analysed by FCM and for CFU/plasmid retention. Similar changes in fluorescence distribution to Section 3.3 were observed. At 0 weeks storage the majority of cells were in a high-fluorescence population (Figure 3.9a) and later in storage (12 weeks) a substantial proportion of cells had transitioned to a lower fluorescence population (Figure 3.9b). For ease of quantification and analysis FCM data were displayed as forward scatter versus green fluorescence plots (Figure 3.9c,d) and gated into 3 populations corresponding to high fluorescence (P1), intermediate fluorescence (P2) and low fluorescence (P3) (corresponding to GFP$^-$ cells and likely small amounts of sample noise). The use of FSC versus FL1-A plots for analysis allowed for clearer differentiation of fluorescence populations than FL1-A histograms that showed overlapping populations due to differences in cell size (i.e. a small cell with a high GFP concentration is likely to have similar fluorescence to a larger cell with a low GFP concentration).

Over 16 weeks storage at 4°C there was a clear, steady decrease in the number of cells in the highly fluorescent population P1 and a concurrent increase of cells in the intermediate fluorescent population P2 (Figure 3.10a). The proportion of cells in the GFP$^-$ population P3 remained low throughout, reaching a peak of 5.2% at 10 weeks. As cells were stored at 4°C and thus growth would be severely limited these data support the conclusion that the relative increase in the proportion of population P2 over time was due to transition of cells from population P1 to P2 as opposed to growth of P2 cells.

135
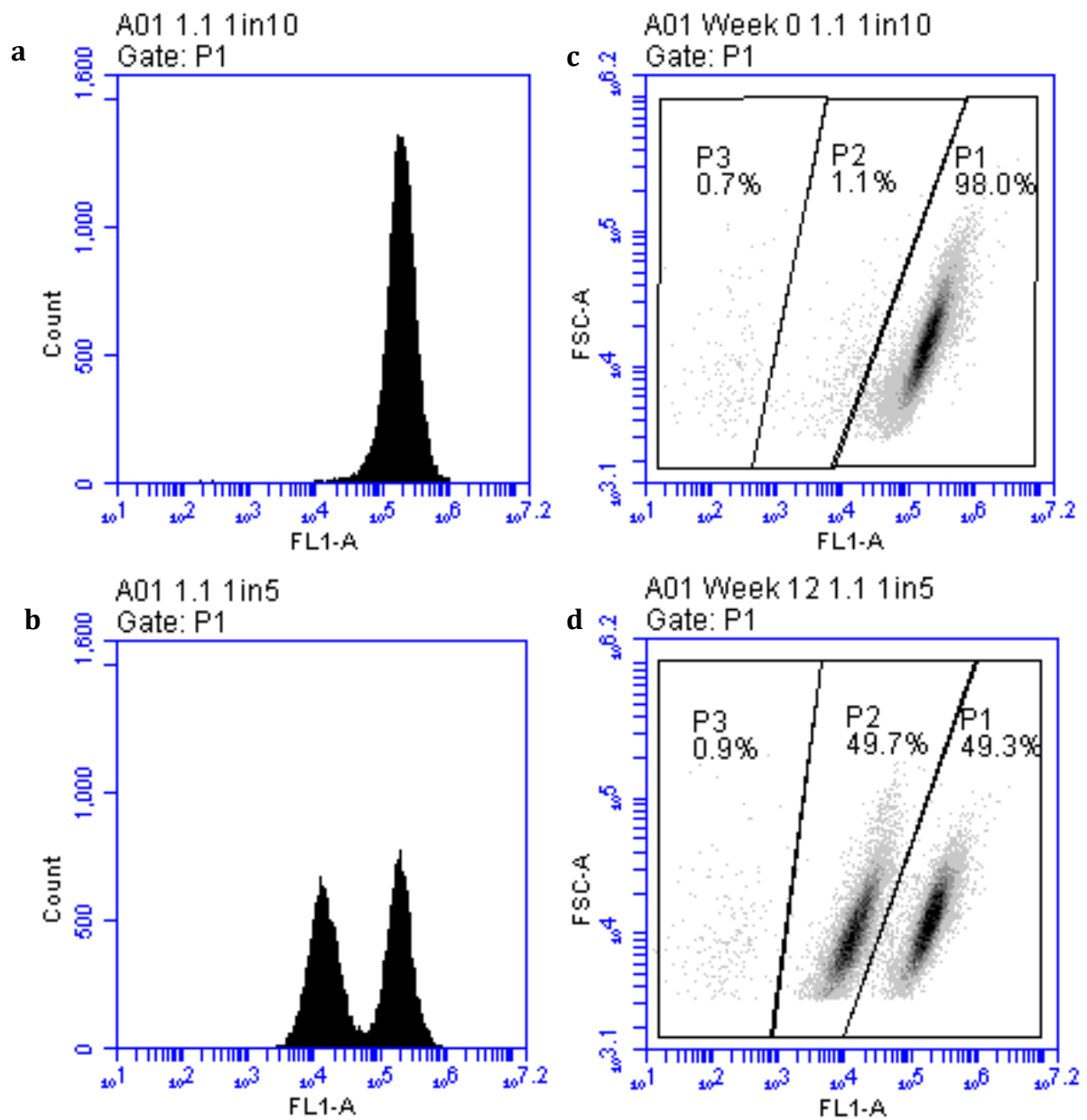
Figure 3.9: FCM gating of fluorescence population shift for transformants during storage

a) Green fluorescence intensity (FL1-A) histogram for a representative transformant at 0 weeks storage. b) As (a) for 12 weeks storage. c) FSC-A versus FL1-A plot of (a) showing typical gating for P1, P2 and P3. d) As (c) for data from (b). Data are representative samples from 4 replicate cultures.
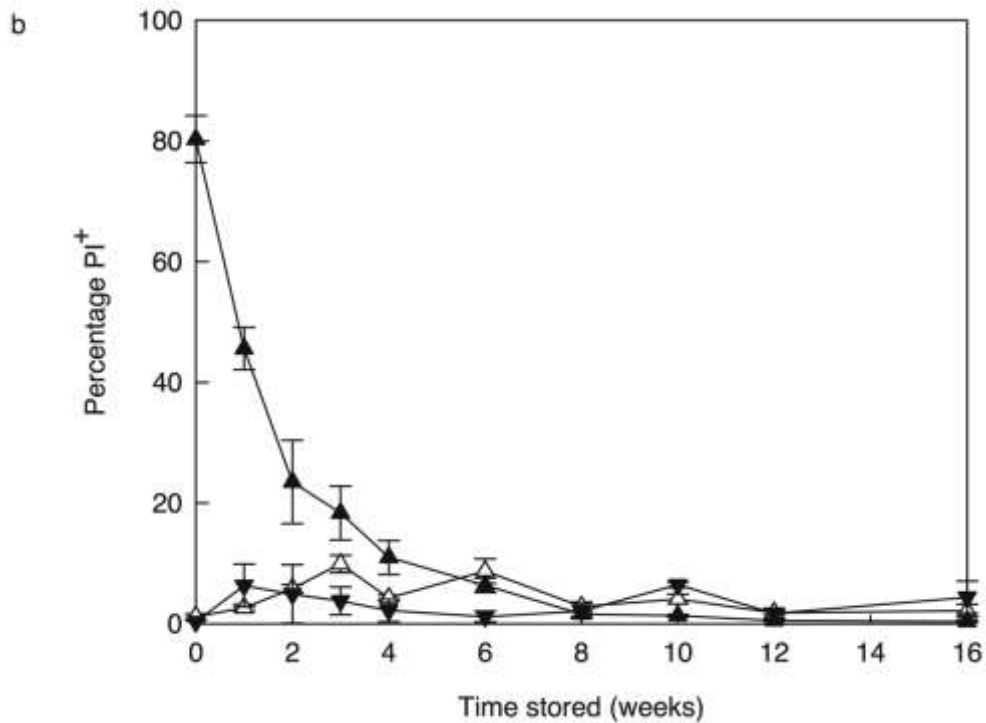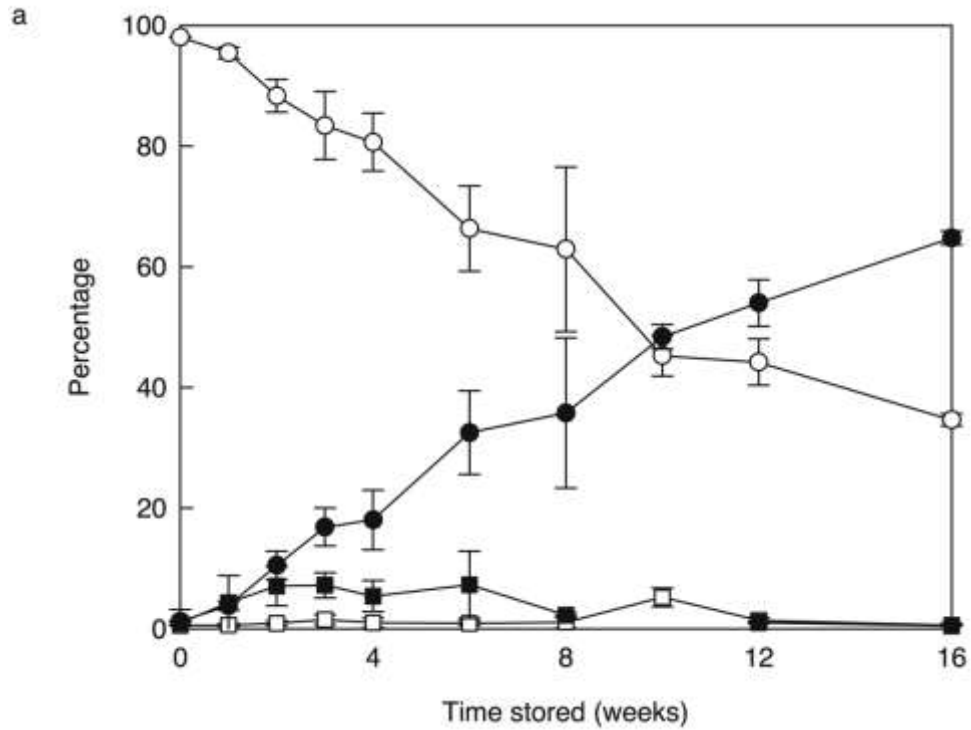
Figure 3.10: FCM analysis of fluorescence population shift for transformants during storage

a) Distribution of cells in P1 (○), P2 (●) and P3 (□); total percentage of PI+ (dead) cells (■) during 16 weeks storage. b) Percentage of PI+ (dead) cells in P1 (△), P2 (▲) and P3 (▽) during 16 weeks storage. Data are mean values from 4 replicate cultures, error bars are ± 1x standard deviation.

Further analysis of FCM data showed that the mean green fluorescence of each population remained relatively constant throughout the 16 weeks (Figure 3.11a) but mean forward scatter appeared to decrease (Figure 3.11b), suggesting that cell size decreased while GFP content remained similar. FSC-A data also showed that population P1 had a consistently higher forward scatter than P2 and that during the course of the experiment the variance of FSC-A decreased, suggesting increased homogeneity of cell size.

PI staining of the entire sample showed that the proportion of PI$^+$ (dead) cells remained below 10% throughout the experiment, reaching peak values between 2-6 weeks (Figure 3.10a). Analysis of PI staining for the 3 individual fluorescence populations however showed that while the proportion of PI$^+$ P1 and P3 cells remained low (<10%) throughout, population P2 was almost 80% PI$^+$ at 0 weeks decreasing to be equivalent to populations P1 and P3 after 6 weeks (Figure 3.10b). This suggests that cells initially in population P2 were predominantly dead but as increasing numbers of cells entered P2 the PI$^+$ proportion diminished, presumably cells entering P2 during storage were more viable than those produced during the growth stage.

3.4.2.2. The effect of subsequent culture and RPP

Throughout the storage period overnight cultures (35 mL LB-carbenicillin, 30°C, 14 h) were inoculated from the agar plate cultures and grown. It was observed that OD$_{650}$ values at 14 h post-inoculation showed an overall decreasing trend with increased storage time, decreasing dramatically between 8 and 12 weeks storage (Figure 3.12a). To determine the cause of this decrease, at regular intervals (0, 4, 8 and 12 weeks) overnight cultures were also grown for an extended period of time (12, 14, 16 and 18 h post-inoculation). Extended incubation resulted in increased OD$_{650}$ values for all cultures (Figure 3.12b) and cultures from 0, 4 and 8 weeks reached similar final
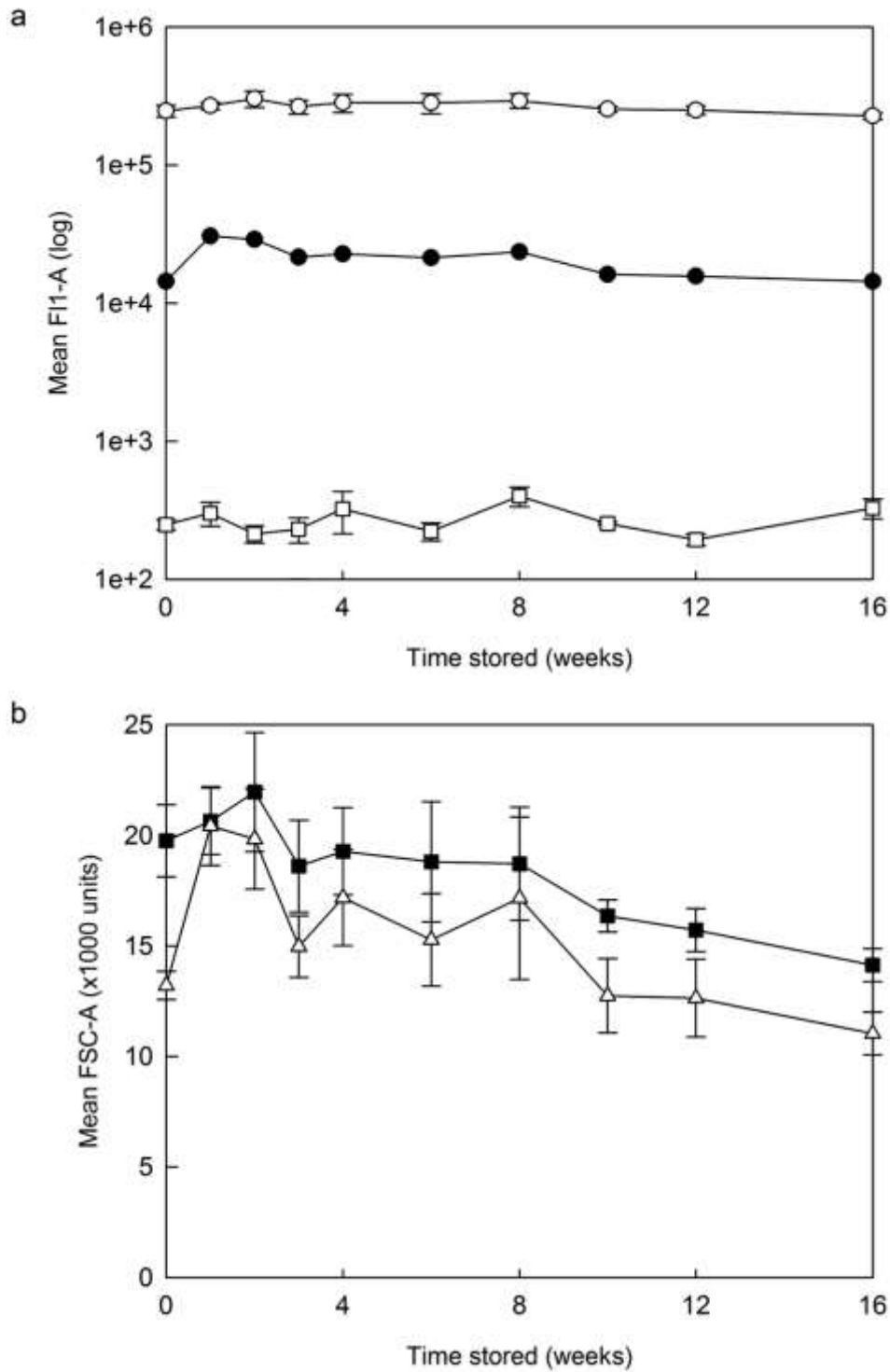
Figure 3.11: FCM analysis of fluorescence populations during storage
a) Mean green fluorescence (FL1-A) of cells in P1 (○), P2 (●) and P3 (□) during 16 weeks storage. b) Mean forward scatter (FSC-A) of cells in P1 (■) and P2 (△) during 16 weeks storage. Data are mean values from 4 replicate cultures, error bars are ± 1x standard deviation.
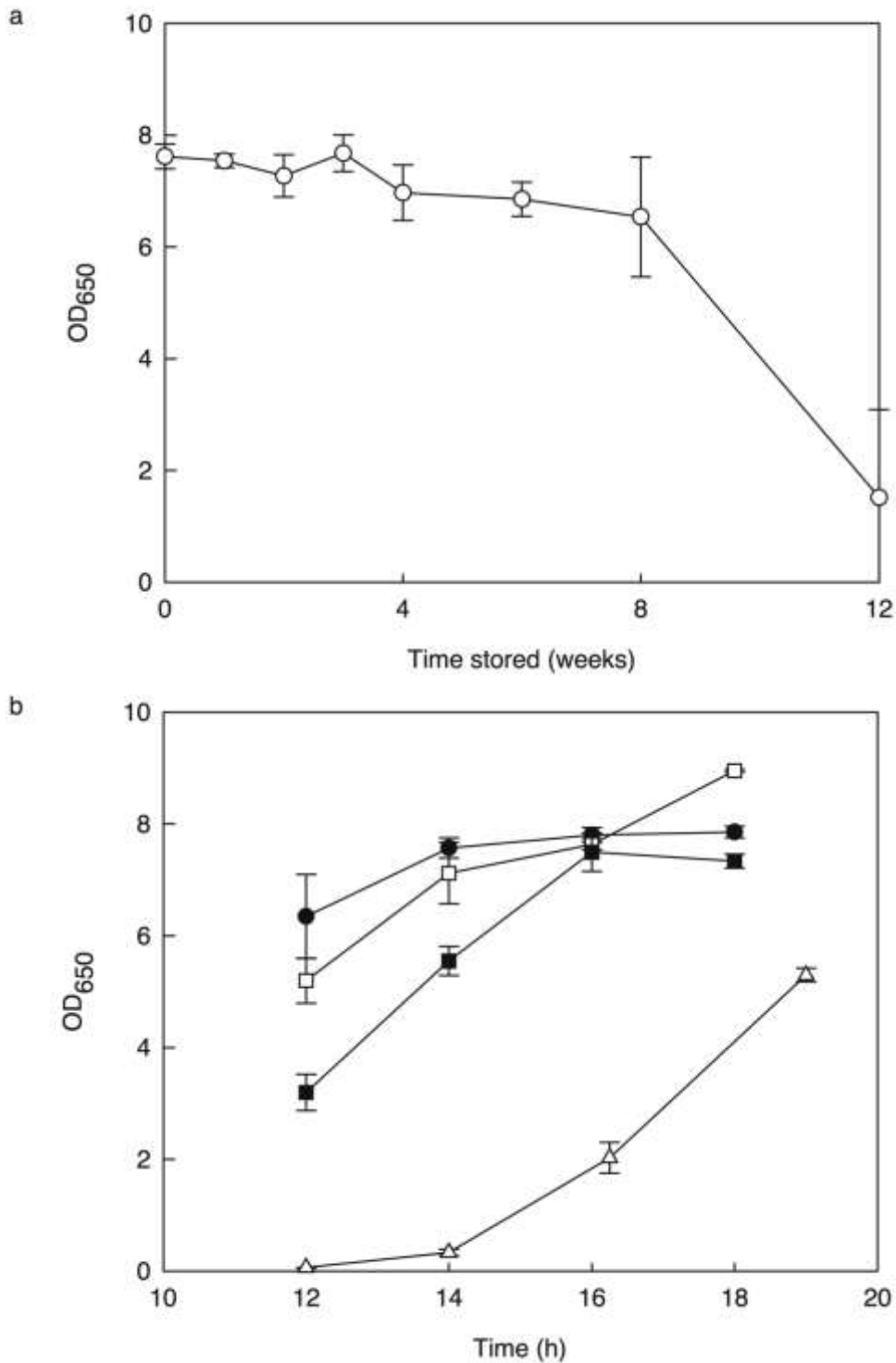
Figure 3.12: Growth of liquid cultures from agar plate cultures during storage
a) OD_{650} data from 14 h overnight cultures during storage ($\bigcirc$) (mean values of 8 replicate cultures). b) OD_{650} data from extended growth of overnight cultures at 0 ($\bullet$), 4 ($\square$), 8 ($\blacksquare$) and 12 ($\triangle$) weeks storage (mean values of 4 replicate cultures). Error bars are ± 1x standard deviation

densities. These data suggest that decreased overnight growth was as a result of either an increased lag period or fewer initial viable cells.

To determine the effects of plate-aging on subsequent RPP 2 plates were taken after 0, 4 and 12 weeks storage overnight cultures and used to inoculate 4 shake-flask RPP cultures (100 mL LB-glucose, protocol SA), grown to 24 h post-induction and analysed by FCM and SDS-PAGE and for OD_{650}. The percentage of GFP+ cells as determined by FCM for all cultures remained above 97% throughout (data not shown) but there was a decreasing trend in mean FL1-A of GFP+ cells at 24 h post-induction with increasing storage (Figure 3.13a). This was accompanied by a decrease in mean FSC-A and therefore it is likely that the decrease in FL1-A was due to a decrease in cell size rather than decreasing CheY::GFP content. Possibly related to the change in cell size, there was also an increase in OD_{650}. SDS-PAGE analysis showed no overall trend in total CheY::GFP accumulation and of CheY::GFP solubility (Figure 3.13b). These data suggest that increased initial culture age and P2 content have a mixed effect on improved RPP, the principal negative effect being the increase in inoculum lag phase.

3.4.2.3. The effect of sealing/gas transfer

Agar plate cultures used above were sealed during storage by gas-barrier film (Nescofilm) in order to limit the likelihood of contamination and to decrease the rate of moisture loss. The experiment was repeated with unsealed plates and it was observed that the rate of fluorescence population change increased substantially for unsealed plates., After 4 weeks storage the percentage of cells in population P1 had decreased to 80% for sealed and 35% for unsealed plates (Figure 3.14). Again, the decrease in population P1 was accompanied by an equivalent increase in P2. The reason for this phenomenon is uncertain as both the gas composition and drying rates will be affected by sealing, but this may provide a route for further study.
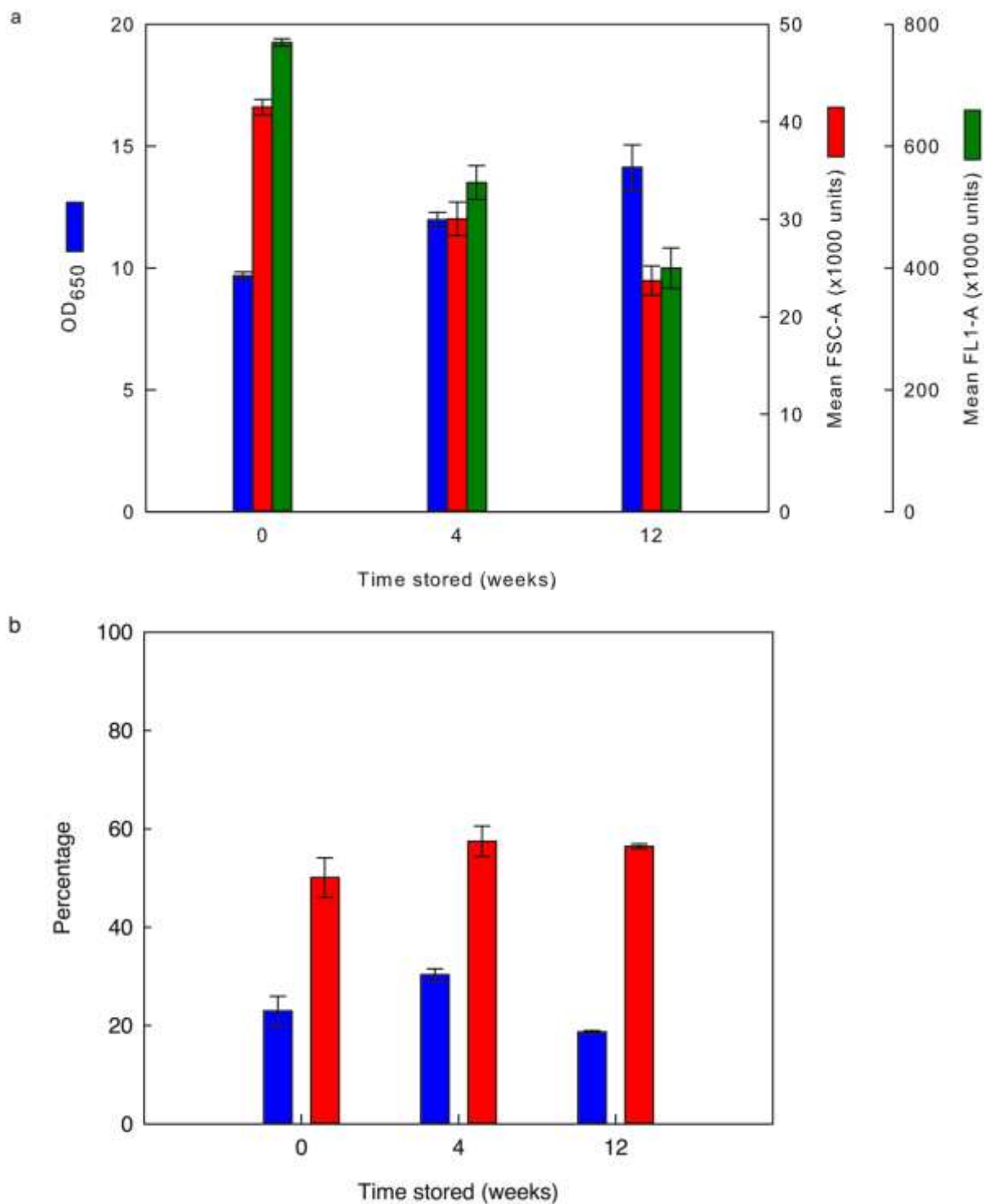
Figure 3.13: Analysis of liquid cultures from agar plate cultures during storage
a) OD_{650} data (blue), mean forward scatter (FSC-A) (red) and mean green fluorescence (FL1-A) values (green) of 24 h RPP cultures during storage (mean of 4 replicate cultures). b) SDS-PAGE data from (a); percentage total protein that was CheY::GFP (blue) and percentage soluble CheY::GFP (red). Error bars are ± 1x standard deviation.
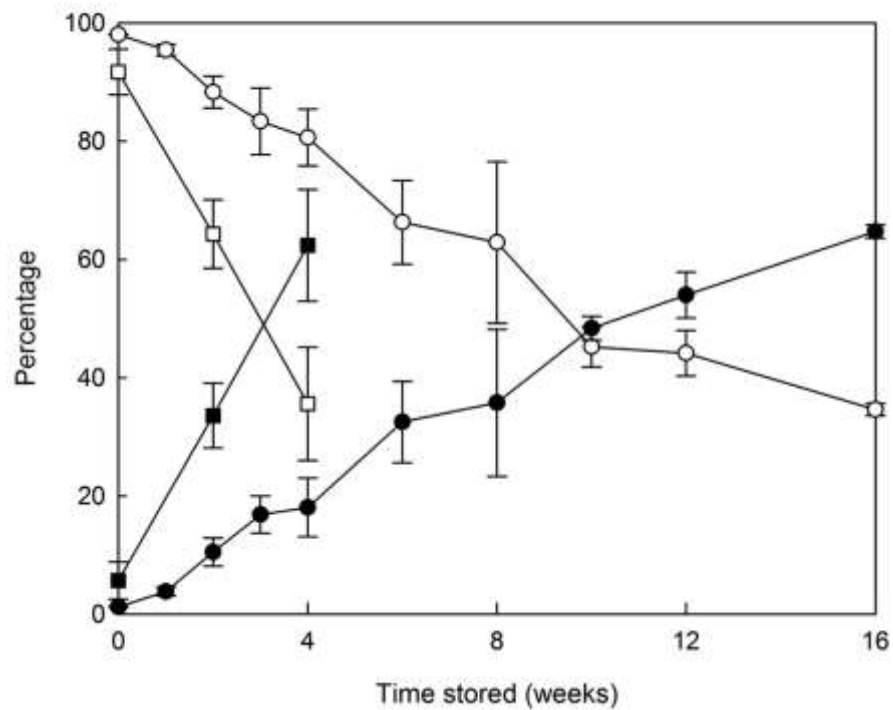
Figure 3.14: FCM analysis of fluorescence population shift for sealed and unsealed transformants during storage

Analysis of fluorescence population shift for agar plate cultures sealed with gas-barrier film; percentage of cells in populations P1 (○) and P2 (●) (mean of 8 replicates). Analysis of fluorescence population shift for unsealed agar plate cultures; percentage of cells from in populations P1 (□) and P2 (■) (mean of 4 replicates). Error bars are ± 1x standard deviation.

3.4.2.4. Determination of plasmid loss

It was initially hypothesised that the shift in fluorescence distribution from population P1 to P2 during plate aging was due to plasmid loss, and that population P2 resulted from P1 cells that had lost the plasmid but retained a moderate amount of fluorescence. This fluorescence was retained as a result of residual CheY::GFP (due to 'leaky' expression) that was then not fully diluted out of the cells due the low growth rates resulting from incubation at 4°C and selective pressure from the presence of carbenicillin, resulting in a GFP intermediate phenotype. This hypothesis was supported by a number of observations: Overnight cultures showed increased lag phases with increasing storage time, if cultures containing carbenicillin were inoculated with increasing proportions of plasmid⁻ cells this would decrease the amount of effectively viable cells and hence increase lag phase. PI staining of plate samples showed population P2 to be initially composed of predominantly dead cells; the proportion of PI⁺ P2 cells then decreases with increasing storage time, presumably as carbenicillin concentrations decreased due to degradation by β-lactamase. Also, mean FSC-A values for P1 were consistently higher than that of P2, consistent with light scatter profiles of untransformed cells. Loss of fluorescence and FCM has been used to monitor loss of FP-encoding plasmids in both *E. coli* (Patkar *et al.*, 2002; Sevastyanovich *et al.*, 2009) and yeast (Ishii *et al.*, 2010). Alternately Bahl *et al.* (2004) used derepression of chromosomally encoded GFP to indicate loss of a LacI-encoding plasmid. If plasmid loss could be established as the cause of P2 formation this would provide a useful diagnostic. CFU-based measures of plasmid retention for agar plate samples during storage did not show a strong correlation to the percentage of P1 cells as would be expected if population P1 were comprised of plasmid⁺ cells (Figure 3.15), although this may have been due to differences in the culturability of the 2 populations. CFU count plates are
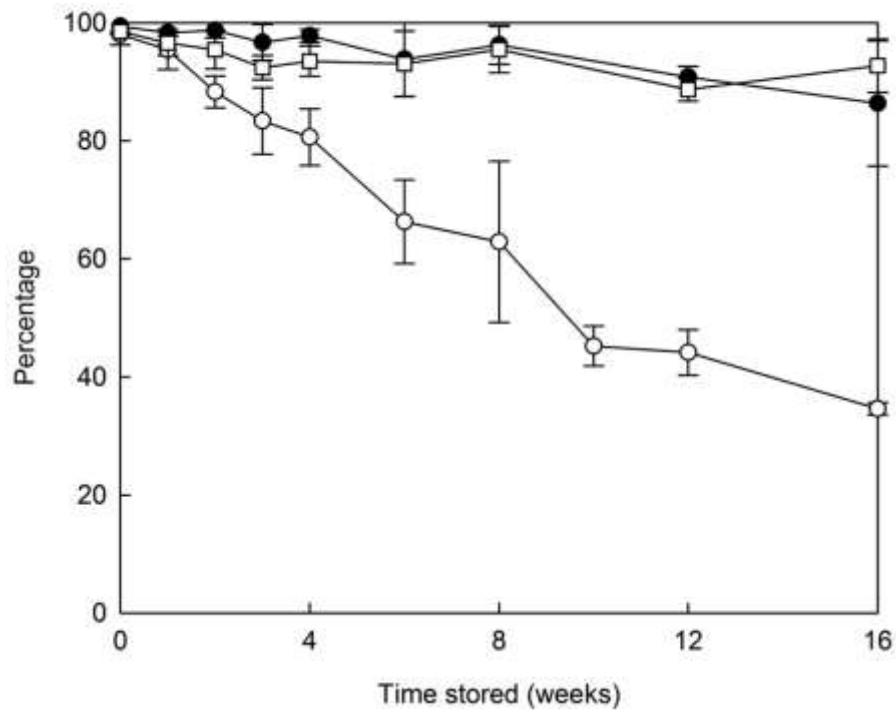
Figure 3.15: Plasmid retention of agar plate cultures during storage
FCM data: percentage of cells in population P1 (○). CFU-based data: percentage of GFP[+]
colonies (●) and percentage of carbenicillin resistant (and hence plasmid[+]) colonies
(□) (mean of 8 samples). Error bars are ± 1x standard deviation.

initially grown without antibiotic selection and hence should allow growth of cells regardless of plasmid presence. However, if the proportion of VBNC cells among the plasmid⁻ population is higher than among the plasmid⁺ population, for example due to the presence of carbenicillin on the storage plate and associated stresses, this would account for the discrepancy.

CFU counts and replica plating of cell suspensions directly from the plate were unable to identify sufficient proportions of plasmid negative cells to correspond to population P2. This was hypothesised to be due to population P2 containing predominantly VBNC cells and P1 predominantly culturable. It was therefore decided to test plasmid retention of P1 and P2 cells separately. Cells were taken from plates with high P1 (short storage time) and high P2 (long storage time) content and sorted with a BD FACSAria FACS at the Functional Genomics, Proteomics and Metabolomics Facility, University of Birmingham by T. Overton. Sorted cells were analysed by FCM (BD-Accuri C6) for absolute cell counting, plated on to non-selective NA then replica-plated onto NA-carbenicillin for plasmid retention. GFP⁺ and plasmid⁺ colonies for sorted P1 cells were above 97% for all samples. It was not possible however to accurately determine plasmid retention rates for P2 samples as CFU counts were too low to provide accurate data at 5-12 colonies per 100 μL; this was determined to be equivalent to <1% of total cells plated. In addition, GFP⁺ phenotypes were observed among the P2-derived colonies. It was therefore concluded that the hypothesis of P2 being due to plasmid loss was not conclusively supported by the data obtained and an alternate hypothesis was formed.

3.4.2.5. Determination of inclusion body formation

It was hypothesised that the formation of population P2 is a result of CheY::GFP aggregation into inclusion bodies. It is known that IBs formed from fluorescent proteins can retain considerable amounts of fluorescence (albeit less than an equivalent amount

of soluble protein) (García-Fruitós *et al.*, 2005), it is reasonable to assume CheY::GFP IBs to be less fluorescent than an equivalent amount of soluble protein and high concentrations of protein can be prone to precipitation/aggregation. In order to test this hypothesis, further unsealed plates of transformants were grown and at intervals were harvested, resuspended in PBS and analysed by FCM (including CR staining) and SDS-PAGE. As larger amounts of cells were required in this instance than for previous work (for SDS-PAGE analysis), for each sample an entire agar plate culture was harvested and resuspened in PBS, hence for this work each data point represents a single replicate culture sampled once as opposed to a culture sampled throughout.

Solubility of CheY::GFP as determined by BugBuster® fractionation followed by SDS-PAGE remained at approximately 55% throughout (Figure 3.16), which did not appear to support the hypothesis. CR staining however, showed a statistically significant correlation with population P2 formation ($r = 0.993$, $p < 0.01$) (Figure 3.17a) suggesting population P2 cells were CR+. It was also observed that after 2 weeks storage the percentage of CR+ cells was higher than that of PI+, strongly suggesting that the suspected issue with permeability of the cell membrane to CR described (Section 3.2) had been overcome. Further analysis of CR-stained cells showed that not only did P2 cells predominantly stain CR+ but also that on staining with CR, green fluorescence decreased substantially (Figure 3.17b,c). It is hypothesised that this was as a result of FRET interactions between the GFP and CR fluorophores and if so supports the conclusion that CR was bound to CheY::GFP IBs as FRET requires the two fluorophores to be in close proximity (typically 1-10 nm (Johnson & Spence, 2010)). In addition to FCM analysis of CR-stained cells it was also possible to visualise CR+ cells by fluorescence microscopy (Figure 3.18), when subject to a modified staining protocol involving 5 times the concentration of CR and 6 times the incubation period (100 µL·mL-
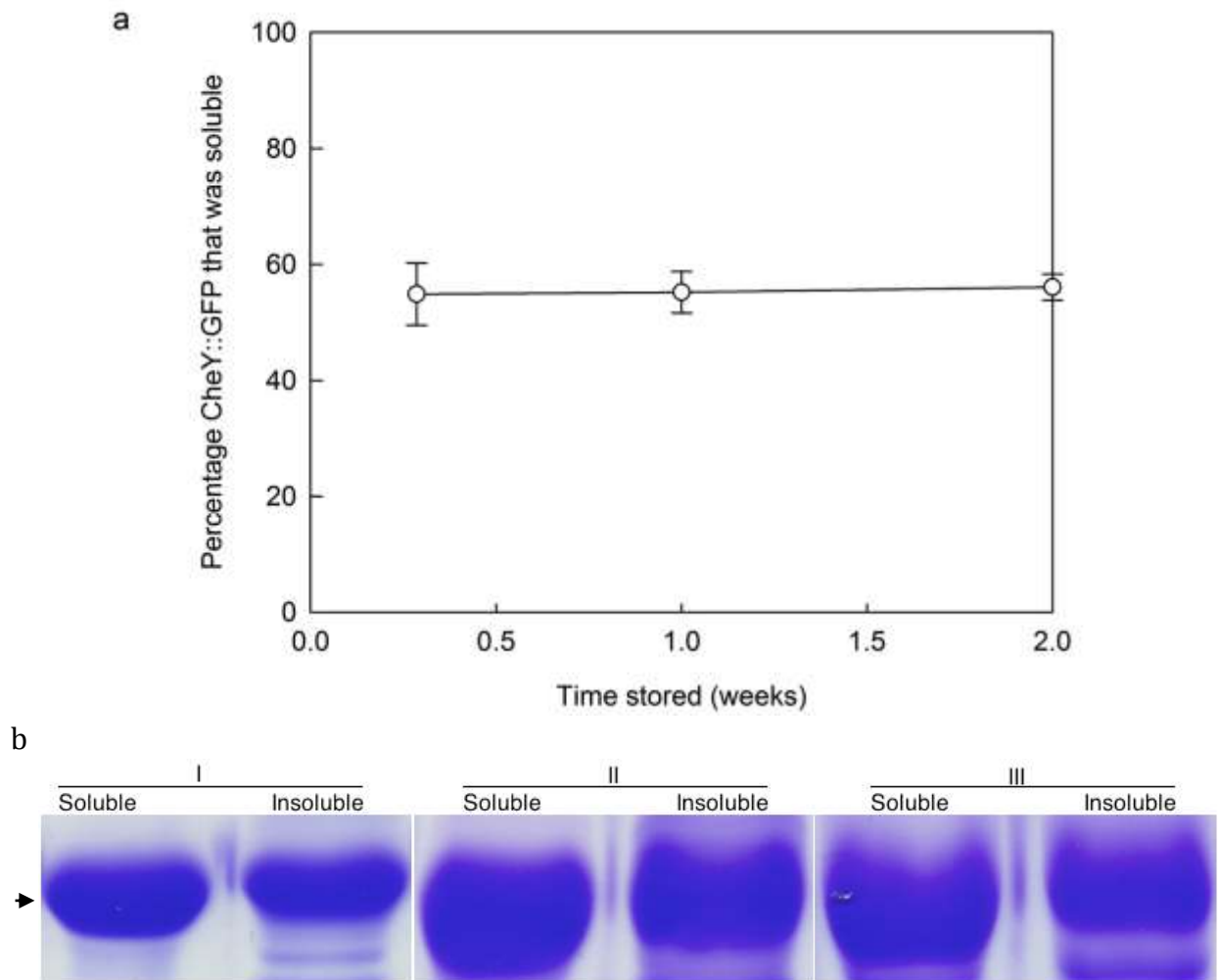
Figure 3.16: SDS-PAGE analysis of agar plate cultures during storage

a) Percentage solubility of CheY::GFP from cultures during storage, mean values from minimum of 2 replicate cultures, error bars are ± 1x standard deviation. b) Representative SDS-PAGE samples from data in (a), arrow indicates approximately 42 kDa.
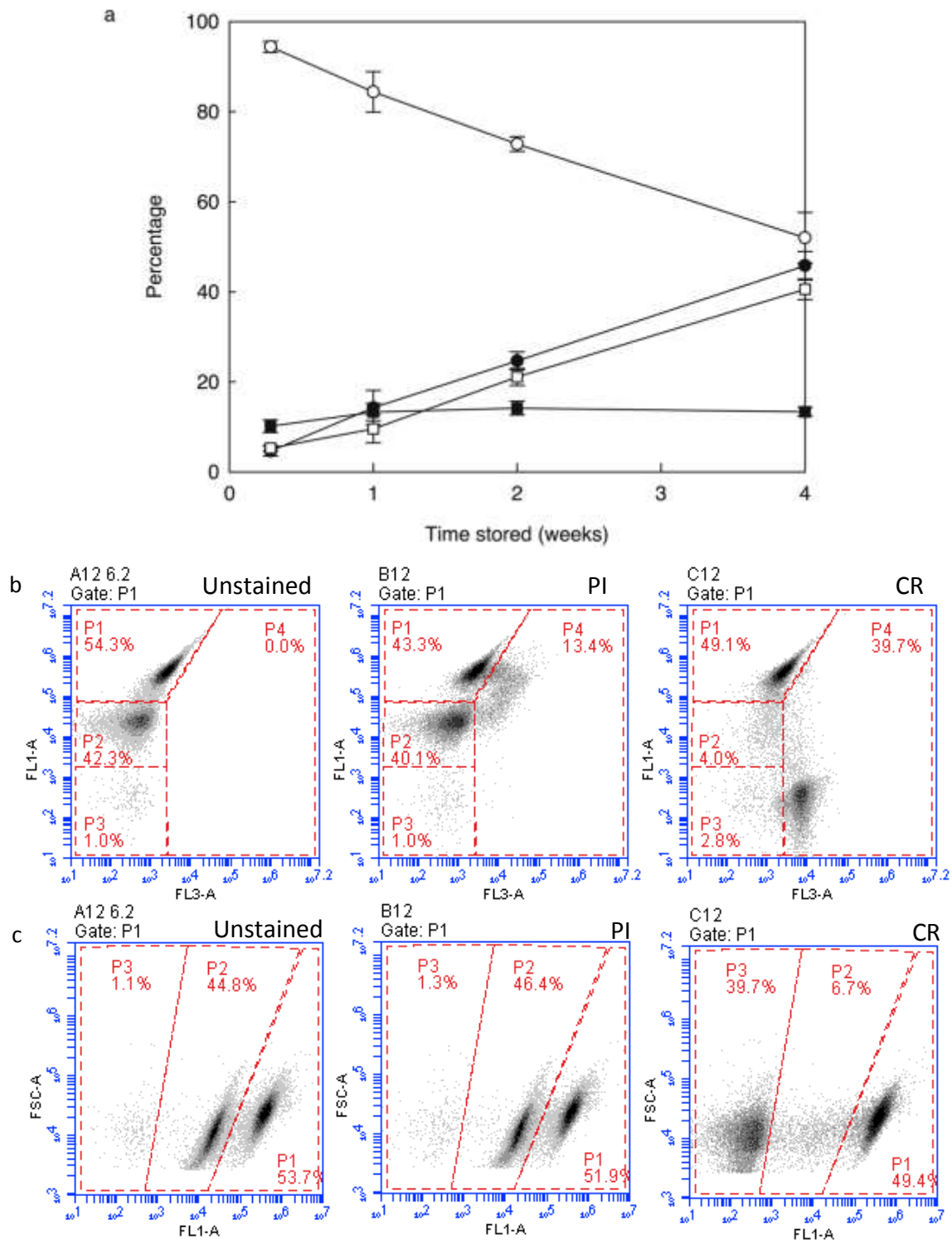
Figure 3.17: FCM detection of inclusion bodies in agar plate cultures during storage

a) Percentage of cells in P1 (○) and P2 (●), percentage of CR⁺ (□) and PI⁺ cells (■). Mean values from minimum of 2 replicate cultures, error bars are ± 1x standard deviation. b) Mean green fluorescence (FL1-A) versus red fluorescence (FL3-A) plots of cells after 4 weeks storage, unstained and stained with PI and CR. c) Mean forward scatter (FSC-A) versus FL1-A plots of data from (b), representative sample from 4 replicate cultures. Gates P1,2 & 3 corresponds to populations P1,2 & 3 respectively, gate P4 corresponds to PI⁺ or CR⁺ cells.

149
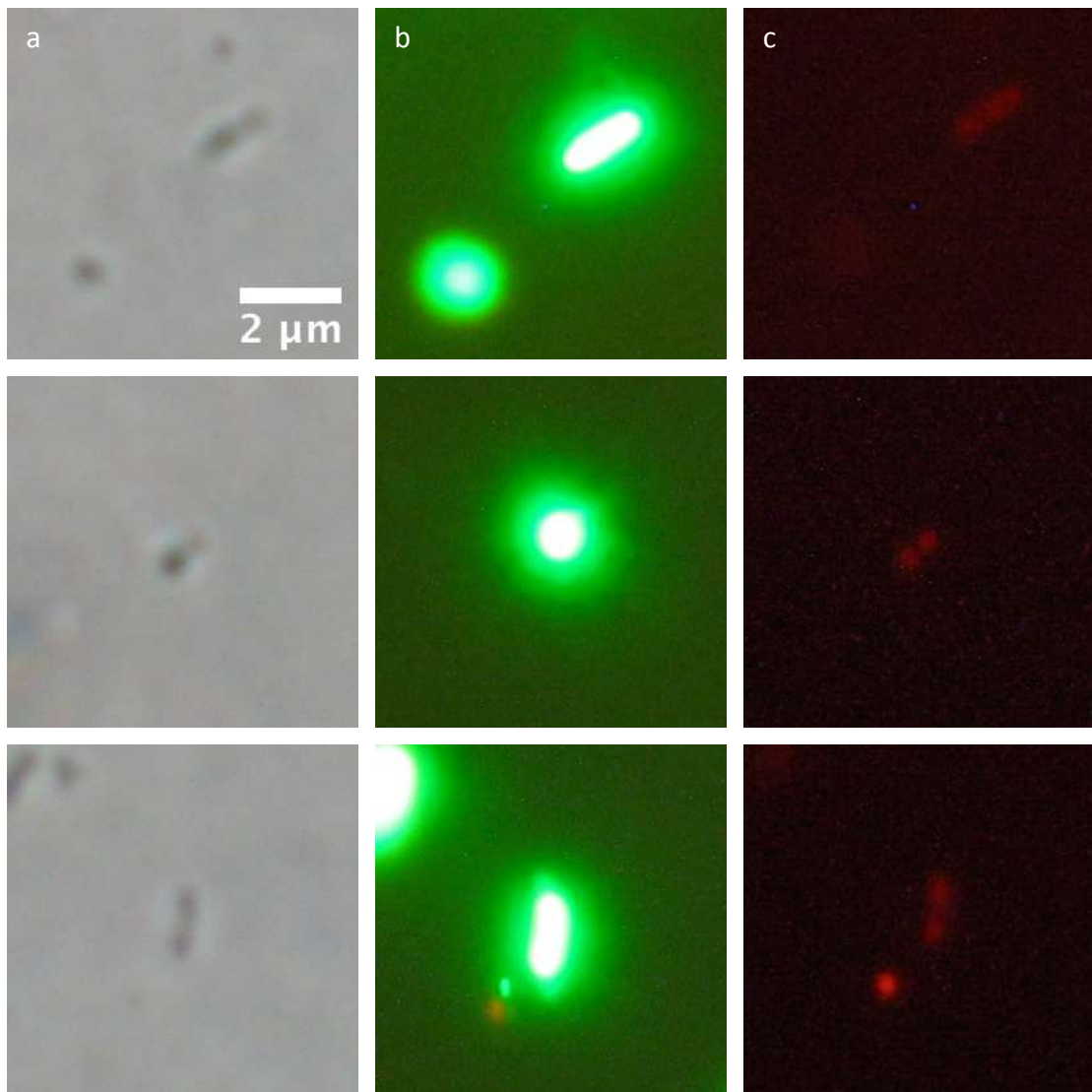
Figure 3.18: Imaging of CR-stained bacteria from agar plate cultures stored for 4 weeks

Photomicrographs of a cell suspension stained with 100 µL of 1 mg·mL^{-1} CR (in DMSO) per 1 mL cell suspension and incubated for 3 h under ambient conditions. Images of 3 individual CR$^+$ cells from 3 separate exposures. a) Phase contrast. b) Green fluorescence (GFP). c) Red fluorescence (CR).

[1] of 1 mg·mL[-1] CR (in DMSO), 3 h). Micrographs clearly showed refractive particles (presumed to be IBs) at the poles of cells under phase contrast, green fluorescence throughout the cell and red fluorescence localised to the poles of the cell, therefore it is concluded that this corresponded to CR-stained IBs.

While SDS-PAGE analysis proved to be inconclusive, CR staining of cells appeared to show that P2 cells possessed increased amyloid character and that was localised to polar refractive particles consistent with IBs. From these data the following explanation is proposed: Agar plate cultures during storage contained an amount of CheY::GFP due to 'leaky' expression from the T7 polymerase expression system that existed both solubly and in IBs. P1 cells contained soluble CheY::GFP and amyloid[-], high fluorescence IBs having been grown under relatively mild conditions (25°C, uninduced). During storage CheY::GFP IBs became amyloid[+] in a stochastic manner, losing fluorescence and causing the cell to enter population P2. Fluorescent CheY::GFP IBs are likely to contain both soluble, folded (hence fluorescent) and partially unfolded protein (Jevševar *et al.*, 2005; Peternel *et al.*, 2008) in close proximity, the presence of partially unfolded protein could provide nucleation points for amyloid and it is a characteristic of amyloid that following the initial nucleation event polymerisation can occur rapidly (García-Fruitós *et al.*, 2011). From these data is it also likely that BugBuster®, the reagent used for the extraction of soluble protein, is only able to release soluble CheY::GFP from the cytoplasm and not from inside IBs.

At 0 weeks storage population P2 was predominantly PI[+], i.e. P2 cells formed during the initial growth phase at 25°C were predominantly dead. This is consistent with data from Section 3.2 that CR[+] cells appeared to be a subset of PI[+] cells and that CheY::GFP amyloid deposits may be antimicrobial (Kagan *et al.*, 2012). The proportion of PI[+] P2 cells decreased during storage, as the overall proportion of P2 cells increased, suggesting that

P2 cells formed during storage were not dead. Sorting data showed that population P2 had very low culturability, even after the proportion of PI$^+$ had decreased to similar levels to P1, strongly suggesting that despite not being dead P2 cells were still under stress. From these observations it appears the suspected cytotoxicity of CheY::GFP amyloid deposits may be related to either growth phase, or more likely, temperature as PI$^+$ P2 cells were formed primarily during plate growth at 25°C and PI$^-$ P2 cells were formed during storage at 4°C where cell growth should have been negligible.

It is obvious from the data presented that population P2 comprises of (depending on culture age) dead or VBNC cells, resulting in increased lag periods in subsequent culture that would be undesirable for reproducibility of inocula in industrial bioprocess. CR$^+$ cell counts appeared to be approximately equivalent to the proportion of population P2 and therefore likely could be used to predict the proportion of population P2. The vast majority of RPP-related bioprocesses will not involve FP fusions and therefore a P2 population would not be observed, but CR staining may be utilised as a rapid and inexpensive diagnostic method.

3.4.3. Conclusion

Here was presented a series of experiments monitoring the aging of agar plate cultures of *E. coli* BL21* producing CheY::GFP and to determine the cause of the formation of an intermediate fluorescence population of unknown origin and low levels of viability and culturability. An initial hypothesis of plasmid loss could not be proved and it was subsequently demonstrated that P2 cells contained amyloid and that staining with the amyloidophilic fluorescent dye CR could be used to estimate P2 content. From these observations it is concluded that CR staining of agar plate cultures prior to inoculation of fermenter inocula might be a useful diagnostic step. To fully establish the utility of this

method it would be useful to establish whether this phenomenon occurs with other RPP products (specifically non-fluorescent proteins) and is therefore not case dependent. It is acknowledged that industrial processes will more often use cell banks to improve reproducibility as opposed to agar plate cultures and therefore it may be of interest to determine if similar phenomena are observed in cell banks.

Chapter 4: Results 2 – Application of 'Improved' Physiological

Stress-Minimised Production Conditions to Industrially Derived

Fed-Batch RPP Protocols

4.1. Introduction

Sevastsyanovich *et al.* (2009) proposed and demonstrated a novel protocol for producing high yields of soluble, cytoplasmic RP in *E. coli* by minimising the physiological stresses inherent to typical RPP protocols using a reduced growth temperature (25°C) and IPTG concentration (8 μM) (Section 1.3.7). While this 'improved protocol' reduced the overall rate of RP synthesis, it also reduced the amount of plasmid-free and dead cells, increasing the effective production period, and hence increased RP yields. In addition, the decreased production rate allowed improved RP folding and hence a greater proportion of RP was available in the soluble fraction. The improved protocol therefore represents a significant development in the theory behind RPP processes.

The principal aim of this study was to further develop and optimise this 'improved' protocol. The logical extension for research into RPP at the laboratory scale is application to industrial practices. Although industrial RPP frequently utilises production in the form of IBs (Section 1.2.1.2) the variable efficacy of refolding procedures leaves substantial opportunity for development of RPP protocols aimed at optimising the production of soluble RP. An examination of the fermentation protocol used by Sevastsyanovich *et al.* (2009) identified three potential limitations for industrial application. First, LB is used as both growth medium and feed. As LB contains tryptone, an animal-derived component, it is unsuitable for cGMP processes. Complex components are also subject to batch variation that has already been demonstrated to affect biomass formation in the *E. coli* BL21* pETCheY::GFP model RPP system (Section 2.10.1). Second, the feed contains large amounts of complex components (100 g·L^{-1} tryptone, 50 g·L^{-1} yeast extract, 200 g·L^{-1} glucose) making it costly and difficult to dissolve. In addition, this may affect culture osmolarity and interfere with downstream processing. Finally the use

of glucose as a carbon source, while often used industrially, can be problematic as it is readily converted to acids by both overflow and fermentative metabolism (hence requiring stringent feed, O_2 and pH control) and for fed-batch production is limited in its solubility, limiting high cell densities. The glucose must also be sterilised separately to nitrogen-containing components then mixed, introducing additional preparation stages and increasing the potential for contamination.

Considering the predicted limitations for industrial use, the 'improved' protocol as it is currently presented it is unlikely to widely influence process development in industry. Industrial biotechnology, particularly the pharmaceutical industry, would be unlikely to adopt dramatic changes to the protocols used, for both economic and regulatory reasons. The work presented in this chapter is therefore aimed primarily to apply the stress-minimisation paradigm to industry-derived high cell density *E. coli* RPP fermentation protocols in order that it is presented in a format more readily applicable to existing industrial practices and hence, it is hoped, will be more widely adopted. This chapter also examines in more detail aspects of stress-minimised RPP including the effects of altering the induction point, carbon source, and growth medium composition.

2 RPP fermentation protocols were chosen for the application of improved production conditions, in both cases the protocols had been used previously in this department as part of industrial collaborations in the production of periplasmically-located Fab and were available in publicly-accessible literature. The first protocol was taken from Want *et al.* (2009) and addresses the concerns outlined above as it uses a semi-defined medium containing yeast extract as its only complex component, reducing potential for batch-variation and removing concerns associated with using animal-derived tryptones. It uses glycerol as carbon source, simplifying medium preparation as this can be autoclaved with nitrogen-containing components. The feed contains only glycerol and

MgSO$_4$, again simplifying preparation and in addition, the concentration of glycerol used is much higher than possible for glucose, facilitating growth to high cell densities. In addition the growth and induction conditions of this protocol (37°C and 100 mM IPTG) are likely to generate high levels of physiological stress and hence are ideal for adaptation for stress minimisation. The second protocol was taken from Humphreys *et al.* (2002) and uses a chemically defined medium removing any variability from biologically-derived components.

4.2. Results & Discussion

In all cases fermentations were set up in a 5 L Electrolab Fermac 310/60 fermentation apparatus containing 1.5 L initial growth medium and grown in fed-batch according to the protocols detailed in Section 2.6. On-line data monitored by the fermentation apparatus (DOT, agitation, pH and temperature) and by gas-MS (CDC, OXC and RQ) and scans of gels used for SDS-PAGE analysis for all fermentations are provided in Appendix 1 (Chapter 7). For fermentations according to protocols A-B3 the data presented are derived from a single representative fermentation from a minimum of 2 replicate cultures, for protocol C fermentations the data are derived from a single culture. A summary of end-point fermentation data for all protocols presented here is provided in Table 4.1, in addition to data from Sevastsyanovich *et al.* (2009) and Alfasi (2010) for comparison purposes.

4.2.1. Production of CheY::GFP by *E. coli* BL21* using Protocol A

The principal aim of this study was to adapt the 'improved' RPP fermentation protocol by Sevastsyanovich *et al.* (2009) and Alfasi (2010) (henceforth referred to as Protocol A) to facilitate industrial applications. It was decided initially to directly replicate this work

Table 4.1: Summary of end-point fermentation data

Protocol	A	B1	B2	B3	A(1)[1]	A(2)[2]	C1	C2	C3
OD$_{650}$	64.5 (71.8)	297	288	167	66.5[3]	144 (150)	81.2 (91.2)	162	277
DCW (g·L^{-1})	30.1	72.8	77.5	42.6	26.6	43.9	32.5[6]	48.4	66.8
Y$_{xs}$ (g·g^{-1})	0.363	0.365	0.378	0.411	0.320	0.529	0.198[6]	0.296	0.408
CFU (cfu·mL^{-1})	2.0x10^{10}	7.0x10^{11}	*	1.1 x10^{10}	*	1.0x10^{11} (1.7x10^{11})	(9.3x10^{10})	1.8x10^{12} (6.1x10^{12})	2.5x10^{10} (4.8x10^{10})
% GFP$^+$ (FCM)	94.3	95.2	73.5	98.16	*	*	93.5	92.3	96.8
% PI$^+$	2.10	9.3	8.82	5.88	*	*	1.6	5.2	6.1
Mean FL1-A (GFP$^+$ cells)	166,736 (176,128)	240,094	369,354	295,534	*	*	62,434	460,722	221,131
Mean FL1-A (All cells)	157,227	228,612	271,631	290,099	*	*	58,399	305,338	214,104
% CV FL1-A	73.3	54.0	82.88	73.43	*	*	94.0	75.0	57.4
Mean FSC-A	8,756	23,920	25,231	22,689	*	*	7,169	24,229	24,829
% CheY::GFP soluble	62.5	55.6	37.4	49.5	*	*	*	60.0	42.2
% CheY::GFP of total protein	13.0	24.5	25.7	30.1	31.0	35.0	*	28.7	27.2
Total CheY::GFP (g·L^{-1})	3.1 / 2.3[4]	12.7 / 10.7[4]	15.7 / 12.0[4]	11.1 / 7.7[4]	5.8[1] / 4.9[4]	9.2[4]	*	10.5 / 7.6[4]	14.2 / 10.9[4]
Total soluble CheY::GFP (g·L^{-1})	2.0 / 1.5[4]	7.1 / 6.0[4]	5.9 / 4.5[4]	5.5 / 3.8[4]	*	*	*	6.3 / 4.5[4]	6.0 / 4.6[4]
DCW/OD$_{650}$ ratio	0.47	0.24	0.27	0.26	*	*	*	0.30	0.24
Length of fermentation (hours)	73.5	41.0	46.75	46	*	94	47[7]	72	55
Cost of growth medium (£·L^{-1})[5]	7.43	4.85	4.85	2.49	7.43	7.43	*	*	*

N.B. Values in parentheses are peak measurements if not final

* - Data not available

1 – From Sevastsyanovich *et al.* (2009)

2 – From Alfasi (2010)

3 – Calculated from stated DCW (assuming OD$_{650}$ 1 = 0.4 g·L^{-1} DCW)

4 – Estimated from DCW and % CheY::GFP of total protein assuming protein comprises 60% of *E. coli* dry cell mass (based on 50-61% estimates from Valgepea *et al.* (2013) and 70% from Sevastsyanovich *et al.* (2009).

5 – Calculated from chemical prices listed on fisher.co.uk or sigmaaldrich.com, correct as of September 2013. Not including essential components (e.g. IPTG, carbenicillin, antifoam, pH control).

6 – Calculated from OD$_{650}$ (assuming OD$_{650}$ 1 = 0.4 g·L^{-1} DCW)

7 – Terminated early

using the facilities and equipment available in order to establish a base-line from which adaptations could be compared and to determine the extent of the limitations predicted above.

A fermenter was set up according to protocol A (Tables 2.6, 2.7). CheY::GFP production was induced at OD_{650} ~0.5 (3.5 h post-inoculation). Feeding began on depletion of initial carbon source as indicated by on-line measurements (DOT, CDC and OXC (Figure 7.2) at ~11 h post-induction according to an stepped linear scheme adapted from that determined by Alfasi (2010) such that the feed rate was increased when on-line monitoring systems indicated that the feed rate had become growth-limiting. The fermentation was terminated when it was apparent that growth had ceased at 70 h post-induction.

Cell density as measured by OD_{650} steadily increased up to approximately 48 h post-induction, and then fluctuated, reaching a peak OD_{650} of 71.2 at 52 h, although by termination at 70 h it had decreased to 64 (most likely due to dilution by feed, acid and base addition) (Figure 4.1a). On-line data also showed a decrease in metabolic activity (O_2 demand and CO_2 production) following complete addition of the feed at 57-58 h confirming that there was no further growth during this period. Final biomass was determined at 30.1 $g \cdot L^{-1}$ DCW, which equates to a biomass yield (Y_{xs}) of 0.36 $g \cdot g^{-1}$ glucose (although this is potentially complicated by additional carbon supplied by the 46 and 23 $g \cdot L^{-1}$ of tryptone and yeast extract respectively). CFU counts increased up to a maximum of 3.5×10^{10} $cfu \cdot mL^{-1}$ at 46 h post-induction then decreased to a final count of 2.1×10^{10} $cfu \cdot mL^{-1}$. The percentage of dead cells, determined by FCM with PI staining, did not increase above 2.6% throughout the fermentation (Figure 4.1b) and therefore did not account for the final decrease in CFU, the decrease in culturability may therefore be possibly due to an increase in VBNC cells.
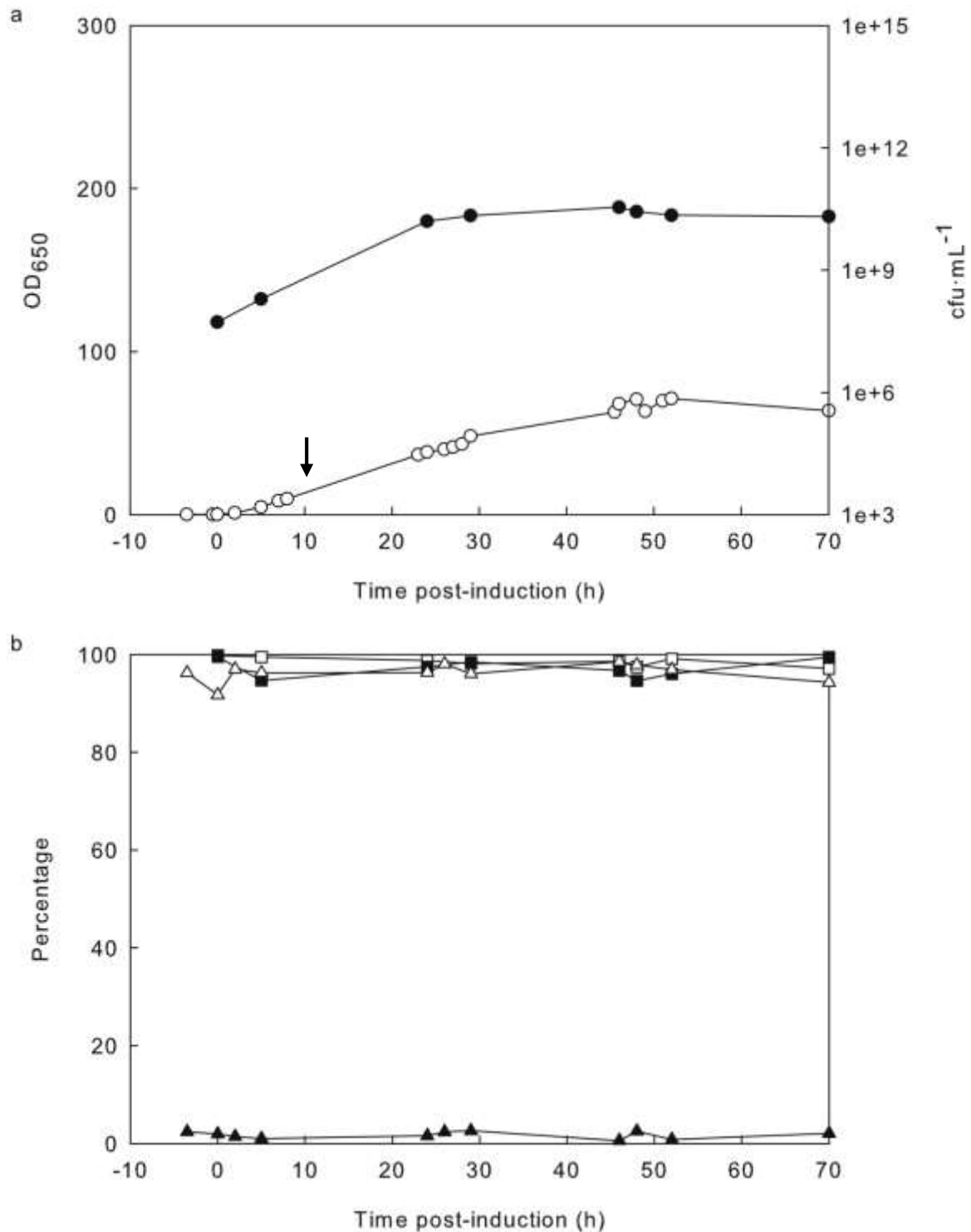
Figure 4.1: Culture growth, plasmid retention and viability data from a protocol A fermentation

a) Growth data: OD_{650} measurements (○), CFU counts (●). b) Plasmid retention and FCM data: Percentage colonies plasmid$^+$ (□), percentage colonies GFP$^+$ (■), percentage cells GFP$^+$ (by FCM) (△), percentage cells PI$^+$ (dead) (by FCM) (▲). Arrow indicates point at which feed was turned on.

Plasmid retention, as determined by both replica plating and GFP$^+$ phenotype of colonies on CFU plates, remained above 94 and 97% respectively throughout the fermentation and GFP$^+$ cells, as determined by FCM, remained above 90% (Figure 4.1b), strongly suggesting that the majority of cells remained productive throughout the fermentation.

Mean cellular green fluorescence increased between inoculation and induction followed by a rapid decrease until 2 h post-induction; this coincided with an increase followed by a decrease in FSC and is therefore interpreted as stationary phase cells from the inoculum increasing in size during initial stages of growth, then decreasing in size as the cells divide with a concomitant decrease in fluorescence as residual CheY::GFP was divided between daughter cells (Figure 4.2a). After 2 h post-induction fluorescence increased rapidly up to 162,000 at 29 h, then increased only slightly up to a maximum of approximately 167,000 at 52 h and remained stable until termination. As there was also little change in FSC during the same period that would indicate changing cell size it appeared that little functional CheY::GFP was accumulated following 29 h. CheY::GFP solubility increased over 20% from induction to 62% at 26 h post-induction, then appeared to fluctuate until the end of the fermentation reaching a final solubility of 63% (Figure 4.2b). The percentage of total protein that was CheY::GFP peaked at 24 h post induction at 17% which decreased to 13% at the end of the fermentation, barely any increase over that obtained pre-induction. Final yield of total CheY::GFP produced was calculated from SDS-PAGE and BCA assay data to be 3.1 g·L^{-1} (g protein per L culture volume), corresponding to a yield of 2.0 g·L^{-1} soluble CheY::GFP.

From the data above it appears that fermentation A progressed in 3 stages. First, accumulation of both biomass and soluble cellular CheY::GFP between 0-28 h post-induction, production of biomass with constant cellular CheY::GFP concentration
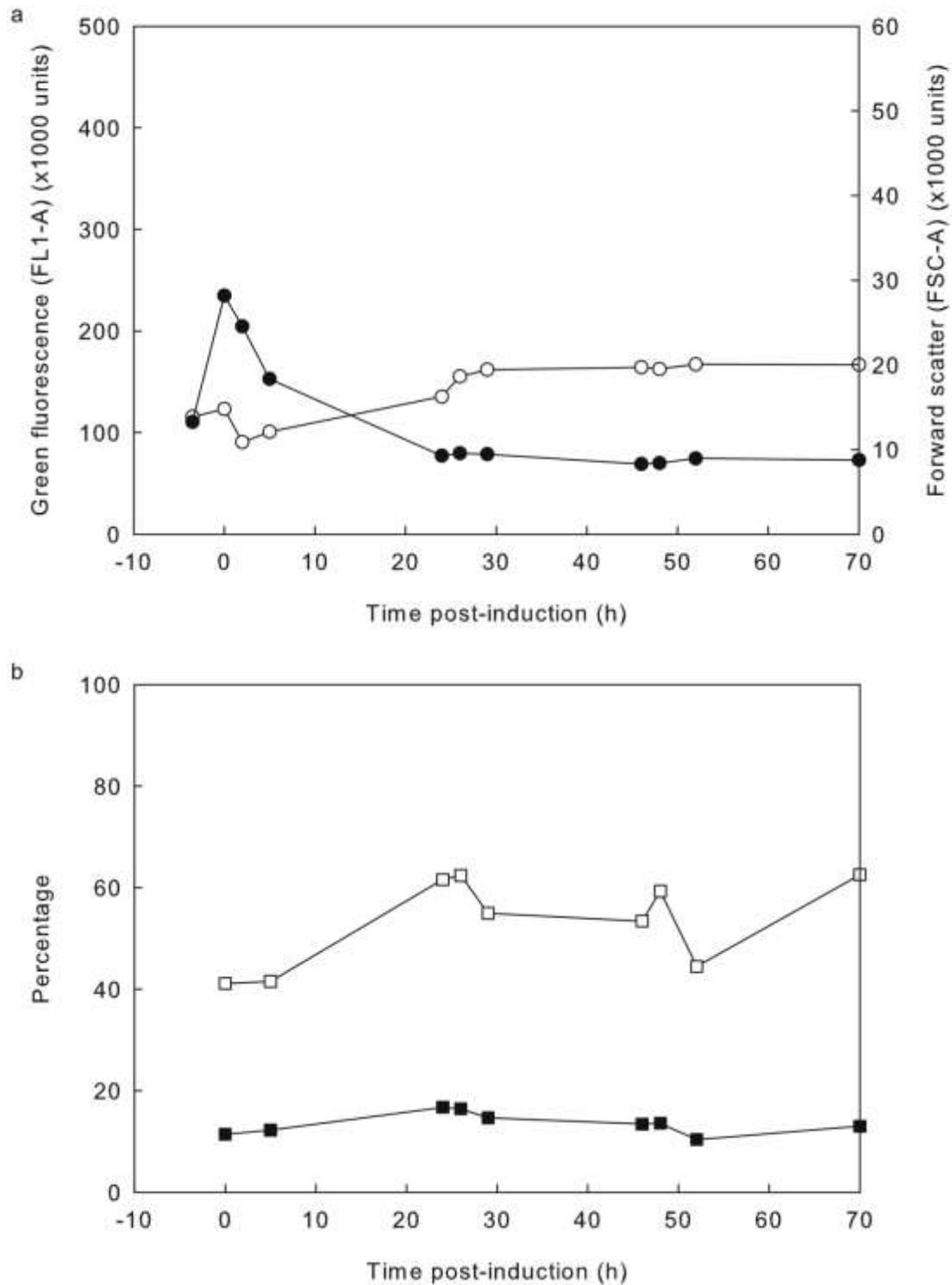
Figure 4.2: FCM and SDS-PAGE data from a protocol A fermentation
a) FCM data: Mean cellular green fluorescence (FL1-A) of GFP+ cells (○), mean cellular forward scatter (FSC-A) of GFP+ cells (●). b) SDS-PAGE data: Percentage of CheY::GFP that was soluble (□), percentage of total cellular protein that was CheY::GFP (■).

between 28-48 h then production of neither biomass nor CheY::GFP between 48 h and termination. It could therefore be argued that harvesting should occur at ~48 h.

Accumulation and solubility of CheY::GFP showed no overall increase after 24-28 h and biomass accumulation ceased after ~48 h. A potential explanation for this observation is as follows: It has been established that CheY::GFP production using this protocol is enhanced by supplementation with serine, threonine and asparagine as concentrations of these can become limiting (Jones, 2007). The media components used are undefined and therefore different sources and batches may have different amino acid distributions, thereby rendering the existing supplementation regime ineffective. If a different amino acid became limiting this could limit CheY::GFP and biomass formation. It is further presumed that limitation would not become sufficiently limiting for growth to cease as feeding tryptone and yeast extract would still provide sub-optimal amounts.

Protocol A is a direct replicate of the original improved fermentation protocol by Sevastsyanovich *et al.* (2009), a comparison of the data discussed above to that reported previously should demonstrate the efficacy of the replication here attempted. There are 2 published data sets using protocol A; Sevastsyanovich *et al.* (2009) and Alfasi (2010) (for convenience hereafter referred to as A(1) and A(2) respectively and summarised in Table 4.1). Both A(1) and A(2) showed similar accumulation of CheY::GFP reaching 31 and 35% respectively, but did not show consistent biomass accumulation with A(1) reaching less than half the terminal OD_{650} of A(2).

Terminal OD_{650} of A was similar to A(1), but much lower than A(2), although due to a higher DCW/OD_{650} the DCW of A was more similar to A(2) than OD_{650} (as A(1) predicted DCW from OD_{650} a comparison to A(1) is superfluous). A also showed much lower levels of RP accumulation than A(1) and A(2). It is worth noting that CheY:GFP yield data from A(1) and A(2) are predicted based on assumed cellular protein content, not determined

biochemically and are therefore of limited utility for comparison (see Section 4.2.9). From these data it appears that attempts to fully replicate the original improved protocol (i.e. high biomass and high CheY::GFP content) were limited in success, despite >10 attempts (Figure 7.1). A precise reason for this is not known, but the following is hypothesised: The large amounts of undefined media components present a large potential for variability and that this at least contributed to the limits in biomass and RP accumulation. CheY::GFP production is known to require amino acid supplementation (Jones, 2007), as alternate sources of medium components may have different amino acid distributions additional development stages may be required to devise an optimal supplementation scheme. In addition it was found later in this work that phosphate may have become limiting for growth (discussed in Section 4.2.8). Other data however were indicative of successful CheY::GFP production under improved conditions (high fluorescence, plasmid retention and CheY::GFP solubility and low cell death). Taken together, this suggests protocol A as previously reported is only partially reproducible, RP quality and the physiological state of the culture were consistent, suggesting that the stress-minimisation paradigm is effective but the transferral of A to an alternate setting placed limitations in both biomass generation (particularly as A(1) and A(2) showed inconsistent OD_{650} results) and overall product formation. This is consistent with the potential limitations predicted in Section 4.1 and supports the decision to adapt the stress-minimisation paradigm to more industrially-relevant (and therefore more likely reproducible) fermentation protocols.

In addition to low biomass and RP accumulation, additional limitations to protocol A were observed: First, the timing of required operator interventions was inefficient, induction occurs at ~3.5 h post-inoculation, an amino acid supplement is administered at 5 h post-induction and the feed must be manually turned on when batch-phase

glucose is consumed, typically between 7-11 h post-induction, requiring up to 14-15 h continual operator monitoring during initial stages. Second, total duration of the fermentation, at over 70 h is rather long for industrial applications as originally noted by Alfasi (2010).

Overall it can be concluded that while the stress-minimisation paradigm showed potential utility, the original improved fermentation protocol did not appear to be readily transferable to alternate settings and included many features that are unsuitable for industrial application, justifying the work subsequently presented.

4.2.2. Production of CheY::GFP by *E. coli* BL21* using Protocol B

The fermentation protocol used by (Want *et al.*, 2009) was selected as an example industrially-derived protocol for use in this study. In order to determine the effects of applying improved production conditions it was first necessary to grow cultures under standard conditions, termed Protocol B. Culture and feed volumes were reduced from the published protocol (3 L media and 1 L feed to 1.5 L media and 0.5 L feed) to remain approximately consistent with protocol A, and to ensure sufficient headspace as the culture is prone to foaming (see Appendix 2 (Chapter 8)). Duplicate fermentations were carried out according to protocol B (Tables 2.6, 2.7) and while there were commonalities in results (predominantly in growth patterns) there was substantial heterogeneity in measures of productivity. As it was not possible to generate reproducible results these data will not be discussed here in detail but are provided in Appendix 1 (Chapter 7, Figures 7.3, 7.4).

Both fermentations showed limited growth post-induction necessitating termination at OD_{650} ~80 (Figure 7.3a), substantially lower than would be expected considering the amount of carbon supplied as glycerol. In addition, both fermentations showed

substantial levels of plasmid loss as compared to A across all analytical methods used (replica-plating of CFU colonies, GFP+ phenotyping of CFU colonies and GFP+ phenotyping of cells by FCM) (Figure 7.3b,c), including a minimum of 40% GFP- cells in the pre-induced period. As growth limitation and plasmid loss are established as being indicative of RPP-related stress, based on these observations alone, it is strongly suggested that protocol B caused physiological stress. It is also likely that heterogeneity in results is indicative of the cellular response to stressful conditions.

4.2.3. Application of 'improved' production conditions to Protocol B (Protocol B1)

Although not conclusive, it appeared likely that protocol B produced RPP under stressful conditions. It was then necessary to determine whether application of improved production conditions could increase process productivity. A derivative protocol was designed where production conditions were modified to conform to the central tenets of the improved protocol (25°C, 8 μM IPTG), termed Protocol B1. Feeding began when on-line measurements indicated exhaustion of glycerol from the initial medium at 20-21 h post-inoculation and induction was concomitant with feeding at OD_{650} 47.8. Fermentations were terminated when it became apparent that growth had ceased at 21 h post-induction.

OD_{650} and on-line data (DOT, agitation, MS) showed slow initial growth, followed by a steady increase during the initial stages of feeding (Figures 4.3a, 7.5). The culture was then terminated at a final OD_{650} of 297. DOT and agitation data also suggested that following 6 h post-induction O_2 demand exceeded maximum aeration capacity until 12 h DOT and MS data showed that oxygen consumption and CO_2 evolution sharply decreased, hypothesised to be due to the feed rate becoming limited for growth during this period, potentially feeding only sufficient for cell maintenance. CFU counts appeared
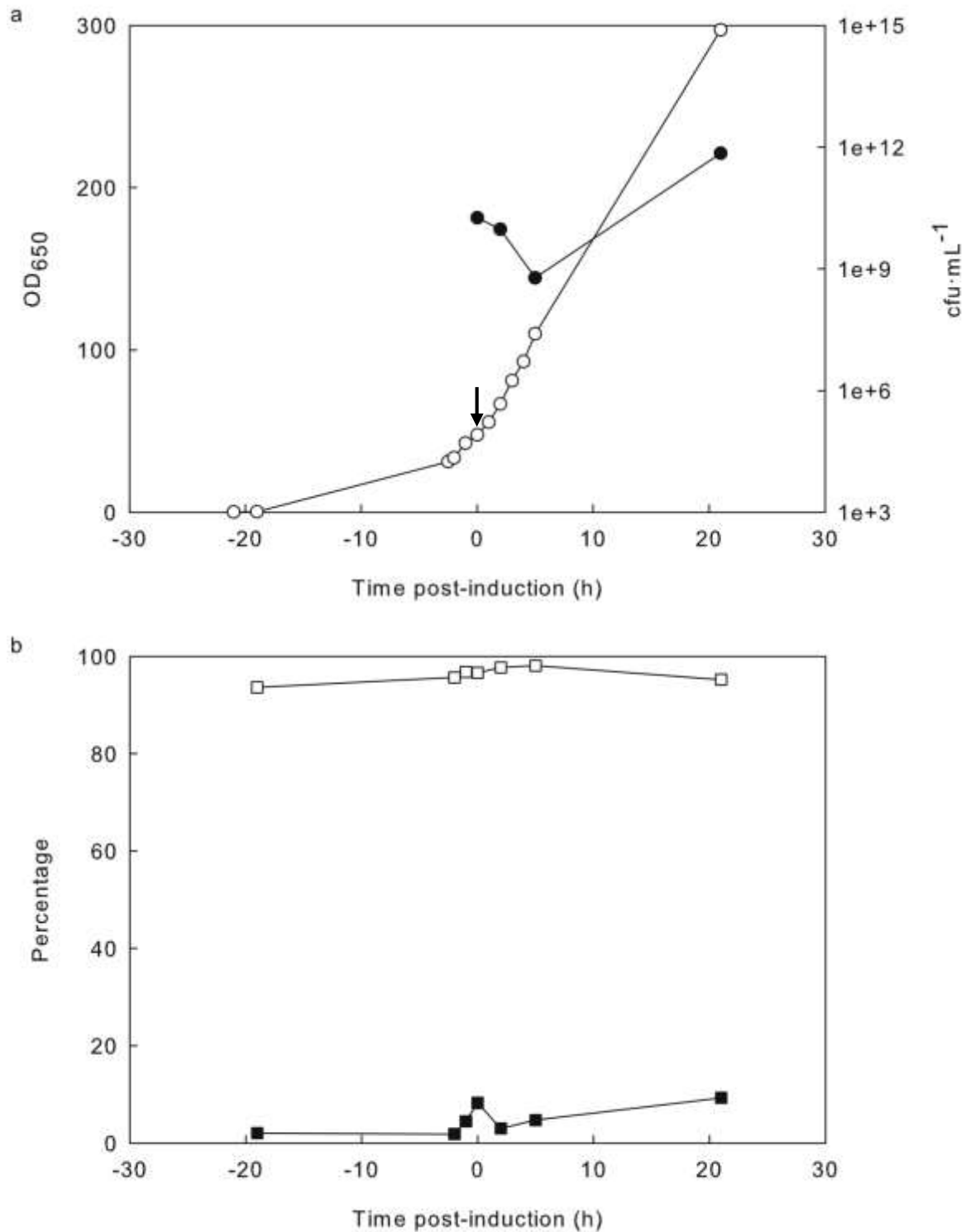
Figure 4.3: Culture growth, plasmid retention and viability data from a protocol B1 fermentation

a) Growth data: OD_{650} measurements ($\bigcirc$), CFU counts ($\bullet$). b) FCM data: percentage cells GFP+ ($\square$), percentage cells PI+ (dead) ($\blacksquare$). Arrow indicates point at which feed was turned on.

to decrease during the initial stages of feeding, then increased to 7.0×10^{11} cfu·mL^{-1} at termination. Although the proportion of dead (PI$^+$) cells remained lower than 10% throughout there were increases immediately before induction and at culture termination, suggesting that cell viability decreased when glycerol was depleted (Figure 4.3b). At termination the culture contained 72.8 g·L^{-1} DCW, corresponding to a Y_{xs} of 0.356 g·g^{-1} glycerol. The DCW/OD$_{650}$ ratio at 0.24 was substantially lower than both that observed for A (0.47) and the typically cited value of 0.4 (Sevastsyanovich *et al.*, 2009), an increase in cellular light scattering may have been an indicator of increased inclusion body formation (Hwang & Feldberg, 1990).

Unexpectedly, RQ data began higher than expected for growth on glycerol, only reaching a value close to that of the oxidation of pure glycerol (0.86) approximately 6 h before induction (Figure 7.5b). This is hypothesised to be due to preferential consumption of trace carbohydrates from the yeast extract in preference to glycerol.

Between 6 h pre-induction and 17 h post-induction a substantial amount of base was required for pH regulation (0.34 L) (Figure 7.5c), although base pump figures are likely to be overestimates as pockets of gas frequently accumulate in the tubing during the course of the fermentation. Addition of a large volume of base during the fermentation is indicative of acid production; this is surprising as glycerol is typically used to limit acid production. In fact, more base was added during this fermentation than for A that used glucose.

The proportion of GFP$^+$ cells as determined by FCM remained above 93% throughout the fermentation (Figure 4.3b). Fluorescence and forward scatter decreased during the initial stages of growth followed by a small increase immediately before induction (Figure 4.4a). Again this is interpreted as the dilution of cellular GFP content and decreasing size during rapid, uninduced growth, followed by an increase in CheY::GFP

168
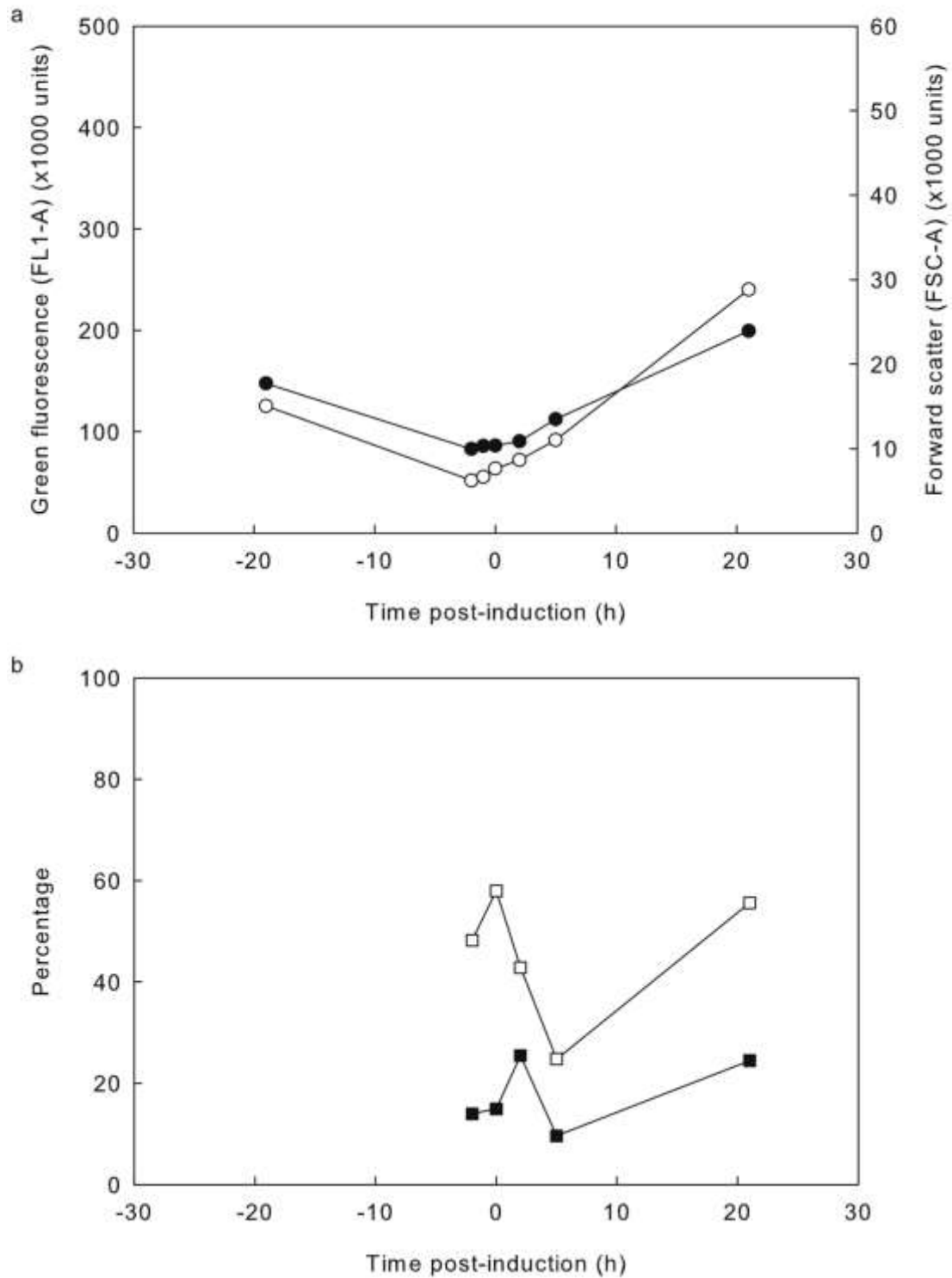
Figure 4.4: FCM and SDS-PAGE data from a protocol B1 fermentation
a) FCM data: Mean cellular green fluorescence (FL1-A) of GFP+ cells (○), mean cellular forward scatter (FSC-A) of GFP+ cells (●). b) SDS-PAGE data: Percentage of CheY::GFP that was soluble (□), percentage of total cellular protein that was CheY::GFP (■).

content and cell size as the growth rate decreased due to consumption of glycerol. Following induction, fluorescence increased to a maximum of 240,000.

RP solubility peaked at the point of induction, then decreased to a low point at 5 h post-induction but recovered to almost the peak value at termination reaching a final solubility of 56% (Figure 4.4b). Total RP accumulation followed a similar pattern although peak accumulation occurred at 2 h post-induction. Final yield of CheY::GFP was calculated to be 12.7 g·L^{-1}, corresponding to a yield of 7.1 g·L^{-1} soluble CheY::GFP. These data suggest that RP accumulation and solubility is lowest during periods of rapid growth.

Although not conclusive, it appeared likely that protocol B produced CheY::GFP under stressful conditions based upon apparent limitations in growth and substantial plasmid loss. A comparison to the data from B1 strongly supports this conclusion as there was a 4-fold increase in terminal OD$_{650}$ and the proportion of GFP$^+$ cells remained above 93% throughout. This suggests that cells in B1, unlike B, remained productive and viable throughout which is indicative of successful application of the stress-minimisation paradigm. The sole limitation observed for industrial application of the modified protocol is a 39% increase in run time (~12 h), but it is thought that the substantial increases in productivity and reproducibility should be sufficient to justify this expense.

The purpose of this series of experiments was to develop a fermentation protocol that combined the productivity of the original improved protocol with the scalability of an industrially-derived protocol as complex growth media are rarely used for HCDC (Lee, 1996). It is therefore also necessary to compare the original improved protocol (A) and the industrially-derived improved protocol developed here (B1) to determine whether B1 showed sufficient improvements in productivity to be recommended over A. B1 showed an almost 5-fold increase in cell density over A, CFU increased by over 1 log and

DCW was over double that of A, although as B1 used a larger amount of glycerol than A did of glucose, biomass yields were almost identical. FCM analysis showed similar levels of GFP+ cells, suggesting similar levels of vector retention, but cells in B1 were 46% more fluorescent and showed higher levels of homogeneity as evidenced by the coefficient of variance (CV) of the fluorescence values. B1 did show higher levels of dead cells however this did not appear to have affected productivity. RP solubility was 11% higher in A, but total RP accumulation in B1 was almost double that of A. These data, combined with the increased biomass resulted in over 4-fold increase in CheY::GFP yield and an over 3-fold increase in yield of soluble CheY::GFP. In addition to improvements in growth and productivity the fermentation run-time was 44% lower in B1 which would represent substantial savings in operating costs for industrial applications and further aided by the fact that growth medium B was 35% cheaper than A.

B1 also addressed a further limitation identified in A; it required fewer operator interventions, only inoculation, induction/feeding and harvesting. In addition, as induction/feeding occurred 20-21 h post-inoculation and harvesting, 21 h post-induction, all necessary interventions can be timed to occur within standard working days.

Overall B1 showed a moderate decrease in product quality to A, this appears to be in agreement with Moore *et al.* (1993) who showed that complex growth medium increased RP solubility. However, biomass formation and RPP productivity increased by sufficiently large amounts that overall amounts of total and soluble CheY::GFP produced were higher. In addition this was produced in a shorter amount of time using a cheaper growth medium and therefore at a lower cost. It is therefore concluded that the improved, industrially-derived protocol developed here represents a viable and attractive alternative to the improved fermentation protocol as originally reported.

An unexpected observation was that B1 required the addition of substantially more base to regulate the culture pH than A. This is in spite of the fact that growth on glucose, as in A, would be expected to generate a greater amount of acid than growth on glycerol, as in B. In addition, B1 was not supplemented with amino acids (serine, threonine and asparagine) unlike A, yet solubility and accumulation of CheY::GFP was not substantially affected. This was despite evidence that supplementation was required for effective production (Jones, 2007) and that amino acid deficiencies can be detrimental to growth and RPP. The initial yeast extract-derived amino acids could not account for total protein production and therefore it is concluded that additional *de novo* amino acid biosynthesis was sufficient for both biomass and CheY::GFP accumulation.

Although it was demonstrated that B1 did show improvements in CheY::GFP production over A the specific reasons for these improvements were unclear as the 2 protocols differ in multiple factors, specifically the point of induction (OD_{650} ~0.5 and ~40), principal carbon source (glucose and glycerol) and the growth medium. It was therefore decided to further adapt protocol B in order to investigate these factors.

4.2.4. The effect of changing induction point

One of the principal differences between protocols A and B is that protocol A induces RPP early in the fermentation, in mid-logarithmic phase (OD_{650} ~0.5) whereas B1 induces later in the fermentation, concurrent with feeding (OD_{650} 40-50). Induction at high biomass (i.e. later in the culture) is a strategy frequently used in industrial processes, separating the growth and production phases. This strategy is highly effective for systems where the production window, the period in which producing cells remain active and viable, is limited for example by the production of toxic protein or to limit the metabolic burden of standard RPP. Improved RPP does not limit the viability of

producing cells and increases the production window, although at the expense of limiting overall RPP rates. It would therefore be worthwhile to determine whether early induction of an improved B-derivative fermentation would increase RPP productivity over B1 due to the increased potential production window. Earlier induction has previously been established as having a positive effect on protein yields and a negative effect on biomass yields (Want *et al.*, 2009). In order to determine whether changing induction point affects improved RPP Protocol B2 was modified so that induction occurred at OD_{650} ~0.5, (as in protocol A); this was termed Protocol B2. A comparison of B1 and B2 will therefore highlight the effects of earlier induction

A fermenter was set up according to protocol B2 and was induced by the addition of 8 μM IPTG when the culture reached an OD_{650} of ~0.5, as in A. Feeding began prior to depletion of carbon source (17.5 h post-induction), was paused to allow consumption of glycerol (18.5 h) and resumed once it was apparent that the glycerol had been consumed (22 h). The fermentation was terminated once it had been determined that growth had ceased (43 h).

Growth of B2 progressed similarly to B1, reaching similar terminal OD_{650} readings at approximately 40 hours post-inoculation (Figure 4.5a). Terminal DCW was slightly higher than B1 at 77.5 g·L^{-1}, corresponding to a slightly higher Y_{xs} of 0.378. As with B1 the DCW/OD_{650} ratio was lower than expected at 0.27. On-line data also showed similar patterns to B1: initial RQ values were again higher than that expected for glycerol; metabolic activity declined after 28 h post-induction, again likely due to the feed rate becoming limiting; and again, a substantial volume of base was required for pH regulation.

The proportion of GFP$^+$ cells as determined by FCM (Figure 4.5b) during the initial stages of the fermentation was lower than expected; at approximately 80%. This was
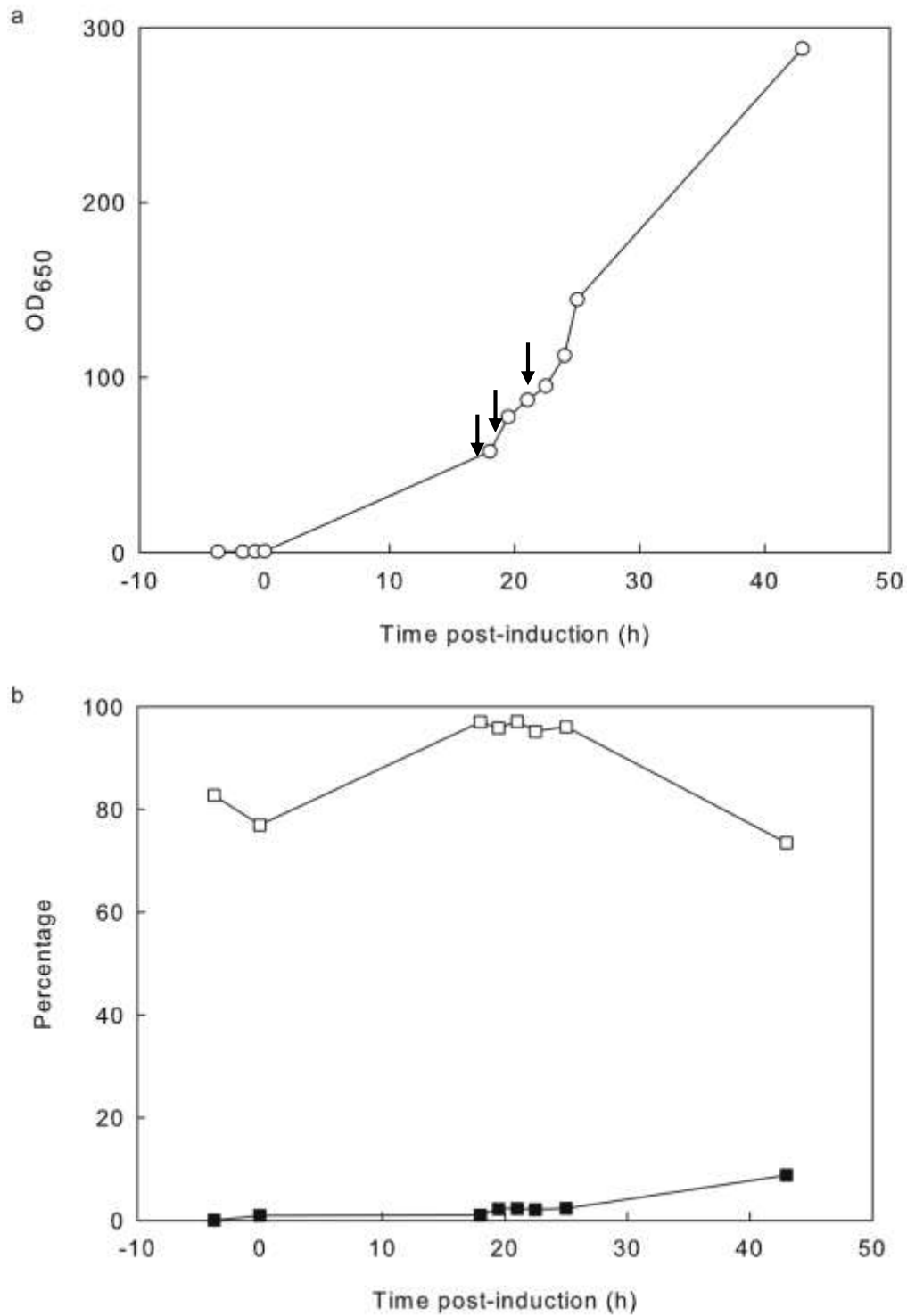
Figure 4.5: Culture growth, plasmid retention and viability data from a protocol B2 fermentation

a) Growth data: OD_{650} measurements ($\bigcirc$). b) FCM data: percentage cells GFP+ ($\square$), percentage cells PI+ (dead) ($\blacksquare$). Arrows indicate points at which (left to right) feed was turned on, paused and resumed.

however not due to the presence of GFP$^-$ cells but due to non-fluorescent particulate matter with a scatter distribution that overlapped that of the cells, observed to be consistent with the presence of antifoam (Figure 4.6). The low cell densities observed during the early stages of this fermentation resulted in a low signal to noise ratio and hence underestimated the proportion of GFP$^+$ cells. During the second day of monitoring, GFP$^+$ cells remained above 95% but at termination the proportion had dropped to 74%. PI staining also showed an increase in the percentage of dead cells at termination up to 8.8%. These data suggest that by termination the culture had become stressed.

The mean green fluorescence and FSC of GFP$^+$ cells during the initial stages of fermentation showed a pattern similar to that of A (Figure 4.7a); and again this is interpreted as an increase followed by a decrease in cell size as stationary-phase cells from the inoculum re-enter growth. After 18 h post-induction both parameters steadily increased to termination reaching final values of 370,000 for FL1-A and 25,000 for FSC-A. The increase in FL1-A during the latter stages of fermentation was greater than that of FSC-A, suggesting accumulation of GFP. In addition, the value of FL1-A at termination was the peak value measured during the fermentation, strongly suggesting effective accumulation of CheY::GFP.

SDS-PAGE analysis (Figure 4.7b) showed an increase in the percentage of total cellular protein that was CheY::GFP throughout; from 16% at the point of induction to 26% at termination. However the percentage solubility showed an overall decrease during the fermentation from 47% at induction to 37% at termination, suggesting that overall product quality had decreased. Final yield of CheY::GFP was calculated at 15.7 g·L^{-1}, corresponding to a yield of 5.9 g·L^{-1} soluble CheY::GFP.
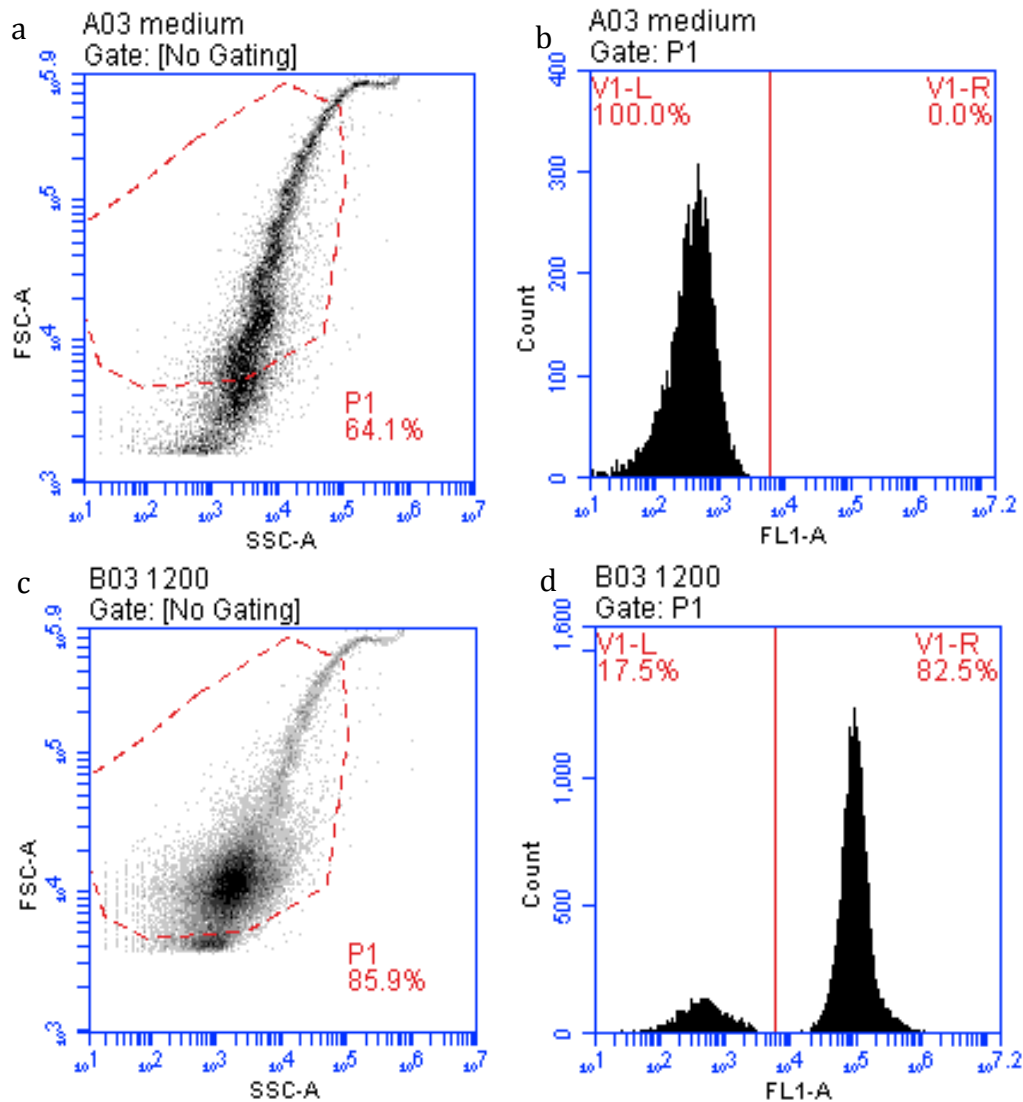
Figure 4.6: FCM analysis of early-stage fermentation samples showing interference by antifoam

a) FSC-A versus SSC-A intensity plot of uninoculated fermentation media from B3 showing particulate matter (antifoam) falling within noise-excluding gate P1. b) Histogram of green fluorescence intensity for (a), vertical gate dividing events into low (V1-L) and high (V1-R) green fluorescence events. c) FSC-A versus SSC-A intensity plot of B3 2 hours post-inoculation, particulate matter can still be seen. d) Histogram of green fluorescence intensity for (c).
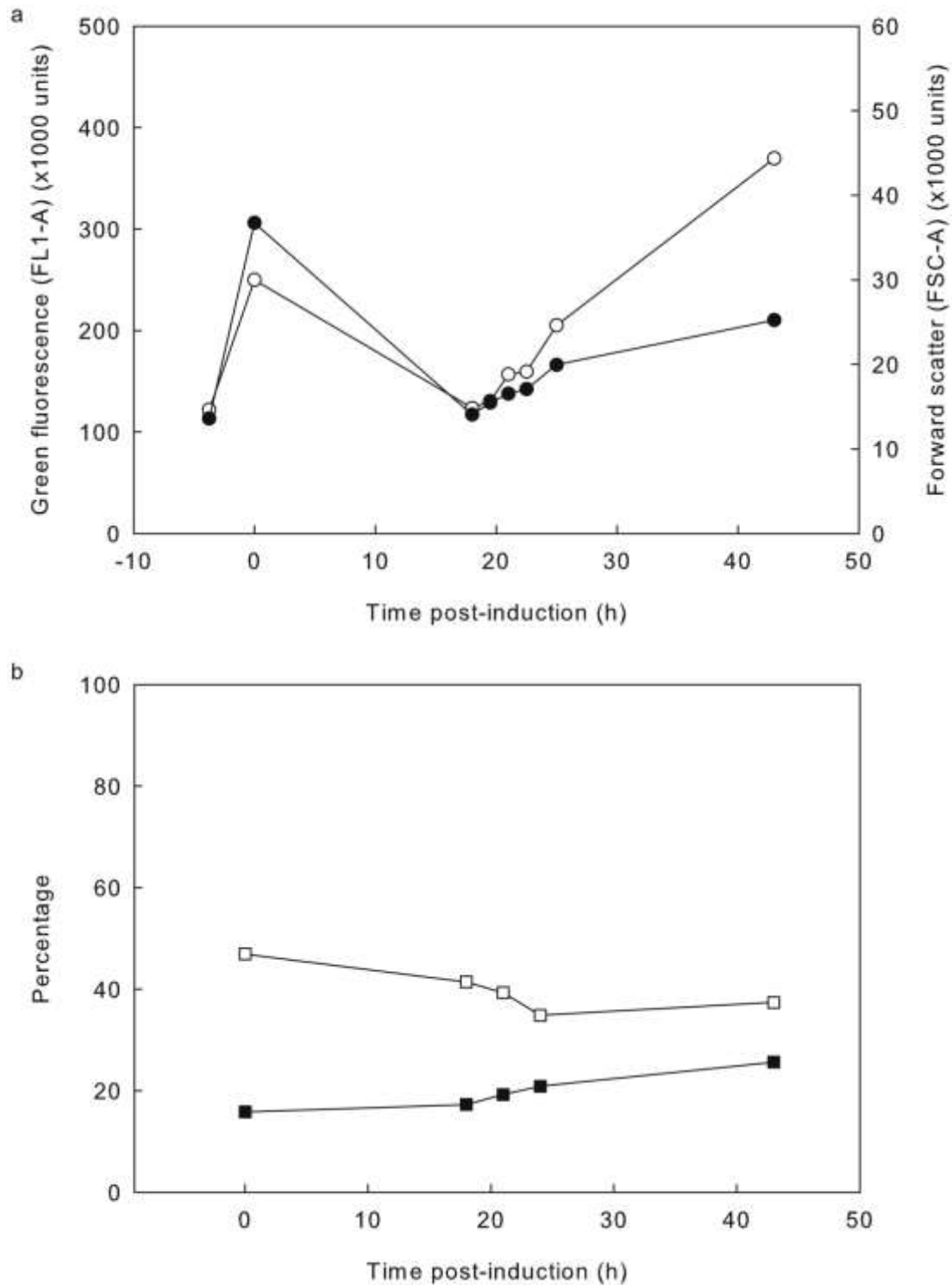
Figure 4.7: FCM and SDS-PAGE data from a protocol B2 fermentation
a) FCM data: Mean cellular green fluorescence (FL1-A) of GFP+ cells (○), mean cellular forward scatter (FSC-A) of GFP+ cells (●). b) SDS-PAGE data: Percentage of CheY::GFP that was soluble (□), percentage of total cellular protein that was CheY::GFP (■).

In terms of growth and biomass generation B1 and B2 showed similar final measurements (Table 4.1). Final OD_{650} was marginally higher for B1, although DCW was slightly higher for B2, resulting in a slightly higher Y_{xs}. In addition, culture viability, as indicated by the percentage of PI^+ cells, was similar. It can therefore reasonably be concluded that culture growth and biomass generation did not appear to be affected by earlier induction.

In terms of productivity the effect of early induction was mixed; The percentage of total protein that was CheY::GFP was similar, but the product quality, as indicated by CheY::GFP solubility was almost 20% lower in B2, resulting in higher total yields of CheY::GFP for B2 but lower yields of soluble CheY::GFP. The proportion of GFP^+ cells was over 20% lower for B2, but the mean green fluorescence of the GFP^+ cells was over 50% higher but more heterogeneous, with a higher CV. It was unexpected that despite similar amounts of CheY::GFP per unit biomass and lower levels of solubility that B2 showed a higher FL1-A value. This can be explained that FL1-A measurements were taken only as the average of GFP^+ cells whereas SDS-PAGE samples are taken from the culture as a whole, B1 contained a larger amount of productive cells than B2, but productive cells in B2 had accumulated more CheY::GFP. In support of this FL1-A values of all cells for B1 and B2 (taking into account non-productive cells) were more similar.

Costs of raw materials were identical but B2 required a longer fermentation, as would be expected due to the extended induction period, increasing overall fermentation costs, with similar yields of total and soluble CheY::GFP it is therefore concluded that B1 is likely to be the more economical choice.

The effects of induction point on RPP under standard conditions was explored by Want *et al.* (2009) who demonstrated that earlier induction resulted in lower yields of biomass but increases in RP yield. It has been demonstrated here that under improved

conditions this is not the case, earlier induction had no overall effect on biomass generation, a slight increase in total fermentation time and a mixed effect on RPP productivity. Standard RPP diverts substantial cellular resources, limiting growth at the expense of RPP (Sevastsyanovich *et al.*, 2009; Bentley *et al.*, 1990), therefore it is likely that earlier induction will divert a greater amount of cellular resources increasing RP yield at the expense of growth. As improved RPP does not limit growth as dramatically, that there was no limitation in biomass generation with earlier induction was not unexpected.

The rate of RP synthesis under improved conditions is slower than under standard conditions and therefore a longer induced period is likely to be required to accumulate a particular RP yield. If an improved fermentation is induced too late, there may be insufficient time to accumulate economically sufficient RP before growth ceases. However it also now appears that substantial extension of the production window for improved RPP causes heterogeneity in productivity and physiological stress that may be undesirable. Determination of the interaction between these opposing factors to determine optimal induction and harvesting points may be a useful future study.

4.2.5. The effect of changing principal carbon source

The second principal difference between protocols A and B is that of the carbon source used; A uses glucose whereas B uses glycerol. The effects of changing carbon source appears to be case-dependent; there are numerous studies that claim RPP is enhanced by growth on both glucose (Carvalho *et al.*, 2012; Tseng & Leng, 2012) and glycerol (Zhang *et al.*, 2010; Pflug *et al.*, 2007; Luo *et al.*, 2006). Protocol B2 was further modified such that glucose replaced glycerol as carbon source (the 35 g·L^{-1} glycerol used in the batch phase of B2 was replaced by 0.5 g·L^{-1} glucose and the feed (0.5 L) was comprised

of 400 g·L^{-1} glucose and 7.4 g·L^{-1} MgSO$_4$·7H$_2$O) and termed Protocol B3. A comparison of B2 with B3 allowed the effects of changing carbon source to be determined.

A fermentation was set up according to Protocol B3 and induced by the addition of 8 μM IPTG when the culture reached an OD$_{650}$ of ~0.5 (6 h post-inoculation). Feeding began when on-line measurements determined that the initial carbon source had been depleted at 9 h post-induction. Initially the feed rate was calculated in order to produce an exponential growth curve for a μ of 0.2, using the following equation (Strandberg *et al.*, 1994):

$$F = \left(\frac{1}{S}\right) \times \left(\frac{\mu}{Y_{XS}} + m\right) \times X_0 \times e^{\mu t}$$

Where: *F* equals the feed rate (L·h^{-1}); X$_0$, total biomass at start of feed (g DCW); μ, specific growth rate (0.2 h^{-1}); *t*, time (h); S, feed glucose concentration (400 g·L^{-1}); Y$_{XS}$, cell yield on glucose (0.622 g biomass·g glucose^{-1}; (Wallace & Holms, 1986)); and m, maintenance coefficient for glucose (0.00468 g glucose·g biomass^{-1}·h^{-1}; (Wallace & Holms, 1986)). This was continued until the feed rate equalled that of B1-2 (45 mL·L^{-1}·h^{-1} relative to initial volume) then feeding remained linear until termination. The fermentation was terminated once it had been determined that growth had ceased at 40 hours post-induction.

Following the initial lag period, growth proceeded steadily up to 26 h post-induction reaching OD$_{650}$ 138 and 1.2x10^{11} cfu·mL^{-1} at 24 h (Figure 4.8a). At termination the OD$_{650}$ had increased further to 167 but CFU counts had decreased to 1.1x10^{10} cfu·mL^{-1}, suggesting that at termination cells had ceased to grow and had likely entered stationary phase. Agitation data suggest that following 25 h O$_2$ consumption began to decrease, possibly indicating that the feed had become limiting and gas-MS data show that after the feed was exhausted at 32 h metabolic activity rapidly decreased. At termination the
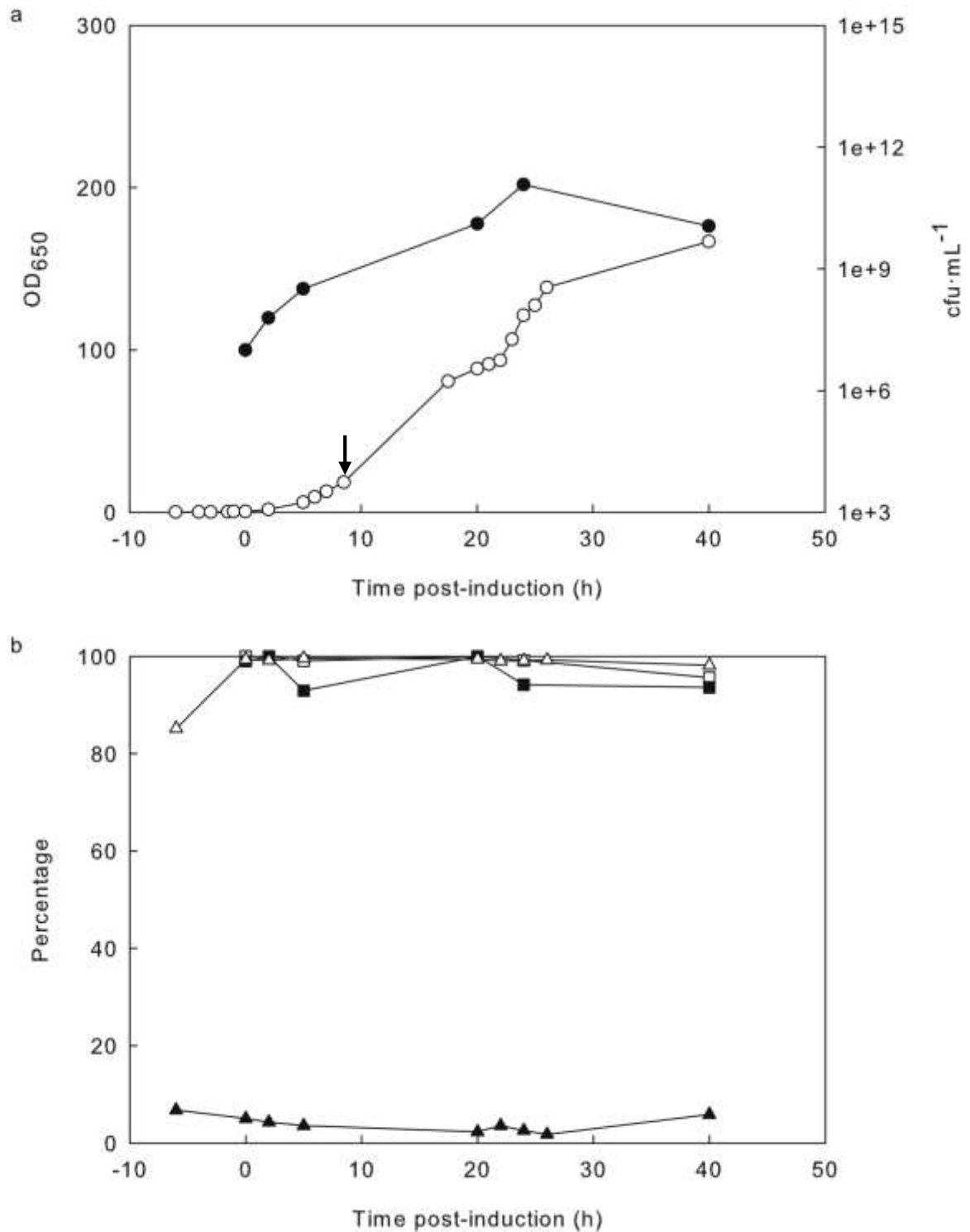
Figure 4.8: Culture growth, plasmid retention and viability data from a protocol B3 fermentation

a) Growth data: OD_{650} measurements ($\bigcirc$), CFU counts ($\bullet$). b) Plasmid retention and FCM data: Percentage colonies plasmid[+] ($\square$), percentage colonies GFP[+] ($\blacksquare$), percentage cells GFP[+] (by FCM) ($\triangle$), percentage cells PI[+] (dead) (by FCM) ($\blacktriangle$). Arrow indicates point at which feed was turned on.

fermentation had achieved a DCW of 42.6 g.L^{-1}, resulting in a Y_{xs} of 0.411 g.g^{-1} and a DCW/OD$_{650}$ ratio of 0.26.

Plasmid retention, as measured by replica plating and GFP$^+$ phenotyping of CFU plates remained above 92% throughout and the percentage of GFP$^+$ cells determined by FCM remained above 98%, except for the initial sample that was 85% (Figure 4.8b). This was again due to high levels of background noise caused by antifoam and low cell density (as Figure 4.6). It can therefore be concluded that there was no appreciable amount of plasmid loss during the fermentation. The percentage of dead cells as determined by FCM (Figure 4.8b) remained at less than 7% throughout, although there was an increase between 26 h (1.8%) and termination (5.9%), possibly suggesting a slight increase in cell stress due to the onset of stationary phase.

Green fluorescence of GFP$^+$ cells and FSC (Figure 4.9a) showed a similar pattern to other early-induced fermentations: During the early stages of fermentation both FL1-A and FSC-A rapidly increased up to the point of induction then gradually decreased. Again this is interpreted as being due to cells initially increasing in size as they re-enter growth phase followed by decreasing cell size as they begin to divide, causing partitioning of GFP between cells and hence decrease in FL1-A. Between 6 and 20 h FSC-A decreased to a greater extent than FL1-A, suggesting that while the cells became smaller, GFP content increased. Between 20 and 26 h FL1-A increased by approximately 30% and FSC-A increased by approximately 20%, suggesting accumulation of GFP despite increasing cell size. At termination FL1-A had increased to 296,000 from 265,000 at 26 h, and FSC-A had increased from 21,800 to 22,700. Again FL1-A had increased to a greater extent than FSC-A, suggesting further accumulation of GFP.

SDS-PAGE data (Figure 4.9b) showed an increase in the percentage of total cellular protein that was CheY::GFP from 2 h post-induction until termination, reaching a final
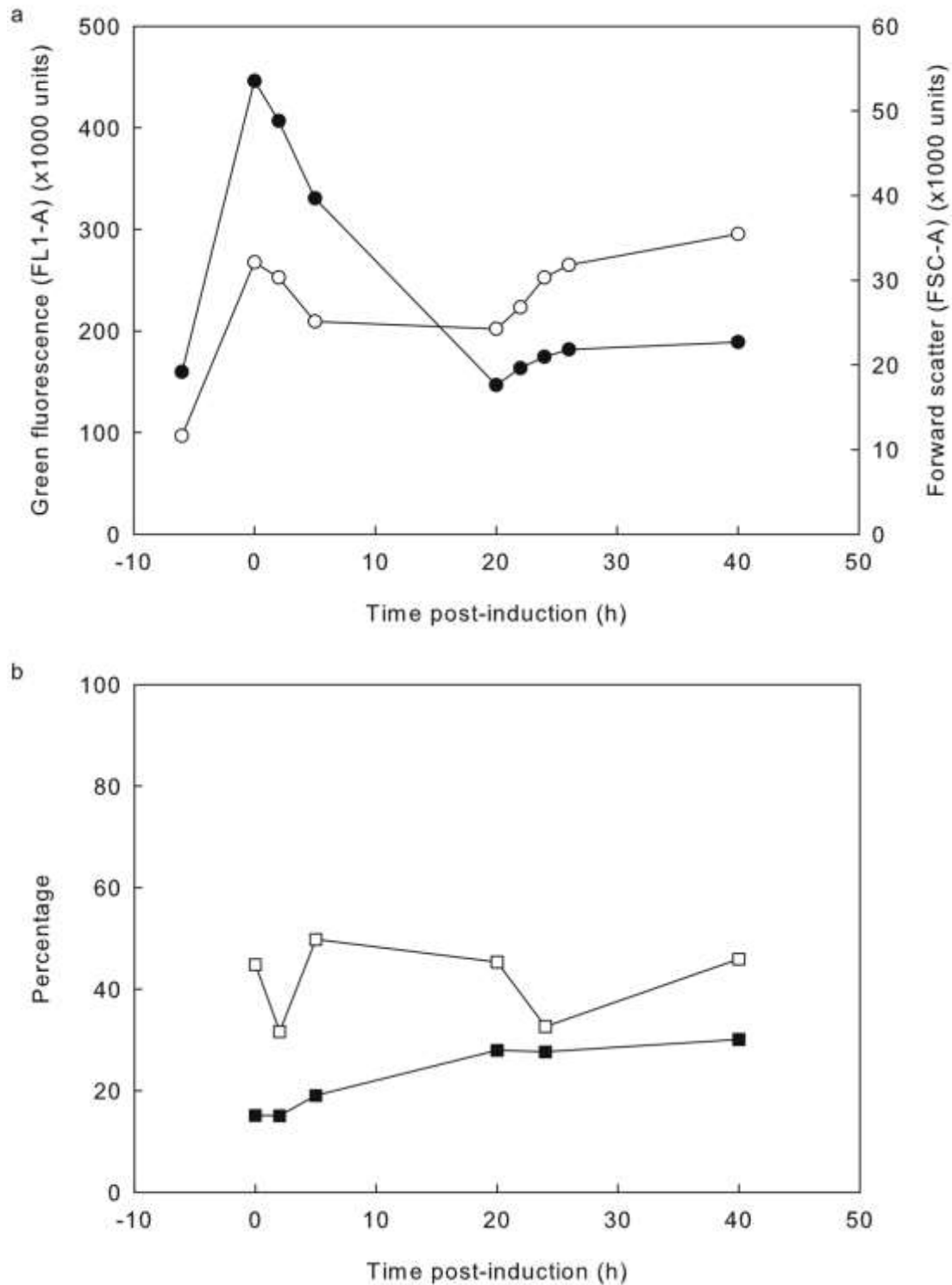
Figure 4.9: FCM and SDS-PAGE data from a protocol B3 fermentation
a) FCM data: Mean cellular green fluorescence (FL1-A) of GFP+ cells (○), mean cellular forward scatter (FSC-A) of GFP+ cells (●). b) SDS-PAGE data: Percentage of CheY::GFP that was soluble (□), percentage of total cellular protein that was CheY::GFP (■).

peak value of 30%. CheY::GFP solubility fluctuated during the fermentation, peak solubility of 50% was observed at 5 h post-induction and final solubility reached 46%. Final yield of CheY::GFP was calculated at 11.1 g·L^{-1}, corresponding to a yield of 5.5 g·L^{-1} soluble CheY::GFP.

B3 showed lower final OD_{650} and DCW than B2, this was expected, as the total mass of glucose added in B3 was substantially lower than that of glycerol in B2. However B3 showed a higher Y_{xs} and therefore the efficiency of converting carbon source to biomass was slightly higher.

FCM data showed a higher proportion of GFP$^+$ cells and lower proportion of PI$^+$ cells in B2 suggesting that B3 was less stressed. The mean green fluorescence of GFP$^+$ cells in B3 was slightly (<20%) lower than B2, as was the mean FSC.

SDS-PAGE analysis showed that both CheY::GFP accumulation and solubility were higher in B3 than B2. Due to lower yield of biomass the yield of total CheY::GFP was 35% lower although the higher solubility of B3 resulted in similar yields of soluble CheY::GFP. This occurred despite B2 showing a higher final fluorescence of GFP$^+$ cells. Again this is interpreted as being the result of B2 having fewer GFP$^+$ cells. This conclusion is supported by final FL1-A values for all cells that were similar for both protocols

Both protocols required similar amounts of time therefore operational costs would also be similar, but growth medium for B3 cost almost half that of B2, due to the higher cost of glycerol and that less glucose was used in B3 than glycerol in B2. This may therefore result in a slight economic advantage to B3 for the production of soluble CheY::GFP as yields of soluble protein were similar.

4.2.6. The effect of changing growth medium

To determine the effects of changing the growth medium from LB to the more defined medium of Want *et al.* (2009), protocol A can be compared to B3, the improved conditions, early-induced, glucose utilising derivative of B.

B3 showed a 2.5 fold increase in OD_{650} over A along with a 41% increase in DCW, and a 13% increase in Y_{xs}. In addition B3 showed a higher final OD_{650} and a similar final DCW to A(2), the most productive reported instance of A. As these values were obtained with the same amount of glucose and a drastically lower amount of complex components it is reasonable to conclude that, in terms of biomass generation, the growth medium from B was more effective than A. This finding is surprising as tryptone and yeast extract contain 4.3 and 16.3% carbohydrates respectively in addition to carbon incorporated into amino acids and therefore the amount of organic carbon available in A was higher.

FCM analysis showed that the amounts of GFP^+ and PI^+ cells were similar between B3 and A, suggesting similar levels of plasmid retention and cell viability. Mean FL1-A was almost 2-fold higher for B3. Although FL1-A was higher for B3, FSC-A was over 2-fold higher and therefore the increase in FL1-A may be, at least partially, due to increasing cell size.

At 30.1% the final percentage of total cellular protein that was CheY::GFP in B3 was the highest achieved during this study, over double that of A and similar to that reported in A(1). Protein solubility however was lower than A at 49.5% which is in agreement with Moore *et al.* (1993) that more complex media favour protein solubility. Finally, B3 required 37% less run-time and the growth medium cost 66% less per litre.

An unexpected additional effect of changing the growth medium was also observed, all fermentations using growth medium B required the addition of a greatly increased amount of base to regulate pH. Base addition in fermentations is usually required due to

production of acidic by-products of carbohydrate metabolism either by overflow metabolism of high concentrations of glucose or anaerobic mixed acid fermentation (Lee, 1996), although less common, fermentation to acid has been documented with glycerol (Korz *et al.*, 1995). Base addition occurred irrespective of the carbon source used (precluding overflow metabolism) and occurred under periods where oxygen should not have been limiting (precluding fermentative metabolism). It is therefore assumedly a result of the growth medium itself. The cause of this is currently unknown; therefore more detailed metabolic analysis of cultures growing on the medium would be a potentially beneficial route of further investigation. This phenomenon is discussed in more detail in Section 4.2.8.

4.2.7. Adaptation of a chemically-defined medium based fermentation protocol for 'improved' RPP

The physiological stress-minimising improved RPP protocol was successfully applied to an industrially-derived HCDC protocol using semi-defined media that would eliminate some limitations of complex media such as LB (Section 4.1). A semi-defined medium however, still relies on some complex components and it may still be preferable to use a chemically-defined medium. Here follows a series of experiments designed to adapt a defined medium HCDC RPP protocol to produce soluble cytoplasmic RP under stress-minimising conditions.

To assess the effects of growth on chemically-defined growth media on 'improved' RPP the protocol of Humphreys *et al.* (2002) was adapted for production under stress-minimised conditions and to be consistent with the fermentations presented above. Initial culture volume was reduced to 1.5 L and the fermentation was grown at 25°C throughout. Both feeding and induction were substantially altered. In Humphreys *et al.*

186

(2002) cultures were initially pulse-fed 2 doses of 45 mL 80% (w/w) glycerol then induced by feeding 20-50 $g \cdot L^{-1}$ lactose when glycerol was exhausted at OD_{600} ~80, this was altered such that feeding was identical to protocol B and RPP was induced by addition of 80 μM ITPG at OD_{650} ~80.

4.2.7.1. Attempt 1 (C1)

On consumption of the initial 31.11 $g \cdot L^{-1}$ glycerol (as indicated by DOT and MS data at 26.5 h post-inoculation) feeding began identically to that of protocols B1-3. RPP was induced once the culture OD had reached a minimum of 80 at 39 h post-inoculation. The fermentation was terminated earlier than expected at 8 h post-induction as OD_{650} readings had indicated that growth had ceased.

Growth, as indicated by OD_{650} (Figure 4.10a) data did not increase substantially before ~20 h pre-induction suggesting an extended lag-phase compared to A and B fermentations, as would be expected for growth on minimal media. The culture grew rapidly between 20-11.5 h pre-induction, reaching OD_{650} 45.7, but on resumption of monitoring the next day the OD_{650} had reached 89.4 at 1 h pre-induction, this increased to a peak of 91.2 at 1 h post-induction after which it decreased steadily to 81.2 at 8 h post-induction presumably due to dilution of the culture by feed and base addition and suggesting that the culture had ceased to grow therefore causing termination. CFU counts (Figure 4.10a) increased from 1.4×10^{10} $cfu \cdot mL^{-1}$ at 12.5 h pre-induction to 8.3×10^{10} at induction then increased only slightly to 9.3x10 at 3 h post-induction. On-line data support OD_{650} data and show that following ~7.5 h pre-induction metabolic activity began to decrease, suggesting this was the point at which the culture ceased to grow. In addition on-line data showed no initial phase of higher RQ as with B1-3 (Figure 7.8b), supporting the assumption that this was due to consumption of trace
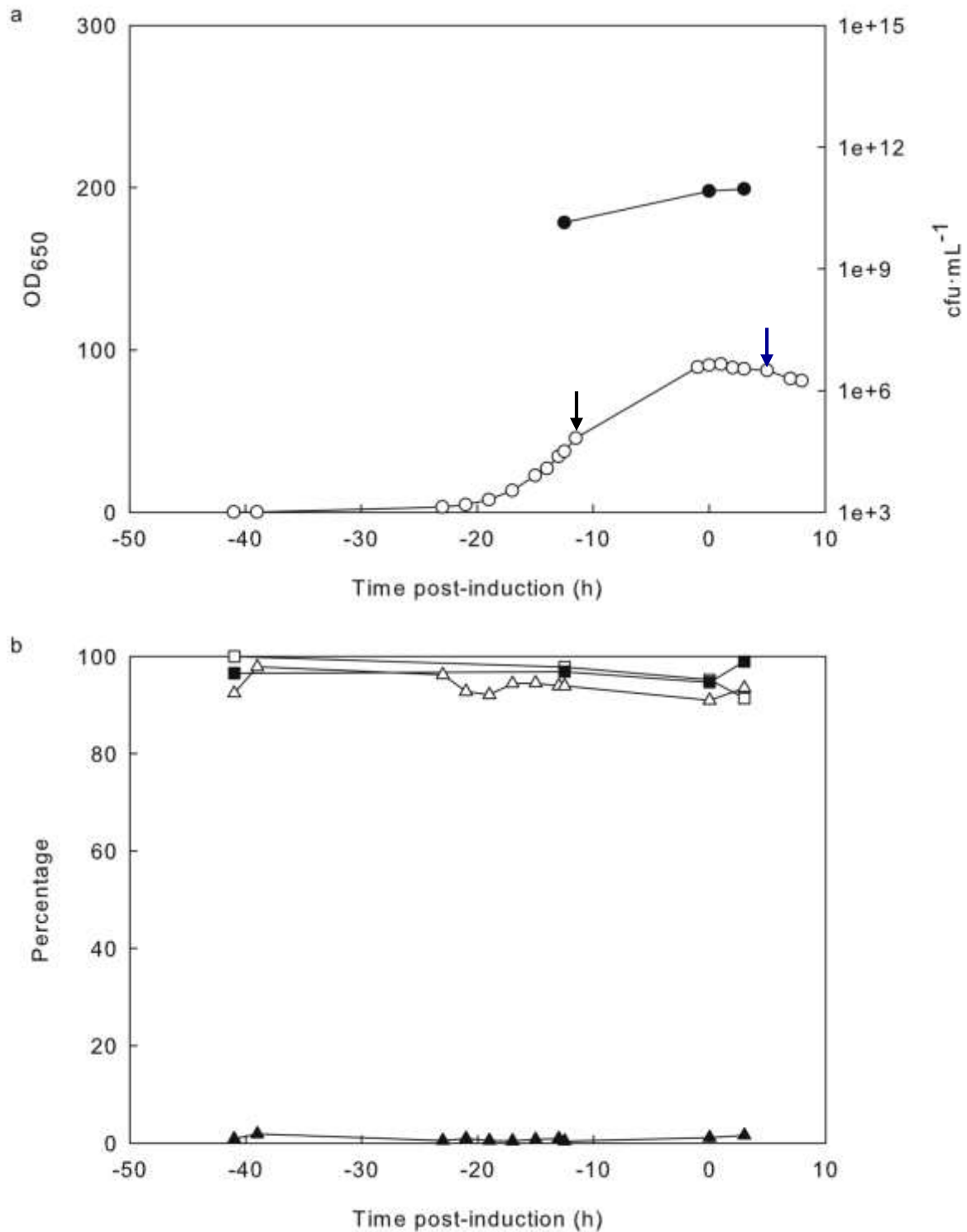
Figure 4.10: Culture growth, plasmid retention and viability data from a protocol C1 fermentation

a) Growth data: OD_{650} measurements ($\bigcirc$), CFU counts ($\bullet$). b) Plasmid retention and FCM data: Percentage colonies plasmid[+] ($\square$), percentage colonies GFP[+] ($\blacksquare$), percentage cells GFP[+] (by FCM) ($\triangle$), percentage cells PI[+] (dead) (by FCM) ($\blacktriangle$). Black arrow indicates point at which feed was turned on, blue arrow indicates addition of 15 g $(NH_4)_2SO_4$.

carbohydrates from the yeast extract and similarly to B1-3, show substantial base addition in order to maintain pH following transition to exponential growth.

As the application of improved conditions to protocol B increased the final OD over that reported by Want *et al.* (2009) it was unexpected that was not the case for protocol C. A potential explanation was formed on the observation that the DOT would rapidly and transiently decrease as base was added to the culture for pH regulation and increasing the pH set point at 6.4 at 3.5 h post-induction (requiring an increase in rate of base addition to maintain) causing an increase in metabolic activity. An increase in metabolic activity in response to base addition, as the base was 10% NH_3, suggested that growth limitation may be due to nitrogen starvation especially as the concentration of $(NH_4)_2SO_4$ in the initial growth medium was 63% lower than B, therefore a supplement of 50 mL 30% $(NH_4)_2SO_4$ was added at 4.75 h post induction. This addition did not cause an increase in OD, assumedly this was because the culture had already begun to enter stationary phase and therefore was terminated prematurely.

As the fermentation was terminated prematurely, measures of productivity were limited. All measures of plasmid retention (percentage of GFP+ colonies, plasmid+ colonies and GFP+ cells by FCM) (Figure 4.10b) remained above 90% throughout and the percentage of PI+ cells remained below 2%, suggesting that despite potential metabolic deficiencies, physiological stresses were not sufficient to impair plasmid retention or viability.

Initial FCM data showed a similar pattern to A, B2 and B3; FL1-A and FSC-A of GFP+ cells (Figure 4.11) both increased initially then decreased as cells entered exponential phase. Between 21-12.5 h pre-induction the FL1-A and FSC-A showed no overall increase, suggesting that during this period of uninduced growth basal levels of CheY::GFP production had been reached. After 12.5 h pre-induction fluorescence increased with a
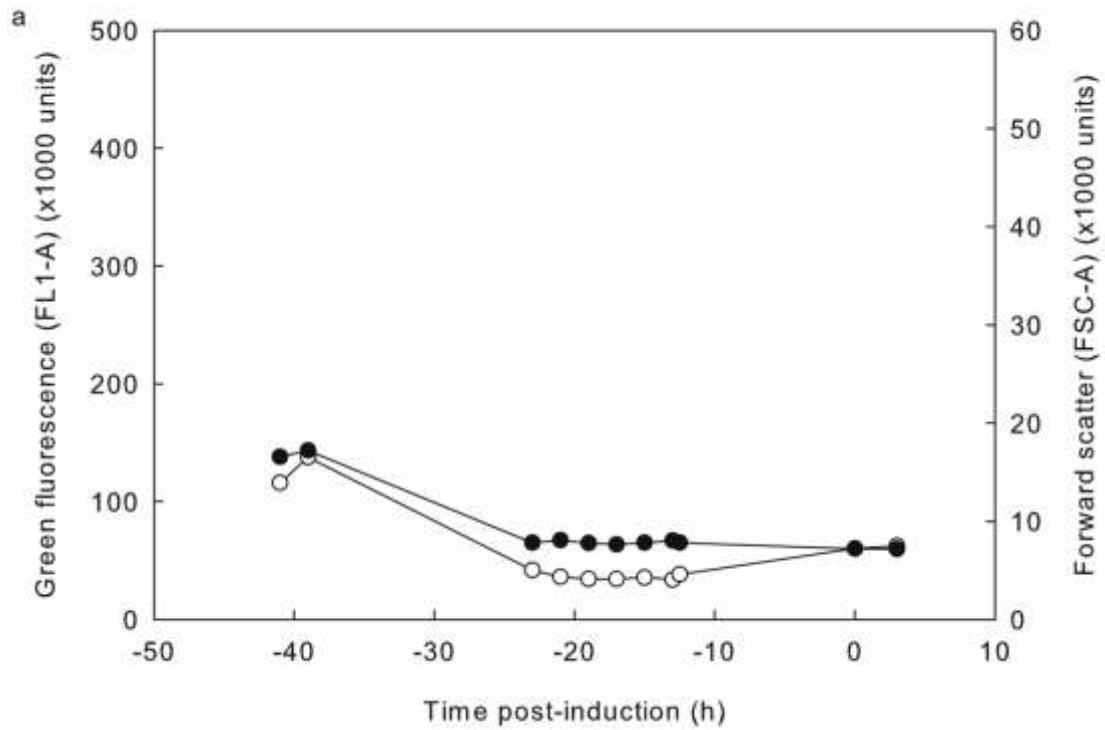
Figure 4.11: FCM data from a protocol C1 fermentation

FCM data: Mean cellular green fluorescence (FL1-A) of GFP+ cells (○), mean cellular forward scatter (FSC-A) of GFP+ cells (●).

slight decrease in FSC-A, suggesting an increase in CheY::GFP content despite most of this period being uninduced. This is likely due to leaky CheY::GFP production during a decreasing growth rate.

Based on the observations during C1 regarding the potential N-deficiency it was therefore decided to repeat the fermentation with additional $(NH_4)_2SO_4$ during the initial fermentation media termed attempt C2.

4.2.7.2. Attempt 2 (C2)

A fermenter was set up as C1 except that the 5.2 $g·L^{-1}$ $(NH_4)_2SO_4$ in the initial growth medium was increased to equal that of protocol B (14 $g·L^{-1}$) and termed C2 (summarised in Table 4.2). Feeding began once on-line data indicated that the initial glycerol had been consumed (26 h post-inoculation), RPP was induced when the culture had reached an OD_{650} of ~80 (79.1 at 31 h post-induction) and the fermentation was terminated when on-line measurements indicated that growth had ceased (41 h post-induction).

OD_{650} data showed a shorter lag period compared to C1 (Figure 4.12a), this is unlikely to be due to the N-supplementation; rather a malfunction of the cooling system at 22 h pre-induction that caused the temperature to increase to 36°C before correction at 8 h pre-induction (Figure 7.9a). This will have increased the growth rate and hence decrease the lag period. After ~11 hours pre-induction metabolic activity increased rapidly suggesting growth and the OD_{650} increased from 28.6 at 7 h pre-induction to 99.2 at 2 h post-induction. Between 1-20 h post-induction metabolic activity showed a steady decline suggesting that growth had predominantly ceased. On resumption of monitoring at 17 h post-induction the OD had increased moderately to 121, then decreased to 107 at 18-20 h, supporting the conclusion that between 2-20 h cell growth had become limited despite feeding. As with C1, growth limitation was not expected and as the

Table 4.2: Summary of modifications to growth medium C

	C1	C2	C3
Growth medium composition	5.2 g·L^{-1} $(NH_4)_2SO_4$ 3.86 g·L^{-1} $NaH_2PO_4·H_2O$ 4.03 g·L^{-1} KCl 1.04 g·L^{-1} $MgSO_4·7H_2O$ 0.25 g·L^{-1} $CaCl_2·2H_2O$ 10 mL·L^{-1} SM6E Trace elements solution 4.16 g·L^{-1} Citric acid monohydrate 31.11 g·L^{-1} Glycerol 0.66 mL·L^{-1} PPG	**14 g·L^{-1} $(NH_4)_2SO_4$** 3.86 g·L^{-1} $NaH_2PO_4·H_2O$ 4.03 g·L^{-1} KCl 1.04 g·L^{-1} $MgSO_4·7H_2O$ 0.25 g·L^{-1} $CaCl_2·2H_2O$ 10 mL·L^{-1} SM6E Trace elements solution 4.16 g·L^{-1} Citric acid monohydrate 31.11 g·L^{-1} Glycerol 0.66 mL·L^{-1} PPG	**14 g·L^{-1} $(NH_4)_2SO_4$** 3.86 g·L^{-1} $NaH_2PO_4·H_2O$ 4.03 g·L^{-1} KCl 1.04 g·L^{-1} $MgSO_4·7H_2O$ 0.25 g·L^{-1} $CaCl_2·2H_2O$ 10 mL·L^{-1} SM6E Trace elements solution 4.16 g·L^{-1} Citric acid monohydrate 31.11 g·L^{-1} Glycerol 0.66 mL·L^{-1} PPG **2 g·L^{-1} KH_2PO_4** **16.5 g·L^{-1} K_2HPO_4** **1.5 mL·L^{-1} Conc. H_3PO_4**
Supplements	15 g $(NH_4)_2SO_4$ (4.75 hours post-induction)	15 g $(NH_4)_2SO_4$ (2 hours post-induction) 20 g Na_2HPO_4 (20 hours post-induction) 10 g NaH_2PO_4 (20 hours post-induction) 1.5 mL *E. coli* sulphur-free salts (20 hours post-induction)	2 mL *E. coli* sulphur-free salts (11.5 hours post-induction) 2 mL SM6E Trace elements solution (11.5 hours post-induction)

N.B: Alterations to basic growth medium in Table 2.6 are in bold.
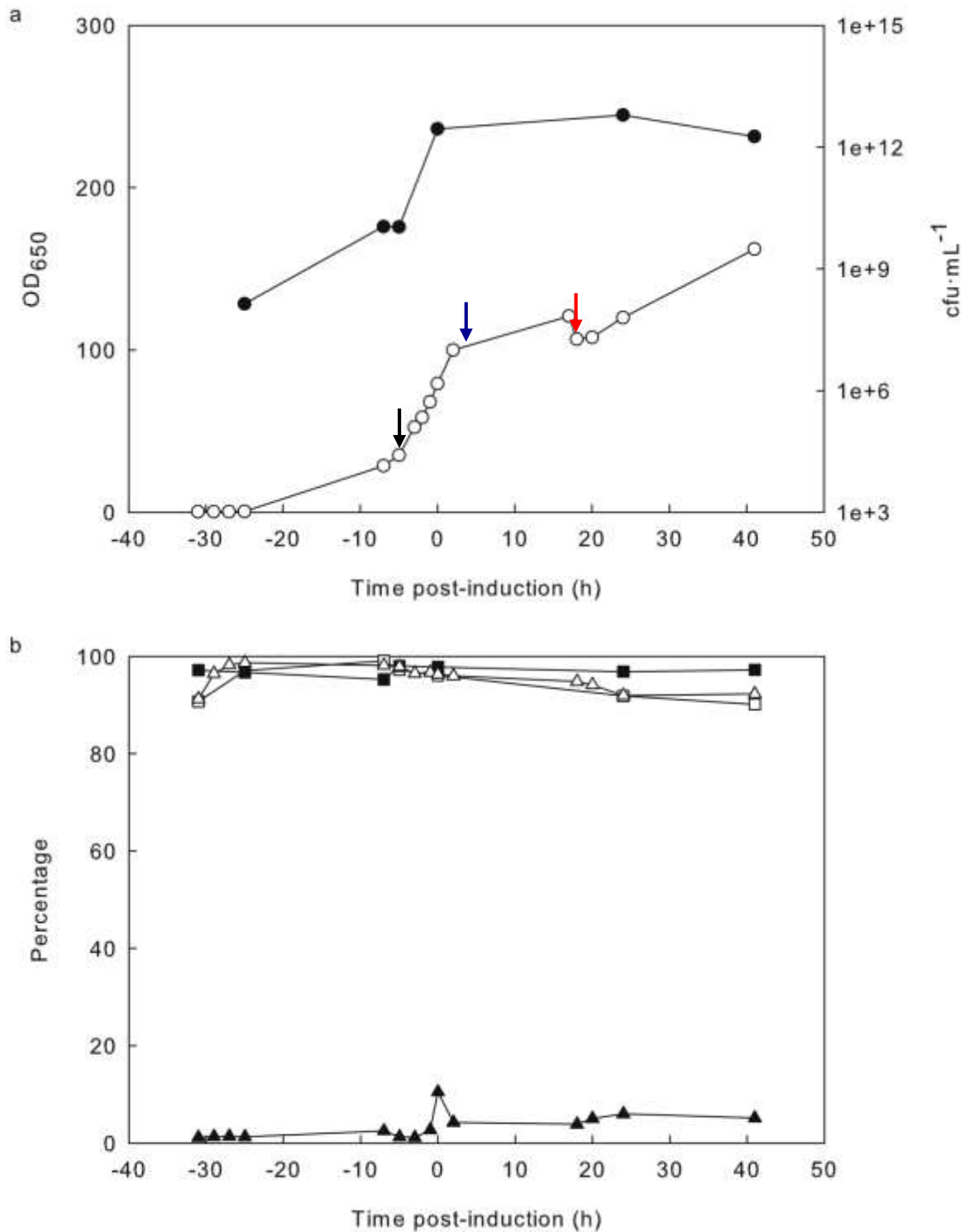
Figure 4.12: Culture growth, plasmid retention and viability data from a protocol C2 fermentation

a) Growth data: OD_{650} measurements ($\bigcirc$), CFU counts ($\bullet$). b) Plasmid retention and FCM data: Percentage colonies plasmid+ ($\square$), percentage colonies GFP+ ($\blacksquare$), percentage cells GFP+ (by FCM) ($\triangle$), percentage cells PI+ (dead) (by FCM) ($\blacktriangle$). Black arrow indicates point at which feed was turned on, blue arrow indicates addition of 15 g $(NH_4)_2SO_4$, red arrow indicates addition of 20 g Na_2HPO_4 and 10 g NaH_2PO_4.

nitrogen content had already been increased (as a precaution an additional supplement of 15 g $(NH_3)_2SO_4$ was also added at 2 h post-induction), N-starvation was unlikely to be the cause. As the amount of phosphates in medium C was also substantially less than B it was hypothesised that this may be the cause and therefore at 20 h post-induction a supplement of 20 g Na_2HPO_4 and 10 g NaH_2PO_4 was added (Table 4.2), along with 1.5 mL of *E. coli* sulphur free salts. This supplement had a pronounced effect on the culture, metabolic activity showed a rapid upward spike and growth resumed at a steady state reaching OD_{650} 162 at termination. CFU data also support the above trends (Figure 4.12a), there was an increase during the initial stages of culture, up to the point of feeding, followed by a slight lag then increasing up to 2.8×10^{12} cfu·mL^{-1} at induction. This increased only slightly to 6.1 $\times 10^{12}$ at 24 h, supporting the conclusion that during this period there was little growth, then decreased to 1.8×10^{12} at termination, the decrease was likely due to the culture having ceased to grow at ~33 h and hence likely to have entered stationary phase. At termination a DCW of 48.4g·L^{-1} was reached, corresponding to a Y_{xs} of 0.296 g·g^{-1}. Similarly to B1-3 the DCW/OD_{650} ratio was lower than expected at 0.3.

Plasmid retention (plasmid$^+$ and GFP$^+$ colonies) remained above 90% throughout although the percentage of GFP$^+$ colonies showed a slight decreasing trend after 7 h pre-induction (Figure 4.12b). This is supported by FCM data that showed the percentage of GFP$^+$ cells remained above 90% throughout, the initial increase from 91% to a peak value of 99% was due, as in Section 4.2.4, to initial samples containing a large proportion of non-fluorescent particulate matter due to the low initial cell numbers, decreasing as the cell count increased. As with the percentage of GFP$^+$ colonies, the proportion of GFP$^+$ cells showed a slight decrease during the fermentation, suggesting there may have been low levels of plasmid loss. The percentage of dead (PI$^+$) cells

remained below 10% except at induction where it reached a peak value of 10.5% (Figure 4.12b). Green fluorescence and forward scatter of GFP⁺ cells (Figure 4.13a) showed the expected initial trend of increasing followed by decreasing, as explained above. From 7 h pre-induction onward, both parameters showed increasing trends, FSC-A appeared to increase with growth rate, during periods of slower growth (2-18 h and 24-41 h post-induction) the rate of increase was lower. FL-1A on the other hand increased steadily throughout the remainder of the fermentation reaching a final peak value of 461,000, the highest value for a fermentation during this study. As FL1-A increased to a greater extent during the fermentation than FSC-A, this suggests that the increase in fluorescence was not simply a product of increasing cell volume with constant CheY::GFP content. Also, as FL1-A increased steadily during the periods where FSC-A increased only slightly and growth rate was depressed this suggests that it is during periods of slower growth that functional protein is principally accumulated.

SDS-PAGE analysis (Figure 4.13b) showed relatively high levels of RP production, the percentage of total cellular protein that was CheY::GFP increased throughout reaching a final peak value of 29% and the solubility of CheY::GFP showed an overall increase reaching a final peak value of 60%. Both values are almost equal to the highest values previously achieved (30% from B3 and 63% for A, respectively). Final yield of CheY::GFP was calculated at 10.5 g·L⁻¹, corresponding to a yield of 6.3 g·L⁻¹ soluble CheY::GFP.

The data above suggest that despite accidental heat/cold shock during the initial stages of production and likely metabolic deficiencies, reasonable amounts of biomass were produced (OD$_{650}$ >100) and high RPP productivity. Based on these observations it was decided that it would be beneficial to test this protocol further, with an increased phosphate concentration, termed attempt C3.
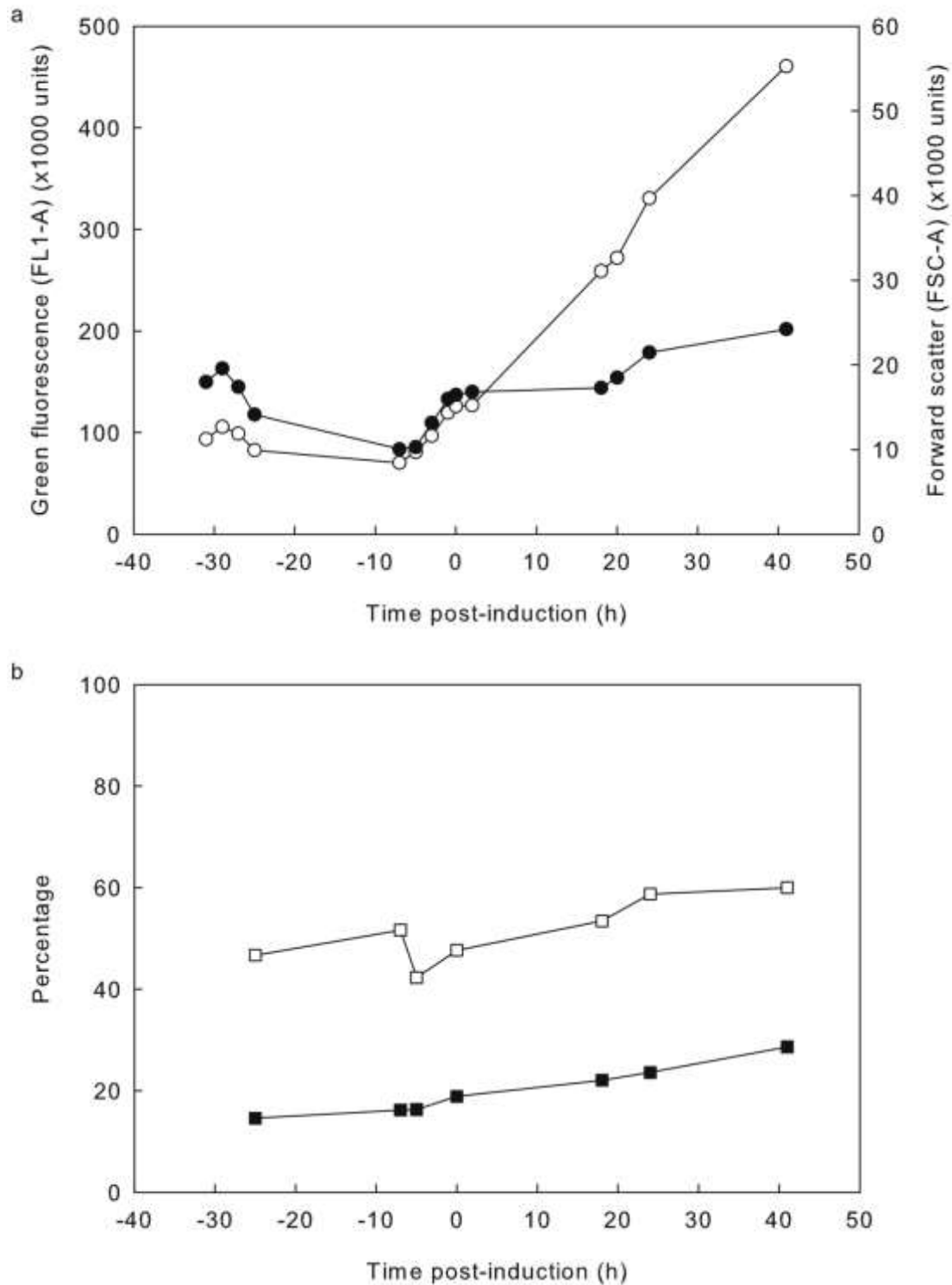
Figure 4.13: FCM and SDS-PAGE data from a protocol C2 fermentation
a) FCM data: Mean cellular green fluorescence (FL1-A) of GFP+ cells (○), mean cellular forward scatter (FSC-A) of GFP+ cells (●). b) SDS-PAGE data: Percentage of CheY::GFP that was soluble (□), percentage of total cellular protein that was CheY::GFP (■).

4.2.7.3. Attempt 3 (C3)

A fermenter was set up as protocol C2 except that the growth medium was supplemented with the phosphate-containing components of medium B (2 g·L^{-1} KH$_2$PO$_4$, 16.5 g·L^{-1} K$_2$HPO$_4$, 1.5 mL·L^{-1} Conc. H$_3$PO$_4$) and termed C3 (Table 4.2). Feeding began when on-line data suggested that batch-phase glycerol had been consumed (33.3 h post-inoculation). RPP was induced when the culture had reached OD$_{650}$ of ~80 (82.0 at 38.5 h post-inoculation). The fermentation was terminated when on-line data indicated growth had ceased (16.5 h post-induction).

OD$_{650}$ data (Figure 4.14a) showed an extended lag period until ~10 h pre-induction then a steady increase reaching 249 at 12 h post-induction. While the OD$_{650}$ data during this period did not suggest any major limitations to growth, the DOT/agitation and CDC/OXC did, showing a gradual decrease in metabolic activity between 2-12 h post-induction. As this media was already assumed to have been deficient in nitrogen and phosphate it was hypothesised that a trace element may have also become limiting and therefore a supplement of 3 mL SM6E trace elements solution and 2 mL *E. coli* sulphur free salts was added at 11.5 h post-induction, but this had no effect. It was found that the decrease in metabolic activity observed was due to the feed rate becoming limiting for growth as an increase in feed rate at 12 h post-induction caused a rapid decrease in the DOT and increases in the CDC and OXC that continued until termination. At termination the OD$_{650}$ had increased to a final peak value of 277 comparable to the 2 highest of this study (B2 at 297 and B3 at 288), the final DCW was 66.8 g·L^{-1}, corresponding to a Y_{xs} of 0.408 g·g^{-1}, comparable to the highest value achieved for this study (0.411 for B4) and a DCW/OD$_{650}$ ratio of 0.24. Once again, despite growth on glycerol a substantial volume of base was required to maintain pH.
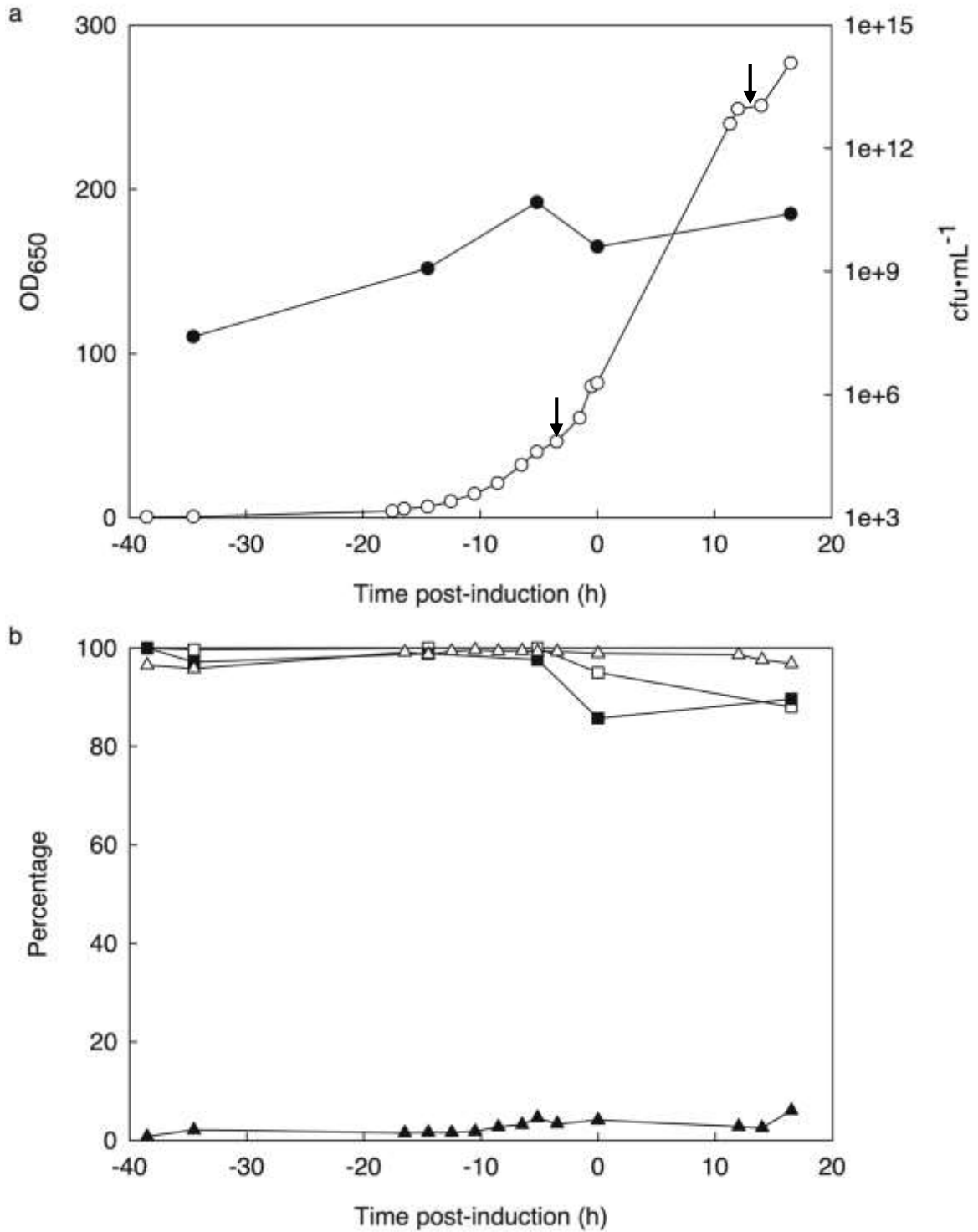
Figure 4.14: Culture growth, plasmid retention and viability data from a protocol C3 fermentation

a) Growth data: OD_{650} measurements ($\bigcirc$), CFU counts ($\bullet$). b) Plasmid retention and FCM data: Percentage colonies plasmid+ ($\square$), percentage colonies GFP+ ($\blacksquare$), percentage cells GFP+ (by FCM) ($\triangle$), percentage cells PI+ (dead) (by FCM) ($\blacktriangle$). Arrows indicate points at which (left to right) feed was turned on or rate of feeding increased.

Unlike C1 and C2, plasmid retention did decrease below 90% between induction and termination, although the percentage of GFP$^+$ cells remained above 95% throughout (Figure 4.14b). As CFU counts decreased between 6.5 h pre-induction and induction this is likely due to differential culturability in plasmid$^+$ and plasmid$^-$ cells. Although CFU data suggested a decrease in culturability the percentage of dead (PI$^+$) cells remained below 10% and hence it is likely that there was an increase in VBNC cells.

FCM analysis of GFP$^+$ cells (Figure 4.15a) showed an initial increase then decrease in both FSC-A and FL1-A, as described above. Following this, both parameters increased slowly up to induction then more rapidly. Peak fluorescence occurred at termination (221,000) and as fluorescence increased to a greater extent than forward scatter it is assumed that this was as due to CheY::GFP accumulation.

SDS-PAGE data (Figure 4.15b) showed the percentage of cellular CheY::GFP increased to a peak (29%) at 12 h post-induction then decreased slightly to 27% at termination. CheY::GFP solubility increased to a peak (49%) at induction, decreased to 40% at 12 h post-induction then increased slightly to 42% at termination. Final yield of CheY::GFP was calculated at 14.2 g·L^{-1}, corresponding to a yield of 6.0 g·L^{-1} soluble CheY::GFP.

It is apparent from these data that the metabolic deficiencies of C1 and C2 were remedied as C3 obtained both a similar biomass to B2 and B3 and a similar Y_{xs} value. The effects on RPP productivity were mixed, as both fermentations generated CheY::GFP to a similar percentage of total cellular protein the higher biomass of C3 resulted in a higher total yield of CheY::GFP but the higher solubility of C2 resulted in similar yields of soluble CheY::GFP.
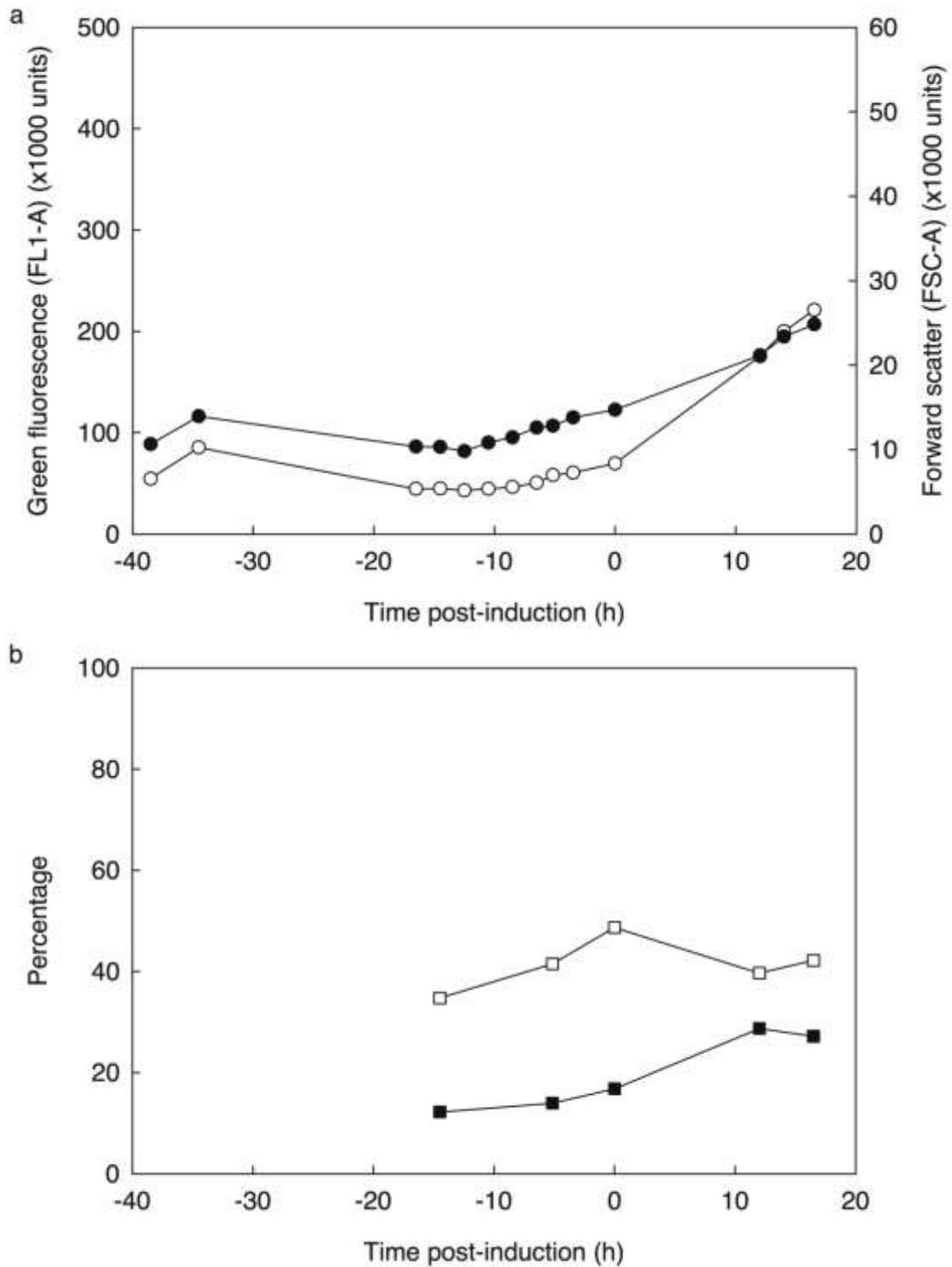
Figure 4.15: FCM and SDS-PAGE data from a protocol C3 fermentation
a) FCM data: Mean cellular green fluorescence (FL1-A) of GFP+ cells (○), mean cellular forward scatter (FSC-A) of GFP+ cells (●). b) SDS-PAGE data: Percentage of CheY::GFP that was soluble (□), percentage of total cellular protein that was CheY::GFP (■).

4.2.8. A reassessment of organic nitrogen and phosphorus content in fermentation media and its impact on 'improved' RPP

4.2.8.1. Limitation of growth due to due to deficiencies in growth media

The most striking and unexpected results from this series of experiments was that the application of improved conditions to the fermentation media of Humphreys *et al.* (2002) did not result initially in high cell density, apparently due to a combination of nitrogen and phosphate limitation. To understand this further, analyses of phosphorus and nitrogen content for all fermentation protocols used in this study were made. As it was not possible to determine the concentration of NH_4^+ or PO_4^{3-} biochemically, theoretical amounts of nitrogen and phosphorus in the fermenters were calculated from media recipes (when complex components were used values were calculated using published typical analyses (Becton, Dickinson and Co., 2006)) and the amounts incorporated into biomass calculated from DCW or OD_{650} (when DCW was not available, assuming OD_{650} 1 = 0.4 $g·L^{-1}$ DCW) data and the typical elemental composition of *E. coli* published in Doran (2013), are summarised in Table 4.3.

During C1, growth limitation was hypothesised as being due to N-starvation but supplementation of $(NH_4)_2SO_4$ did not cause further growth. Analysis of phosphorus content provided an explanation: at termination the predicted amount of phosphorus incorporated into biomass had already exceeded that available in the starting medium and therefore further growth would be unlikely. The discrepancy between these values is interpreted as being due to fluctuations in the relationship between OD_{650} and DCW used and in cellular phosphorus content as a result of strain and RPP. In addition to a discrepancy in phosphorus content, the predicted nitrogen content was higher than the amount in the growth medium but this is more simply explained as the 10% ammonia added as base is likely to have been incorporated. Although it should theoretically be

Table 4.3: Summary of fermentation nitrogen and phosphorus compositions

Fermentation		Amount in growth medium (g)		Amount in biomass (g)
A	N	22.2		10.525
	P	1.6		2.25
B1	N	7.72		21.8
	P	7.28		4.66
		Base medium	Supplemented	
C1	N	1.66	4.84	10.22[1]
	P	1.3	1.3	2.18[1]
C2	N	4.46	7.64	12.32
	P	1.3	16.48	2.64
C3	N	4.46		18.74
	P	6.66		4.16

[1] – Calculated from OD_{650} (assuming OD_{650} 1 = 0.4 $g \cdot L^{-1}$ DCW) as final DCW not determined for C1

possible to calculate the amount of nitrogen added from the volume of base recorded, this is complicated by ammonia being gaseous under ambient conditions and hence readily coming out of solution. It was observed that the base addition tubing would frequently contain large pockets of gas (presumably NH_3), reducing the accuracy of volume monitoring and decreasing the concentration in solution, therefore these data were omitted from calculations. Humphreys *et al.* (2002) stated that fermentations were harvested at OD 90-100, this is consistent with the peak OD_{650} achieved by C1 (91.2). It is therefore obvious from these data that the unmodified growth is suited specifically to the growth profile and RPP protocol it is used for; a short burst of RPP very late in the fermentation and final ODs of 90-100.

It has previously been noted that for both protocols B and C the volume of base added to the cultures was excessive, as the use of glycerol should limit acid formation. However, analysis of nitrogen content for protocols B and C showed that in all cases the predicted amount of biomass-incorporated nitrogen far exceeded that provided in the growth media. While variation from the typical values in Doran (2013) may account for some of this discrepancy it is unlikely to account for all of it. Protocol A, that contained a greater concentration of nitrogen in the growth medium than in the final biomass, required considerably less base. It is therefore hypothesised that base addition became, effectively, a nitrogen feed. This is also supported by base addition data from C2; following supplementation of $(HN_4)_2SO_2$ at 2 h post-induction the base addition rate was substantially lower (Figure 7.10). There does not appear to be a direct mechanism by which nitrogen starvation would result in acid formation but the following is hypothesised: The principal mechanism by which ammonia is assimilated in *E. coli*, is via glutamate dehydrogenase (GDH) and glutamate synthase (GOGAT) (Tyler, 1978; Umbarger, 1978). Starvation of ammonia could result in accumulation of precursors

such as 2-oxoglutarate and glutamate, both of which would be acidic under physiological pH. It is also established that *E. coli* BL21 and derivatives have differences in carbon metabolism (TCA cycle, glyoxalate shunt and acetyl-CoA synthetase) that may also have an effect (Noronha *et al.*, 2000; Phue & Shiloach, 2004). It may therefore be advised for future work to increase the amount of bioavailable nitrogen in both protocols B and C. It may also be of interest to investigate further the effects on pH of N-starvation, possibly through metabolomic analysis.

It was observed in Section 4.2.1 that the final biomass of protocol A was lower than would be expected based on the medium composition, but compositional analysis provides a potential explanation: Similarly to C1 the predicted final concentration of biomass-incorporated phosphorus in A exceeded the concentration in the growth medium and therefore it can be concluded that phosphorus was likely exhausted by the end of the fermentation, which would be expected to result in growth arrest. In addition, if the complex components used in A(1) and A(2) contained higher proportions of phosphorus, this would explain the higher biomasses achieved. It may therefore be useful for future work to run further fermentations using protocol A supplemented with additional phosphates to investigate whether this is the case.

4.2.8.2. The effects of nutrient limitation on RPP

In addition to effects on growth and biomass formation, growth medium formulation is known to affect RPP productivity. Complex growth media have been known to increase protein solubility (Moore *et al.*, 1993; Fahnert, 2012), potentially the cause of increased CheY::GFP solubility in protocol A. Wang *et al.* (2009b) observed that RP yields can be increased by the addition of ammonia, even for complex media where there should be sufficient organic nitrogen available. As nitrogen is a necessary component of amino acids the underpinning logic is obvious, although this may not be universally true;

Schroeckh *et al.* (1999) successfully used an RPP-induction system based on the *E. coli glnA* promoter to induce at high cell density (OD$_{550}$ 120) only when the concentration of ammonia became limiting (<1 mM). Cultures were induced by replacement of ammonia as base addition by KOH. Control of phosphate concentration can also influence RPP productivity; Huber *et al.* (2011) demonstrated that the total cellular amount of RP for a T7 polymerase expression system increased under conditions where cells were limited for phosphate but not for carbon source. Phosphate is necessary for cell growth (e.g. for DNA replication), but not directly for protein synthesis (unlike nitrogen and carbon) and phosphate-limited cells can remain metabolically active. Phosphate limitation can therefore stall cell growth, allowing redirection of cellular resources towards protein synthesis. Decreased growth rate and increased RP synthesis rates can therefore cause an overall increase in cellular productivity. This is also likely to be the reason for successful use of RPP systems using the promoter for alkaline phosphatase (PhoA) that is activated under phosphate starvation by the 2 component regulator PhoR/PhoB (Lübke *et al.*, 1995; Makino *et al.*, 1986).

An examination of fermentation productivity and nitrogen/phosphorus content through this study suggest that the above effects may have had a role, unfortunately no definitive conclusions may be made. For example, A showed high solubility but low RP yields, was likely phosphate and therefore growth-limited but in light of Wang *et al.* (2009b) may have also been ammonia-limited despite large quantities of organic nitrogen available. For protocols B and C the initial media were incontrovertibly nitrogen limited, but the addition of ammonia throughout the fermentations in base addition prevented any accurate determination of the extent of N-limitation. It would be potentially worthwhile to investigate in greater detail the effects of these limitations on improved RPP, for

example by repeating the fermentation protocols with only a single limiting factor by providing a non-growth limiting excess or additional feed of the others.

While it is not currently possible to ascertain specific effects of medium composition on RPP, a comparison of C2 and C3 suggests an effect of growth rate. During the 2 periods of growth limitation in C2 (2-18 h post-induction during phosphate-limitation and 24-41 h as growth ceased due to the end of feeding) FL1-A increased at a faster rate than that of FSC-A. This was interpreted as an increase in cellular content of folded, mature CheY::GFP during periods of decreased growth rate. In addition, overall growth rate between induction and termination was lower for C2 than C3, but measures of RPP productivity (fluorescence, total CheY::GFP content, CheY::GFP solubility) were consistently higher for C2. These data overall suggest that decreasing growth rates by growth medium manipulation favour both CheY::GFP accumulation and quality.

4.2.9. Analysis of fermentation diagnostic techniques

4.2.9.1. Using the DCW/OD$_{650}$ ratio as a culture diagnostic

It was observed that the ratio between terminal culture DCW and OD$_{650}$ for A was higher than the typically cited value for *E. coli* of 0.4 and lower for B1-C3. This ratio has been shown to be affected by the formation of inclusion bodies i.e. that IB-containing cultures show a lower ratio (Hwang & Feldberg, 1990). This relationship is logical as IBs are highly diffractive particles and hence IB-containing cells will show increased light scatter. To determine whether this was true for the system used here terminal DCW/OD$_{650}$ ratios for A, B1-3 and C2&3 were correlated against the percentage solubility of CheY::GFP. The correlation was found to not be statistically significant (r = 0.611, p =0.198). However the DCW/OD$_{650}$ ratio was found to have significant correlations with 2 other observed parameters; the mean FSC-A of GFP[+] cells (r = -0.956,

p >0.01) and the percentage of total protein that was CheY::GFP (r = -0.877, p =0.021). The former relationship is not unexpected as FSC-A is itself a measure of light scattering and has been used previously to predict biomass (Section 1.5.4.1). The latter relationship is more interesting, suggesting that for the system used here it is the total cellular content of CheY::GFP that affects culture turbidity as opposed to its solubility state. This is still in itself a potentially useful diagnostic as DCW measurements typically require less time than SDS-PAGE analysis.

4.2.9.2. The use of typical cellular composition values for fermentation yield calculations

As demonstrated above the relationship between DCW and OD_{650} data during this study varied substantially from the typically cited ratio of 0.4, likely dependent on cellular RP content. DCW is often predicted from OD_{650} data as opposed to directly measuring DCW (as was done by Sevastsyanovich *et al.* (2009)), but predicted biomass yields may be highly inaccurate; from the data here presented OD_{650} 100 could correspond to a DCW of 24-47 g·L^{-1} and therefore where possible DCW should be measured directly.

In addition to predicting biomass yields from turbidometric data, RP yields can also be predicted from biomass data and typical cellular protein composition, as was done by Sevastsyanovich *et al.* (2009) and Alfasi (2010) assuming *E. coli* contains 70% protein and in the previously published version of the data here presented (Wyre & Overton, 2014b) assuming *E. coli* contains 60% protein. Subsequent reanalysis of the fermentations above to determine total protein concentrations derived using BCA analysis showed that there was a range of protein contents achieved from 71-86%, showing that the 60% figure previously used to predict yields on this data and in fact the 70% figure used by Sevastsyanovich *et al.* (2009) and Alfasi (2010) were in both cases substantial underestimates of the actual values. The resultant yield data were therefore

also all higher than the predicted values. Although the total cellular protein content varied somewhat this did not appear to be sufficient to alter the trends in the resultant yield data from those predicted previously (Figure 4.16). These observations suggest that the use of typical protein content figures to predict RP yield data may be useful as an interim measure to compare on-going work but can introduce substantial potential for error when stating final yields and therefore biochemically determining protein content would be highly advised.

4.3. Conclusions

Preliminary analysis of the RPP-fermentation protocol by Sevastsyanovich *et al.* (2009) incorporating the novel application of a physiological stress-minimisation paradigm for 'improved' RPP identified a number of potential limitations for its application to industrial practices, principally regarding the complex fermentation medium used. Attempts to replicate this work (A) appeared to confirm the preliminary analysis that the heavy reliance on complex medium components with inherent variation in content place limitations on the wider applicability of this protocol. Despite evidence that high-quality, soluble CheY::GFP was produced and cell physiology remained relatively unstressed, as would be expected from successful use of the stress-minimisation method, biomass and RP accumulation appeared to be limited compared to published instances, severely limiting product yields. Although it is likely that these effects may be relieved by supplementation of the growth medium and feed with, for example, alternate amino acids, phosphate and ammonium, determining an optimal strategy would be time consuming, costly and must be repeated for any change to batch or supplier, therefore strongly precluding industrial utility of this protocol.
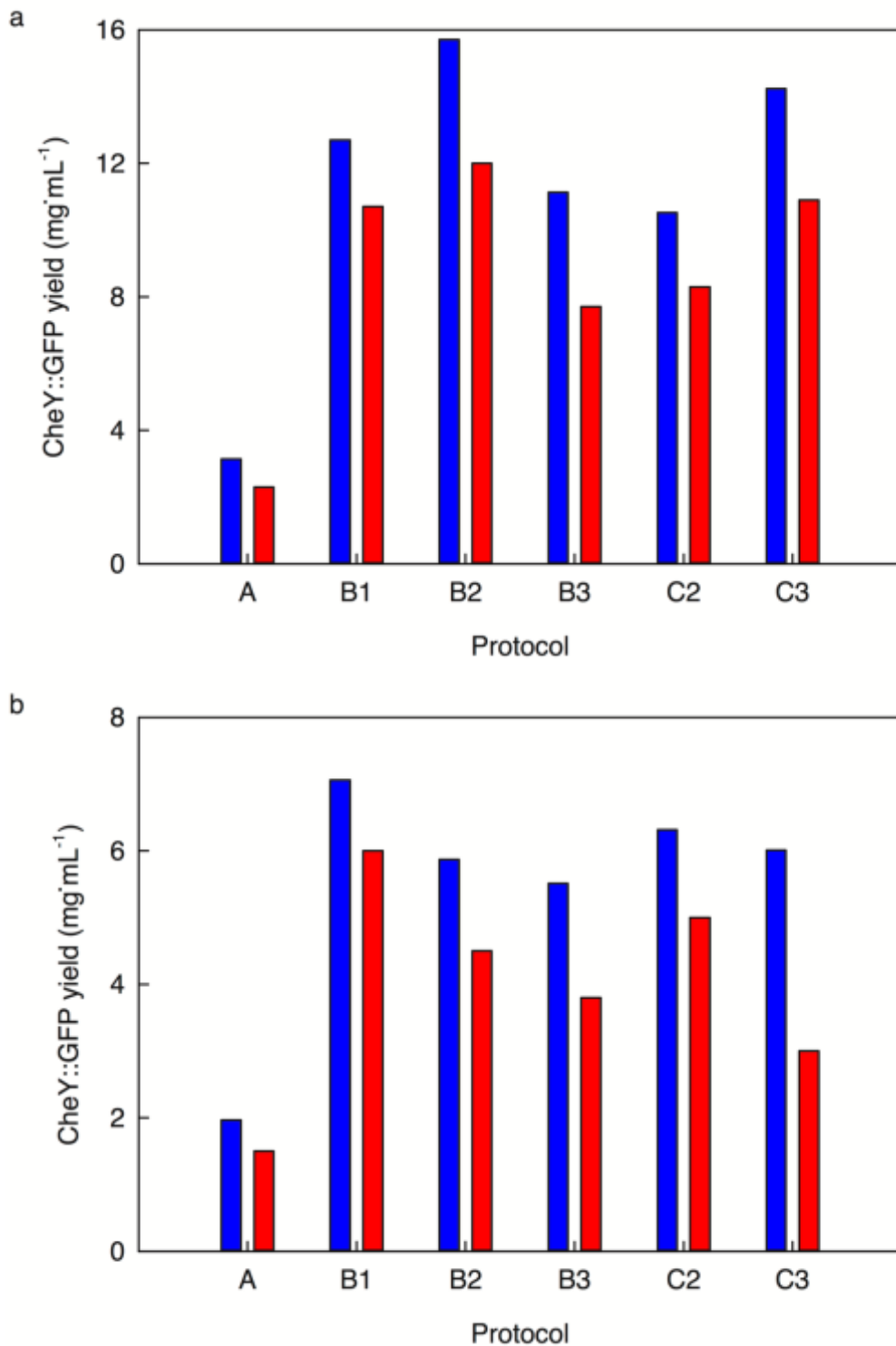
Figure 4.16: Comparison of BCA assay-derived and predicted end-point CheY::GFP yield data

a) Yields of total CheY::GFP at termination for fermentation protocols A, B1-3, C2 & C3 calculated using experimentally-derived total protein content by BCA assay (blue) and estimated from biomass data using typical published values for *E. coli* protein content (red). b) As (a) for yields of soluble CheY::GFP.

Stress-minimisation was found to be readily applicable to an industrially-derived fermentation protocol using semi-defined growth medium, increasing reproducibility and homogeneity of the culture, biomass yields and cell productivity over the unmodified protocol. In addition, it was shown that the improved, industrially derived fermentation protocol (B1) provides an attractive alternative to the original improved protocol (A). B1 showed higher yields of both total and soluble CheY::GFP and improved reproducibility, while being more favourable to industrial applications due to a decreased run time, lower cost growth medium, the elimination of animal-derived medium components and an overall reduction in the amount of complex medium components used.

An investigation into the specific effects of the principal differences between protocols A and B revealed that the largest effects came from changing the growth medium, resulting in increases in biomass generation and in the yields of both total and soluble CheY::GFP, although also resulting in a decrease in CheY::GFP solubility (insufficient to decrease overall soluble yields) and an unexpected culture acidity. Despite the simplicity of using glycerol as carbon source it appeared that there might be some benefits to retaining glucose as the efficiency of conversion to biomass was higher and media costs were lower. Despite evidence that it does under standard conditions, changing the point of induction was shown to have little overall effect on improved RPP although earlier induction (and hence a longer production window) appeared to favour culture heterogeneity. This suggests that, as would be expected, the decreased rates of RPP and increased productive period of cells under improved conditions, drastically alters the productivity window of improved RPP processes over standard.

In all cases modified B-type fermentations (B1-3) gained higher yields of biomass, total and soluble CheY::GFP over A, A(1) and A(2), in shorter time periods and using lower

cost growth medium and in a manner substantially more applicable to current industrial practices. It can therefore be concluded that the application of stress-minimisation to protocol B was successful, that the methods detailed here could be readily applied to wider use in industrial RPP-bioprocesses and that this work represents a worthwhile and novel development for the stress-minimisation methodology.

This chapter also demonstrated the development of an industrially-derived HCDC fermentation using improved RPP conditions and defined growth medium that would fully eliminate reliance on inherently variable complex medium components. Following adaptation of the growth medium to allow greater biomass formation by the supplementation of ammonium and phosphate, high yields of biomass, total and soluble CheY::GFP were generated, comparable to the most productive instances from A-B3, albeit in separate fermentations. It is concluded that with further development it may be possible to devise a protocol with comparable results to B1-3 and hence be a viable alternative.

During these fermentations issues regarding limitation of key nutrients were observed, prompting a reassessment of all 3 fermentation protocols used. It was found that all 3 protocols became limited to varying extents for both phosphate and nitrogen, potentially influencing both growth and RPP productivity and to which, further investigation would be advised. A possible effect of growth rate, as controlled by growth medium content on RPP productivity was also observed and again further investigation is advised.

Chapter 5: General Conclusions & Further Work

5.1. General conclusions

The production of recombinant proteins in *E. coli* is a widely used technique in modern bioprocessing particularly in the biopharmaceutical industry requiring production at industrial scale. For an RPP process to be effective at industrial scale it must attempt to maximise product yields for minimal costs in a reliable, reproducible manner while obeying all necessary regulations. Optimisation of an industrial-sale RPP process will often involve a compromise between multiple factors including biomass generation, product yield, product quality and process costs and any proposed modification must provide an aggregate improvement over all factors for it to be effective. For example, an improvement to product yield will not be effective if it causes a sufficiently reduced biomass yield (a phenomenon that was noted by (Want *et al.*, 2009)) or increased process costs. Similarly any modification to the analysis of an industrial process must present a valid cost-benefit for the overall process for example in replacing a more costly existing method or increasing reliability. These concerns are substantially different from those involved in RPP at the laboratory-scale that will typically require small single batches of product without such rigorous economic pressures.

This work has presented the development of both analytical and production methods for RPP in *E. coli* aimed particularly at industrial applications showing benefits to the analysis of the currently little-studied initial period of RPP culture growth and in the use of production conditions that have to this point not shown widespread adoption.

5.1.1. Novel Applications of Flow Cytometry for Bioprocess Monitoring and Control

Flow cytometry is a technique of great utility for the monitoring of RPP bioprocesses at the industrial scale with the potential for substantial further exploitation (Section 1.5). FCM is a rapid technique that requires very low sample volumes and can assess

numerous cellular parameters (Shapiro, 2005). In addition, recent innovations have reduced both the costs and complexity of both instrumentation and operation (Section 1.5.5) and raised the potential for automated analysis (Gatza *et al.*, 2012). Development of novel FCM applications for RPP monitoring will increase the potential utility of the technique and hence increase the likelihood of its routine use in industry.

The amyloidophilic fluorescent dye Congo red was used to identify and enumerate by FCM cells containing 'classical' IBs with high amyloid content (Section 3.2). The use of FCM and amyloid staining dyes has many benefits for monitoring RPP. When producing proteins solubly the ability to detect IB formation would allow the determination of sub-optimal production conditions. Additionally when producing proteins as IBs for refolding, as is often done industrially (Section 1.2.1.2), it may also be possible to use CR staining to monitor product formation or to enumerate non-productive cells. One limitation of CR-staining for IB detection is that it is dependent on the amyloid character of 'classical' IBs, as shown by CheY::GFP IBs formed at 25°C failing to stain and therefore its efficacy will be case-dependent. It may however be possible to exploit this phenomenon and use amyloid detection to probe IB structure *in vivo* during deliberate IB production. When producing IBs for refolding increased classical character is likely to increase resistance to solubilisation (de Groot & Ventura, 2006; Upadhyay *et al.*, 2012) and when producing IBs as biologically active nanoparticles (Sans *et al.*, 2012; Villaverde *et al.*, 2012) amyloid character will be indicative of reduced biologically activity (García-Fruitós *et al.*, 2007a). In both cases amyloid formation will be detrimental and therefore the ability to rapidly detect it is of great potential.

The initial stages of culture for RPP, particularly on agar plates, are little studied (Wyre & Overton, 2014a). FCM is ideally suited to analysis of these cultures due to the low cell numbers required and multivariate analysis available. This work presented what is

thought to be the first analysis by FCM of RPP bacteria on agar plates (Sections 3.3, 3.4). FCM screening of *E. coli* BL21* transformants expressing CheY::GFP on agar plates allowed the identification of abnormal, less productive cells at the earliest possible stage of culture. This would present a substantial time-saving and hence be of use during industrial strain development and in the production of cell banks. During storage agar plate cultures were also shown to progressively generate a population of cells with intermediate fluorescence and low culturability (population P2), shown by CR staining to be due to the stochastic formation of amyloid deposits within existing non-classical IBs that were likely cytotoxic (González-Montalbán *et al.*, 2005; Lindner *et al.*, 2008; Kagan *et al.*, 2012). As population P2 cells were of low culturability, increasing P2 content reduced initial growth rates in subsequent liquid culture. Determination of P2 content prior to inoculation will therefore allow for more accurate prediction of subsequent growth rates and hence likely be beneficial for process reproducibility. It is likely that this phenomenon will occur in other (non-GFP fusion) proteins and therefore FCM analysis of agar plate cultures prior to inoculation particularly with staining for amyloid may be of wide application for RPP processes.

5.1.2. Application of 'Improved' Physiological Stress-Minimised Production Conditions to Industrially Derived Fed-Batch RPP Protocols

The development of the improved RPP protocol by Sevastsyanovich *et al.* (2009) represents a significant development in the theory behind RPP processes. The fermentation protocol it was applied to however, showed substantial limitations for application in industry, predominantly due to the heavy reliance on complex growth medium components that can incur regulatory concerns and batch variation of which

was shown here to negatively affect culture growth using the *E. coli* BL21*/CheY::GFP model RPP system (Section 2.10.1).

Replication of the original improved protocol (A) showed high levels of plasmid retention, culture viability and CheY::GFP solubility indicative of successful stress-minimised RPP but identified further limitations for wider applications (Section 4.2.1). Both total biomass accumulation and the percentage of total protein that was CheY::GFP were substantially lower than previously published data (Sevastsyanovich *et al.*, 2009; Alfasi, 2010) demonstrating reproducibility issues that would substantially limit industrial use. This was ascribed to the aforementioned reliance on complex medium components, that despite preliminary attempts to optimise the sources and batches of complex components (Section 2.10.1) this was not completely effective and that there remained compositional inadequacies for high cell density production, for example in the amino acid distributions or phosphate content.

Stress-minimised growth conditions were shown to be readily applicable to an industrially-derived RPP fermentation protocol using semi-defined growth medium (B) with improved reproducibility and increases in yields of biomass and RP (both soluble and total) over the original stress minimised RPP protocol and the unmodified industrially-derived protocol (Section 4.2.3). It can be therefore concluded that the physiological stress-minimising paradigm possesses substantial potential for industrial exploitation and that the novel protocol developed presents this concept in a format more easily adapted towards existing practices.

Having established that stress-minimisation can be applied to industrially-relevant RPP protocols the effects of 3 specific differences between protocols A & B were examined in further detail; growth medium, induction point and carbon source. The improvements in biomass and RP generation observed between A and B-derivatives were principally due

to changing the growth medium (Section 4.2.6). This is consistent with preliminary analysis of protocol A that identified its reliance on large amounts of complex medium components as the principal limitation for industrial application. The sole benefit of using a complex medium was an increase in RP solubility, as has already been established (Moore *et al.*, 1993). However increases in both total RP and biomass yields using medium B resulted in a higher yield of soluble CheY:GFP, despite lower percentage solubility. Overall the data strongly supports the prevailing practice among industry that limitation of complex medium components is beneficial for RPP process efficacy and that this practice is equally valid for stress-minimised RPP.

Use of both glycerol and glucose was shown to be effective for stress-minimised RPP (Section 4.2.5). While glycerol allowed higher amounts of biomass to be formed and was simpler to use, glucose allowed greater efficiency of conversion to biomass and was cheaper than glycerol. This suggests that, in line with current practices (Section 1.4.3.2), both glucose and glycerol can be used as carbon source for *E. coli* fermentation using stress-minimised RPP and that the choice will be case-dependent.

There was little overall effect of changing the induction point apart from a slight increase in culture heterogeneity following early induction (Section 4.2.4), unlike studies on standard RPP that suggest an inverse relationship between biomass and RPP yields dependent on induction point (Want *et al.*, 2009). As RPP induction under stress-minimised conditions does not substantially reduce viability of the culture, changing the point of induction predominantly changed the length of induction period. From this it can be concluded that once formed, CheY::GFP is relatively stable and that the culture has a broad productivity window. The increase in heterogeneity observed following early induction (i.e. a longer induced period) would be detrimental for industrial use,

therefore later induction (i.e. minimisation of the induced period) will likely be beneficial for industrial applications of stress-minimised RPP.

Stress-minimised production conditions were also applied to a second industry-derived fermentation protocol (C) using a chemically-defined growth medium in order to completely eliminate the need for complex medium components (Section 4.2.7). This was met with only limited success. The growth medium did not contain sufficient nitrogen and phosphate to support cell densities OD_{650} >100. On supplementation of nitrogen (i.e. phosphate limited) high RPP productivity but low biomass was achieved and on supplementation of both nitrogen and phosphate high biomass was achieved but with poorer RPP productivity. From these observations it was concluded that with further development this protocol would likely be as effective as the improved B-derivatives. It was also hypothesised that phosphate limitation may be beneficial for RPP accumulation by limiting cell growth but not protein synthesis.

Prompted by the discovery of nitrogen and phosphate limitations in medium C nitrogen and phosphate contents were assessed for all fermentations and a number of observations were made. Medium A was found to be phosphate limited and this may have contributed to the low biomass accumulation and high CheY::GFP solubility observed.

It was demonstrated that both media B and C were nitrogen limited and that this may have caused unexpected increases in culture acidity by accumulation of intermediates in the ammonium assimilation pathways (Tyler, 1978; Umbarger, 1978). It is hypothesised that this is likely linked to the atypicalities of central carbon metabolism in derivatives of *E. coli* B (Noronha *et al.*, 2000; Phue & Shiloach, 2004), as these strains are often used industrially to limit acetate formation in fermentations (Overton, 2014) this is likely a highly pertinent avenue of further study.

5.1.3. Summary

The aims of this work were first 'to investigate and develop additional methods by which FCM can be used for the monitoring and analysis of RPP cultures in *E. coli*' and second 'to optimise protocols for RPP in *E. coli* using the physiological stress-minimising improved protocol of Sevastsyanovich *et al.* (2009), specifically regarding its application to industrially-relevant fermentation conditions'. Chapter 3 demonstrated the development of a novel FCM staining protocol for RPP monitoring and applied FCM monitoring to RPP process stages that it appears to have not previously been applied to. Chapter 4 demonstrated the successful application of improved RPP conditions to an industry-derived fermentation protocol and partial success for a second, in addition to exploring the effects of individual process variables and of growth medium composition on 'improved' RPP. From this it can be concluded overall that the aims of this work were successfully met.

5.2. Further work

During this work many opportunities for future work were identified, including the following:

5.2.1. Further development of FCM protocols

The efficacy of the CR staining protocol developed here is likely to be case-dependent and therefore validation with further RP products is necessary, particularly those of industrial interest. Proteins produced both solubly and as IBs would be equally valid targets for this, the latter to determine whether increasing classical IB (i.e. amyloid) character may affect refolding efficiency. In addition to the general CR-staining protocol, the use of CR staining to identify less-culturable amyloid[+] cells in agar plate cultures would benefit from validation with additional RP products. Industrial RPP frequently

uses cryogenically stored cell banks to inoculate fermentations as opposed to agar plate cultures therefore it may be of use to determine whether the aging-associated amyloid formation observed here also occurs in cell banks.

5.2.2. Further optimisation of fermentation protocols.

This study demonstrated deficiencies in nitrogen or phosphate for all fermentation media used with possible effects on both RPP productivity and biomass generation. It would be of great use to determine the specific effects of these deficiencies and to determine optimal formulations or supplementation regimes for each medium.

RP and biomass yield limitations in A may also have been attributable to sub-optimal amino acid distributions in the complex components. Redevelopment of an amino acid supplementation regime would be time-consuming and costly but as Wang *et al.* (2009b) suggested complex media may still show N-limitation and hence benefit from ammonium supplementation. It may therefore be of interest to investigate this further as a more generic optimisation method for complex media-based protocols.

Nitrogen limitation in media B and C was thought to be the cause of high culture acidity. It may be of use to determine whether increased nitrogen content decreases acid production and hence base addition and if so, metabolomic analysis of *E. coli* BL21* to determine the biochemical cause would be of great interest. As this phenomenon may potentially be linked to the metabolic atypicality of B-derivative strains (Noronha *et al.*, 2000; Phue & Shiloach, 2004) comparison to a K12-derivative strain would also be of use.

Protocols A and C2 were shown to be phosphate limited and this appeared to cause lower biomass yields but increased RP solubility, likely due to limiting growth rates. Phosphate supplementation increased biomass generation in C3, therefore this may be a

strategy to increase biomass in A. It may also be worthwhile to investigate further whether growth rate control by phosphate limitation can improve CheY::GFP production as hypothesised, possibly by decreasing the initial concentration of phosphates in media B and feeding limiting concentrations during the induced period.

Moving the point of induction later in the culture (from OD_{650} ~0.5 to 40-50) resulted in a more homogeneous culture and hence may likely be beneficial. It would possibly be of interest to determine the latest point in a fermentation that would allow for satisfactory levels of RP accumulation and hence define the limits of the productivity window.

 A major limitation for the work presented here is that it was developed using a single model protein. To further demonstrate the utility of the protocols here developed they must be validated by the use of additional proteins, particularly those of industrial relevance such as antibody fragments and hormones (see Table 1.1). Proteins that require translocation and PTM would also be useful for further work as stress-minimised RPP can prevent overload of the export and modification pathways and hence improve production (Sevastsyanovich *et al.*, 2009) and industrial products frequently require export and modification.

The experimental methodology used in this study for fermentation protocol development changed a single factor at a time, it is acknowledged that this approach has since been superseded by more systematic methods (Bora *et al.*, 2012; Papneophytou & Kontopidis, 2014) (Section 1.4.5). This section has outlined numerous factors for further development that would require assessment at a range of values and therefore it would be recommended when investigating these factors to adopt a DoE methodology.

Finally, the fermentations in this study were substantially lower in volume (≤2.5 L) than those used industrially, therefore it would be logical to determine performance of the protocols developed here during scale-up, at least to pilot-scale.

5.2.3. Further investigations into the effects of oxygen-limitation on recombinant GFP production

Although FPs and FP-fusions are not RP targets of particular industrial relevance, they are often exploited during research due to the relationship between fluorescence and RP synthesis (Vizcaino-Caston *et al.*, 2012) (Section 1.6.5). Observations made during this work suggested that the oxygen limitation typical of high cell density fermentation may affect the use of fluorescent model proteins during RPP research due to the requirement of O_2 for maturation of the *av*GFP fluorophore (Figure 1.7a).

Preliminary investigations into the effects of oxygen limitation on CheY::GFP fluorescence proved inconclusive. While one test determined that insufficient imCheY::GFP had accumulated to affect FCM fluorescence measurements (Section 2.10.2.2) an earlier observation showed substantial increase in fluorimetry readings following culture reoxygenation (summarised in Appendix 2, Chapter 8). Further investigation in this phenomenon would be of interest, for example into the effects of varying the length and extent of oxygen limitation.

Chapter 6:   References

Abrahmsén, L., Moks, T., Nilsson, B. & Uhlén, M. (1986). Secretion of heterologous gene products to the culture medium of *Escherichia coli*. *Nucleic Acids Res* **14**, 7487–7500.

Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics Int* **11**, 36–43.

Ades, S. E. (2004). Control of the alternative sigma factor sigmaE in *Escherichia coli*. *Curr Opin Microbiol* **7**, 157–162.

Adrio, J.-L. L. & Demain, A. L. (2010). Recombinant organisms for production of industrial products. *Bioeng Bugs* **1**, 116–131.

Al-Fageeh, M. B. & Marchant, R. J. (2006). The cold-shock response in cultured mammalian cells: Harnessing the response for the improvement of recombinant protein production. *Biotechnol Bioeng* **93**, 829-835.

Alba, B. M. & Gross, C. A. (2004). Regulation of the *Escherichia coli* sigma-dependent envelope stress response. *Mol Microbiol* **52**, 613–619.

Alfasi, S. N. (2010). *Physiological aspects underpinning recombinant protein production in Escherichia coli.* PhD Thesis, University of Birmingham.

Alfasi, S., Sevastsyanovich, Y., Zaffaroni, L., Griffiths, L., Hall, R. & Cole, J. (2011). Use of GFP fusions for the isolation of *Escherichia coli* strains for improved production of different target recombinant proteins. *J Biotechnol* **156**, 11–21.

Amann, E., Brosius, J. & Ptashne, M. (1983). Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in *Escherichia coli*. *Gene* **25**, 167–178.

Andersen, J. B., Sternberg, C., Poulsen, L. K., Bjorn, S. P., Givskov, M. & Molin, S. (1998). New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl Environ Microbiol* **64**, 2240–2246.

Andersson, L., Yang, S., Neubauer, P. & Enfors, S. O. (1996). Impact of plasmid presence and induction on cellular responses in fed batch cultures of *Escherichia coli*. *J Biotechnol* **46**, 255–263.

Aristidou, A. A., San, K. Y. & Bennett, G. N. (1999). Improvement of biomass yield and recombinant gene expression in *Escherichia coli* by using fructose as the primary carbon source. *Biotechnol Prog* **15**, 140–145.

Arsène, F., Tomoyasu, T. & Bukau, B. (2000). The heat shock response of *Escherichia coli*. *Int J Food Microbiol* **55**, 3–9.

Assenberg, R., Wan, P. T., Geisse, S. & Mayr, L. M. (2013). Advances in recombinant protein expression for use in pharmaceutical research. *Curr Opin Struct Biol* **23**, 393–402.

Bahl, M. I., Sørensen, S. J. & Hestbjerg Hansen, L. (2004). Quantification of plasmid loss in *Escherichia coli* cells by use of flow cytometry. *FEMS Microbiol Lett* **232**, 45–49.

Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol* **10**, 411–421.

Baneyx, F. & Mujacic, M. (2004). Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol* **22**, 1399–1408.

Barnhart, M. M. & Chapman, M. R. (2006). Curli biogenesis and function. *Annu Rev Microbiol* **60**, 131–147.

Battesti, A., Majdalani, N. & Gottesman, S. (2011). The RpoS-mediated general stress response in *Escherichia coli*. *Annu Rev Microbiol* **65**, 189–213.

Becton, Dickinson & Company (2006). *BD Bionutrients™ Technical Manual.* 3rd ed.

Bentley, W. E. & Kompala, D. S. (1990). Optimal induction of protein synthesis in recombinant bacterial cultures. *Ann N Y Acad Sci* **589**, 121–138.

Bentley, W. E., Mirjalili, N., Andersen, D. C., Davis, R. H. & Kompala, D. S. (1990).

Plasmid-encoded protein: the principal factor in the 'metabolic burden' associated with recombinant bacteria. *Biotechnol Bioeng* **35**, 668–681.

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). *Methods Mol Biol* **866**, 115–127.

Broger, T., Odermatt, R. P., Huber, P. & Sonnleitner, B. (2011). Real-time on-line flow cytometry for bioprocess monitoring. *J Biotechnol* **154**, 240–247.

Brosius, J., Erfle, M. & Storella, J. (1985). Spacing of the -10 and -35 regions in the *tac* promoter. Effect on its *in vivo* activity. *J Biol Chem* **260**, 3539–3541.

Burgess-Brown, N. A., Sharma, S., Sobott, F., Loenarz, C., Oppermann, U. & Gileadi, O. (2008). Codon optimization can improve expression of human genes in *Escherichia coli*: A multi-gene study. *Protein Expr Purif* **59**, 94–102.

Butler, K., English, A. R., Ray, V. A. & Timreck, A. E. (1970). Carbenicillin: chemistry and mode of action. *J Infect Dis* **122**, Suppl:S1–8.

Calderone, T. L., Stevens, R. D. & Oas, T. G. (1996). High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in *Escherichia coli*. *J Mol Biol* **262**, 407–412.

Calos, M. (1978). DNA sequence for a low-level promoter of the *lac* repressor gene and an 'up' promoter mutation. *Nature* **274,** 762-765.

Carvalho, R. J., Cabrera-Crespo, J., Tanizaki, M. M. & Gonçalves, V. M. (2012). Development of production and purification processes of recombinant fragment of pneumococcal surface protein A in *Escherichia coli* using different carbon sources and chromatography sequences. *Appl Microbiol Biotechnol* **94**, 683–694.

Chalfie, M. (1995). Green fluorescent protein. *Photochem Photobiol* **62**, 651-656.

Chatterji, D. & Ojha, A. K. (2001). Revisiting the stringent response, ppGpp and starvation signaling. *Curr Opin Microbiol* **4**, 160–165.

Choi, J. H. & Lee, S. Y. (2004). Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Appl Microbiol Biotechnol* **64**, 625–635.

Choi, J., Keum, K. & Lee, S. (2006). Production of recombinant proteins by high cell density culture of *Escherichia coli*. *Chem Eng Sci* **61**, 876-885.

Cormack, B. P., Valdivia, R. H. & Falkow, S.(1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33–38.

Coutard, B., Gagnaire, M., Guilhon, A.-A. A., Berro, M., Canaan, S. & Bignon, C. (2008). Green fluorescent protein and factorial approach: an effective partnership for screening the soluble expression of recombinant proteins in *Escherichia coli*. *Protein Expr Purif* **61**, 184–190.

Craggs, T. D. (2009). Green fluorescent protein: structure, folding and chromophore maturation. *Chem Soc Rev* **38**, 2865–2875.

Cranenburgh, R. M., Hanak, J. A., Williams, S. G. & Sherratt, D. J. (2001). *Escherichia coli* strains that allow antibiotic-free plasmid selection and maintenance by repressor titration. *Nucleic Acids Res* **29**, E26.

Cunha, T. & Aires-Barros, R. (2002). Large-scale extraction of proteins. *Mol Biotechnol* **20**, 29–40.

Daegelen, P., Studier, F. W., Lenski, R. E., Cure, S. & Kim, J. F. (2009). Tracing ancestors and relatives of *Escherichia coli* B, and the derivation of B strains REL606 and BL21(DE3). *J Mol Biol* **394**, 634–643.

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.

Davey, H. M. & Kell, D. B. (1996). Flow cytometry and cell sorting of heterogeneous microbial populations: the importance of single-cell analyses. *Microbiol Rev* **60**, 641–696.

Day, R. N. & Davidson, M. W. (2009). The fluorescent protein palette: tools for cellular imaging. *Chem Soc Rev* **38**, 2887–2921.

De Boer, H. A., Comstock, L. J. & Vasser, M. (1983). The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proc Natl Acad Sci USA* **80**, 21–5.

De Giorgi, F., Brini, M., Bastianutto, C., Marsault, R., Montero, M., Pizzo, P., Rossi, R. & Rizzuto, R. (1996). Targeting aequorin and green fluorescent protein to intracellular organelles. *Gene* **173**, 113–117.

De Groot, N. S. & Ventura, S. (2006). Protein activity in bacterial inclusion bodies correlates with predicted aggregation rates. *J Biotechnol* **125**, 110–113.

De Groot, N. S., Sabate, R. & Ventura, S. (2009). Amyloids in bacterial inclusion bodies. *Trends Biochem Sci* **34**, 408–416.

De Marco, A., Deuerling, E., Mogk, A., Tomoyasu, T. & Bukau, B. (2007). Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli*. *BMC Biotechnol* **7**, 32.

Delic, M., Mattanovich, D. & Gasser, B. (2010). Monitoring intracellular redox conditions in the endoplasmic reticulum of living yeasts. *FEMS Microbiol Lett* **306**, 61–66.

Demain, A. L. & Vaishnav, P. (2009). Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv* **27**, 297–306.

Díaz, M., Herrero, M., García, L. & Quirós, C. (2010). Application of flow cytometry to industrial microbial bioprocesses. *Biochem Eng J* **48**, 385-407.

Dinnis, D. M. & James, D. C. (2005). Engineering mammalian cell factories for improved recombinant monoclonal antibody production: lessons from nature? *Biotechnol Bioeng* **91**, 180-189.

Doekel, S., Eppelmann, K. & Marahiel, M. A. (2002). Heterologous expression of nonribosomal peptide synthetases in B. subtilis: construction of a bi-functional *B. subtilis*/*E. coli* shuttle vector system. *FEMS Microbiol Lett* **216**, 185–191.

Dong, H., Nilsson, L. & Kurland, C. G. (1995). Gratuitous overexpression of genes in *Escherichia coli* leads to growth inhibition and ribosome destruction. *J Bacteriol* **177**, 1497–1504.

Doran, P. M. (2013). *Bioprocess engineering principles, 2nd Ed.*, Academic Press, Oxford.

Drepper, T., Eggert, T., Circolone, F., Heck, A., Krauss, U., Guterl, J.-K. K., Wendorff, M., Losi, A., Gärtner, W. & Jaeger, K.-E. E. (2007). Reporter proteins for in vivo fluorescence without oxygen. *Nat Biotechnol* **25**, 443–445.

Du, Z., Treiber, D., McCoy, R. E., Miller, A. K., Han, M., He, F., Domnitz, S., Heath, C. & Reddy, P. (2013). Non-invasive UPR monitoring system and its applications in CHO production cultures. *Biotechnol Bioeng* **110**, 2184–2194.

Dürrschmid, K., Reischer, H., Schmidt-Heck, W., Hrebicek, T., Guthke, R., Rizzi, A. & Bayer, K. (2008). Monitoring of transcriptome and proteome profiles to investigate the cellular response of *E. coli* towards recombinant protein expression under defined chemostat conditions. *J Biotechnol* **135**, 34–44.

Eibl, R., Kaiser, S., Lombriser, R. & Eibl, D. (2010). Disposable bioreactors: the current state-of-the-art and recommended applications in biotechnology. *Appl Microbiol Biotechnol* **86**, 41–49.

Elowitz, M. B., Surette, M. G., Wolf, P. E., Stock, J. & Leibler, S. (1997). Photoactivation turns green fluorescent protein red. *Curr Biol* **7**, 809–812.

Elvin, C. M., Thompson, P. R., Argall, M. E., Hendry, P., Stamford, N. P., Lilley, P. E. & Dixon, N. E. (1990). Modified bacteriophage lambda promoter vectors for overproduction of proteins in *Escherichia coli. Gene* **87**, 123–126.

Eron, L. & Block, R. (1971). Mechanism of initiation and repression of *in vitro* transcription of the *lac* operon of *Escherichia coli. Proc Natl Acad Sci USA* **68**, 1828–1832.

Espargaró, A., Sabate, R. & Ventura, S. (2012). Thioflavin-S staining coupled to flow cytometry. A screening tool to detect in vivo protein aggregation. *Mol Biosyst* **8**, 2839–2844.

Fahnert, B. (2012). Using folding promoting agents in recombinant protein production: a review. *Methods Mol Biol* **824**, 3–36.

Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J. L., Vázquez, E. & Villaverde, A. (2009). Microbial factories for recombinant pharmaceuticals. *Microb Cell Fact* **8**, 17.

Ferry, R. M., Farr, L. E. & Hartman, M. G. (1949). The Preparation and Measurement of the Concentration of Dilute Bacterial Aerosols. *Chem Rev* **44**, 389-417.

Flamm, E. L. (1991). How FDA approved chymosin: a case history. *Nat Biotechnol* **9**, 349–351.

French, C., Keshavarz-Moore, E. & Ward, J. M. (1996). Development of a simple method for the recovery of recombinant proteins from the *Escherichia coli* periplasm. *Enzyme Microb Tech* **19**, 332–338.

Frommer, W. B., Davidson, M. W. & Campbell, R. E. (2009). Genetically encoded biosensors based on engineered fluorescent proteins. *Chem Soc Rev* **38**, 2833–2841.

Fulwyler, M. J. (1965). Electronic separation of biological cells by volume. *Science* **150**, 910–911.

Gamer, J., Multhaup, G., Tomoyasu, T., McCarty, J. S., Rudiger, S., Schonfeld, H. J., Schirra, C., Bujard, H. & Bukau, B. (1996). A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the *Escherichia coli* heat shock transcription factor sigma32. *EMBO J* **15**, 607-617.

Garcia, J. R., Cha, H. J., Rao, G., Marten, M. R. & Bentley, W. E. (2009). Microbial nar-GFP cell sensors reveal oxygen limitations in highly agitated and aerated laboratory-scale fermentors. *Microb Cell Fact* **8**, 6.

García-Fruitós, E., González-Montalbán, N., Morell, M., Vera, A., Ferraz, R. M., Arís, A., Ventura, S. & Villaverde, A. (2005). Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microb Cell Fact* **4**, 27.

García-Fruitós, E., Martínez-Alonso, M., Gonzàlez-Montalbán, N., Valli, M., Mattanovich, D. & Villaverde, A. (2007a). Divergent genetic control of protein solubility and conformational quality in *Escherichia coli. J Mol Biol* **374**, 195–205.

García-Fruitós, E., Arís, A. & Villaverde, A. (2007b). Localization of functional polypeptides in bacterial inclusion bodies. *Appl Environ Microbiol* **73**, 289–294.

García-Fruitós, E., Sabate, R., de Groot, N. S., Villaverde, A. & Ventura, S. (2011). Biological role of bacterial inclusion bodies: a model for amyloid aggregation. *FEBS J* **278**, 2419–2427.

Gasser, B., Saloheimo, M., Rinas, U., Dragosits, M., Rodríguez-Carmona, E., Baumann, K., Giuliani, M., Parrilli, E., Branduardi, P., Lang, C., Porro, D., Ferrer, P., Tutino, M. L., Mattanovich, D. & Villaverde, A. (2008).** Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microb Cell Fact* **7**, 11.

Gatza, E., Pena, P. V., Srienc, F., Overton, T., Lavarreda C. A., Rogers C. E. (2012). *Bioprocess Monitoring with the BD Accuri™ C6 Flow Cytometer.* Becton, Dickinson and Company.

Gentry, D. R., Hernandez, V. J., Nguyen, L. H., Jensen, D. B. & Cashel, M. (1993). Synthesis of the stationary-phase sigma factor sigma s is positively regulated by ppGpp. *J Bacteriol* **175**, 7982–9.

Georgiou, G. & Segatori, L. (2005). Preparative expression of secreted proteins in bacteria: status report and future prospects. *Curr Opin Biotechnol* **16**, 538–545.

Gill, R. T., Valdes, J. J. & Bentley, W. E. (2000). A comparative study of global stress gene regulation in response to overexpression of recombinant proteins in *Escherichia coli. Metab Eng* **2**, 178–189.

Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K. & Riggs, A. D. (1979). Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *Proc Natl Acad Sci USA* **76**, 106–110.

Goldman, E., Rosenberg, A. H., Zubay, G. & Studier, F. W. (1995). Consecutive low-usage leucine codons block translation only when near the 5' end of a message in *Escherichia coli. J Mol Biol* **245**, 467–473.

González-Montalbán, N., Carrió, M. M., Cuatrecasas, S., Arís, A. & Villaverde, A. (2005). Bacterial inclusion bodies are cytotoxic *in vivo* in absence of functional chaperones DnaK or GroEL. *J Biotechnol* **118**, 406–412.

González-Montalbán, N., Villaverde, A. & Aris, A. (2007). Amyloid-linked cellular toxicity triggered by bacterial inclusion bodies. *Biochem Biophys Res Commun* **355**, 637–642.

Grimwade, L., Gudgin, E., Bloxham, D., Bottley, G., Vassiliou, G., Huntly, B., Scott, M. A. & Erber, W. N. (2012). Detection of cytoplasmic nucleophosmin expression by imaging flow cytometry. *Cytometry A* **81**, 896–900.

Gucker, F. T., Pickard, H. B. & O'Konski, C. T. (1947). A photoelectric instrument for comparing the concentrations of very dilute aerosols, and measuring low light intensities. *J Am Chem Soc* **69**, 429–438.

Gucker, F. T. & O'Konski, C. (1949). Electronic Methods of Counting Aerosol Particles. *Chem Rev* **44**, 373-388 .

Guisbert, E., Herman, C., Lu, C. Z. & Gross, C. A. (2004). A chaperone network controls the heat shock response in *E. coli. Genes Dev* **18**, 2812–2821.

Guisbert, E., Yura, T., Rhodius, V. A. & Gross, C. A. (2008). Convergence of molecular, modeling, and systems approaches for an understanding of the *Escherichia coli* heat shock response. *Microbiol Mol Biol Rev* **72**, 545–554.

Gumpert, J. & Hoischen, C. (1998). Use of cell wall-less bacteria (L-forms) for efficient expression and secretion of heterologous gene products. *Curr Opin Biotechnol* **9**, 506–509.

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. *J Bacteriol* **177**, 4121–4130.

Haddadin, F. T. & Harcum, S. W. (2005). Transcriptome profiles for high-cell-density recombinant and wild-type *Escherichia coli. Biotechnol Bioeng* **90**, 127–153.

Haldimann, A., Daniels, L. L. & Wanner, B. L. (1998). Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the *Escherichia coli* phosphate regulon. *J Bacteriol* **180**, 1277–1286.

Hansen, M. C., Palmer, R. J., Udsen, C., White, D. C. & Molin, S. (2001). Assessment of

GFP fluorescence in cells of *Streptococcus gordonii* under conditions of low pH and low oxygen concentration. *Microbiol* **147**, 1383–1391.

Harcum, S. W. & Haddadin, F. T. (2006). Global transcriptome response of recombinant *Escherichia coli* to heat-shock and dual heat-shock recombinant protein induction. *J Ind Microbiol Biotechnol* **33**, 801–814.

Hedhammar, M., Stenvall, M., Lönneborg, R., Nord, O., Sjölin, O., Brismar, H., Uhlén, M., Ottosson, J. & Hober, S. (2005). A novel flow cytometry-based method for analysis of expression levels in *Escherichia coli*, giving information about precipitated and soluble protein. *J Biotechnol* **119**, 133–146.

Heim, R., Prasher, D. C. & Tsien, R. Y. (1994). Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci USA* **91**, 12501–12504.

Heim, R., Cubitt, A. B. & Tsien, R. Y. (1995). Improved green fluorescence. *Nature* **373**, 663–664.

Herrera, G., Martinez, A., Blanco, M. & O'Connor, J.-E. E. (2002). Assessment of *Escherichia coli* B with enhanced permeability to fluorochromes for flow cytometric assays of bacterial cell function. *Cytometry* **49**, 62–69.

Hess, J. F., Oosawa, K., Kaplan, N. & Simon, M. I. (1988). Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell* **53**, 79–87.

Hewitt, C. J. & Nebe-Von-Caron, G. (2001). An industrial application of multiparameter flow cytometry: assessment of cell physiological state and its application to the study of microbial fermentations. *Cytometry* **44**, 179–187.

Huber, R., Roth, S., Rahmen, N. & Büchs, J. (2011). Utilizing high-throughput experimentation to enhance specific productivity of an *E. coli* T7 expression system by phosphate limitation. *BMC Biotechnol* **11**, 22.

Humphreys, D. P., Carrington, B., Bowering, L. C., Ganesh, R., Sehdev, M., Smith, B. J., King, L. M., Reeks, D. G., Lawson, A. & Popplewell, A. G. (2002). A plasmid system for optimization of Fab' production in *Escherichia coli*: importance of balance of heavy chain and light chain synthesis. *Protein Expr Purif* **26**, 309–320.

Hussain, H., Maldonado-Agurto, R. & Dickson, A. J. (2014). The endoplasmic reticulum and unfolded protein response in the control of mammalian recombinant protein production. *Biotechnol Lett* **36**, 1581–1593.

Huang, C.-J. J., Lin, H. & Yang, X. (2012). Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J Ind Microbiol Biotechnol* **39**, 383–399.

Hwang, S. O. & Feldberg, R. S. (1990). Effect of inclusion body production on culture turbidity and cell dry weight in growing bacterial cultures. *Biotechnol Prog* **6**, 48–50.

Inouye, S. & Tsuji, F. (1994). *Aequorea* green fluorescent protein. *FEBS Lett* **341**, 277-280.

Ishii, S., Tago, K. & Senoo, K. (2010). Single-cell analysis and isolation for microbiology and biotechnology: methods and applications. *Appl Microbiol Biotechnol* **86**, 1281–1292.

Islam, R. S., Tisi, D., Levy, M. S. & Lye, G. J. (2007). Framework for the rapid optimization of soluble protein expression in *Escherichia coli* combining microscale experiments and statistical experimental design. *Biotechnol Prog* **23**, 785–793.

Itakura, K., Hirose, T., Crea, R., Riggs, A. D., Heyneker, H. L., Bolivar, F. & Boyer, H. W. (1977). Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science* **198**, 1056–1063.

Jalalirad, R. (2010). *Improving the recovery of "difficult to release" periplasmically-*

expressed products from recombinant E. coli. PhD thesis, University of Birmingham.

Japrung, D., Chusacultanachai, S., Yuvaniyama, J., Wilairat, P. & Yuthavong, Y. (2005). A simple dual selection for functionally active mutants of *Plasmodium falciparum* dihydrofolate reductase with improved solubility. *Protein Eng Des Sel* **18**, 457–464.

Jeong, H., Barbe, V., Lee, C. H., Vallenet, D., Yu, D. S., Choi, S.-H. H., Couloux, A., Lee, S.-W. W., Yoon, S. H. & other authors. **(2009).** Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *J Mol Biol* **394**, 644–652.

Jeong, K. J. & Lee, S. Y. (1999). High-level production of human leptin by fed-batch cultivation of recombinant *Escherichia coli* and its purification. *Appl Environ Microbiol* **65**, 3027–3032.

Jevševar, S., Gaberc-Porekar, V., Fonda, I., Podobnik, B., Grdadolnik, J. & Menart, V. (2005). Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. *Biotechnol Prog* **21**, 632–639.

Jin, K., Thomas, O. R. & Dunnill, P. (1994). Monitoring recombinant inclusion body recovery in an industrial disc stack centrifuge. *Biotechnol Bioeng* **43**, 455–460.

Johnson, I. S. (1983). Human insulin from recombinant DNA technology. *Science* **219**, 632–637.

Johnson, I. & Spence, M. T. Z. (2010). *Molecular Probes Handbook, A Guide to Fluorescent Probes and Labeling Technologies, 11th Edition*, Life Technologies Corporation.

Jones, J. J. (2007). *Green fluorescent protein as an analytical tool to dissect the physiology of recombinant protein production in fermenters*. PhD Thesis, University of Birmingham.

Jones, J. J., Bridges, A. M., Fosberry, A. P., Gardner, S., Lowers, R. R., Newby, R. R., James, P. J., Hall, R. M. & Jenkins, O. (2004). Potential of real-time measurement of GFP-fusion proteins. *J Biotechnol* **109**, 201–211.

Junker, B. (2007). Foam and its mitigation in fermentation systems. *Biotechnol Prog* **23**, 767–784.

Jürgen, B., Lin, H. Y., Riemschneider, S., Scharf, C., Neubauer, P., Schmid, R., Hecker, M. & Schweder, T. (2000). Monitoring of genes that respond to overproduction of an insoluble recombinant protein in *Escherichia coli* glucose-limited fed-batch fermentations. *Biotechnol Bioeng* **70**, 217–224.

Kagan, B. L., Jang, H., Capone, R., Teran Arce, F., Ramachandran, S., Lal, R. & Nussinov, R. (2012). Antimicrobial properties of amyloid peptides. *Mol Pharm* **9**, 708–717.

Kane, J. F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Curr Opin Biotechnol* **6**, 494–500.

Kapust, R. B. & Waugh, D. S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* **8**, 1668–1674.

Katzen, F., Chang, G. & Kudlicki, W. (2005). The past, present and future of cell-free protein synthesis. *Trends Biotechnol* **23**, 150–156.

Kensy, F., Engelbrecht, C. & Büchs, J. (2009a). Scale-up from microtiter plate to laboratory fermenter: evaluation by online monitoring techniques of growth and protein expression in *Escherichia coli* and *Hansenula polymorpha* fermentations. *Microb Cell Fact* **8**, 68.

Kensy, F., Zang, E., Faulhammer, C., Tan, R.-K. K. & Büchs, J. (2009b). Validation of a high-throughput fermentation system based on online monitoring of biomass and

fluorescence in continuously shaken microtiter plates. *Microb Cell Fact* **8**, 31.

Kober, L., Zehe, C. & Bode, J. (2012). Development of a novel ER stress based selection system for the isolation of highly productive clones. *Biotechnol Bioeng* **109**, 2599–2611.

Koch, A. L., Robertson, B. R. & Button, D. K. (1996). Deduction of the cell volume and mass from forward scatter intensity of bacteria analyzed by flow cytometry. *J Microbiol Meth,* **27,** 49-61.

Korz, D. J., Rinas, U., Hellmuth, K., Sanders, E. A. & Deckwer, W. D. (1995). Simple fed-batch technique for high cell density cultivation of *Escherichia coli. J Biotechnol* **39**, 59–65.

Kroll, J., Steinle, A., Reichelt, R., Ewering, C. & Steinbüchel, A. (2009). Establishment of a novel anabolism-based addiction system with an artificially introduced mevalonate pathway: complete stabilization of plasmids as universal application in white biotechnology. *Metab Eng* **11**, 168–177.

Kroll, J., Klinter, S., Schneider, C., Voss, I. & Steinbüchel, A. (2010). Plasmid addiction systems: perspectives and applications in biotechnology. *Microb Biotechnol* **3**, 634–657.

Kurland, C. & Gallant, J. (1996). Errors of heterologous protein expression. *Curr Opin Biotechnol* **7**, 489–493.

Kurland, C. G. & Dong, H. (1996). Bacterial growth inhibition by overproduction of protein. *Mol Microbiol* **21**, 1–4.

Kusano, K., Waterman, M. R., Sakaguchi, M., Omura, T. & Kagawa, N. (1999). Protein synthesis inhibitors and ethanol selectively enhance heterologous expression of P450s and related proteins in *Escherichia coli. Arch Biochem Biophys* **367**, 129–136.

Lee, S. Y. (1996). High cell-density culture of *Escherichia coli. Trends Biotechnol* **14**, 98–105.

Lehtinen, J., Nuutila, J. & Lilius, E.-M. M. (2004). Green fluorescent protein-propidium iodide (GFP-PI) based assay for flow cytometric measurement of bacterial viability. *Cytometry A* **60**, 165–172.

Lewis, G., Taylor, I. W., Nienow, A. W. & Hewitt, C. J. (2004). The application of multi-parameter flow cytometry to the study of recombinant *Escherichia coli* batch fermentation processes. *J Ind Microbiol Biotechnol* **31**, 311–322.

Life Technologies Corp. (2010). *BL21 Star™(DE3) One Shot® BL21 Star™(DE3)pLysS One Shot® Chemically Competent Cells, User manual*

Lilie, H., Schwarz, E. & Rudolph, R. (1998). Advances in refolding of proteins produced in *E. coli. Curr Opin Biotechnol* **9**, 497–501.

Lindner, A. B., Madden, R., Demarez, A., Stewart, E. J. & Taddei, F. (2008). Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc Natl Acad Sci USA* **105**, 3076–3081.

Lizak, C., Fan, Y.-Y. Y., Weber, T. C. & Aebi, M. (2011). N-Linked glycosylation of antibody fragments in *Escherichia coli. Bioconjug Chem* **22**, 488–496.

Lopez, P. J., Marchand, I., Joyce, S. A. & Dreyfus, M. (1999). The C-terminal half of RNase E, which organizes the *Escherichia coli* degradosome, participates in mRNA degradation but not rRNA processing in vivo. *Mol Microbiol* **33**, 188–199.

Lübke, C., Boidol, W. & Petri, T. (1995). Analysis and optimization of recombinant protein production in *Escherichia coli* using the inducible pho A promoter of the *E. coli* alkaline phosphatase. *Enzyme Microb Technol,* **17,** 923-928.

Lucigen Corp. (2013). *ClearColi™ BL21(DE3) Electrocompetent Cells, Technical manual*.

Luo, Q., Shen, Y.-L. L., Wei, D.-Z. Z. & Cao, W. (2006). Optimization of culture on the

overproduction of TRAIL in high-cell-density culture by recombinant *Escherichia coli*. *Appl Microbiol Biotechnol* **71**, 184–191.

Makino, K., Shinagawa, H., Amemura, M. & Nakata, A. (1986). Nucleotide sequence of the phoB gene, the positive regulatory gene for the phosphate regulon of *Escherichia coli* K-12. *J Mol Biol* **190**, 37–44.

Makrides, S. C. (1996). Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev* **60**, 512–538.

March, J. C., Rao, G. & Bentley, W. E. (2003). Biotechnological applications of green fluorescent protein. *Appl Microbiol Biotechnol* **62**, 303–315.

Märkl, H., Zenneck, C., Dubach, A. C. & Ogbonna, J. C. (1993). Cultivation of *Escherichia coli* to high cell densities in a dialysis reactor. *Appl Microbiol Biotechnol* **39**, 48–52.

Marshall, J., Molloy, R., Moss, G. W., Howe, J. R. & Hughes, T. E. (1995). The jellyfish green fluorescent protein: a new tool for studying ion channel expression and function. *Neuron* **14**, 211–215.

Martínez, J. L., Liu, L., Petranovic, D. & Nielsen, J. (2012). Pharmaceutical protein production by yeast: towards production of human blood proteins by microbial fermentation. *Curr Opin Biotechnol* **23**, 965–971.

Martínez-Alonso, M., González-Montalbán, N., García-Fruitós, E. & Villaverde, A. (2008). The functional quality of soluble recombinant polypeptides produced in *Escherichia coli* is defined by a wide conformational spectrum. *Appl Environ Microbiol* **74**, 7431–7433.

Martínez-Alonso, M., García-Fruitós, E., Ferrer-Miralles, N., Rinas, U. & Villaverde, A. (2010). Side effects of chaperone gene co-expression in recombinant protein production. *Microb Cell Fact* **9**, 64.

Matos, C. F., Branston, S. D., Albiniak, A., Dhanoya, A., Freedman, R. B., Keshavarz-Moore, E. & Robinson, C. (2012). High-yield export of a native heterologous protein to the periplasm by the tat translocation pathway in *Escherichia coli*. *Biotechnol Bioeng* **109**, 2533–2542.

Matsui, T., Yokota, H., Sato, S. & Mukataka, S. & Takahashi, . (1989). Pressurized culture of *Escherichia coli* for a high concentration. *Agr Biol Chem Tokyo* **53**, 2115-2120.

Mattanovich, D. & Borth, N. (2006). Applications of cell sorting in biotechnology. *Microb Cell Fact* **5**, 12.

Mattanovich, D., Gasser, B., Hohenblum, H. & Sauer, M. (2004). Stress in recombinant protein producing yeasts. *J Biotechnol* **113**, 121–135.

Merck KGaA. (2013a). *69659 pLysS DNA, Technical manual.*

Merck KGaA. (2013b). *70837 Origami™ B(DE3) Competent Cells, Technical manual.*

Merck KGaA. (2013c). *71397 Rosetta™ 2(DE3) Competent Cells, Technical manual.*

Merck KGaA. (2013d). *71136 Rosetta-gami B(DE3) Competent Cells, Technical manual.*

Mergulhão, F. J., Summers, D. K. & Monteiro, G. A. (2005). Recombinant protein secretion in *Escherichia coli*. *Biotechnol Adv* **23**, 177–202.

Miao, H., Ratnasingam, S., Pu, C. S., Desai, M. M. & Sze, C. C. (2009). Dual fluorescence system for flow cytometric analysis of *Escherichia coli* transcriptional response in multi-species context. *J Microbiol Methods* **76**, 109–119.

Miroux, B. & Walker, J. E. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **260**, 289–298.

Moks, T., Abrahmsén, L., Holmgren, E., Bilich, M., Olsson, A., Uhlén, M., Pohl, G.,

Sterky, C., Hultberg, H. & Josephson, S. (1987). Expression of human insulin-like growth factor I in bacteria: use of optimized gene fusion vectors to facilitate protein purification. *Biochemistry* **26**, 5239–5244.

Moore, J. T., Uppal, A., Maley, F. & Maley, G. F. (1993). Overcoming inclusion body formation in a high-level expression system. *Protein Expr Purif* **4**, 160–163.

Mori, H., Yano, T., Kobayashi, T. & Shimizu, S. (1979). High density cultivation of biomass in fed-batch system with DO-stat. *J Chem Eng Japan* **12**, 313–319.

Morin, J. G. & Hastings, J. W. (1971). Energy transfer in a bioluminescent system. *J Cell Physiol* **77**, 313–318.

Nebe-von-Caron, G., Stephens, P. J., Hewitt, C. J., Powell, J. R. & Badley, R. A. (2000). Analysis of bacterial function by multi-colour fluorescence flow cytometry and single cell sorting. *J Microbiol Methods* **42**, 97–114.

New England Biolabs Inc. (2013). *Restriction endonucelases, Technical guide 2013/14.*

Ni, Y. & Chen, R. (2009). Extracellular recombinant protein production from *Escherichia coli. Biotechnol Lett* **31**, 1661–1670.

Nielsen, J. (2013). Production of biopharmaceutical proteins by yeast: advances through metabolic engineering. *Bioengineered* **4**, 207–211.

Nishimura, A., Morita, M., Nishimura, Y. & Sugino, Y. (1990). A rapid and highly efficient method for preparation of competent *Escherichia coli* cells. *Nucleic Acids Res* **18**, 6169.

Noronha, S. B., Yeh, H. J., Spande, T. F. & Shiloach, J. (2000). Investigation of the TCA cycle and the glyoxylate shunt in *Escherichia coli* BL21 and JM109 using (13)C-NMR/MS. *Biotechnol Bioeng* **68**, 316–327.

Okabe, M., Ikawa, M., Kominami, K., Nakanishi, T. & Nishimune, Y. (1997). 'Green mice' as a source of ubiquitous green cells. *FEBS Lett* **407**, 313–319.

Olempska-Beer, Z. S., Merker, R. I., Ditto, M. D. & DiNovi, M. J. (2006). Food-processing enzymes from recombinant microorganisms--a review. *Regul Toxicol Pharmacol* **45**, 144–158.

Overman, R. C., Debreczeni, J. E., Truman, C. M., McAlister, M. S. & Attwood, T. K. (2014). Completing the structural family portrait of the human EphB tyrosine kinase domains. *Protein Sci* **23**, 627–638.

Overton, T. W. (2014). Recombinant protein production in bacterial hosts. *Drug Discov Today* **19**, 590–601.

Palomares, L. A., Estrada-Moncada, S. & Ramírez, O. T. (2004). Production of recombinant proteins. *Methods Mol Biol* **267**, 15-52.

Pandhal, J., Ow, S. Y., Noirel, J. & Wright, P. C. (2011). Improving N-glycosylation efficiency in *Escherichia coli* using shotgun proteomics, metabolic network analysis, and selective reaction monitoring. *Biotechnol Bioeng* **108**, 902–912.

Pandhal, J., Desai, P., Walpole, C., Doroudi, L., Malyshev, D. & Wright, P. C. (2012). Systematic metabolic engineering for improvement of glycosylation efficiency in *Escherichia coli. Biochem Biophys Res Commun* **419**, 472–476.

Papaneophytou, C. P. & Kontopidis, G. (2014). Statistical approaches to maximize recombinant protein expression in *Escherichia coli*: a general review. *Protein Expr Purif* **94**, 22–32.

Patil, C. & Walter, P. (2001). Intracellular signaling from the endoplasmic reticulum to the nucleus: the unfolded protein response in yeast and mammals. *Curr Opin Cell Biol* **13**, 349–355.

Patkar, A., Vijayasankaran, N., Urry, D. W. & Srienc, F. (2002). Flow cytometry as a useful tool for process development: rapid evaluation of expression systems. *J*

Biotechnol **93**, 217–29.

Peternel, S., Grdadolnik, J., Gaberc-Porekar, V. & Komel, R. (2008). Engineering inclusion bodies for non denaturing extraction of functional proteins. *Microb Cell Fact* **7**, 34.

Peternel, S. & Komel, R. (2010). Isolation of biologically active nanomaterial (inclusion bodies) from bacterial cells. *Microb Cell Fact* **9**, 66.

Petsch, D. & Anspach, F. B. (2000). Endotoxin removal from protein solutions. *J Biotechnol* **76**, 97–119.

Pflug, S., Richter, S. M. & Urlacher, V. B. (2007). Development of a fed-batch process for the production of the cytochrome P450 monooxygenase CYP102A1 from *Bacillus megaterium* in *E. coli. J Biotechnol* **129**, 481–488.

Phadtare, S., Alsina, J. & Inouye, M. (1999). Cold-shock response and cold-shock proteins. *Curr Opin Microbiol* **2**, 175–180.

Phelps, M. R., Hobbs, J. B., Kilburn, D. G. & Turner, R. F. (1995). An autoclavable glucose biosensor for microbial fermentation monitoring and control. *Biotechnol Bioeng* **46**, 514–524.

Phue, J.-N. N. & Shiloach, J. (2004). Transcription levels of key metabolic genes are the cause for different glucose utilization pathways in *E. coli* B (BL21) and *E. coli* K (JM109). *J Biotechnol* **109**, 21–30.

Pinske, C., Bönn, M., Krüger, S., Lindenstrauss, U. & Sawers, R. G. (2011). Metabolic deficiences revealed in the biotechnologically important model bacterium *Escherichia coli* BL21(DE3). *PLoS ONE* **6**, e22830.

Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier, M. J. (1992). Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* **111**, 229–233.

Ramdzan, Y. M., Polling, S., Chia, C. P., Ng, I. H., Ormsby, A. R., Croft, N. P., Purcell, A. W., Bogoyevitch, M. A., Ng, D. C., Gleeson, P. A. & Hatters, D. M. **(2012).** Tracking protein aggregation and mislocalization in cells with flow cytometry. *Nat Methods* **9**, 467–470.

Ramírez, D. M. & Bentley, W. E. (1995). Fed-batch feeding and induction policies that improve foreign protein synthesis and stability by avoiding stress responses. *Biotechnol Bioeng* **47**, 596–608.

Randers-Eichhorn, L., Albano, C. R., Sipior, J., Bentley, W. E. & Rao, G. (1997). On-line green fluorescent protein sensor with LED excitation. *Biotechnol Bioeng* **55**, 921–926.

Reid, B. G. & Flynn, G. C. (1997). Chromophore formation in green fluorescent protein. *Biochemistry* **36**, 6786–6791.

Rieseberg, M., Kasper, C., Reardon, K. F. & Scheper, T. (2001). Flow cytometry in biotechnology. *Appl Microbiol Biotechnol* **56**, 350–360.

Riesenberg, D., Schulz, V., Knorre, W. A., Pohl, H. D., Korz, D., Sanders, E. A., Ross, A. & Deckwer, W. D. (1991). High cell density cultivation of *Escherichia coli* at controlled specific growth rate. *J Biotechnol* **20**, 17–27.

Rouyi, C., Baiya, S., Lee, S.-K. K., Mahong, B., Jeon, J.-S. S., Ketudat-Cairns, J. R. & Ketudat-Cairns, M. (2014). Recombinant expression and characterization of the cytoplasmic rice β-glucosidase Os1BGlu4. *PLoS ONE* **9**, e96712.

Rumbold, K., van Buijsen, H. J. J., Gray, V. M., van Groenestijn, J. W., Overkamp, K. M., Slomp, R. S., van der Werf, M. J. & Punt, P. J. (2010). Microbial renewable feedstock utilization: a substrate-oriented approach. *Bioeng Bugs* **1**, 359–366.

Rutten, T. P., Sandee, B. & Hofman, A. R. (2005). Phytoplankton monitoring by high

performance flow cytometry: a successful approach? *Cytometry A* **64**, 16–26.

Sachdev, D. & Chirgwin, J. M. (1998a). Order of fusions between bacterial and mammalian proteins can determine solubility in *Escherichia coli. Biochem Biophys Res Commun* **244**, 933–937.

Sachdev, D. & Chirgwin, J. M. (1998b). Solubility of proteins isolated from inclusion bodies is enhanced by fusion to maltose-binding protein or thioredoxin. *Protein Expr Purif* **12**, 122–132.

Saïda, F., Uzan, M., Odaert, B. & Bontems, F. (2006). Expression of highly toxic genes in *E. coli*: special strategies and genetic tools. *Curr Protein Pept Sci* **7**, 47–56.

Sample, V., Newman, R. H. & Zhang, J. (2009). The structure and function of fluorescent proteins. *Chem Soc Rev* **38**, 2852–2864.

Samuelsson, E., Moks, T., Nilsson, B. & Uhlen, M. (1994). Enhanced *in vitro* refolding of insulin-like growth factor I using a solubilizing fusion partner. *Biochemistry* **33**, 4207–4211.

Sans, C., García-Fruitós, E., Ferraz, R. M., González-Montalbán, N., Rinas, U., López-Santín, J., Villaverde, A. & Álvaro, G. (2012). Inclusion bodies of fuculose-1-phosphate aldolase as stable and reusable biocatalysts. *Biotechnol Prog* **28**, 421–427.

Scheidle, M., Jeude, M., Dittrich, B., Denter, S., Kensy, F., Suckow, M., Klee, D. & Büchs, J. (2010). High-throughput screening of *Hansenula polymorpha* clones in the batch compared with the controlled-release fed-batch mode on a small scale. *FEMS Yeast Res* **10**, 83–92.

Schmidt, F. R. (2004). Recombinant expression systems in the pharmaceutical industry. *Appl Microbiol Biotechnol* **65**, 363–372.

Schmidt, M., Babu, K. R., Khanna, N., Marten, S. & Rinas, U. (1999). Temperature-induced production of recombinant human insulin in high-cell density cultures of recombinant *Escherichia coli. J Biotechnol* **68**, 71–83.

Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671–675.

Schneider, K., Schütz, V., John, G. T. & Heinzle, E. (2010). Optical device for parallel online measurement of dissolved oxygen and pH in shake flask cultures. *Bioprocess Biosyst Eng* **33**, 541–547.

Schröder, M. & Kaufman, R. (2005). ER stress and the unfolded protein response. *Mutat Res* **569**, 29-63.

Schroeckh, V., Wenderoth, R., Kujau, M., Knüpfer, U. & Riesenberg, D. (1999). The use of elements of the *E. coli* Ntr-system for the design of an optimized recombinant expression system for high cell density cultivations. *J Biotechnol* **75**, 241–250.

Sevastsyanovich, Y., Alfasi, S., Overton, T., Hall, R., Jones, J., Hewitt, C. & Cole, J. (2009). Exploitation of GFP fusion proteins and stress avoidance as a generic strategy for the production of high-quality recombinant proteins. *FEMS Microbiol Lett* **299**, 86–94.

Sevastsyanovich, Y. R., Alfasi, S. N. & Cole, J. A. (2010). Sense and nonsense from a systems biology approach to microbial recombinant protein production. *Biotechnol Appl Biochem* **55**, 9–28.

Sevastsyanovich, Y. R., Leyton, D. L., Wells, T. J., Wardius, C. A., Tveen-Jensen, K., Morris, F. C., Knowles, T. J., Cunningham, A. F., Cole, J. A. & Henderson, I. R. (2012). A generalised module for the selective extracellular accumulation of recombinant proteins. *Microb Cell Fact* **11**, 69.

Shaner, N. C. (2013). The mFruit collection of monomeric fluorescent proteins. *Clin Chem* **59**, 440–441.

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E. & Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma sp.* red fluorescent protein. *Nat Biotechnol* **22**, 1567–1572.

Shapiro, H. M. (2005). *Practical flow cytometry, 4th Ed.*, John Wiley & Sons, USA.

Shiloach, J. & Fass, R. (2005). Growing *E. coli* to high cell density--a historical perspective on method development. *Biotechnol Adv* **23**, 345–357.

Shimomura, O., Johnson, F. H. & Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea. J Cell Comp Physiol* **59**, 223–239.

Shimomura, O., Johnson, F. H. & Morise, H. (1974). Mechanism of the luminescent intramolecular reaction of aequorin. *Biochemistry* **13**, 3278–3286.

Siller, E., DeZwaan, D. C., Anderson, J. F., Freeman, B. C. & Barral, J. M. (2010). Slowing bacterial translation speed enhances eukaryotic protein folding efficiency. *J Mol Biol* **396**, 1310–1318.

Singh, S. M. & Panda, A. K. (2005). Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng* **99**, 303–310.

Skerra, A. (1994). Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli. Gene* **151**, 131–135.

Soini, J., Ukkonen, K. & Neubauer, P. (2008). High cell density media for *Escherichia coli* are generally designed for aerobic cultivations - consequences for large-scale bioprocesses and shake flask cultures. *Microb Cell Fact* **7**, 26.

Soriano, E., Borth, N., Katinger, H. & Mattanovich, D. (1999). Flow cytometric analysis of metabolic stress effects due to recombinant plasmids and proteins in *Escherichia coli* production strains. *Metab Eng* **1**, 270–274.

Stock, J. B., Rauch, B. & Roseman, S. (1977). Periplasmic space in *Salmonella typhimurium* and *Escherichia coli. J Biol Chem* **252**, 7850–7861.

Strandberg, L., Andersson, L. & Enfors, S. O. (1994). The use of fed batch cultivation for achieving high cell densities in the production of a recombinant protein in *Escherichia coli. FEMS Microbiol Rev* **14**, 53–56.

Studier, F. W. & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113–130.

Studier, F. W., Daegelen, P., Lenski, R. E., Maslov, S. & Kim, J. F. (2009). Understanding the differences between genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3) and comparison of the *E. coli* B and K-12 genomes. *J Mol Biol* **394**, 653–680.

Su, W. W. (2005). Fluorescent proteins as tools to aid protein production. *Microb Cell Fact* **4**, 12.

Su, Y.-C. C., Lim, K.-P. P. & Nathan, S. (2003). Bacterial expression of the scFv fragment of a recombinant antibody specific for *Burkholderia pseudomallei* exotoxin. *J Biochem Mol Biol* **36**, 493–498.

Sundström, H., Wållberg, F., Ledung, E., Norrman, B., Hewitt, C. J. & Enfors, S.-O. O.(2004). Segregation to non-dividing cells in recombinant *Escherichia coli* fed-batch fermentation processes. *Biotechnol Lett* **26**, 1533–1539.

Sung, Y. C., Anderson, P. M. & Fuchs, J. A. (1987). Characterization of high-level expression and sequencing of the *Escherichia coli* K-12 cynS gene encoding cyanase. *J Bacteriol* **169**, 5224–5230.

Sunya, S., Delvigne, F., Uribelarrea, J.-L. L., Molina-Jouve, C. & Gorret, N. (2012). Comparison of the transient responses of *Escherichia coli* to a glucose pulse of various

intensities. *Appl Microbiol Biotechnol* **95**, 1021–1034.

Sunya, S., Bideaux, C., Molina-Jouve, C. & Gorret, N. (2013). Short-term dynamic behavior of *Escherichia coli* in response to successive glucose pulses on glucose-limited chemostat cultures. *J Biotechnol* **164**, 531–542.

Swalley, S., Fulghum, J. & Chambers, S. (2006). Screening factors effecting a response in soluble protein expression: Formalized approach using design of experiments. *Anal Biochem* **351,** 122-127.

Sweet, C. R. (2003). Expression of recombinant proteins from *lac* promoters. *Methods Mol Biol,* **235,** 277-288.

Świeżawska, B., Jaworski, K., Pawełek, A., Grzegorzewska, W., Szewczuk, P. & Szmidt-Jaworska, A. (2014). Molecular cloning and characterization of a novel adenylyl cyclase gene, HpAC1, involved in stress signaling in *Hippeastrum x hybridum*. *Plant Physiol Biochem* **80**, 41–52.

Sørensen, H. P. & Mortensen, K. K. (2005). Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb Cell Fact* **4**, 1.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol* **72**, 211–222.

Thattai, M. & van Oudenaarden, A. (2004). Stochastic gene expression in fluctuating environments. *Genetics* **167**, 523–530.

Thastrup, O., Tullin, S., Poulsen, L.K., & Bjørn S. P. 2001. Novel variants of green fluorescent protein, GFP. US Patent 6172188

Tseng, C.-L. L. & Leng, C.-H. H. (2012). Influence of medium components on the expression of recombinant lipoproteins in *Escherichia coli*. *Appl Microbiol Biotechnol* **93**, 1539–1352.

Tsien, R. Y. (1998). The green fluorescent protein. *Annu Rev Biochem* **67**, 509–544.

Turner, S., Reid, E., Smith, H. & Cole, J. (2003). A novel cytochrome c peroxidase from *Neisseria gonorrhoeae*: a lipoprotein from a Gram-negative bacterium. *Biochem J* **373**, 865–873.

Tyler, B. (1978). Regulation of the assimilation of nitrogen compounds. *Annu Rev Biochem* **47**, 1127–1162.

Um, S.-H. H., Kim, J.-S. S., Kim, K., Kim, N., Cho, H.-S. S. & Ha, N.-C. C. (2013). Structural basis for the inhibition of human lysozyme by PliC from *Brucella abortus*. *Biochemistry* **52**, 9385–9393.

Umbarger, H. E. (1978). Amino acid biosynthesis and its regulation. *Annu Rev Biochem* **47**, 532–606.

Upadhyay, A. K., Murmu, A., Singh, A. & Panda, A. K. (2012). Kinetics of inclusion body formation and its correlation with the characteristics of protein aggregates in *Escherichia coli*. *PLoS ONE* **7**, e33951.

Valderrama-Rincon, J. D., Fisher, A. C., Merritt, J. H., Fan, Y.-Y. Y., Reading, C. A., Chhiba, K., Heiss, C., Azadi, P., Aebi, M. & DeLisa, M. P. (2012). An engineered eukaryotic protein glycosylation pathway in *Escherichia coli*. *Nat Chem Biol* **8**, 434–436.

Valgepea, K., Adamberg, K., Seiman, A. & Vilu, R. (2013). *Escherichia coli* achieves faster growth by increasing catalytic and translation rates of proteins. *Mol Biosyst* **9**, 2344–2358.

Vallejo, L. F. & Rinas, U. (2004). Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb Cell Fact* **3**, 11.

VanBogelen, R. A. & Neidhardt, F. C. (1990). Ribosomes as sensors of heat and cold

shock in *Escherichia coli. Proc Natl Acad Sci USA* **87**, 5589–5593.

Veal, D. A., Deere, D., Ferrari, B., Piper, J. & Attfield, P. V. (2000). Fluorescence staining and flow cytometry for monitoring microbial cells. *J Immunol Methods* **243**, 191–210.

Vera, A., Arís, A., Carrió, M., González-Montalbán, N. & Villaverde, A. (2005). Lon and ClpP proteases participate in the physiological disintegration of bacterial inclusion bodies. *J Biotechnol* **119**, 163–171.

Vera, A., González-Montalbán, N., Arís, A. & Villaverde, A. (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnol Bioeng* **96**, 1101–6.

Villaverde, A. & Carrió, M. M. (2003). Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol Lett* **25**, 1385–1395.

Villaverde, A., García-Fruitós, E., Rinas, U., Seras-Franzoso, J., Kosoy, A., Corchero, J. L. & Vazquez, E. (2012). Packaging protein drugs as bacterial inclusion bodies for therapeutic applications. *Microb Cell Fact* **11**, 76.

Vives-Rego, J., Lebaron, P. & Nebe-von Caron, G. (2000). Current and future applications of flow cytometry in aquatic microbiology. *FEMS Microbiol Rev* **24**, 429–448.

Vizcaino-Caston, I. (2012). *Monitoring bacterial physiology during recombinant protein production using reporter gene technology.* PhD Thesis, University of Birmingham.

Vizcaino-Caston, I., Wyre, C. & Overton, T. W. (2012). Fluorescent proteins in microbial biotechnology - new proteins and new applications. *Biotechnol Lett* **34**, 175–186.

Votyakova, T. V., Kaprelyants, A. S. & Kell, D. B. (1994). Influence of viable cells on the resuscitation of dormant cells in *Micrococcus luteus* cultures held in an extended stationary phase: The population effect. *Appl Environ Microbiol* **60**, 3284–3291.

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. & de Gier, J.-W. (2008). Tuning *Escherichia coli* for membrane protein overexpression. *Proc Natl Acad Sci USA* **105**, 14371–14376.

Waldo, G. S., Standish, B. M., Berendzen, J. & Terwilliger, T. C. (1999). Rapid protein-folding assay using green fluorescent protein. *Nat Biotechnol* **17**, 691–695.

Wall, J. & Solomon, A. (1999). Flow cytometric characterization of amyloid fibrils. *Methods Enzymol* **309**, 460-466.

Wall, J. S. (2002). Flow cytometric charecterization of amyloid fibrils. US Patent 6423270

Wallace, R. J. & Holms, W. H. (1986). Maintenance coefficients and rates of turnover of cell material in *Escherichia coli* ML308 at different growth temperatures. *FEMS Microbiol Letts* **37**, 317-320.

Wållberg, F., Sundström, H., Ledung, E., Hewitt, C. J. & Enfors, S.-O. O. (2005). Monitoring and quantification of inclusion body formation in *Escherichia coli* by multi-parameter flow cytometry. *Biotechnol Lett* **27**, 919–926.

Walls, D. & Loughran, S. T. (2011). Tagging recombinant proteins to enhance solubility and aid purification. *Methods Mol Biol* **681**, 151–175.

Walsh, G. (2010a). Biopharmaceutical benchmarks 2010. *Nat Biotechnol* **28**, 917–924.

Walsh, G. (2010b). Post-translational modifications of protein biopharmaceuticals. *Drug Discov Today* **15**, 773–780.

Walsh, G. & Jefferis, R. (2006). Post-translational modifications in the context of therapeutic proteins. *Nat Biotechnol* **24**, 1241–1252.

Wang, L., Maji, S. K., Sawaya, M. R., Eisenberg, D. & Riek, R. (2008). Bacterial inclusion bodies contain amyloid-like structure. *PLoS Biol* **6**, e195.

Wang, H., Nakata, E. & Hamachi, I. (2009a). Recent progress in strategies for the creation of protein-based fluorescent biosensors. *Chembiochem* **10**, 2560–77.

Wang, H., Wang, F. & Wei, D. (2009b). Impact of oxygen supply on rtPA expression in *Escherichia coli BL21* (DE3): ammonia effects. *Appl Microbiol Biotechnol* **82**, 249–259.

Wang, Y., Xuan, Y., Zhang, P., Jiang, X., Ni, Z., Tong, L., Zhou, X., Lin, L., Ding, J. & Zhang, Y. (2009c). Targeting expression of the catalytic domain of the kinase insert domain receptor (KDR) in the peroxisomes of *Pichia pastoris. FEMS Yeast Res* **9**, 732–741.

Wang, M., Weiss, M., Simonovic, M., Haertinger, G., Schrimpf, S. P., Hengartner, M. O. & von Mering, C. (2012). PaxDb, a database of protein abundance averages across all three domains of life. *Mol Cell Proteomics* **11**, 492–500.

Want, A., Thomas, O. R., Kara, B., Liddell, J. & Hewitt, C. J. (2009). Studies related to antibody fragment (Fab) production in *Escherichia coli* W3110 fed-batch fermentation processes using multiparameter flow cytometry. *Cytometry A* **75**, 148–154.

Want, A. J. (2010). *Physiological studies on bacterial fermentations using multi-parameter flow cytometry.* PhD thesis, University of Birmingham.

Westers, L., Westers, H. & Quax, W. J. (2004). *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochim Biophys Acta* **1694**, 299–310.

Wilks, J. C. & Slonczewski, J. L. (2007). pH of the cytoplasm and periplasm of *Escherichia coli*: rapid measurement by green fluorescent protein fluorimetry. *J Bacteriol* **189**, 5601–5607.

Williams, S. G., Cranenburgh, R. M., Weiss, A. M., Wrighton, C. J., Sherratt, D. J. & Hanak, J. A. (1998). Repressor titration: a novel system for selection and stable maintenance of recombinant plasmids. *Nucleic Acids Res* **26**, 2120–2124.

Wittmann, C., Kim, H. M., John, G. & Heinzle, E. (2003). Characterization and application of an optical sensor for quantification of dissolved O_2 in shake-flasks. *Biotechnol Lett* **25**, 377–380.

Wood, W. B. (1966). Host specificity of DNA produced by *Escherichia coli*: bacterial mutations affecting the restriction and modification of DNA. *J Mol Biol* **16**, 118–133.

Yamanaka, K. (1999). Cold shock response in *Escherichia coli. J Mol Microbiol Biotechnol,* **1,** 193-202.

Wyre, C. & Overton, T. W. (2014a). Flow cytometric analysis of *E. coli* on agar plates: implications for recombinant protein production. *Biotechnol Lett* **36**, 1485–1494.

Wyre, C. & Overton, T. W. (2014b). Use of a stress-minimisation paradigm in high cell density fed-batch *Escherichia coli* fermentations to optimise recombinant protein production. *J Ind Microbiol Biotechnol* **41**, 1391–1404.

Yu, Z. W. & Quinn, P. J. (1994). Dimethyl sulphoxide: a review of its applications in cell biology. *Biosci Rep* **14**, 259–281.

Yura, T. & Nakahigashi, K. (1999). Regulation of the heat-shock response. *Curr Opin Microbiol* **2**, 153–158.

Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M. G. & Alon, U. (2006). A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli. Nat Methods* **3**, 623–628.

Zawada, J. F., Yin, G., Steiner, A. R., Yang, J., Naresh, A., Roy, S. M., Gold, D. S., Heinsohn, H. G. & Murray, C. J. (2011). Microscale to manufacturing scale-up of cell-

free cytokine production--a new approach for shortening protein production development timelines. *Biotechnol Bioeng* **108**, 1570–1578.

Zhang, C., Xing, X.-H. H. & Lou, K. (2005). Rapid detection of a GFP-marked *Enterobacter aerogenes* under anaerobic conditions by aerobic fluorescence recovery. *FEMS Microbiol Lett* **249**, 211–218.

Zhang, J.-D. D., Li, A.-T. T. & Xu, J.-H. H. (2010). Improved expression of recombinant cytochrome P450 monooxygenase in *Escherichia coli* for asymmetric oxidation of sulfides. *Bioprocess Biosyst Eng* **33**, 1043–1049.

Zimmer, M. (2009). GFP: from jellyfish to the Nobel prize and beyond. *Chem Soc Rev* **38**, 2823–2832.

Chapter 7: Appendix 1 – Supplementary Data to Chapter 4
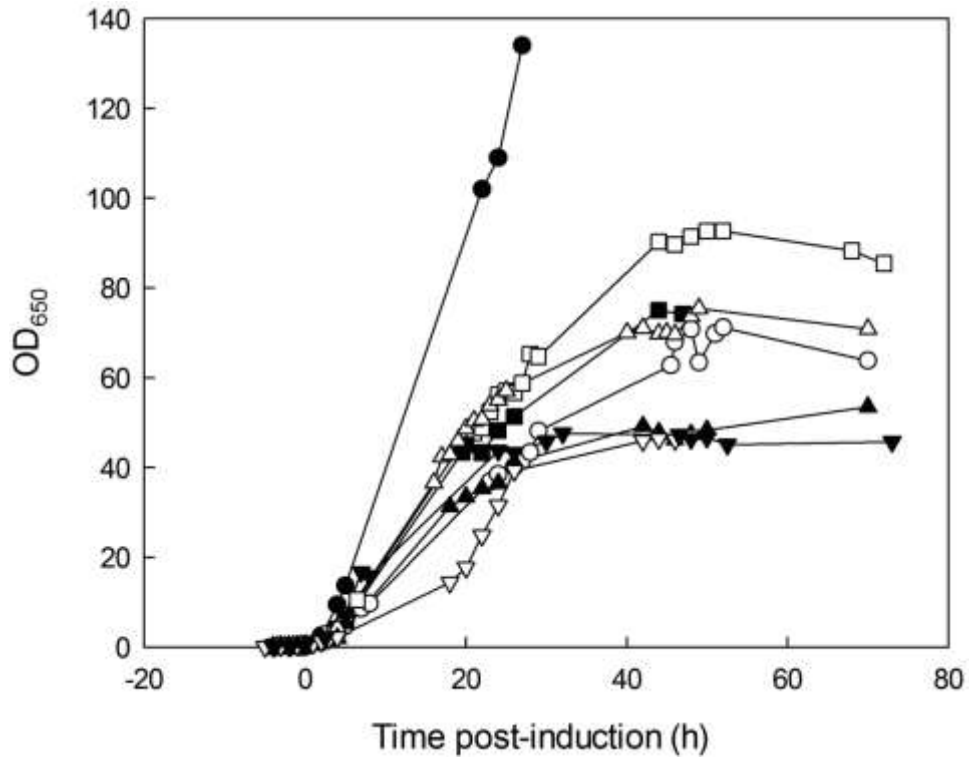
Figure 7.1: Culture growth data from repeat protocol A fermentations

OD_{650} data during additional protocol A fermentations showing an inability to replicate the high biomass yields of A(1) and A(2): Fermentation presented in Chapter 4 (○), fermentation using equipment and chemicals used in Sevastsyanovich *et al.* (2009) and Alfasi (2010) (in collaboration with L. Zaffaroni & Prof J.A. Cole, School of Biosciences, University of Birmingham) (fermentation terminated prematurely due to excessive culture foaming) (●), 7 replicate fermentations protocol A to that presented in Chapter 4 (□,■,△,▲,▽,▼).
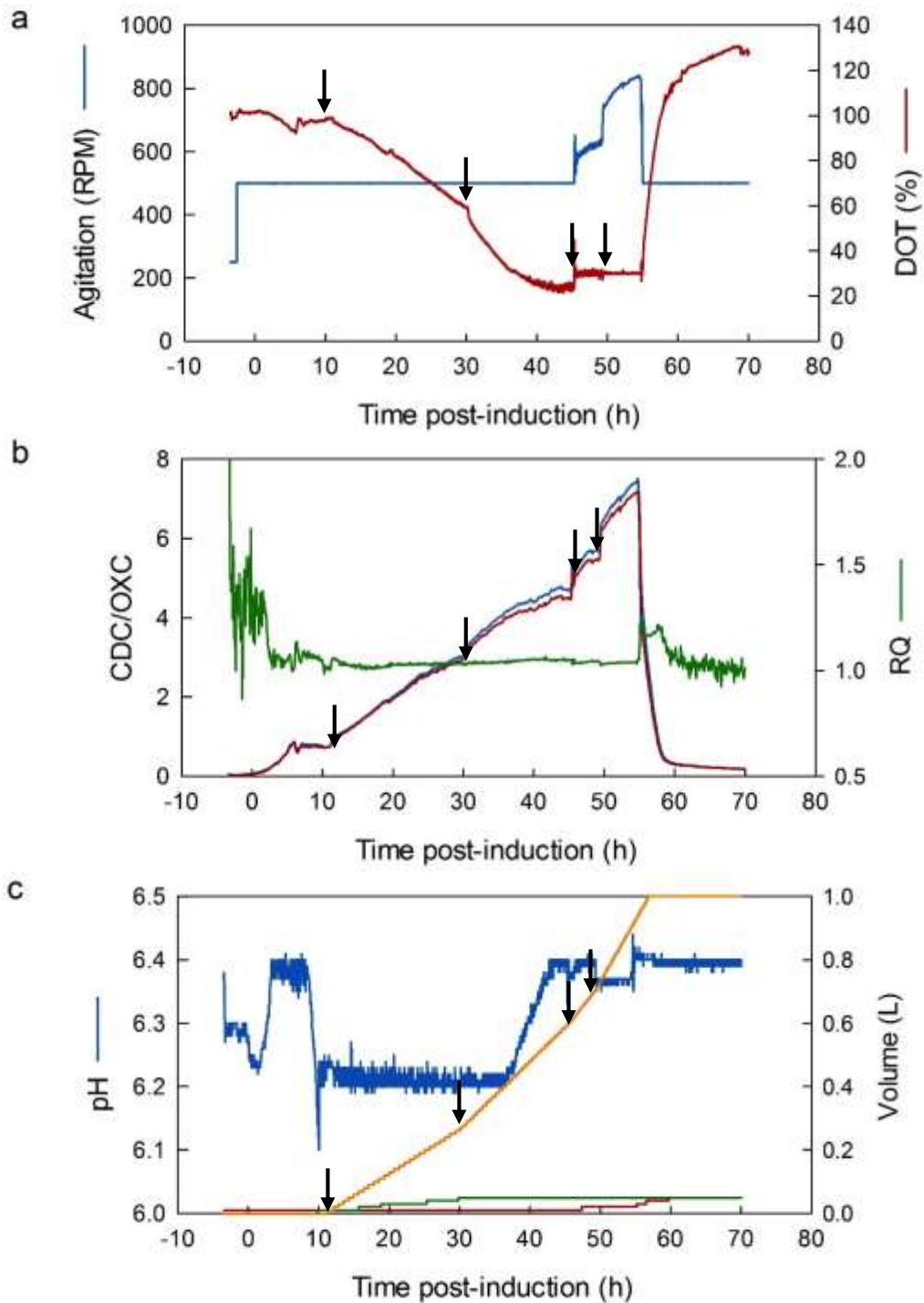
Figure 7.2: Online data from a protocol A fermentation (Section 4.2.1)
a) Agitation (blue) and DOT (red) data. b) Gas-MS data; CDC (blue) OXC (red), RQ (green). c) pH (blue), volumes of acid (red), base (green) and feed (orange) added to vessel. Arrows indicate points at which feed was turned on or rate of feeding increased.
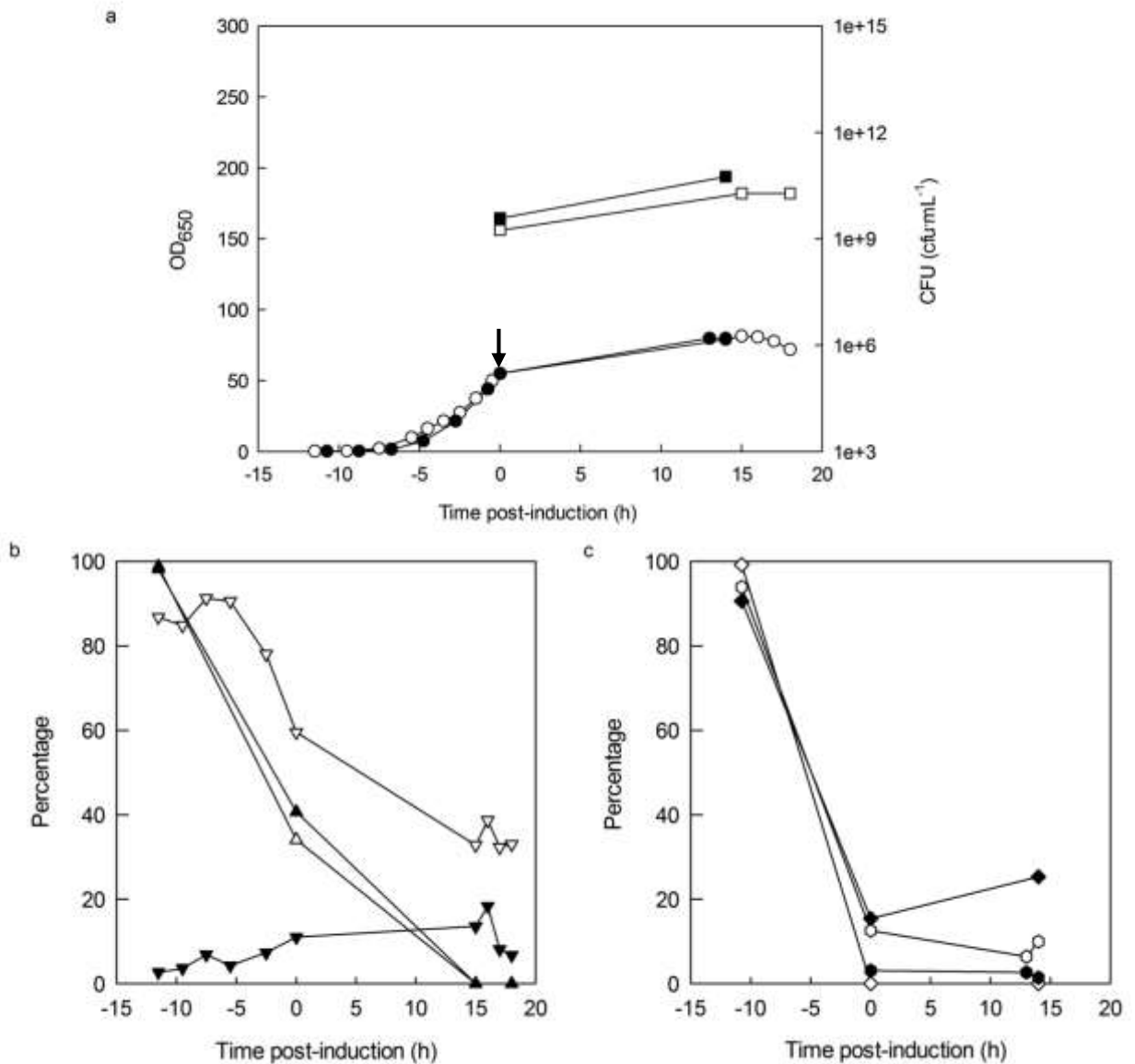
Figure 7.3: Culture growth, plasmid retention and viability data from 2 protocol B fermentations (Section 4.2.2)

a) Growth data: OD_{650} data for repeat 1 (○) and repeat 2 (●), CFU counts for repeat 1 (□) and repeat 2 (■). b) Plasmid retention and viability data for repeat 1: percentage colonies plasmid+ (△), percentage colonies GFP+ (▲), percentage cells GFP+ (by FCM) (▽) and percentage cells dead (PI+, by FCM) (▼). c) Plasmid retention and viability data for repeat 2: percentage colonies plasmid+ (◇), percentage colonies GFP+ (◆), percentage cells GFP+ (by FCM) (○) and percentage cells dead (PI+, by FCM) (◉) for repeat 2. Arrow indicates point at which feed was turned on.
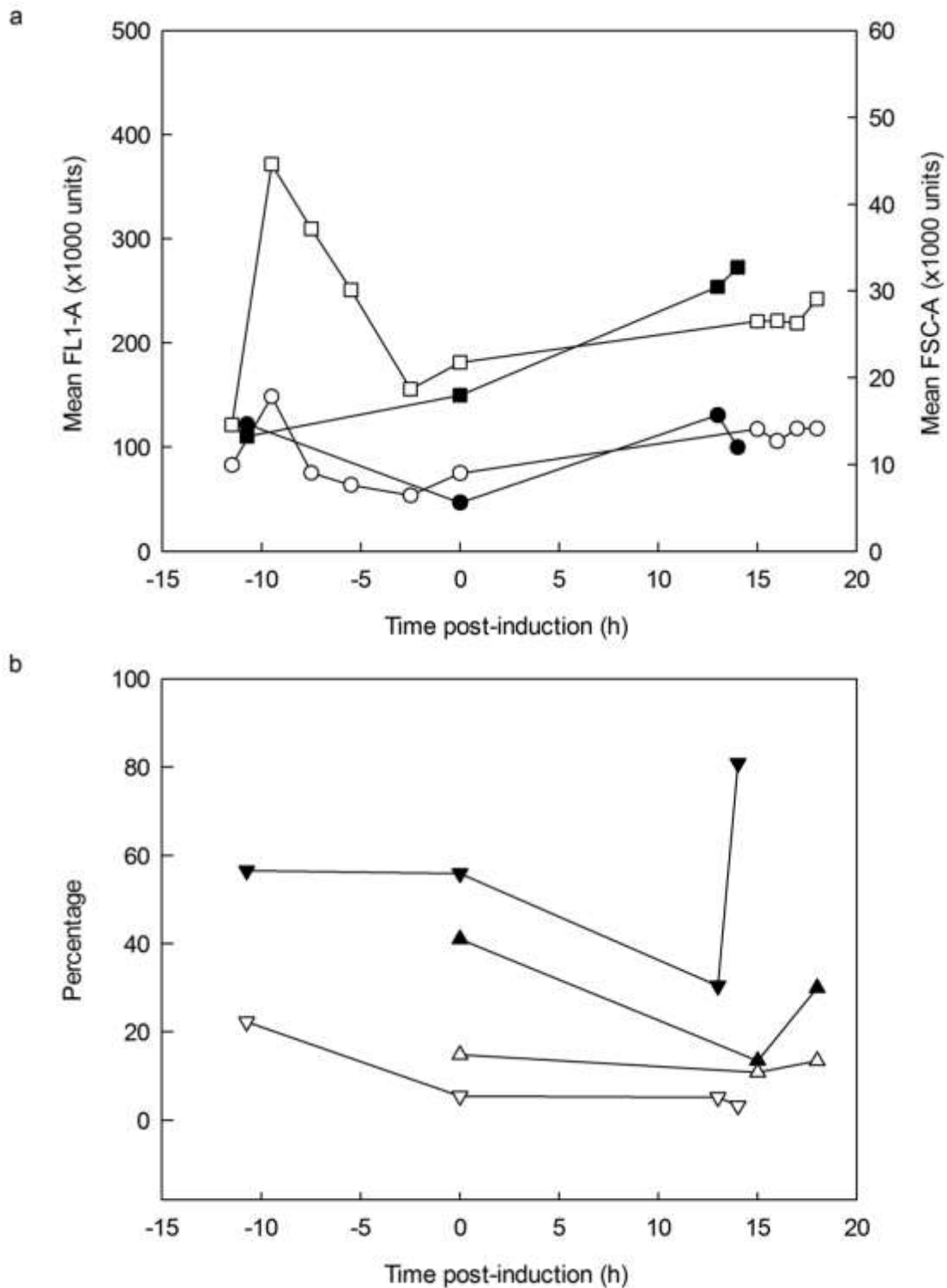
Figure 7.4: FCM and SDS-PAGE data from 2 protocol B fermentations (Section 4.2.2)

a) FCM data: Mean cellular green fluorescence (FL1-A) of GFP+ cells for repeat 1 (○) and repeat 2 (●), mean cellular forward scatter (FSC-A) of GFP+ cells for repeat 1 (□) and repeat 2 (■). b) SDS-PAGE data: Percentage of CheY::GFP that was soluble for repeat 1 (△) and repeat 2 (▲), percentage of total cellular protein that was CheY::GFP for repeat 1 (▽) and repeat 2 (▼).
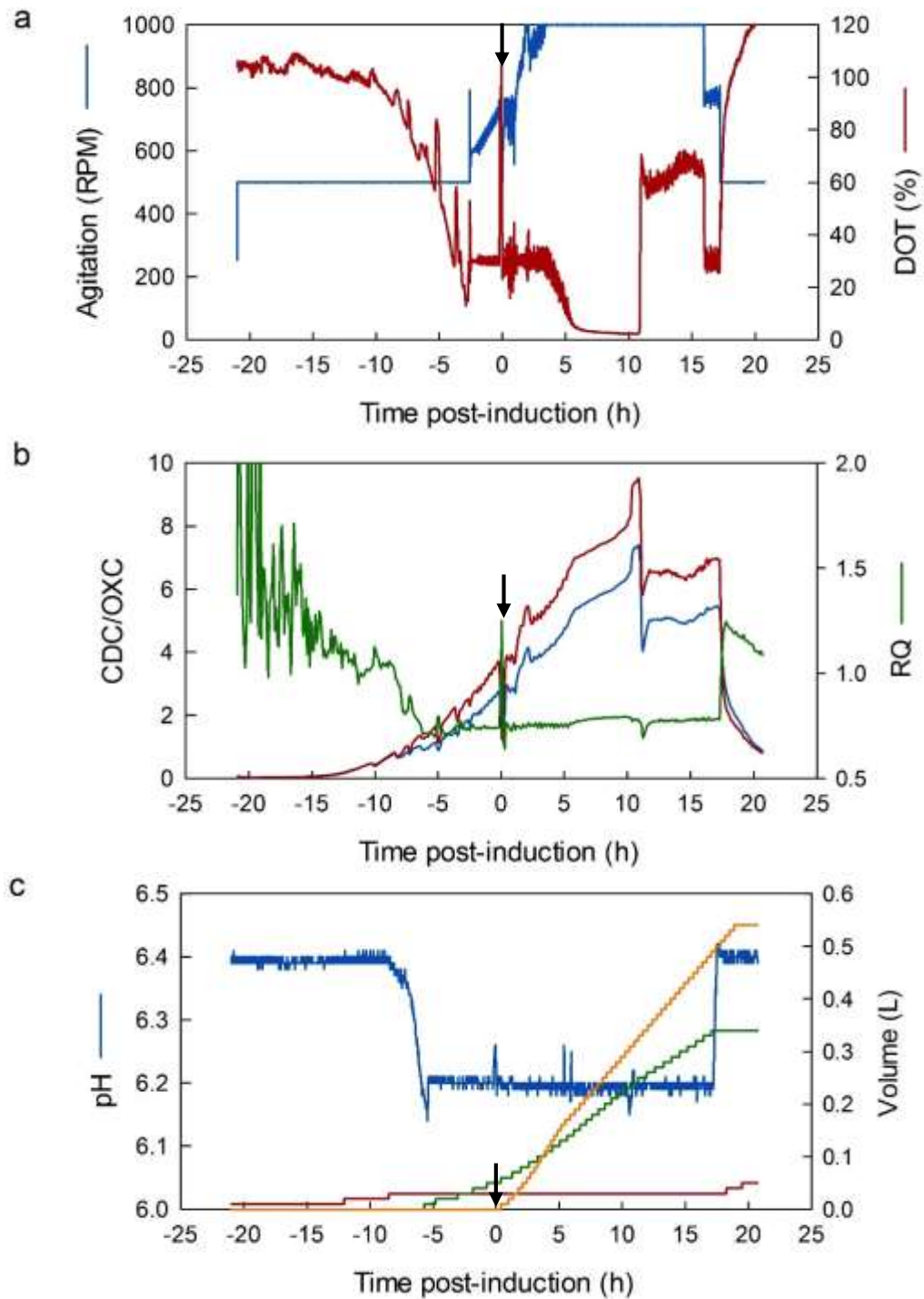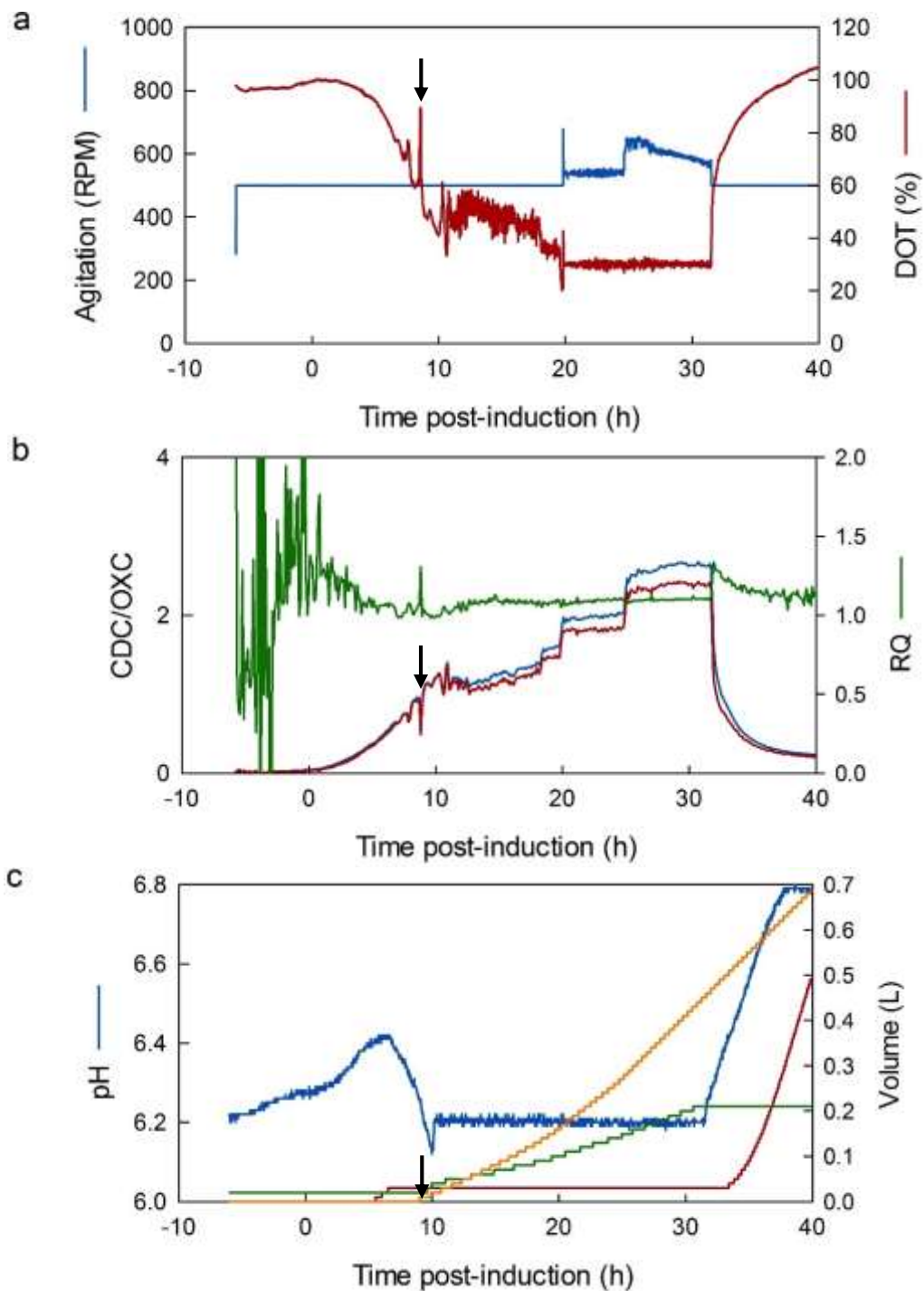
Figure 7.5: Online data from a protocol B1 fermentation (Section 4.2.3)
a) Agitation (blue) and DOT (red) data. b) Gas-MS data; CDC (blue) OXC (red), RQ (green). c) pH (blue), volumes of acid (red), base (green) and feed (orange) added to vessel. Arrows indicate point at which feed was turned on.
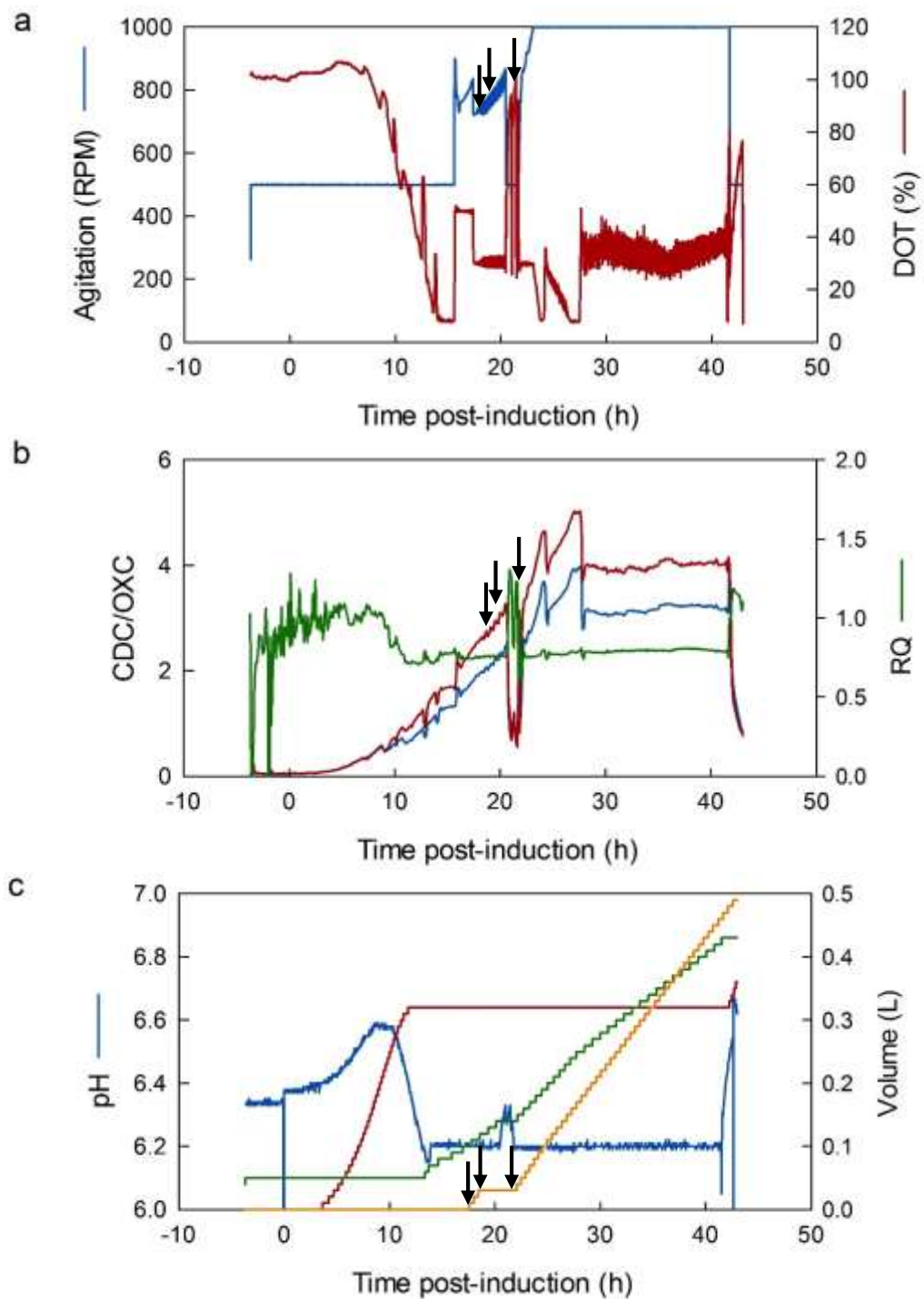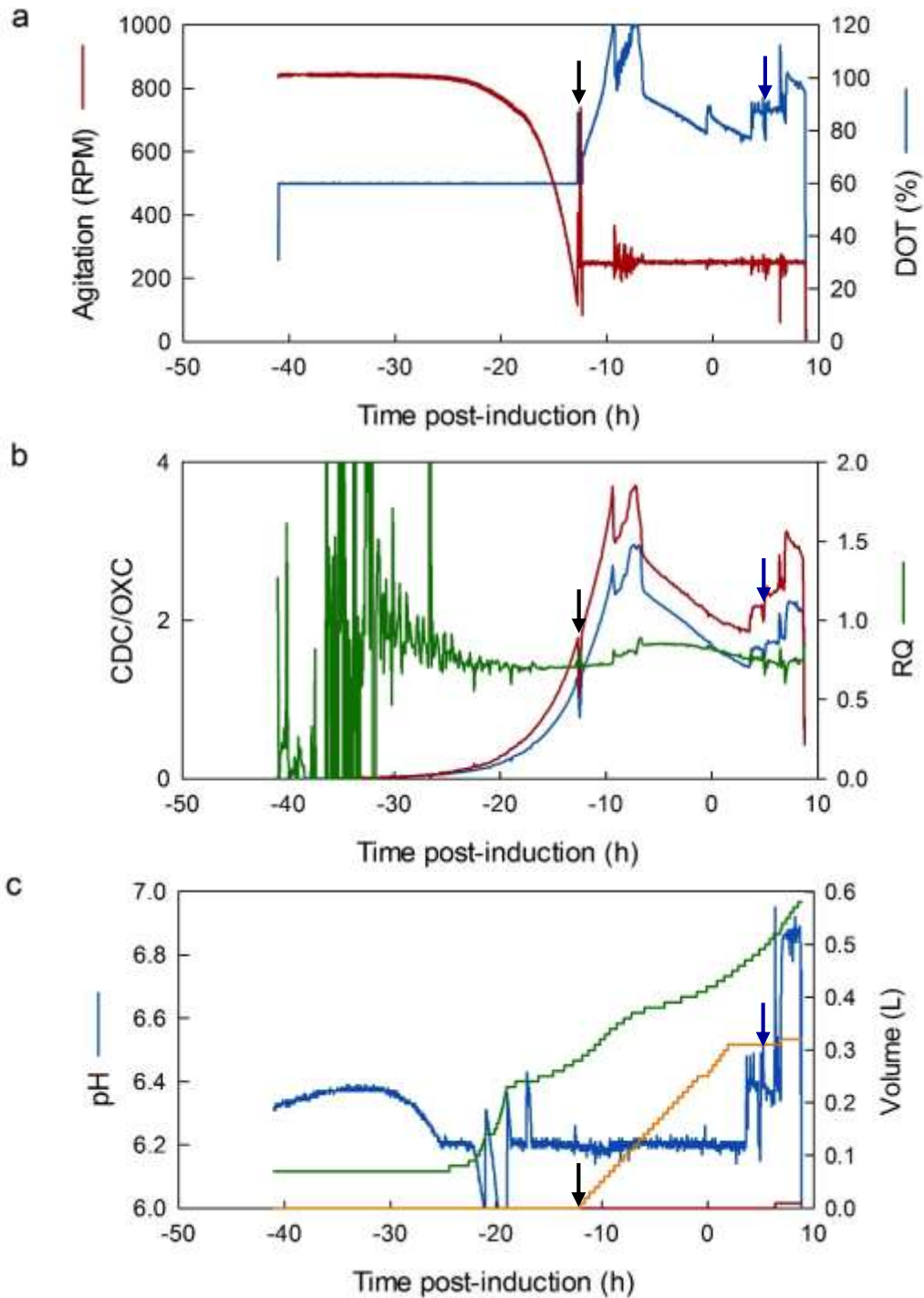
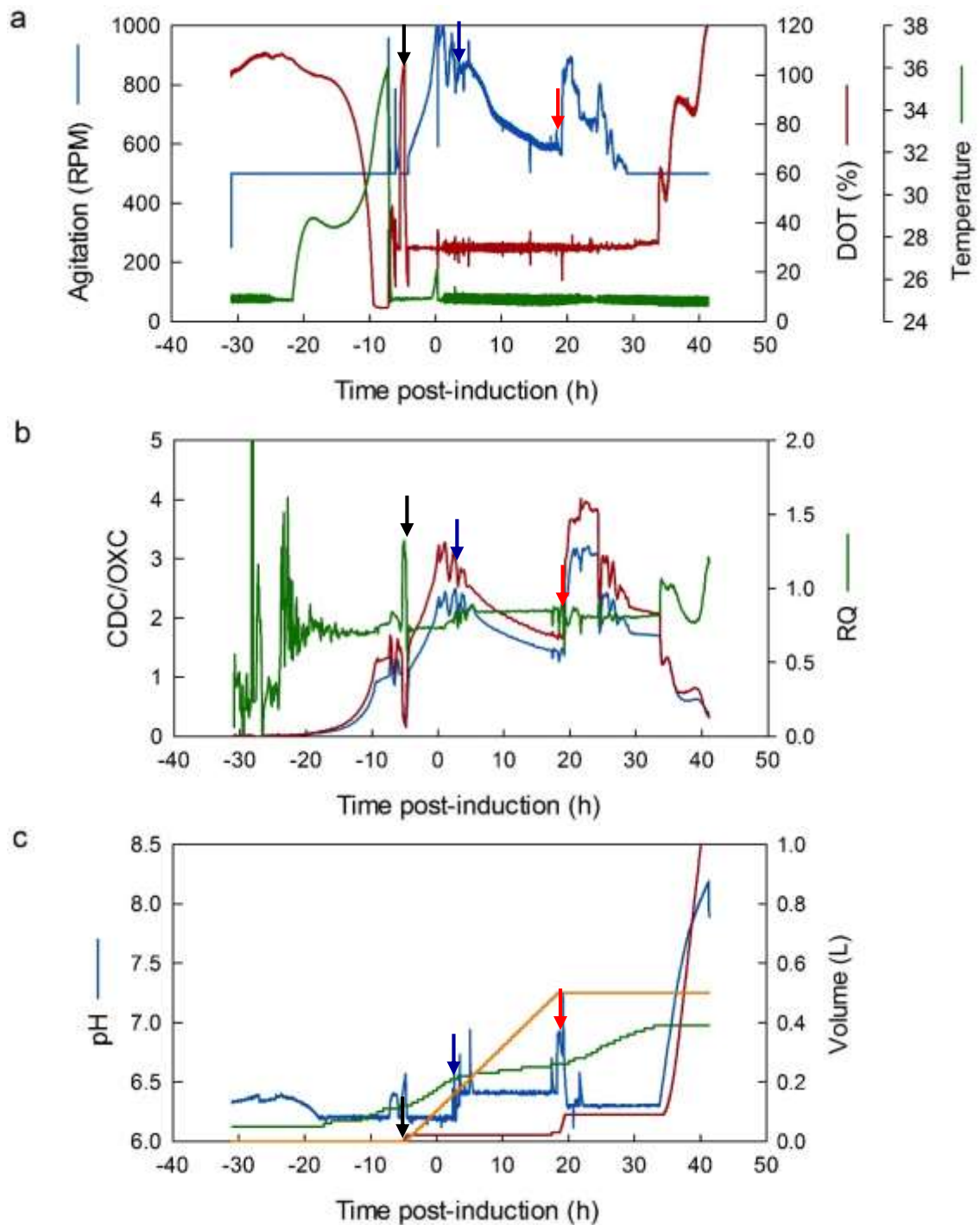**Figure 7.6: Online data from a protocol B2 fermentation (Section 4.2.4)**
a) Agitation (blue) and DOT (red) data. b) Gas-MS data; CDC (blue) OXC (red), RQ (green). c) pH (blue), volumes of acid (red), base (green) and feed (orange) added to vessel. Arrows indicate points at which the feed was turned on, paused and resumed.

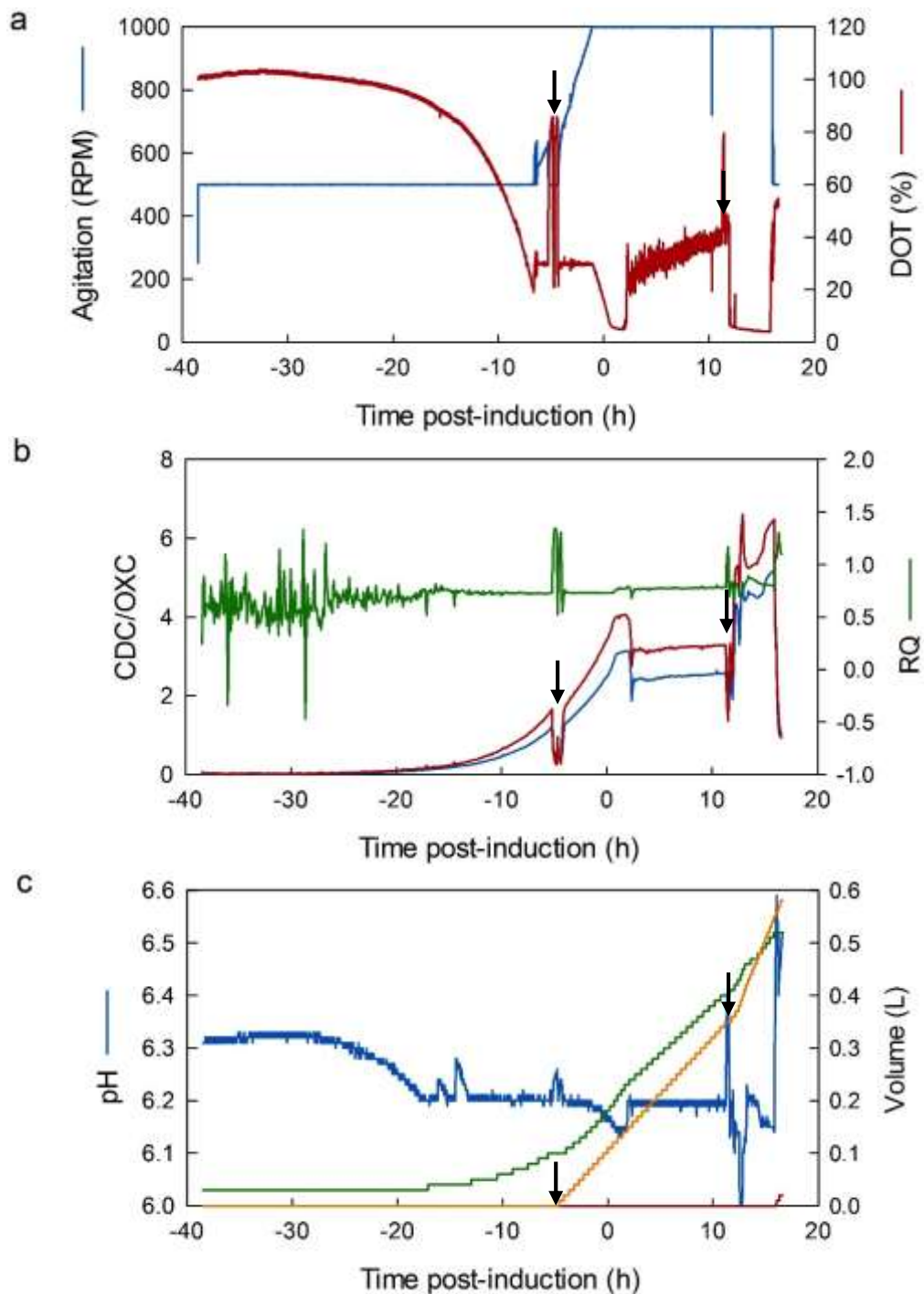**Figure 7.7: Online data from a protocol B3 fermentation (Section 4.2.6)**
a) Agitation (blue) and DOT (red) data. b) Gas-MS data; CDC (blue) OXC (red), RQ (green). c) pH (blue), volumes of acid (red), base (green) and feed (orange) added to vessel. Arrows indicate point at which feed was turned on.

Figure 7.8: Online data from a protocol C1 fermentation (Section 4.2.7.1)
a) Agitation (blue) and DOT (red) data. b) Gas-MS data; CDC (blue) OXC (red), RQ
(green). c) pH (blue), volumes of acid (red), base (green) and feed (orange) added to
vessel. Black arrows indicate point at which feed was turned on, blue arrows indicate
addition of 15 g $(NH_4)_2SO_4$.

Figure 7.9: Online data from a protocol C2 fermentation (Section 4.2.7.2)
a) Agitation (blue), DOT (red) and temperature (green) data. b) Gas-MS data; CDC (blue) OXC (red), RQ (green). c) pH (blue), volumes of acid (red), base (green) and feed (orange) added to vessel. Black arrows indicate point at which feed was turned on, blue arrows indicate addition of $(NH_4)_2SO_4$, red arrows indicate addition of 20 g Na_2HPO_4 and 10 g NaH_2PO_4.

Figure 7.10: Online data from a protocol C3 fermentation (Section 4.2.7.3)
a) Agitation (blue) and DOT (red) data. b) Gas-MS data; CDC (blue) OXC (red), RQ (green). c) pH (blue), volumes of acid (red), base (green) and feed (orange) added to vessel. Arrows indicate points at which feed was turned on or rate of feeding increased.
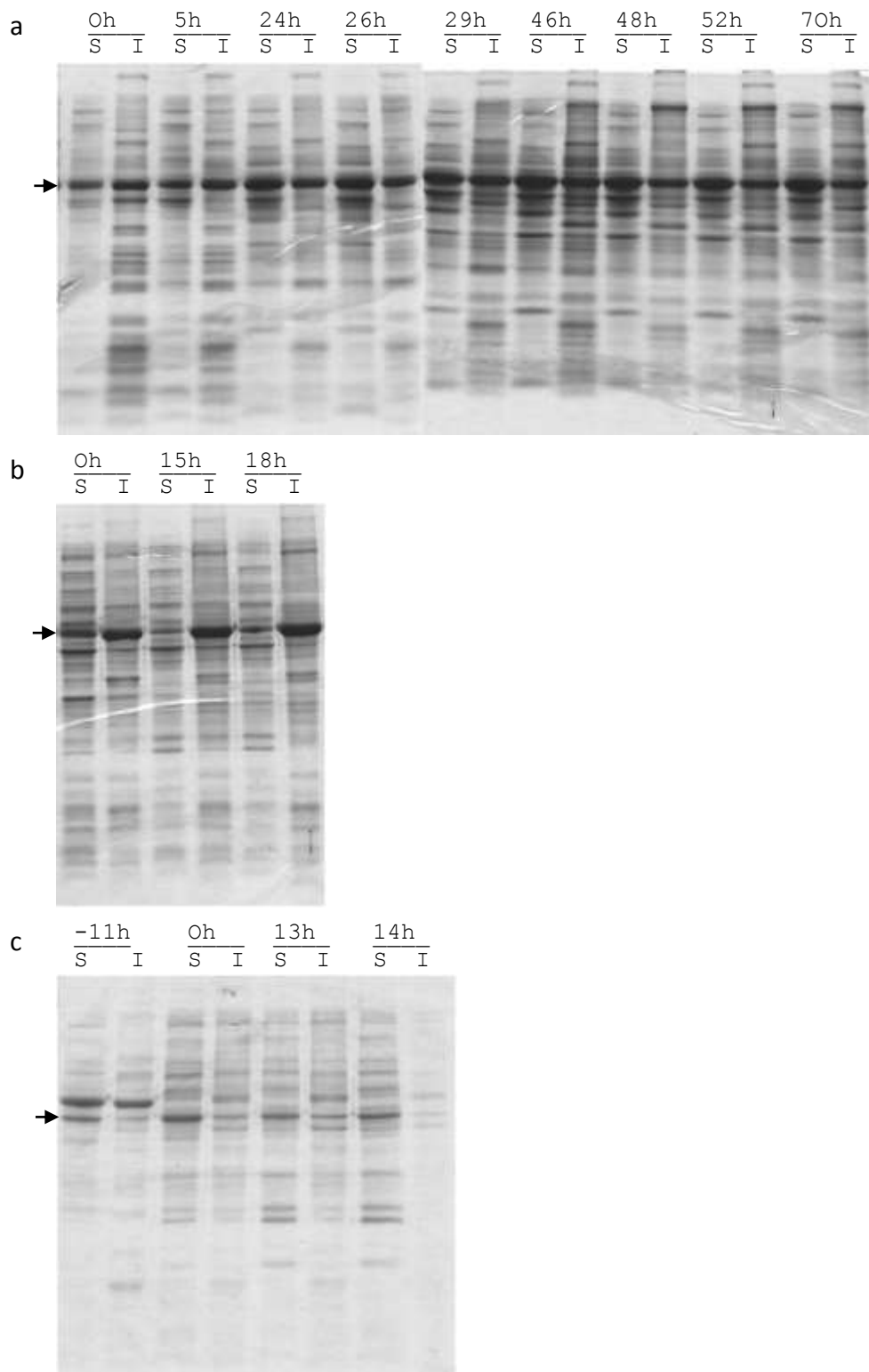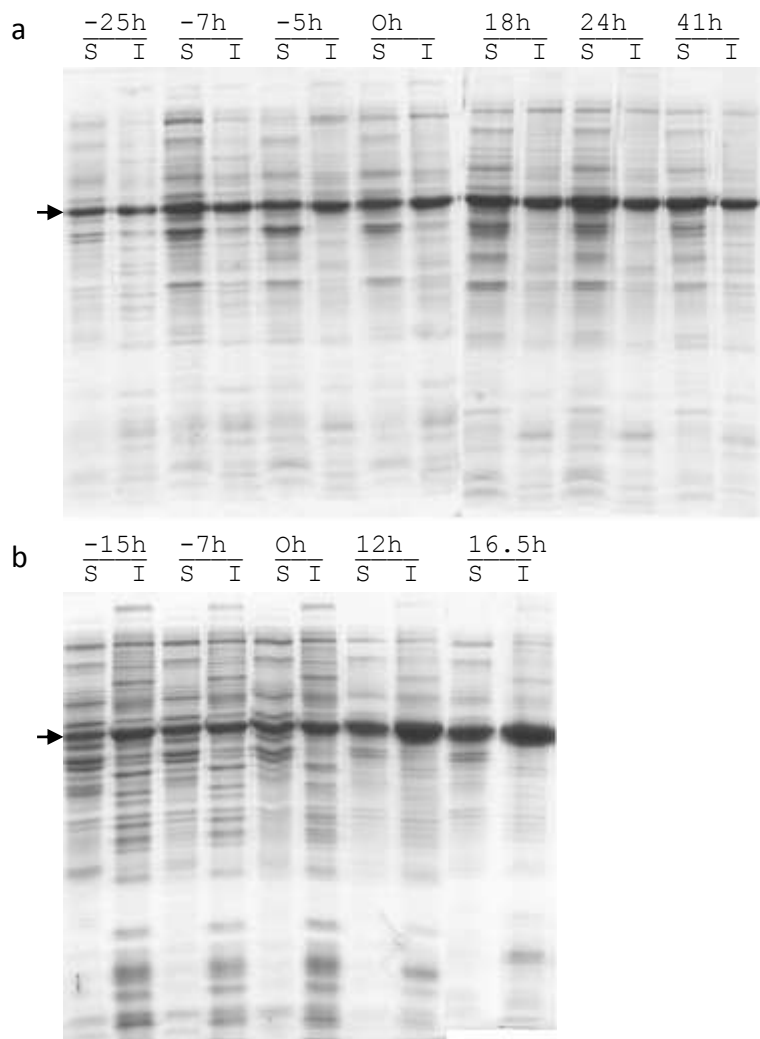
Figure 7.11: SDS-PAGE gel images from protocol A & B fermentations
SDS-PAGE gel images of BugBuster®-fractionated fermentation cell samples, 'S' and 'I' indicate soluble and insoluble fractions respectively. Time of sampling is indicated above lanes, all times in h post-induction. Arrows indicate approx. 42 kDa. a) Protocol A (Section 4.2.1). b) Protocol B replicate 1 (Section 4.2.2). c) Protocol B replicate 2 (Section 4.2.2).
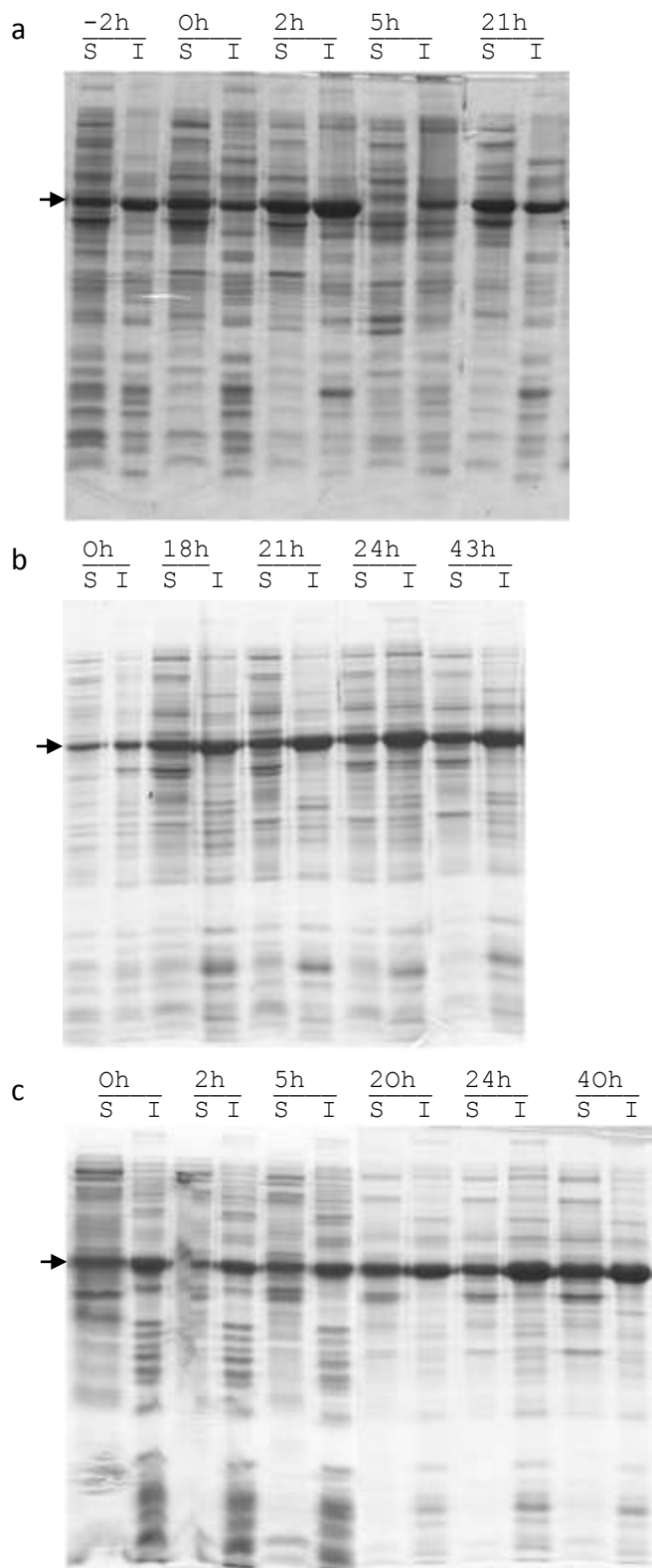
Figure 7.12: SDS-PAGE gel images from protocol B1-B3 fermentations
SDS-PAGE gel images of BugBuster®-fractionated fermentation cell samples, 'S' and 'I' indicate soluble and insoluble fractions respectively. Time of sampling is indicated above lanes, all times in h post-induction. Arrows indicate approx. 42 kDa. a) Protocol B1 (Section 4.2.3). b) Protocol B2 (Section 4.2.4). c) Protocol B3 (Section 4.2.4).

Figure 7.13: SDS-PAGE gel images from protocol C2 & C3 fermentations
SDS-PAGE gel images of BugBuster®-fractionated fermentation cell samples, 'S' and 'I' indicate soluble and insoluble fractions respectively. Time of sampling is indicated above lanes, all times in h post-induction. Arrows indicate approx. 42 kDa. a) Protocol C2 (Section 4.2.7.2). b) Protocol C3 (Section 4.2.7.3).

Chapter 8: Appendix 2 - Observations Regarding Excessive

Foaming in Fermentations

8.1. Introduction

A frequent problem observed with the model BL21*/CheY::GFP-RPP system used in this work is that it is highly prone to excessive culture foaming, often sufficient to necessitate termination of the experiment (terminal foaming) due to blockage of the exhaust filter and leakage of culture from the vessel. The reason for this is at present unknown. The occurrence of terminal foaming during this study ceased following the replacement of silicon antifoam by PPG, but prior to this a phenomenon of potential interest occurred. The majority of terminal foaming incidents occurred during the night while the vessel was not directly monitored and due to the absence of a foam sensor it was not possible to determine exactly when foaming occurred in order to correlate the incident to on-line data and hence determine a cause. On a single instance however, terminal foaming occurred while the vessel was being monitored, allowing not only a sufficiently limited time-frame to examine on-line data but also off-line data before and after the foaming incident.

8.2. Results & Discussion

E. coli BL21* pETCheY::GFP was grown in a 1.5 L fed-batch fermentation (protocol B3, silicone antifoam), terminal foaming was observed to have occurred between 37 and 41 h post-induction when the fermentation was terminated. Samples were taken during the fermentation, including before and after foaming at 37 and 41 h and analysed for OD_{650} and fluorescence (by fluorimetry). While there was a small decrease in OD the most dramatic changes occurred in fluorescence (Figure 8.1a,b). Following foaming, specific fluorescence (fluorescence per unit OD) had almost tripled and background fluorescence (absolute fluorescence of 0.22 μm-filtered culture broth) increased by over 60%, suggesting rapid accumulation of mature GFP and leakage of protein into
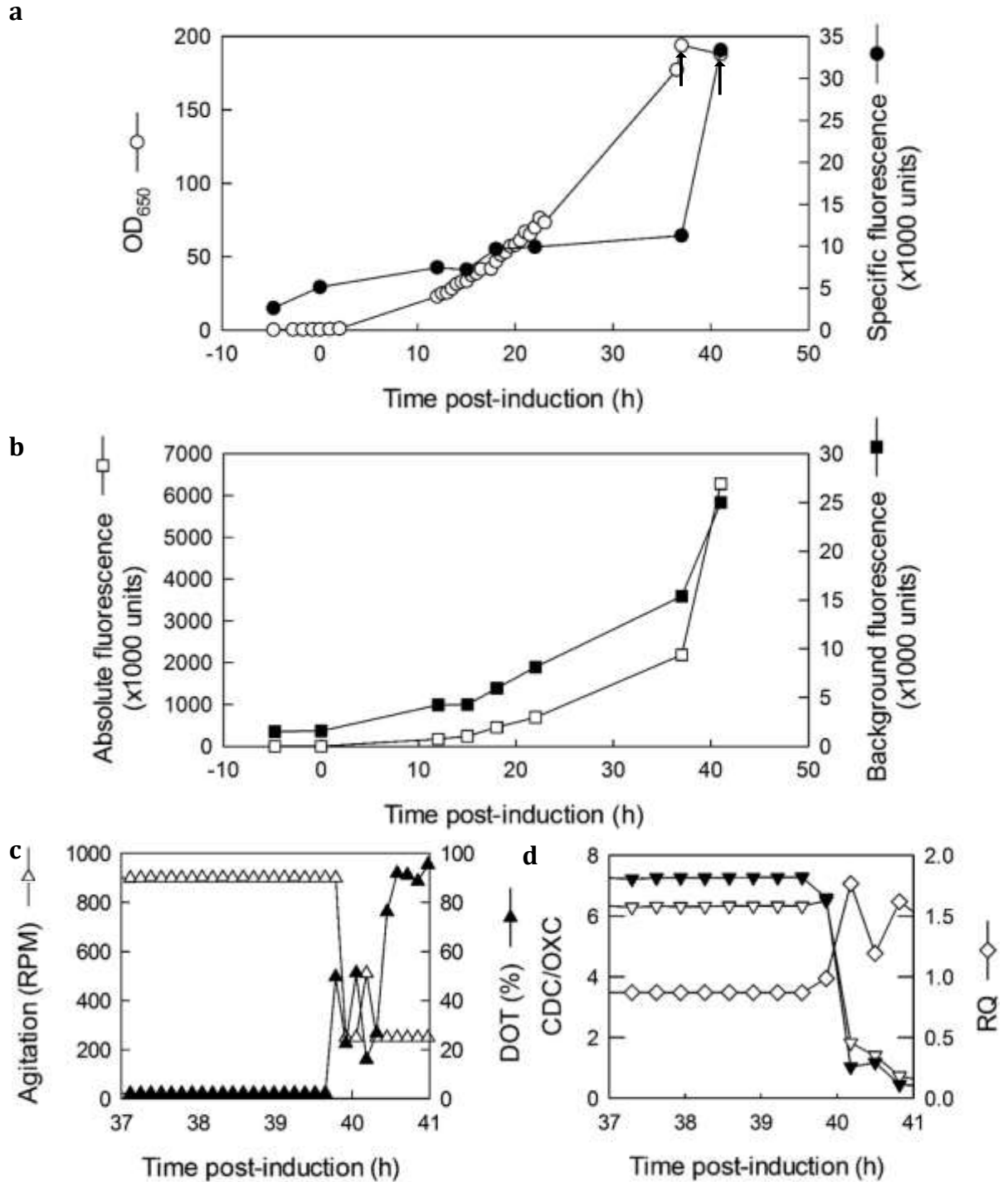
Figure 8.1: Data from terminal foaming of a fermentation
a) OD650 (○) and specific fluorescence (fluorescence per unit OD650) (●) throughout the fermentation. b) Background fluorescence (□) and absolute fluorescence (■) throughout the fermentation. c) Agitation (△) and DOT (▲) data during foaming incident. d) CDC (▽), OXC (▼) and RQ (◇) data during foaming incident. Sample points before and after foaming incident are indicated by arrows.

the extracellular milieu. This coincided with a rapid increase in the DOT (and consequent decrease in agitation rate along with decreases in CDC and OXC readings) following the addition of the entire 0.5 L feed (at approximately 36 h), which is presumed to be indicative of exhaustion of carbon-source and cessation of growth (Figure 8.1c,d).

Based on these data, the following hypothesis to explain the high frequency of terminal foaming for the BL21*/CheY::GFP model RPP system was proposed. Increasing the agitation rate can cause foaming (Junker, 2007) but is unlikely to be the cause in this case. Prior to foaming the DOT was exhausted due to the high oxygen demands of the culture, but on consumption of the carbon source oxygen demands decreased and the DOT rapidly increased. Although it is acknowledged that this does possibly contradict the findings in Section 2.10.2 if sufficient imGFP did in fact form during the oxygen-limited period, then rapid oxygenation could have triggered a burst of GFP maturation. This is supported by the dramatic increase in fluorescence observed following reoxygenation as it is unlikely as large an increase would result from *de novo* CheY::GFP synthesis alone. As GFP maturation produces hydrogen peroxide (Figure 1.7) (Tsien, 1998) a burst of GFP maturation could result in an equivalent pulse of intracellular H_2O_2. If the pulse of H_2O_2 is sufficient to trigger oxidative stress and cell death/lysis, this would result in an increase in extracellular protein concentration that is known to cause foaming in later-stage bioreactor cultures (Junker, 2007). This is supported by the increase in fluorescence of the culture medium, as this suggests an increase in leakage of cytoplasmic proteins.

Section 2.10.2 concluded that insufficient immature CheY::GFP accumulated during oxygen-limited HCDC to require routine additional sample processing by AFR, whereas the data presented here suggests the accumulation of large quantities of immature

CheY::GFP. It is thought that this discrepancy is likely to result from variability in the oxygen-limited phase such as the length of oxygen-limitation, the rate of CheY::GFP production, cell density and the growth rate. As such it is concluded that while routine use of AFR would be discouraged the possibility that accumulation of immature CheY::GFP may occur should not be discounted. If this should occur it is likely that it could be identified, as here, by comparison to the final measurement where the DOT had returned to ~100%.

8.3. Conclusions & further work

As this phenomenon was only observed on a single instance it was not possible to investigate further and therefore make any firm conclusions. This does, however provide a potential route of further study, dependent on replication of the conditions and therefore the foaming incident. Quantification of PI[+] cells by FCM of samples pre and post-foaming ought to identify any increase in cell death and comparison of absolute cell counts may allow determination of cell lysis. It may be possible to directly detect intracellular H_2O_2 through use of ROS-responsive fluorescent dyes such as the CELLROX® Deep Red and Orange Reagents (Life Technologies Corp.) (other ROS-responsive dyes such as H_2DCFDA would not be of use in this case as their fluorescence spectra overlap that of GFP) (Johnson & Spence, 2010). It may also be possible to test for a cellular response to the presence of H_2O_2. *E. coli* responds to peroxide stress by production of catalase (KatG), which can be detected and quantified biochemically and either gene reporter studies (such as the pUA66 *katG* reporter (Zaslaver *et al.*, 2006)) or qRT-PCR could be used to detect upregulation of genes involved in the oxidative stress response.

Chapter 9: Appendix 3 – Publications Derived from this Work

3 journal articles were published durng the course of this work, containing elements of Chapters 1, 3 & 4.

Chapter 1:
  Vizcaino-Caston, I., Wyre, C. & Overton, T. W. (2012). Fluorescent proteins in microbial biotechnology--new proteins and new applications. *Biotechnol Lett* **34**, 175–86. doi: 10.1007/s10529-011-0767-5.

Chapter 3:
  Wyre, C. & Overton, T. W. (2014a). Flow cytometric analysis of *E. coli* on agar plates: implications for recombinant protein production. *Biotechnol Lett* **36**, 1485–94. doi: 10.1007/s10529-014-1511-8.

Chapter 4:
  Wyre, C. & Overton, T. W. (2014b). Use of a stress-minimisation paradigm in high cell density fed-batch *Escherichia coli* fermentations to optimise recombinant protein production. *J Ind Microbiol Biotechnol* **41**, 1391–404. doi: 10.1007/s10295-014-1489-1.