OPTIMISATION OF RECOMBINANT PROTEIN PRODUCTION IN *Escherichia coli*

by

NURUL ASMA HASLIZA ZULKIFLY

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 Science

University of Birmingham

September 2023

ABSTRACT

Currently, recombinant protein production (RPP) has become more crucial in biopharmaceutical industry and *E. coli* has become the bacteria of choice. This research has been done to find the optimum condition for RPP using *E. coli* as the host, and looking for an opportunity to be applied industrially. Green fluorescent protein (GFP) was used as a marker to indicate the production of green fluorescence-proteins, which in this study, was applied to TNFα-GFP. TNFα-GFP is recombinant human TNFα fused to GFP. Two promoter systems were used, namely pORT(T7) plasmid which was inserted into bacterial host BL21(T7) and p*araBAD* plasmid which was recombined into BL21(A) strain. Arabinose was selected to be used as the inducer in this study to enhance protein production. The parameters investigated in this study were the carbon source and its concentration (glycerol and glucose), the growth temperature (30°C and 37°C), and the concentration of inducer (arabinose). Flow cytometry was routinely used to monitor the quantity and folding state of the TNFα-GFP fusion and cell physiology and viability on a single cell level. Further analysis identified three subpopulations with different fluorescence intensity and productivities, existed in the fermentation medium when using pORT(T7) system. In contrast to p*araBAD* system which had shown the production of only one single population throughout the fermentation. The optimum condition for RPP measured from this study was the usage of glycerol with the concentration of 0.4% (v/v), incubated at 30°C and arabinose concentration induced was 0.015%. The optimum conditions obtained in these experiments were used for the production of GFP and TNFα-GFP in fed-batch fermentation at 5 L scale. It was concluded that with further adaptation, it is possible to use these optimal conditions to be applied for RP production on an industrial scale. The present study also involved the production of soluble and functional Fumarase B (FumB) enzyme, which was achieved via the use of the restriction-free (RF) cloning method to insert the *fumB* gene into the pET41c plasmid.

DEDICATION

Dedicated with lots of loves and thanks;

Mohd Faisal bin Sulong, Hubby

Nurul Afia Faziyya binti Mohd Faisal, Daughter

Muhammad Afif Fawwaz bin Mohd Faisal, Son

Zulkifly bin Ahmad, Abah

Wan Zaliti binti Wan Ngah, Mama

ACKNOWLEDGEMENT

First and foremost, thanks to God Almighty that has given me hope, strength and wisdom to complete my thesis successfully.

I also would like to express my thanks to my supervisor, Dr. Tim Overton for his outstanding guidance, patience, support, supervision, cooperation as well as his precious knowledge throughout my study.

Later, I would like to thank to all my colleagues in the lab who help and support me during my academic year, especially Ikhlaas Kasli, James Leech, Iqram Mohd Fauzi and Gina. Not to forget, I thank Elaine Mithchell, David French, Ronnie Baglin and Janine Fenton for their exceptional technical assistance in the lab. From the School of Biosciences, I would like to thank Prof Jeff Cole and his team for their visions and support.

Dedicated with many loves and thanks to my husband, Mohd Faisal bin Sulong, my daughter, Nurul Afia Faziyya, my son, Muhammad Afif Fawwaz, and my family especially my parents, Zulkifly bin Ahmad and Wan Zaliti binti Wan Ngah as well as my father-in-law, Sulong bin Mohd, and all my beloved siblings and siblings-in-law. Thank you so much for being very patience waiting on my success and thanks for all the supports and prayers.

From financial perspective, I would like to acknowledge the support of the Ministry of Higher Education of Malaysia (Malaysia's MoHE) who funded my PhD, a chance which not many people can have it. Also, not to forget Universiti Sultan

Zainal Abidin (UniSZA) who funded the allowance and expenses fees for me during my study in University of Birmingham.

Special thanks to Cobra Biologics for supplying me the plasmid to work on in this project, and University of Birmingham for providing me with a nice place to carry out my research and providing all the equipment required, and more important, thank you to all the lovely people who work there.

Last but not least, I must thank my whole family who giving me endless support throughout my study. Without them, I could not have done this.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

µM	Micromolar
APS	Ammonium per sulphate
Ara	Arabinose
BOX	Bis-oxonol
bp	Base pairs
BPS	Blue protein standard
CFU	Colony forming units
DCW	Dry cell weight
ddH_2O	Double-distilled water
DNA	Deoxyribonucleic acid
DO	Dissolved oxygen
DOT	Dissolved oxygen tension
DSB	Disulphide bond
DSP	Downstream processing
FACS	Fluorescence-activated cell sorting
FCM	Flow cytometry
FDA	United States Food & Drugs Administration
FM	Fluorescence microscopy
FP	Fluorescent protein
FSC	Forward scatter
g/L	Gram per litre
GF	Green fluorescence
GFP	Green fluorescent protein
GST	Glutathione S-transferase
h	hour/ hours
HCDC	High cell density culture
hGH	Human growth hormone
IAA	Indoleacrylic acid
IB/IBs	Inclusion body/bodies

IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
LB	Luria Bertani
mAb/ Mab	Monoclonal antibody
mg/mL	Milligram per millilitre
min	Minutes
NA	Nutrient agar
$OD_{xxx\ nm}$	Optical density at XXX nanometre
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PI	Propidium iodide
PPG	Polypropylene glycol
PTM	Post-translational modification
RF	Restriction-free
rhTNFα	Recombinant human Tumour necrosis factor alpha
RNA	Ribonucleic acid
RNAP	RNA polymerase
RP	Recombinant protein
rpm	Revolutions per minute
RPP	Recombinant protein production
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SSC	Side scatter
TB	Terrific Broth
TCP	Total cellular protein
TEMED	Tetramethyl ethylenediamine
TNFα	Tumour necrosis factor alpha
VBNC	Viable but non-culturable cells
µg/mL	Microgram per millilitre
≈	Approximately

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

This chapter presents an overview of the scientific literature studies on recombinant protein production (RPP) using *Escherichia coli* as the host strain, its optimisation conditions as well as scale-up strategies approached utilising bioreactors to synthesize the highest amount of green fluorescence protein (GFP). The procedures applied in this research study are also discussed and ultimately, the research objectives are described.

1.1 Recombinant protein production in the biopharmaceutical industry

Biopharmaceuticals offer great achievements of modern science. These drugs are progressively being utilised in almost all disciplines of medicine. Since its introduction in 1982, biopharmaceutical drugs have revolutionized the way many various diseases are treated, such as those used to treat cancers and metabolic disorders. Due to the enormous demand for these drugs, the market for biopharmaceuticals has expanded much faster recently than the entire drug market and is thought to have significant potential for further dynamic growth (Kesik-Brodacka, 2018).

Proteins are structural components of biological assemblies that play important roles in intracellular and intercellular interactions. They act as catalysts for metabolic reactions primarily responsible for essential cell signalling events in the body. Hence, the lack of specific polypeptide production in the body generally results in mild to severe pathological disorders. These diseases are usually cured through the clinical administration of externally sourced proteins to attain normal systemic or tissue level concentrations. Although most human proteins have high

commercial value and pharmaceutical importance, these proteins are not naturally obtainable from the human body (Ferrer-Miralles et al., 2009). Unlike most of other pharmaceutical, proteins cannot be synthesized chemically due to their complex structure and function. Therefore, these proteins are produced in biological processes, normally inside microorganisms as host.

Drugs typically refer to low molecular weight molecules, while biopharmaceuticals denote drugs with high molecular weight and are characterised by nucleotide (DNA or RNA) or amino acid (proteins and peptides) polymers. DNA vaccines and gene therapies are some examples of biopharmaceutical products that contain nucleic acids and have good prospects in the biopharmaceutical industry. To date, only a handful of nucleic acid-based drugs have been clinically approved as therapeutic drugs used as medicine (Jozala et al., 2016). This study focuses on proteins and peptides as they represent a major source of biopharmaceuticals (Walsh, 2018).

1.1.1 An overview of RPP

Recombinant DNA is those specific protein-encoding DNA that has been prior modified either by combining or has been engineered, thus the proteins resulting from such modification are called recombinant proteins. Genetic engineering allows these DNA-encoding genes to be introduced into specific plasmids for cloning. Such transgenic cells would play a role as the host and act as bio-factories to express recombinant genes under optimal conditions and subsequently secrete recombinant proteins (Rodríguez-Carmona et al., 2010).

In the late 1970's, recombinant DNA (rDNA) technologies were developed using *Escherichia coli* (*E. coli*) as a biological system. This technology offers an efficient platform to manufacture the desired human proteins on an impressive scale with a reasonably inexpensive approach. Additionally, this technology uses available microbial cells such as bacteria and yeasts, which can be cultivated through uncomplicated procedures and appliances (Khan et al., 2016).

Boyer and Robert Swanson founded Genentech in 1976 (Kaur et al., 2018), which subsequently commercialised synthetic recombinant human insulin in 1978, produced by genetically modified *E. coli* (Goeddel et al., 1979). In 1982, the United States Food and Drugs Administration (FDA) authorised the commercial production of Humulin® recombinant insulin expressed in *E. coli* for diabetes treatment. It was the pioneer recombinant pharmaceutical product that was sold to the pharmaceutical industry prior to the production of over 400 other recombinant proteins (Mühlmann et al., 2017).

Pharmaceutical, biotechnological, and research applications

The adaptability and the potential expansion of recombinant protein production (RPP) create new and diverse marketable drug products for biopharmaceutical industries. After the approval of the recombinant human insulin, new recombinant DNA products have been commercially available in the market and this has led to the enhancement of numerous heterologous protein expression systems. This advancement has created various strains derived from numerous microbial species for protein production (Jozala et al., 2016). For most of the recombinant

proteins available in the biopharmaceutical market, *E. coli* is the main source of microbial species used for recombinant protein production.

The production of protein-based biopharmaceuticals such as insulin is one of the most recognised applications of RPP. Insulin was approved for diabetic treatment using *E. coli* as the expression host in 1982. Later, in 1994, the bovine growth hormone (bGH) was endorsed by FDA and it set a new model for the recombinant DNA production of heterologous proteins for therapeutic applications using *E. coli* as the host expression system (Baeshen et al., 2015).

Furthermore, RPP is also important in biological studies to produce enzymes for laboratory use such as restriction enzymes and thermostable DNA polymerases used in recombinant DNA technology (Wyre and Overton, 2014a). A brief timeline of the biotechnology history is presented in **Figure 1.1.**

1.1.2 The future of biopharmaceutical and bioengineering

The initial stages of small scale bioprocessing are crucial for the faster and cheaper development of new biopharmaceutical products (Betts et al., 2014). Biopharmaceutical products have been used since the 19th century such as insulin and diphtheria antitoxin therapy. Significant advancements in the production processes and clinical approval of recombinant protein products, including growth hormones and interferons (IFN-α, β, and γ), in the biopharmaceutical industry occurred during the 1980s. The output of biopharmaceutical products, also known

as bioprocesses, refers to vast procedures (Jozala et al., 2016). In biotechnology, bioprocessing is integral in generating biopharmaceutical products.

Most biopharmaceutical products are recombinant therapeutic proteins obtained from biotechnological experiments and procedures. These products derived from a range of biological sources, including microorganisms, genetically engineered organisms, or even animal fluids, organs, and tissues (Costa et al., 2014).

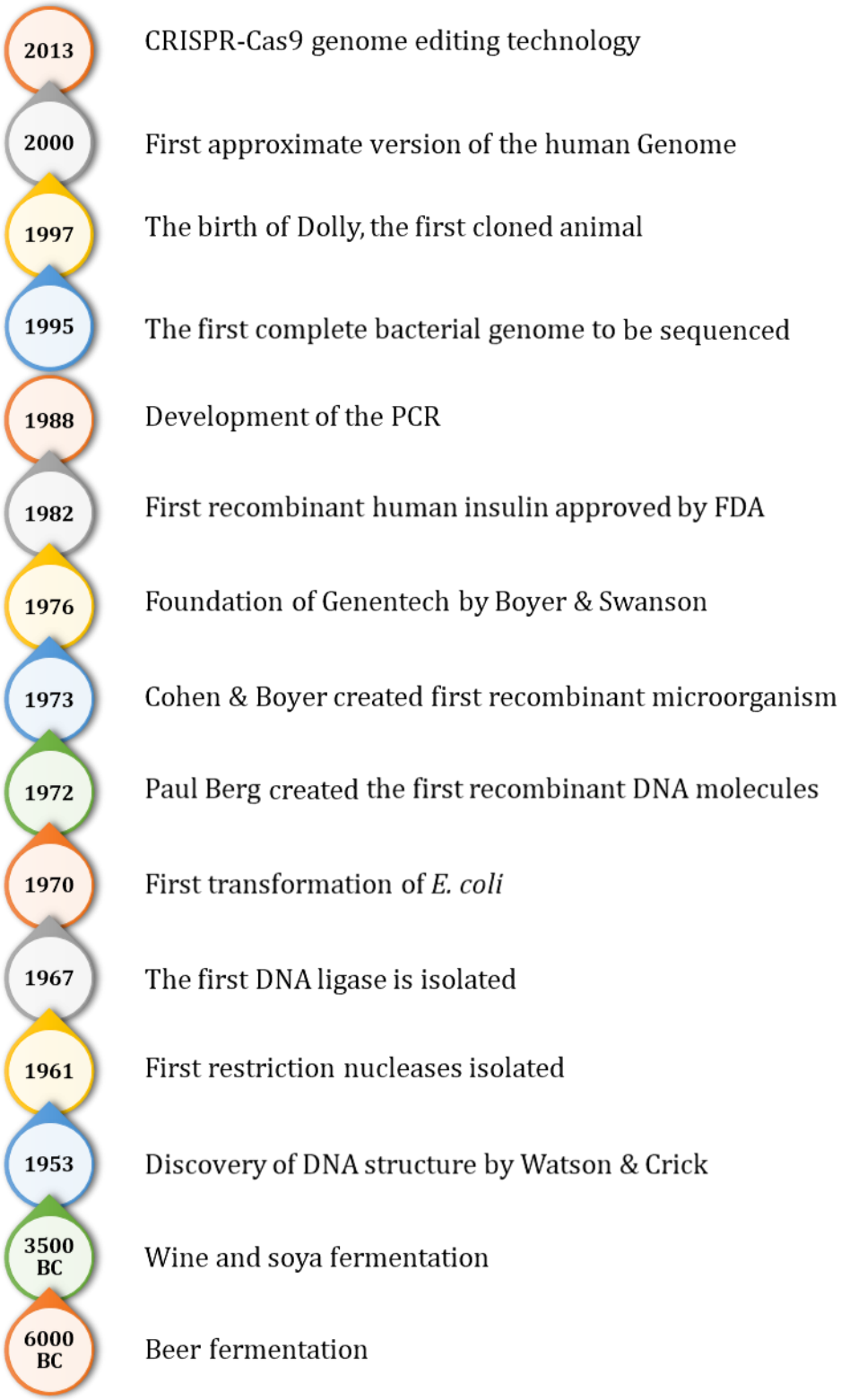


Figure 1.1: A timeline of significant developments in the history of biotechnology industry (Adapted from Buchholz and Collins, 2013)

Moreover, new technological procedures are constantly improving the production of microorganism-derived biopharmaceuticals, despite the availability of expression systems using mammalian cell lines, insects, and plants. This is mainly due to the accessibility of diverse host strains, adaptability of plasmid vectors, and profitability using microbial-based expression systems as compared to other expression systems (Jozala et al., 2016). Additionally, genetic engineering technology has not only revolutionised the pharmaceutical industry, but also significantly impacted other sectors such as the environmental, agriculture, biotechnology, and chemical industries.

Economic impact on the biotechnology industry

To date, the use of biopharmaceutical products has increased substantially around the world. In 2016, around 1357 products were approved by the FDA and the European Medicine Agency (EMA). Among these products, more than 130 products had different pharmaceutical formulations (or reference products) while 737 were biosimilars and the remaining products were categorised as biobetters (http://www.biopharma.com). More importantly, between 2013 and 2016, 73 biopharmaceutical products were clinically approved for human use, wherein 23 monoclonal antibodies (mAbs) were approved for treating inflammatory diseases and neoplastic tumours in human, as well as for multiple diagnostic procedures (http://www.biopharma.com). The EMA had also authorised two new pharmaceutical products for gene therapy (corrective gene insertion to generate a protein in one's genomic DNA in order to treat genetic disorders) applied in human therapeutic procedures. Although biopharmaceutical products are beneficial for

the control or cure of diseases, the treatment costs can reach as high as USD 1 million per patient (Jozala et al., 2016).

In 2016, , the total market sales of recombinant microbial products are estimated to be around USD 50 billion, representing approximately 33 % of biopharmaceutical products available in the market (Jozala et al., 2016). Selection of microorganisms to produce biopharmaceutical products is based on numerous factors which include inexpensive production, easy manipulation and propagation of cells as well as molecular biology procedures. Several examples of the most important biopharmaceutical products acquired from *E. coli* are listed in **Table 1.1**.

Following the first approval of recombinant human insulin; Humulin, in 1982, eight recombinant proteins for human use had been approved over the rest of that decade such as the human growth hormone, monoclonal antibodies, and interferons (Walsh, 2014). During that time frame, the pharmaceutical industry became very progressive in establishing "conventional" microbial processes. As the result, several new microbial products were developed and sold to the biopharmaceutical market in the late 1980s and early 1990s. **Table 1.2** highlights some recombinant proteins that have been approved for therapeutic functions and their production systems (Stanbury et al., 2017).

Table 1.1: Therapeutic recombinant proteins produced using heterologous gene expression in *E. coli* (Jozala et al., 2016)

Biopharmaceutical	*Commercial names*	*Clinical use*
Anakinra (interleukin 1 (IL1) receptor antagonist)	Antril, Kineret	Rheumatoid arthritis
Aldesleukin (interleukin 2)	Proleukin	Melanoma and renal cancers
Calcitonin (salmon calcitonin)	Fortical	Post-menopausal osteoporosis
Denileukin diftitox	Ontak	T-cell lymphoma
Filgrastim pegylated	Neulasta	Neutropenia treatment (as a consequence of AIDS, chemotherapy, bone-marrow transplantations etc.)
Filgrastim (analogue to the granulocyte colony-stimulating factor)	Neupogen	Neutropenia (as a consequence of AIDS, chemotherapy, bone-marrow transplantations etc.)
Glucagon	Glucagon	Hypoglycaemia
Growth hormone (GH)	Genotropin, Norivitropin, Nutropin, Humatrope, Siazen, Serostim	Turner and Prader-Willi and syndromes
Insulin (inhalation)	Exubera	Diabetes mellitus
Insulin (zinc extended)	Lente, Ultralente	Diabetes mellitus
Insulin (fast-acting)	Lispro	Diabetes
Interferon-α2a	Roferon-A	Chronic hepatitis C, hairy cell leukaemia, chronic myelogenous leukaemia, and Kaposi's sarcoma
Interferon-α2b pegylated	Peg-intron	Chronic hepatitis C, hairy cell leukaemia, chronic myelogenous leukaemia, and Kaposi's sarcoma
Interferon-α2b	Intron A	Chronic hepatitis C, hairy cell leukaemia, chronic myelogenous leukaemia, and Kaposi's sarcoma
Interferon-γ1b	Actimmune	Chronic granulomatous disease, severe osteopetrosis
Interferon-β1b	Betaseron	Multiple sclerosis

Table 1.2: Major recombinant protein groups developed as therapeutic agents (Stanbury et al., 2017)

Therapeutic Agent	*Recombinant Protein(s)*	*Clinical Usage*	*Production System/Hosts*
Anticoagulants	Hirudin	Anticoagulant	*S. cerevisiae*
Blood factors *Thrombolytics*	Factors VIII Tissue plasminogen activator	Anaemia Clot lysis	Mammalian cells Mammalian cells, *E. coli*
Cytokines	Interferon-alpha Interferon-beta	Cancers, leukaemia, and hepatitis B Cancers, genital warts, lateral sclerosis, amyotrophic	*E. coli* *E. coli*
Growth factors	Erythropoietin Granulocyte-colony stimulating factor Granulocyte-macrophage colony stimulating factor	Anaemia Cancer Bone marrow transplantation	Mammalian cells *E. coli,* Mammalian cells *E. coli*
Hormones	Human growth hormone Insulin Glucagon Follicle-stimulating hormone	Hypopituitary dwarfism Diabetes Type 2 diabetes Infertility	*E. coli, S. cerevisiae* *E. coli, S. cerevisiae* *E. coli, S. cerevisiae* Mammalian cells

1.1.3 Expression systems and features

Recombinant proteins synthesized from recombinant DNA are often carried out in host cells of a different species than their original host (Sevastsyanovich et al., 2012). Currently, various host expression systems for RPP are available, ranging from microbes and yeast to higher eukaryote species, including insects, filamentous fungi, plants and mammalian cells. A vast range of expression system is necessary to select protein for manufacturing purposes. The characteristics of recombinant protein and its usage are essential for host system selection.

Most of the biopharmaceutical products synthesis was previously focused on using mammalian cells. However, due to advantages of microorganism, the biopharmaceutical industry is slowly turning its focus on the production of heterologous proteins using recombinant DNA technologies in microorganisms. Owing to the recent advances in recombinant DNA and molecular biology protocols, human proteins are generated using heterologous expression based on bacterial systems, such as *E. coli*. The human insulin gene was first isolated and the human protein was produced later using *E. coli* as the host expression system (**Figure 1.2**).

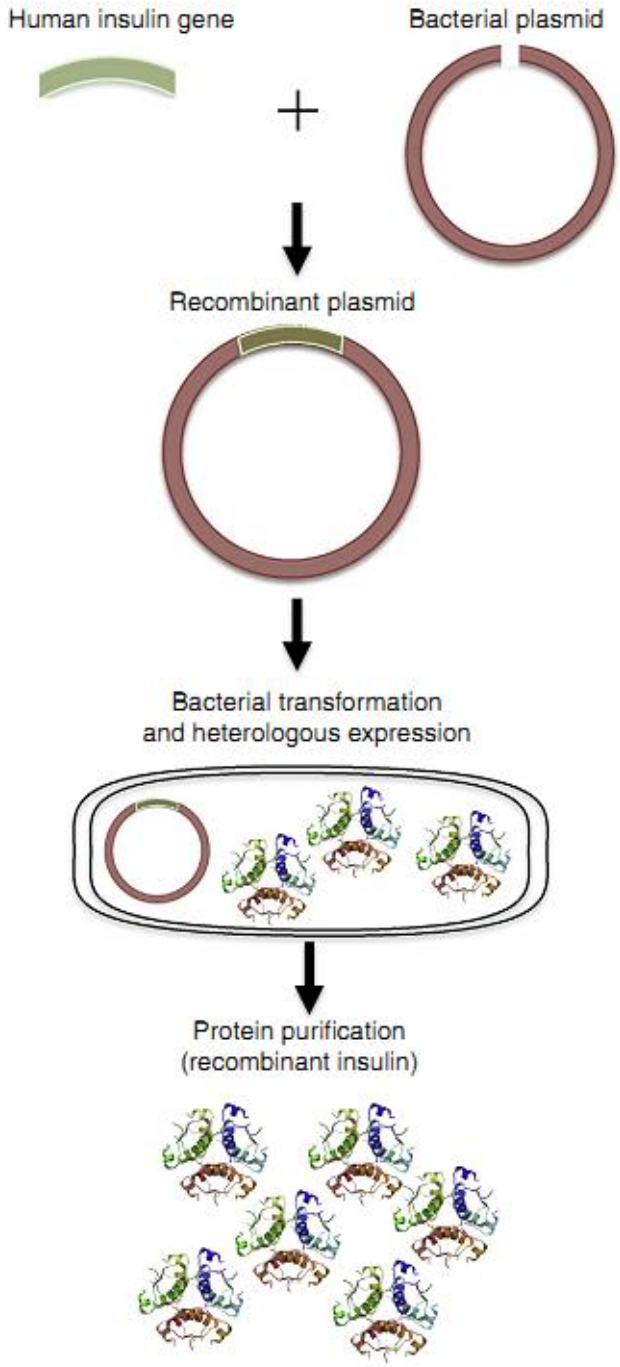


Figure 1.2: Recombinant protein productions in biopharmaceutical industry (Jozala et al., 2016).

The human target gene can be isolated and ligated to a plasmid (bacterial vector) using recombinant DNA technology techniques. Bacterial cells are transformed using the recombinant plasmid that carries the human target gene, hence capable of producing a large amount of recombinant proteins.

One of the commonly used promoter systems for efficient recombinant protein production in *E. coli* is the pET system, engineered from the *E. coli lac* promoter (**Figure 1.3**). In the pET system, the plasmid vector that contains the target recombinant gene is regulated by the T7 promoter. Naturally occurring RNA polymerase enzyme in *E. coli* (*Ec* RNAP) does not stimulate the T7 promoter.

The T7 RNA polymerase enzyme (T7 RNAP), which is encoded on a portion of the host cell genome called DE3 locus, needs to be synthesized to allow the expression of plasmid vector for the production of the desired protein (Overton, 2014).

1.1.1 Host expression system for RPP

Over the past four decades, there has been a significant increase in the manufacture of recombinant products for approval in the biopharmaceutical market. There were 212 licensed biopharmaceuticals in total on the US and EU markets in 2014 (Walsh, 2014). In fact, that number has been increased to 374 in 2018. However, due to withdrawal of products from the market, current active products are 316 (Walsh, 2018).

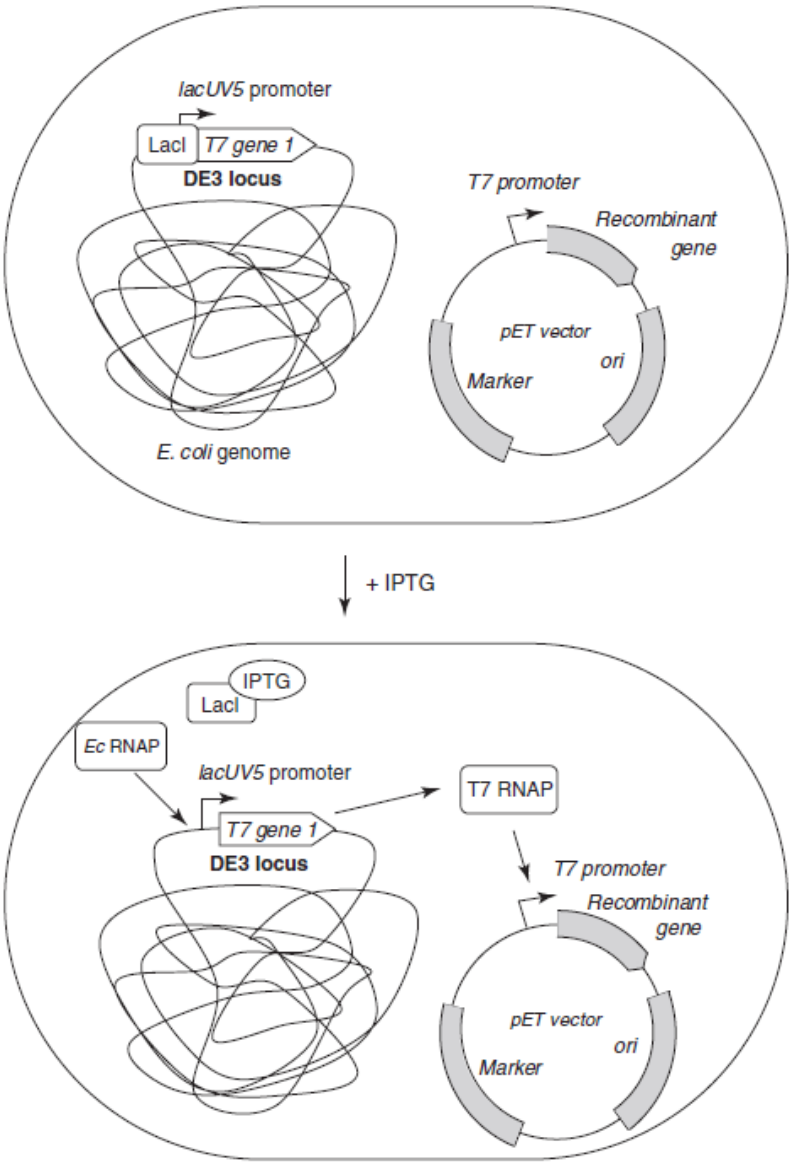


Figure 1.3: The pET system (Overton, 2014)

The *lacUV5* promoter regulates the expression of the T7 RNA polymerase (T7 RNAP). However, the *lacUV5* promoter is repressed by the Lac repressor LacI. The addition of an inducer, isopropyl b-D-1-thiogalactopyranoside (IPTG) to the system will engage to the Lac repressor LacI, therefore vacates the *lacUV5* promoter, allowing T7 gene 1 to be transcribed and synthesise T7 RNAP. This T7 RNAP will activate the T7 promoter to transcribe the recombinant gene in the pET vector expression.

Mammalian and microbial-based expression systems are among the most prominent expression systems that predominate over others. In 2010, microbial systems played a significant role in the production of insulin products, while mammalian systems were used to produce mAbs. The market for mAb-based products increased in 2016, with Chinese hamster ovary (CHO) cell-based systems becoming the most popular mammalian expression system. *Escherichia coli* continues to be the most preferred non-mammalian-based production cell type, despite a slight fall in its use as the expression system. *Saccharomyces cerevisiae* and *Pichia pastoris* are two examples of eukaryotic microbial expression systems based on yeast, and they are equally essential, just behind *E. coli*. Human cell lines are applied in a small number of products to meet particular needs, for example replacement enzymes with specific post-translational modifications such as Replagal (rh-alpha galactosidase; Shire). Other mammalian-based production cells include baby hamster kidney cells (Walsh, 2014). **Figure 1.4** represents the cumulative percentage of product approvals in various production systems from 1982 to 2014.

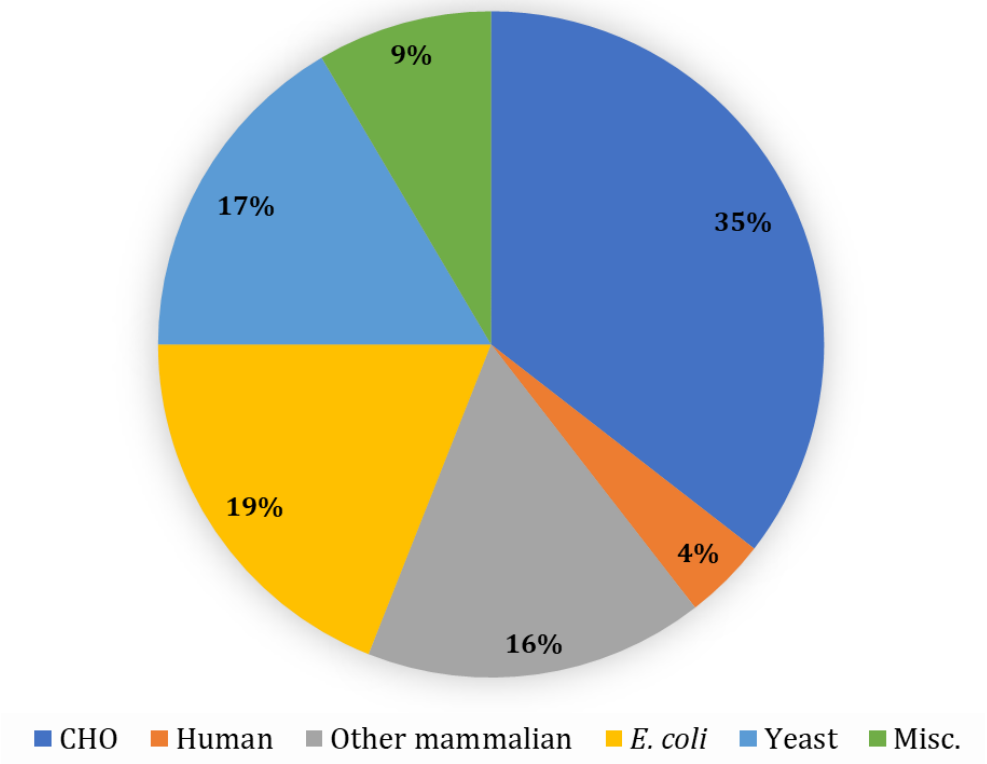


Figure 1.4: Biopharmaceutical products approval, in cumulative percentage, from 1982 to 2014, in different production systems (Walsh, 2014).

1.1.1.1 Bacterial systems

Microbial hosts have become the preferred choice for the recombinant protein production through the recombinant DNA expression. Recombinant protein expression using microbial hosts are currently well optimized, thus contributing up to 40% of current biopharmaceutical products on the market (Walsh, 2014). A wide range of microbial hosts, including bacteria (particularly *E. coli*) and yeast, are ideal for the establishment of different and complicated recombinant proteins.

The selection of an appropriate host strain is very important in the design of recombinant proteins bioprocesses to ensure industry acceptance and widespread use. Mainly bacteria, especially *E. coli* which have a wide range of different physiologies, have attracted tremendous attention to be explored for this reason. Various exploitations have been performed to these organisms including adaptation to different types of substrates for growth and to assess the extent to which they can process and secrete certain proteins (Marschall et al., 2016).

Recombinant proteins should be simply extracted from the cells or their surroundings using an uncomplicated method. Additionally, it is crucial to ensure that the strain itself is safe to use as a host and the applicability of the recombinant proteins generated. All of these factors need to be taken into account in order to create economically superior bioprocess capable of overcoming existing processes.

Bacteria have numerous focal points including their simplicity and capability to grow rapidly in simple bioreactors utilising inexpensive growth media for cost

effectiveness. Its fast growth to high cell density hence simply produces high yields (Ahmad et al., 2018). In other words, the use of microbial hosts is highly advantageous for the production of recombinant proteins and can greatly enhance huge profits to the industry. Thus, bacteria, in particular *E. coli*, demand special consideration for large-scale protein bioproduction.

However, bacterial recombinant protein production is still limited by a number of factors that occasionally prohibit processes from being effective. Bacteria do have certain restrictions, especially when it involved the production of eukaryotic proteins that are much more complex than prokaryotes. For instance, the main drawbacks of *E. coli* as the bacterial host include its lacking of many post-translational modifications (PTM) that typically occurred in eukaryotic proteins, such as glycosylation (Walsh, 2010). This process occurs in the endoplasmic reticulum, involving the bonding of specific oligosaccharides molecules on the original proteins. Glycosylation is very important to ensure proper protein folding for accurate protein functioning. The production of authentic eukaryotic proteins often depends on several stages of PTM, in which the modification involved is particularly glycosylation, yet in addition including hydroxylation, carboxylation, sulphation, and amidation (Walsh, 2010).

The post-translational modification procedure that occurs in most eukaryotic cells is the basis for functional protein production and is likely to be important to prevent immunological reactions when applied for pharmaceutical use. Lack of glycosylation or incorrect glycosylation often results in low protein solubility and

reduced function. Therefore, eukaryotic hosts will be used for recombinant proteins that need to be glycosylated (Overton, 2014).

In addition, bacteria are known to have the ability to produce smaller proteins of size less than 30 kDa (Demain and Vaishnav, 2009). Consequently, bacteria also face the challenge of integrating larger recombinant proteins. However, there has been success in producing 210 kDa protein in the microbial host *E. coli* (Doekel et al., 2002). Nevertheless, due to its simplicity and low cost, *E. coli* remains the ideal choice, both for the laboratory and the industrial scale.

1.1.1.2 Yeast and mammalian cells

Yeast strains, particularly *Saccharomyces cerevisiae* and *Pichia pastoris* have been often use in the biotechnology sector and are one of the industrial host choices for RPP. Similar to bacteria, yeast can be grown in inexpensive culture media in order to generate cells with high density. Yeast as part of eukaryotic organisms is capable of undergoing all the post-translational modifications necessary for functionality, therefore producing correctly folded, even functional, recombinant proteins (Gonçalves et al., 2013). Nevertheless, the yeast-generated glycosylation patterns are dissimilar from those in human proteins as the latter generates *N*-linked oligosaccharides that mainly consist of hyperglycosylation mannose residues (Demain and Vaishnav, 2009). Meanwhile, filamentous fungi such as *Aspergillus niger* are applied for various biotechnological applications to produce enzymes including amylase, cellulase and lactase (Jozala et al., 2016).

The trend towards mammalian cell lines as expression host has dramatically increased over the past few years (Walsh, 2018). Mammalian cells are widely used especially for the generation of monoclonal antibodies since they are able to correctly fold and assemble complex protein structures as well as generate PTM that imitate human cells. Human Embryonic Kidney (HEK-293), Chinese Hamster Ovary (CHO), Baby Hamster Kidney (BHK), and mouse myeloma-derived cells (NS0) are instances of mammalian cells widely used (Croset et al., 2012; Walsh, 2018). **Table 1.3** summarizes the comparison of different expression systems for human proteins production.

Table 1.3: Comparison of different expression systems for human proteins production (Adapted from Gonçalves et al., 2013).

	E. coli	***Yeast***	***Mammalian cells***
Protein productivity	Low	Medium	Optimal
Post-translational modification	Poor	Medium	High
Secretion	Low	Medium	Native & Active
Time and Cost	Very Low	Low	High

1.2 Recombinant Protein Production in *E. coli*

Investigation to the generation of RP is highly rapid in expediting the development of various biopharmaceuticals, especially for heterologous bacterial hosts, in line with the high demand to these products. To manufacture the essential RP, the process needs to be successful to ensure rapid and bulk production of RP and high yield. One of the main emphasis in structural genomics and proteomics is the production of target proteins in bulk quantities. The use of microbial hosts such as *E. coli* has been robotized and optimized considerably to succeed in recombinant protein production and gene cloning measures (Sandomenico et al., 2020). In addition to low cost, the quality level of RP generated is the most priority and is greatly emphasised over quantity and production density alone (Sevastsyanovich et al., 2012).

The use of *E. coli* as microbial host has represented approximately 20 % of therapeutic proteins that have been approved in the market, thus making them the most reliable microorganism to produce heterologous proteins for therapeutic use. The special features of *E. coli* that contribute to the scenario are basically rapid growth and easy large-scale production. This is not only economical, but is also capable of producing high-quality products, especially non-glycosylated proteins for therapeutic usage (Baeshen et al., 2015). The presence of diverse *E. coli* strains and expression vectors, as well as the rapid advancement of bioprocess technology, make them particularly attractive for industrial applications.

Due to its long use in recombinant DNA technology, a diversity of relevant knowledge has been successfully developed for genetic tailoring including gene cloning expressions and the creation of useful vectors (Sandomenico et al., 2020). All collected information related to bacterial metabolism allows for the identification of simple protein folding mechanisms under optimal conditions of media and culture, as well as fast heterologous growth that is very convenient to meet industry needs. The high-level production of functional proteins in *E. coli*, especially those from eukaryotic sources, often quite challenging. **Table 1.4** summarizes some of the general approaches employed to produce heterologous proteins (Jana and Deb, 2005).

Table 1.4: Techniques for optimizing heterologous protein overproduction in *Escherichia coli* (Jana and Deb, 2005).

Techniques	***Comments***
Codon usage (E. coli)	Utilizing the optimal *E. coil* codons often improves yield
Growth conditions/ media	Growth conditions, growth rate, oxygen levels, carbon source and fermenter configuration affect yield
Host strain	Host strain choices impacts expression
Stability of mRNA	mRNA stability impacts yield. Secondary structure, especially at the 5′ end of the message, often plays a critical role
Plasmid copy number	Gene dosage (manipulated by plasmid copy number) affects expression
Promoter	Inducible/constitutive, strong/weak; promoter and regulation are a major influence on protein expression, which is also affected by relative strength and orientation of the plasmid promoters
Selection antibiotic	Antibiotic resistance choice on the expression plasmid can influence heterologous protein expression
Transcription termination	Spacing and effectiveness of transcription terminators affect expression
Translation signal	The ribosome-binding site affects the ribosome loading & clearance level, and hence, expression. Secondary structure at the 5′ end of the message can affect the accessibility of the ribosome binding site
Temperature	Temperature has a pronounced effect on folding and stability of protein, also improves soluble yields

1.2.1 *Escherichia coli* as microbial expression host for RPP

E. coli is a rod-shaped and Gram-negative bacterium, as well as a strain-dependent pathogen, which is commonly found in the gastrointestinal tracts of mammals. In the late 19th century, *E. coli* was isolated by Theodor Escherich and became one of the extensively studied bacteria in the biological disciplines. With the knowledge accessible, it is sensible that *E. coli* should be a preferred choice as a microbial host for recombinant protein production and is recognized at the industrial level for its ease of growth and easily manipulated genetics (Overton, 2014). For years, *E. coli* is the common non-mammalian cell that has been used as host in the biopharmaceutical industry to produce comparatively simple recombinant proteins, for example, insulin, growth hormone and Granulocyte-Colony Stimulating Factor (Castiñeiras et al., 2018b).

E. coli is the most popular bacterium used in recombinant protein production in the biotechnology industry (Huang et al., 2012). Recently, there are many recombinant proteins that has been produced in *E. coli* with high productivity in line with the rapid development of genetic engineering technology (Wurm et al., 2018). One of the main platforms to produce high level recombinant protein depend on genetically modified microorganisms such as *E. coli* because of its well-known genetics and easy genetic manipulation. Its genetics is comparatively simple, properly characterized and many molecular tools are available for this purpose (Mühlmann et al., 2017; Kangwa et al., 2015; Gopal and Kumar, 2013). *E. coli* is one of the most preferable microbial host for numerous recombinant

proteins expression due to its speedy growth rate as well as high rate of recombinant protein synthesis.

E. coli contributes significantly to bacterial RPP. The growth of *E. coli* is exceptionally quick; its growth is capable of growing rapidly to high cell density up to 100 g/L of dry cell weight with dividing time can be as fast as 20 minutes under ideal conditions utilising appropriate media (Wyre, 2014). Compared to other bacterial species, the simplicity of *E. coli* allows its physiology, metabolism and behaviour to be very well understood. *E. coli* can generate disulphide bonds (Gupta and Shukla, 2016), even though it cannot produce recombinant proteins with 'human-like' glycosylation (Kasli et al., 2019).

The widespread availability of *E. coli* in various RPP applications can be observed through the number of approved high-value RP pharmaceuticals in the market (Walsh, 2018). Extensive studies have long been conducted on *E. coli* as the popular host for RPP. All of this information promotes the development of RPPs where access to the selection of expression system and strain adapted for RPP is obtained (Section 1.2.3 and 1.2.4). In addition to the rapid advances in recombinant DNA technology, many protein molecules with various characteristics can be produced in *E. coli* through the numerous engineering tools offered (Kamionka, 2011). In contrast, there are also some issues of *E. coli* weaknesses as the RP microbial host involving poor protein secretion (Section 1.2.2) and RP aggregation into inclusion bodies (Section 1.3.1.1.).

1.2.2 Routes of production

1.2.2.1 Subcellular localisation

Basically, proteins are expressed only in the cytoplasm. However, *E. coli* can store the produced recombinant proteins (RP) in three subcellular places, which are cytoplasm, periplasm and extracellular milieu. In addition to the storage that is maintained at the site of production of the cytoplasm, the accumulation of proteins that occur in the periplasm and extracellular milieu is considered as a form of secretion.

Advantages and disadvantages of cytoplasmic production

Cytoplasmic accumulation is capable of allowing the synthesis of abundant RP even with the simplest expression system. For this reason, this strategy is a superior option for RPP and is recognized in *E. coli*. In contrast, there are some deficiencies when the accumulation of RP occurs in the cytoplasm, which is where it is produced. The cytoplasm environment is less conducive for the formation of disulphide bonds (DSB).

In addition, proteins extraction from the cytoplasm also poses a major challenge for downstream processing (DSP) as it involves the disruption of whole cells. This step requires a large energy input and certainly produces high heat (Wyre and Overton, 2014b). This scenario can cause damage to the protein to be extracted. Furthermore, the high protease activity in the cytoplasm causes the extracted material to have a complex and impure RP mixture. The mixture has the potential

to contain other cellular proteins, endotoxins, nucleic acids and fragments of broken cells (Overton, 2014).

Advantages and disadvantages of periplasmic and extracellular milieu secretion

In contrast to the cytoplasmic production, the expression of proteins secreted into the periplasm or the extracellular milieu has a more significant advantage in terms of stability and biological activity. RP can be easily secreted to the periplasm part of *E. coli* cells in most secretory events. In fact, there are a number of events, in which proteins that have been secreted in the periplasm can penetrate the outer membrane of the cell and subsequently enter the aqueous medium (Bao et al., 2016). This secretion known as extracellular milieu secretion, is very significant for DSP because protein purification can be done effectively, for which the process of cell breakage is not necessary.

This can prevent a harsh process of cell disruption which potentially causes a lot of severe damage to the cell and in turn, affects the protein product to be extracted. Moreover, the space for the extracellular secretion medium is also spacious and expansive, allowing for greater protein expression as it is not obstructed by a narrow space as in the periplasm (Bao et al., 2016). Even the entire cell itself can be isolated from the aqueous media environment by centrifugation, thus the process of protein extraction in the extracellular medium becomes smoother (Sevastsyanovich et al., 2012). And much like periplasm accumulation, protease activity is much lower outside the cell, as *E. coli* does not always release proteins

extracellularly, and yet protein folding can also be improved there, then the potential for protein contamination is reduced (Ytterberg et al., 2019).

However, despite all the advantages mentioned, the fact that, by conventional culturing methods, the accumulation of RP results in the periplasm and extracellular milieu which is much lower than the cytoplasm in most cases, requires a much more complex expression system to meet the challenges. This tends to create additional bottlenecks in the production process. Various approaches have been undertaken to date to greatly enhance the amount of protein secretion such as optimization of the promoter and single peptide (Anuar et al., 2012; Low et al., 2013), as well as the selection and design of the genetically modified expression vectors (Jia and Jeon, 2016). Culturing conditions are also optimized to improve its production, including modifying the inducer concentration (for example, IPTG and arabinose) and optimizing the temperature (Bao et al., 2016).

1.2.2.2 Cytoplasmic recombinant protein production

Cytoplasmic remains the superior choice for RPP

RP secretion in the periplasmic and extracellular milieu essentially contribute significant economic savings for the production of industrial RP. It is able to guarantee simple steps for DSP especially in terms of product purification, which in fact reduces overall cost (Burdette et al., 2018). However, cytoplasmic accumulation remains an option due to the complexity of the system involved in periplasmic secretion.

1.2.2.3 Conformational state

RPs can be soluble or insoluble

The formation of RP inside the host cells can be soluble or insoluble. Under natural conditions, native proteins are produced by *E. coli* with stable folding, soluble and functional, and can therefore be considered as the 'default' creation choice. The RP that can be produced in a soluble form can be of great benefit as it has great potential to indicate the structure and capacity of the original protein.

Insoluble RPs are referred to as inclusion bodies (IB). Inclusion bodies occur during high level expression of recombinant protein, which normally confined in cytoplasm or periplasm of expression host. Under electron microscopes, it can be seen as dense aggregates, which contains unfolded and folded protein in varying proportions. During production of recombinant protein, there will be no assurance that the protein will fold correctly. Hence, the misfolded protein will form inclusion bodies (Upadhyay et al., 2012).

Advantages and disadvantages of IBs

RP is generally required in the form of their active end product, which is soluble. In order to get high translational rate of desired protein, it required high temperature, high inducer concentration as well as strong promoters. Hence, this condition will contribute to the formation of folded and misfolded protein aggregates that accumulated in the cytoplasm as IBs (Singh et al., 2015). IBs place a

major challenge to produce and purify the recombinant protein using *E. coli*. These includes isolation from cell, solubilisation, refolding and purification to yield bioactive protein.

However, the IBs production also brings many advantages to the process. IBs that are large, relatively pure and have solid particles allow significant purification of their RP obtained from crude cell lysate by single centrifugation IBs are proteolysis-resistant aggregates, which protect them from breakdown and damage (Huang et al., 2012). IBs represent a potential pathway for the synthesis of proteins that can be successfully refolded. It is necessary to thoroughly denature IB proteins, for example, using urea or guanidinium hydrochloride, before they can refold into a functional state after the denaturant is removed (Overton, 2014). In the event that refolding is unsuccessful, IBs are worthless and might even be toxic to the host cells due to its dense particle form in inactive state (Hannig and Makrides, 1998).

Previously, RPs tended to aggregate to IBs since many conventional RPP protocols did not tolerate RP production because they stand beyond the cell's ability to produce properly folded proteins. IBs appeared to face a major issue of re-folding their constituent RPs to form its natural dynamic structure, for which efficient and robust handling is required to achieve those objectives. Protein refolding from the IB form needs to go through several stages before its return to the correct form is gradually allowed.

Some RPs are made industrially as IBs which are subsequently recovered and refolded easily, since IBs can be easily isolated after cell disruption due to their

high density. Using this method, the protein of interest is expressed as IBs during fermentation and the IBs are then washed and solubilised prior to the refolding and purification of the recombinant protein. For certain proteins, purification can be a very simple process as the purification and refolding of proteins are performed in a single step (Babu et al., 2000). Many commercial products are produced using this method including recombinant human insulin (Baeshen et al., 2014; Sevastsyanovich et al., 2010).

This process, while it can be very interesting and often used in the industry, is not only an extra processing that incurs additional cost but also requires significant handling to ensure proper folding and functioning (Slouka et al., 2019). Section 1.3.1.1 describe the development of IBs and its common challenges as well as processing approach in details.

1.2.3 Common RPP systems in *E. coli*

RPP has long been thought to contribute poor cell development and subsequent decrease in productivity of *E. coli*. This may be due to RPP-related stress or the metabolic burden caused by the switch to RPP. Toxicity, inclusion body formation, mRNA instability and protein inactivity are just a few of the problems that have driven the development of various vectors, tags and new host strains in recent years (Jia and Jeon, 2016).

In microorganisms, the production of proteins is strictly regulated. Any exploitation or modification of this control can enhance its productivity and

efficiency. Such exploitation of the control systems includes the addition of inducers to the medium, removal of repression control by mutation techniques and increasing number of gene copies coding for the enzyme by recombinant DNA techniques (Stanbury et al., 2017).

1.2.3.1 Expression vectors

Various expression vectors are now commercially available, with most of these vectors using moderate to high copy number of plasmids that could increase protein expression through active selection pressure using antibiotics. To enable protein expression, expression cassettes or low copy number of plasmids can be embedded into chromosome. Constructing the expression vector plasmid demands a few components that should be configured carefully to attain the highest level of protein expression.

Most of the expression systems described in industrial and research studies apply plasmids as expression vectors to synthesise recombinant proteins. The significant impact of plasmids on productivity is dictated by the segregational stability, the structural stability and the copy number of plasmid (Ali et al., 2015).

Expression plasmids are commonly used to clone the coding sequence for the desired protein. The effectiveness of these plasmids has progressed a long way in recent years. The plasmid needs a promoter and a translation initiation region to regulate the expression of the coding sequence, an origin of replication and a selectable marker (such as antibiotic resistance). The presence of fusion tags that are transcribed in-frame with the target gene is another crucial element of these

expression vectors. The recombinant protein, for instance, could be fused to a fluorescent protein. The fundamental assumption is that if the marker protein functions, the resulting heterologous protein must also fold correctly (Rosano et al., 2019).

Sequences encoding for affinity tags and fusion partners are two examples of genetic elements that may be present in a vector in order to facilitate in the protein's detection, purification, and solubilization. Emphasizing genetic elements are also very important, not only for facilitating plasmid maintenance during cell division, but also to minimize the risk of plasmid loss across generations (Peubez et al., 2010). The promoters, translation initiation region, and fusion tags are among those various elements that most significantly affecting transcription, protein yields, solubility, and purification (Jia and Jeon, 2016).

A selection marker can ensure plasmid maintenance, in which only cells containing the plasmid can survive under the selection pressure. Overproducing heterologous protein can cause metabolic burden to the strains and non-producing cells eventually overtake the culture. This will result in yield decrease over time. In most expression vectors, the selection marker consists of an antibiotic resistance gene. Supplementation of antibiotics into the media is convenient and cost-effective, thus becoming the most commonly used strategy for laboratory-scale cultivations. In contrast, its use on a larger scale raises several issues related to environmental pollution and regulatory restrictions (Rosano et al., 2019). For example, Kanamycin and Tetracycline are still tolerable by public health authorities (Peubez et al., 2010).

1.2.3.2 Inducible expression systems

The high level of gene expression at the beginning of fermentation also potentially inhibit growth and cause plasmid instability. This is the reason why, over the past few decades, a number of strategies have been developed and improved to allow for more effective induction of expression in *E. coli*.

Addition of inducer to encourage the production

An integral part of any expression system is the promoter used to control the production of the recombinant gene of interest. Chemical inducers are commonly used to regulate promoters and initiate transcription when introduced to the culture. Inducing promoter expression with IPTG or arabinose is an effective strategy (Faust et al., 2015; Wandrey et al., 2016). However, when applying this approach, it is crucial to monitor the cell culture in order to distribute the inducer at the right time. Indeed, different recombinant proteins have widely varying induction points. When inducers are added in low volumes, there are additional technical issues to overcome. In addition, several inducers present toxicity limitations, which are especially important when manufacturing human therapeutic proteins (Lecina et al., 2013).

On introduction of induction to the cells, RNA polymerase acts to bind to the promoter. Therefore, transcription can be initiated and RP can be generated. Hypothetically, any promoter of *E. coli* can be used for this purpose. However, there is evidence that not all promoters are effective, thus only a limited number are routinely used.

An effective promoter must have several desirable features to make it effective for RPP; tight regulation (low level of basal expression), simple, easy and ready to be transferred to other strains, independent from external stimuli as well as cost-effective induction (Briand et al., 2016). The ideal promoter for recombinant protein expression in *E. coli* is highly regulated to minimise toxicity and cell metabolic burden, apart from controlling efficient transcription for high-level protein synthesis. By adjusting the induction time and inducer concentration, these inducible recombinant protein expression systems permit the growth and production phases to be independently controlled (Mühlmann et al., 2017). A number of frequently used promoters to produce recombinant proteins in *E. coli* for large-scale industrial production are summarised in **Table 1.5.**

1.2.3.3 Lactose-inducible promoters (*lac* promoter-based systems)

E. coli recombinant protein production typically employs a number of inducible expression systems whose components can all be found in the *lac* operon. Allolactose is an intermediary in lactose metabolism that binds the LacI repressor, disengaging it from the operator and allowing expression to occur. IPTG has been frequently used to replace lactose as an inducer due to its chemical inertness and insensitivity to the presence of glucose. Since lactose must be taken in by the cells (through the *lacY*-encoded lactose permease) and then metabolised to the active inducer allolactose, the inducing effect is delayed compared to IPTG application (Mayer et al., 2014).

Table 1.5: Common promoter systems for RPP regulation in *E. coli* (Overton, 2014).

Promoter system	***Source***	***Basis of regulation***	***Notes***
pET system (DE3/T7)	Engineered from *E. coli lac* promoter, T7 RNA polymerase gene and T7 promoters	See main text (Section 1.1.3 and Section 1.2.3.4)	Frequently leaky expression
Other T7 systems	Various	Similar to pET, but the T7 RNA polymerase is expressed in response to signals other than IPTG	Leaky expression
lac systems (Plac, PlacUV5)	Natural or modified versions of *E. coli lac* promoter	Expression is inhibited by the lacI repressor. Addition of lactose or IPTG can withdraw the suppression (Section 1.2.3.3)	Leaky expression
pBAD	*E. coli a*rabinose operon	araBAD promoter is repressed by the AraC repressor. Addition of arabinose will withdraw the suppression (Section 1.2.3.5)	Frequently tight expression
λ P_L	Promoter and repressor cI from λ phage	P_L promoter is repressed by cI repressor protein. A temperature-sensitive version of cI (cI857) is stable at 30 °C but unstable and, thus, lifts repression at 42 °C	Induction at higher temperature; not beneficial for correct protein folding

The T7 promoter system requires the presence of T7 RNA–polymerase in the host strain. The gene is usually integrated into the chromosome with the DE3 phage and controlled via the lacUV5 promoter. Tight regulation can be achieved with the particular genes present on an additional plasmid (Horga et al., 2018). Therefore, many components need to be adjusted properly for reproducible and large-scale gene expression. In addition, large amounts of mRNA generated using the T7 promoter can lead to growth inhibition and high translation rates, thus enhancing inclusion-body formation (Zhang et al., 2015).

1.2.3.4 T7 expression system (T7 promoter-based systems)

The *E. coli* T7 system has been widely applied for high-level expression of genes. This system is composed of lambda DE3 lysogenic *E. coli* strain that carries chromosomally-integrated copies of T7 RNA polymerase gene (gene 1) and is regulated by *lacUV5* promoter. It is similar to pET system (Section 1.1.3), although synthesis of the T7 RNA polymerase is stimulated by signals other than IPTG, for example arabinose. The pET expression system using the T7 promoter is by far the most extensively utilised system for heterogeneous expression in *E. coli*. The T7 promoter, perhaps one of the most potent promoters ever known (Rosano et al., 2019), has substantial activity and the recombinant protein synthesized can accumulate up to 50 percent of the total amount of cellular protein (Jia and Jeon, 2016).

When compared to other heterologous gene expression systems that apply host RNA polymerases, the use of appropriate T7 systems is able to yield much higher protein amounts as the bacteriophage RNA polymerase demonstrates enhanced

processivity (Katzke et al., 2010). For instance, T7 RNA polymerase, has strong specificity in its cognate promoter, which cannot be efficiently recognised by *E. coli* RNA polymerase. Moreover, it is approximately five-fold more active than RNA polymerase found in *E. coli*. Therefore, leading structural genomics experts prefer *E. coli* T7 expression system for high-level expression of globular, soluble, as well as stably-folded prokaryotic and eukaryotic proteins (Gräslund et al., 2008).

The T7 RNA polymerase directs the expression of the target gene, in which the gene coding for T7 polymerase is regulated by the *lac* promoter in *E. coli* chromosome. In the presence of inducer, T7 RNA polymerase is expressed to recognise T7 promoter in the plasmid, thus enabling genes transcription. However, the expression of T7 system is frequently leaky, since the T7 promoter is not entirely repressed by LacI. Some T7 RNA polymerase molecules are nevertheless expressed continuously even without an inducer, therefore generating a large amount of target mRNA. One of the strategy to overcome this problem is by positioning the gene coding for T7 polymerase under *araBAD* promoter regulation, thus allowing tighter control of the T7 expression system (Wycuff and Matthews, 2000; Overton, 2014; Rosano et al., 2019).

1.2.3.5 Arabinose expression system (*araBAD* promoter-based systems)

Apart from glucose or glycerol, *E. coli* can utilise alternative carbon sources such as arabinose. AraC protein has the ability to positively regulate the expression of specific enzymes that metabolise the arabinose molecule into a usable sugar form for the pentose phosphate pathway. The protein is a transcription factor that

induces the expression of the associated enzymes that can enable arabinose uptake (AraE, AraF, and AraG) as well as catabolism (AraB, AraA, and AraD). In the absence of arabinose, AraC regulatory protein binds to two DNA half-sites O1 and O2 regulatory regions, which are located 210 bp from each other on the *araBAD* promoter (**Figure 1.5**). The DNA loop makes a specific binding of AraC dimer that blocks RNA polymerase from approaching the P_{BAD} and P_C promoters. When arabinose is present, AraC can bind to the adjacent DNA half-sites, O1 and I, because the arabinose molecule attaches to AraC causing a conformational change. Hence, the RNA polymerase can now transcribe the genes located downstream of the promoter since AraC is no longer bound to the loop between the two DNA sequences (Sørensen and Mortensen, 2005; Clark and Pazdernik, 2016).

The pBAD system is built on top of the arabinose operon. This arabinose expression system is widely used for its tight regulation (Rosano et al., 2019). The pBAD system is modulated on how much arabinose is supplied in a culture. More arabinose is added if a large quantity of recombinant protein is required. When the recombinant protein is harmful to the host cells, less arabinose is supplied, and hence less recombinant protein is produced. The low level of basal expression can be accomplished using the *araBAD* promoter which is highly desirable for the synthesis of "toxic proteins". "Toxic proteins" is a term that been used to describe recombinant proteins that are hard to produce, as they give bad effects to bacterial host (Overton, 2014).

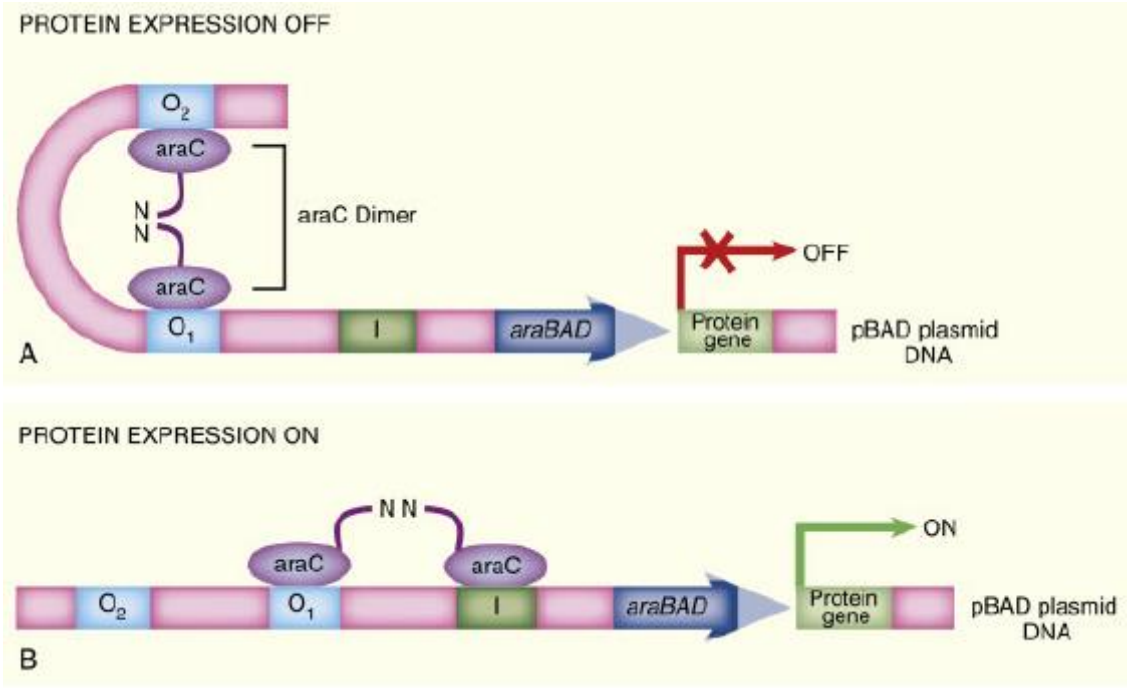


Figure 1.5: pBAD expression system (Clark and Pazdernik, 2016).

The arabinose operon serves as the basis for the pBAD system. When arabinose is induced, it attaches to the AraC regulatory protein, triggering the process. Following activation, AraC moves from the O2 site to the I site, where it binds. This starts the cloned gene's transcription process. (A) Expression of recombinant proteins is inhibited by the AraC protein dimer binding to the pBAD plasmid's O1 and O2 regulatory regions. (B) AraC exhibits a conformational change upon presence of arabinose, in which it now binds to the O1 and I site. The recombinant protein transcription is boosted by this conformation.

While the Arabinose promoter is considered strong, it pales in comparison to the T7 promoter (Jia and Jeon, 2016). However, T7 expression control is frequently leaky, due to low levels of T7 RNA polymerase production in the absence of inducers or low levels of expression from the T7 promoter on the plasmid vector caused by *E. coli* RNA polymerase. It becomes problematic when the recombinant protein produced causes cell growth depression (Overton, 2014). Because of this, alternatives like the araBAD promoter are preferable. Controlling the amount of recombinant protein generated can be very useful in cases of protein toxicity, and the araBAD promoter that allows for such control is a viable alternative (Rosano et al., 2019).

1.2.4 *E. coli* strains and its derivatives

Escherichia coli appears to be the most frequently studied prokaryotic organism that plays a significant role in industrial microbiology and bioengineering applications (Castiñeiras, 2017). Using *E. coli* as the bacterial host system for the expression of recombinant proteins is the top choice in many studies. It is the main tool employed to generate non-glycosylated recombinant protein products, with one-third of medically-approved protein therapeutics expressed using this host system (Walsh, 2018). The identification of *E. coli*'s complete chromosome sequence and the development of various genetic tools enable almost any genetic modification of the host organism. A previous study described an approach that enabled the precise DNA insertion into the

chromosome of the host organism without affecting resistance to antibiotics or other markers (Valešová et al., 2004).

The *E. coli* K-12 strain and its derivatives have also been commonly applied in biotechnology and are classified as risk group 1 for biosafety. The K-12 strains include *E. coli* strains MG1655, DH5α, DH1 and W3110 (Huang et al., 2012). K-12 strains, however, usually have relatively high acetate production in the presence of high glucose concentrations in the culture.

E. coli BL21(DE3) is the most extensively used *E. coli* strain which is known for its lower number of proteases and prevention of plasmid loss throughout the process (Jia and Jeon, 2016; Yu et al., 2015; Rosano and Ceccarelli, 2014). The strain is frequently applied with T7 promoter-based pET expression system and is induced by IPTG that has strong induction capabilities (Bashir et al., 2016; Jia and Jeon, 2016; Wurm et al., 2016; Rosano and Ceccarelli, 2014). Nevertheless, IPTG causes high metabolic stress on *E. coli* (Dvorak et al., 2015; Haddadin and Harcum, 2005), thus promoting inclusion bodies formation. The *E. coli* BL21 strain, on the other hand, has been applied in various protein expression studies (Sørensen and Mortensen, 2005). The strains are employed for RPP due to its lower accumulation of acetate despite the presence of high glucose concentrations that function as carbon source.

However, the drawback of using *E. coli* as an expression host organism for RPP refers to increased generation of endotoxin. Lipopolysaccharides, also known as endotoxins, represent the major constituent of the outer cellular membrane of Gram-negative bacteria which is considered to be pyrogenic for humans. Hence,

removing them from the finished protein product is an essential requirement for protein production (Rinas et al., 2017).

1.3 Common problems associated with RPP

Despite its widespread use due to its valuable features that are capable of bringing diverse profits, particularly to the biopharmaceutical market, the use of *E. coli* also poses some key barriers, including those involving protein folding, plasmid maintenance and RPP-induced stress, as detailed below.

1.3.1 Protein folding

The synthesis of functional proteins using *E. coli* seeks delicate equilibrium among the processes of gene transcription, protein translation, and protein folding. In many cases, folding the proteins accurately demands glycosylation processes, which are notably absent in *E. coli*. The majority of the post-translational modification (PTM) pathways found in higher organisms such as glycosylation cannot be carried out in *E. coli*. as it is unable to glycosylate its proteins. Even though *E. coli* has become the preferred microorganism for the expression of large-scale therapeutic proteins, the use of *E. coli* has been restricted to the production of recombinant biopharmaceutical products due to the lack of these cellular machineries. The *N*-linked glycosylation of complex proteins is one of the most essential PTM in eukaryotic organisms (Gupta and Shukla, 2016). Several forms of these modifications are vital processes for protein functional activity.

Furthermore, the formation of disulphide bond which is necessary for the correct folding of most eukaryotic proteins is prevented by the reduced condition of the cytoplasm in *E. coli*. Hence, the human therapeutic proteins expressed in *E. coli* tends to generate high-level misfolded and unfolded proteins, thus aggregating cytoplasm and forming inclusion bodies (Section 1.2.2.3 and 1.3.1.1). Protein misfolding is a significant problem for the synthesis of biologically active polypeptides, thus limiting the availability of proteins generated, especially in the industrial-biopharmaceutical sector (Villaverde and Carrió, 2003). Therefore, the lack of PTM in *E. coli* restrict its use for the expression of slightly complex recombinant protein products. Several technological advancements have been developed in the *E. coli* expression system to fulfil the requirements of the biotechnology industry (Baeshen et al., 2015).

Another reason for the large number of issues observed in *E. coli* when expressing the human therapeutic proteins is the formation of intermediate forms of the target proteins that are inaccurately folded. This issue induces general stress reactions as well as other overlapping stress responses in the cell (Hoffmann et al., 2004; Gasser et al., 2008). If the sole problem is with the recombinant protein's folding, then the overexpression of folding chaperones such as pre-inducing RpoH regulon RNA polymerase with heat shock treatment may be adequate to achieve success (Hoffmann et al., 2004; Sevastsyanovich et al., 2009a). The difficulty in secreting recombinant proteins into the culture medium are another major issue. As a result, many recombinant proteins end up being discarded as inclusion bodies in the cytoplasm and need subsequent steps to be refolded (Kamionka, 2011).

1.3.1.1 Inclusion bodies (IBs)

The formation of inclusion bodies (IBs) has become a major challenge in bacteria such as *E. coli*, particularly when it is used to produce recombinant proteins that are highly soluble for both laboratory research and industrial applications. IBs have previously been considered as a waste product as they are involved in misfolding and have the tendency to aggregate. They were traditionally believed to be the irretrievable deposits of partially folded protein intermediates that occurred passively in the cell (García-Fruitós, 2010; Rodríguez-Carmona et al., 2010). Hence, various attempts have been made to eliminate or at least partially reduce the formation of inclusion bodies.

While many studies suggest that *E. coli* inclusion bodies are wasteful aggregates of misfolded proteins that should be avoided as much as possible in RPP, it is important to note that many crucial biopharmaceutical protein products are actually made industrially as IBs. Growth factors, insulin, hormones and interleukins are all examples of currently biopharmaceuticals products from *E. coli* that are typically manufactured as IBs and then subjected to solubilization and refolding processes to obtain soluble target products before being entered to the market. (Schmidt, 2004; Eiberle and Jungbauer, 2010; Wurm et al., 2018). For obtaining correctly folded soluble products, IBs should be harvested, washed and solubilised prior to the refolding and purification of the recombinant protein (Choi et al., 2006; Upadhyay et al., 2014). Nevertheless , IBs have long been considered as aggregates of inactive products and thus remained highly undesired by the industry for decades (Choi et al., 2006; García-Fruitós, 2010).

Hence, several efforts were performed recently to develop a new method for the purification of IBs after realising their importance and potential in various biotechnological applications. Increasing evidence suggests that the aggregation of proteins in bacteria is a reversible process that occurs to cope with thermal stress. In addition, there is also evidence that IBs are dynamic and reversible structures. The cellular machinery is thought to support its formation, whereby misfolded polypeptides would be stored when operated under stress conditions until they can be further refolded or proteolysed (Vallejo and Rinas, 2004; Villaverde and Carrió, 2003). However, many RP cannot be refolded from IBs and must be produced in a soluble form. (Wurm et al., 2018).

Recently, IBs have demonstrated a number of benefits, including higher yields of production, simple separation procedures from cells, great purity, and resistance to proteolytic degradation (Ramon et al., 2014; Yamaguchi and Miyazaki, 2014). Accumulation of IB aggregates in the cytoplasm at a much higher concentration than the soluble protein is able to produce a higher yield. Some proteins can even be purified and refolded in a single step, making the purification process incredibly straightforward (Babu et al., 2000). In general, IBs contain recombinant proteins as well as contaminants such as DNA, RNA, chaperones, and lipids. Although the exact proportion of pure, unfolded recombinant proteins in IBs varies, it is usually far over 90%. IBs are stable protein aggregates that provide additional protection to the protein of interest from proteolytic activity (Vallejo and Rinas, 2004). In the downstream processing of RPP, IBs are easily isolated after cell disruption and the resulting IB solution can be stored frozen for a few months, thus providing

manufacturing flexibility. Bacterial IBs are submicron-sized protein clusters that are fully biocompatible and have been lately characterised as inert nanoparticulate materials that can be used as scaffolds for tissue engineering with potential application in material sciences and biomedicine (Rodríguez-Carmona et al., 2010).

Altogether, these features enable IBs to be produced at high yields, in addition to its simple and minimal effort required for isolation and purification. However, this method has certain disadvantages, in which the refolding process of IBs to form active proteins is challenging as high-yielding and efficient refolding requires significant optimisation steps for each target protein. Moreover, the use of chaotropic agents for the re-solubilisation of IBs may affect the integrity of these proteins. Nevertheless, acceptable protein recovery can be attained at a large industrial scale using various established strategies.

Although functional proteins can be recovered by preparative refolding from IBs *in vitro*, the optimisation process for a certain protein species is time-consuming with results that are inconsistent and not so profitable commercially (Clark, 2001; Hadj Sassi et al., 2017). Hence, while a straightforward soluble production process is highly desirable, the lack of knowledge regarding the molecular nature of protein aggregation restricts the use of genetic engineering approaches for successful protein folding. Thus, improved knowledge of protein aggregation in cell physiology will permit the efficient reduction of inclusion body formation to achieve higher yield. Furthermore, it will also enable a detailed understanding of the complexity of cellular mechanisms that are committed to handling protein misfolding (Villaverde and Carrió, 2003).

It is hoped that the knowledge enhancement in relation to the nature of IBs, the nature of its formation, and its association with cellular stress responses will provide further avenues for improved recombinant protein production. In recent decades, the knowledge of inclusion bodies has been increasing rapidly. As a result, its characteristics and potential biotechnological applications have been changed dramatically as well, thus leading to the advancement of purification procedures to conserve its properties (García-Fruitós, 2010).

1.3.1.2 Soluble and insoluble proteins

The production of soluble bioactive proteins without the refolding process can be accomplished in the cytoplasm of *E. coli*. The two common approaches that have been used are decreasing the growth temperature and improving the folding environment to enhance the recombinant protein solubility without altering its protein arrangement. Reducing the growth temperature also diminishes the rate of protein production and inhibits the build-up of folding intermediates in the host cytoplasm (Huang et al., 2012). Moreover, it diminishes protein aggregation by reducing the inter- and intra-atomic hydrophobic interactions that assist the arrangement of IBs. This technique seems to be effective in enhancing the solubility of various proteins including Fab fragments, interferon α-2, and the human growth hormone (Jensen and Carlsen, 1990; Vasina and Baneyx, 1997).

Folding modulators in the cytoplasm such as folding chaperones (GroEL and DnaK) and disaggregating chaperones (ClpB) play a significant yet unique role in the maintenance of correct protein folding. The co-expression of these molecular chaperones is thought to enhance the solubility of various RPs. For instance, with

the co-expression of GroEL/GroES, approximately 65% of complete proteins created against B-type natriuretic single-chain peptide immune response (scFv) was soluble, thus representing a 2.4-fold increase as compared to the strain without co-expressed chaperones. Nevertheless, chaperone co-expression does not ensure improved protein solubility and may exert metabolic stress on the cells, thus resulting in low protein efficiency (Huang et al., 2012).

Various *E. coli* strains have been created to enhance protein folding and solubility in the host cytoplasm. The disturbance of two genes, namely *gor* (glutaredoxin reductase) and *txrB* (thioredoxin reductase) in *E. coli*, enhanced the folding of the recombinant tissue plasminogen activator by encouraging disulphide bond development and isomerisation in the cytoplasm. Moreover, overexpression of the periplasmic disulphide bond isomerase (DsbC) in the cytoplasm of a similar strain additionally improved disulphide bond arrangement. Despite the technological advances, the soluble expression of recombinant protein therapeutics in the cytoplasm can still result in protein misfolding, moderately low protein yields, time-consuming downstream handling processes, and vulnerability to proteolysis.

1.3.2 Plasmid maintenance

The utilisation of genetic elements that permit plasmid maintenance during the cell division process and restrict the likelihood of plasmid loss over several generations should be carefully considered (Peubez et al., 2010). Plasmid copy number is integral to analyse the expression system. Plasmid structural stability denotes acquisition of correct base sequence in all plasmid. Cell progress in the

absence of plasmid can cause substantial loss in recombinant protein productivity (Jana and Deb, 2005). Commonly, multiple copy plasmids are used for maximum *E. coli* recombinant gene expression. Nevertheless, attaining maximum productively may be hindered due to the metabolic burden as a consequence of multiple plasmid copies in some processes of metabolic engineering (Lecina et al., 2013). Plasmids with a high copy number can increase the burden of the cell and induce metabolic stress, thus resulting in low protein synthesis or even cell death (Rosano and Ceccarelli, 2014). Hence, balancing the ratio of biomass growth to product synthesis, which in turn reduces cell stress, is crucial for maximising productivity.

1.3.3 Minimising RPP-induced stress

Numerous parameters must be individually optimised to reduce RPP-induced stress. This includes not only the selection of promoter and vector systems, but also the medium composition, cultivation temperature, oxygen supply, and induction conditions (Mühlmann et al., 2017; Wurm et al., 2018). Typically, thermal stress leads to protein aggregation in the bacteria. In recombinant bacteria, over-expressed plasmid-encoded genes can trigger the transcription of stress response genes such as heat-shock genes, thus resulting in the accumulation of encoded proteins as IBs (Villaverde and Carrió, 2003).

Recombinant protein synthesis in *E. coli* is also influenced by several factors that inhibit its effective processes such as difficulties in expressing a foreign gene and solubility of recombinant proteins for over-expression (Gopal and Kumar, 2013). Greater knowledge of protein aggregation in the cell physiology of bacteria such as

stress reduction can effectively minimise the formation of IBs for a higher soluble yield.

To obtain a high concentration of soluble protein products in *E. coli* cells, a delicate balance must be achieved between the transcription of genes and protein folding. If the protein folding machinery becomes overburdened, the folded secondary protein structures cannot be shaped and this leads to the aggregation of IBs (Gatti-Lafranconi et al., 2011; Marschall et al., 2016).

It is widely known that bacteria overexpressing either prokaryotic or eukaryotic recombinant proteins experience different stresses, essentially conformational and metabolic stresses. In these conditions, protein processing leads to the production of wasteful products mainly due to the lack of solubility, in which many proteins tend to produce high amounts of IBs. By adjusting certain parameters such as promoter strength, inducer concentration, and plasmid copy numbers, variable quantities of the target protein can be expressed as soluble forms. However, many eukaryotic proteins are accumulated in the IBs and seem to display resistance to the solubility enhancement procedures (Ferrer-Miralles et al., 2009).

Lowering the growth temperature is one of the frequent methods used to prevent the formation of IBs by correcting the folding of the protein molecule to improve recombinant protein solubility. A lower growth temperature reduces the protein synthesis rate and stalls the accumulation of folding intermediate structures in the host's cytosol. However, this results in prolonged cultivation periods. Several studies have indicated that RPP can be effectively balanced with biomass amassing by growing microbes at a lower temperature and instigating production at a lower

rate by utilising weaker promoters or lowering inducer concentrations, thus permitting higher biomass concentrations. As the RP is gradually created, effective folding is improved and soluble protein production is enhanced (Wyre and Overton, 2014b).

As reported by Castiñeiras et al. (2018), several parameters affecting RPP were identified during preliminary laboratory-scale batch culture studies. Based on the studies, high yields of recombinant human tumour necrosis factor-α (rhTNFα) proteins has been attained by using industrial fed-batch fermentation. It was observed that lowering the inducer concentration and growth temperatures as well as increasing the synthesis period were much more efficient for improving RPP yields as opposed to altering the growth phase at which the protein production was induced. Specifically, high quantities of rhTNFα (up to 5 g/L) were achieved with minimal plasmid loss despite using synthetic growth media (absence of animal-derived components), thus being fully compliant with the proposed regulatory requirements. The majority of the products were biologically active and soluble. Therefore, minimising stress was an effective method to optimise the synthesis of rhTNFα (Castiñeiras et al., 2018b).

1.4 Fermentation and Scale-Up Production

Recombinant protein production in *E. coli* is a significant aspect of the global biotechnology industries. Production of biopharmaceuticals or relevant industrial proteins, for example, often requires high quantities up to several kilograms per year. RPP fermentation processes suitable for a wide range of scales, from

laboratory use requiring milligram to gram amounts of the target protein to industrial applications. For instance, culture volumes of RPP may range from several millilitres to 100,000s of litres in scale. Hence, there is extensive flexibility in terms of processing for RPP growth in large batches, including growth strategy, medium composition, and culture vessel design.

1.4.1 Design of culture vessel

The design of the RPP culture vessel depends on the quantity of culture that is intended to be produced and the specific growth requirements of the cultured organism. Typically, it can range from the use of a simple shake-flask apparatus to 1000 L-sized industrial bioreactors.

1.4.1.1 Shake-flask cultures

The shake flask is the most basic culture vessel for RPP. Screening studies favour these vessels because of their small volume and other characteristics; they are inexpensive and simple to implement. However, industrial RPP is unachievable due to the small volume. Further, the removal of the culture vessel from the incubator for sampling or adding nutrients during fermentation is not something that can be easily automated, thus potentially disturbing its growth. These limitations prevent the use of real-time monitoring during fermentation.

Scale-down models such as the shake flask is the most common classical platform used for cell development commercially (Rameez et al., 2014). They have the capability of performing high-throughput fermentation experiments but without

offering the ability to control the fermentation process parameters such as dissolved oxygen (DO), agitation rate and pH. These parameters play a vital role in developing robust and efficient cell development processes that impact product yield and quality, thus limiting its ability to replicate bioreactor conditions. In the last decade, a large number of attempts have focused on the development of miniaturised bioreactor systems that are able to mimic the performance of stirred tank systems.

1.4.1.2 Bioreactor (Large-scale fermentation)

Fermentation technology in bioreactor is the best method that is commonly applied to improve cell density and obtain functional recombinant proteins in large quantities. It provides control over important biological, physical and chemical parameters that influence cell growth as well as RPP. These parameters include dissolved oxygen transfer (DOT), pH, temperature, nutrient supply, and many more (Huang et al., 2012). A robust and efficient industrial fermentation process also ensures that all the components used are cost-effective. Additionally, various feeding strategies and scale-up processes need to be well identified.

Additional control mechanisms are often included in the industrial fermentation process as most bioreactors contain several ports that are attached to several analytical probes, such as DOT and pH probe. The physiological state of the growing culture can be assessed using these parameters. Data from these probes is transmitted to an externally-located control unit, where it is used to direct the machinery or adjust the measurements in accordance with predetermined parameters value. DOT can be regulated by altering the aeration rate and agitation

speed, and pH can be maintained by the automatic addition of sterile, concentrated acid and base chemical solutions.

Agitation speed, aeration rates, DOT, pH and temperature are among the parameters whose readings are fed into a software system for further analysis. There are also ports on the vessels that allow for the introduction of inoculation, feeding, and/or chemical inducers. Additionally, foaming can be avoided by the use of an automated antifoam solution addition. This is crucial in order to prevent air filter blockage caused by culture leaking from the vessels.

1.4.2 Recombinant microbial cultivation techniques

1.4.2.1 Batch growth

Batch fermentation process is a simple and easy method of culturing cells within a very short period. However, due to its limited productivity capabilities in comparison to fed-batch culture techniques, it is typically utilised to acquire relatively modest amounts of proteins for protein characterisation or toxicological studies. For therapeutic proteins used in low-dose therapies or as an orphan drug, the batch process is a suitable choice for the reliable and efficient production of proteins (Huang et al., 2012). In the context of therapeutic industrial manufacturing operations, however, its application is extremely uncommon.

1.4.2.2 Fed-batch growth

A fed-batch growth approach, as opposed to batch fermentation, can achieve high cell density culture (HCDC) by the continuous addition of nutrients required to control and sustain cell growth for high productivity. Upon the indication of sharp dissolved oxygen (DO) and pH increases in the batch phase, a gradual amount of nutrients is subsequently fed into the bioreactor.

A medium enriched with a high concentration of glucose was used in order to obtain high cell densities. Nonetheless, in aerobic condition, the availability of excessive carbon source can contribute to the production of acetate which known to impede cell growth. Overflow metabolism is the term used to describe this (Retamal et al., 2018). When the glucose concentrations are greater than the cell's needs, which are roughly 5 g/L (Lee, 1996), an overflow metabolic process is triggered and acetate is produced. Acetate inhibits the growth of *E. coli*, yet it is wasteful for RPP processes to divert the carbon supply to acetate (Anane et al., 2017).

Acetate is known to be generated as a by-product when *E. coli* is grown under oxygen-limiting or anaerobic conditions. However, the growth of *E. coli* in the presence of excess glucose is also known to generate acetate even under aerobic conditions. The cells produce acetate when the carbon flux into the central metabolic pathway is higher than the biosynthetic demands of the cells. A high acetate concentration reduces biomass yield, growth rate, and maximum attainable cell densities in HCDCs. Moreover, acetate has been shown to be more

detrimental to recombinant cells as compared to non-recombinant cells, in which RPP is significantly reduced by the accumulation of acetate.

The two main categories of feeds are the carbon source and the nitrogen source, both of which also contain other essential trace elements and minerals. In fed-batch growth, an ideal feeding process is guaranteed to prevent overflow metabolism and sustain cell growth and effective protein production. Hence, various feeding strategies have been employed for HCDC, including DO-stat control, pH–stat control and feeding based on specific growth rates such as limiting the addition of necessary substrates like glucose.

1.4.3 Optimisation of experiment

1.4.3.1 Growth media

The *E. coli* cultivation media usually require essential components such as carbon source, nitrogen source, minerals, essential salts and certain growth factors for optimal cell density. Generally, three types of media, namely complex media, semi-defined (SD) and chemically-defined (CD) are utilised to enhance the growth of bacteria. The complex medium is composed of compounds from natural origins like protein hydrolysate and yeast extract, the compositions of which are not totally known, in contrast to the CD media, which contains ingredients with known identities and defined concentrations. Last but not least, a semi-defined medium has primarily specified ingredients with the addition of some complex compounds.

Use of media with a well-formulated combination of all the necessary nutritional ingredients delivered at optimal concentrations is essential for achieving robust cell development in the culture medium and higher RPP. Complex and semi-defined media are currently commonly employed in the industry due to their versatility as well as high cell density and protein yields in most of the fermentation processes.

1.4.3.2 Carbon source

Although alternative carbon sources are widely available, glucose and glycerol are the two most prevalent carbon sources employed for the development of *E. coli* growth in industrial applications. Glucose is the preferred carbon source for *E. coli* growth (Choi et al., 2006). It is inexpensive and *E. coli* cells favourably consumed glucose as opposed to other carbon sources (Lee, 1996; Overton, 2014). However, glucose has drawbacks in that it undergoes a mixed-acid fermentation reaction in the presence of low oxygen levels, leading to the formation of a wide range of by-products, including acetate, formate, lactate, succinate, ethanol and hydrogen. The overall effects of these conditions lead to culture acidification and reduced productivity of biomass conversion, thus cell growth and protein production are negatively affected (Lee, 1996). Glucose's reaction with amine groups in a series of Maillard reactions is another factor to consider about while utilising glucose. Hence, it must be prepared and autoclaved separately from the growth medium, and it must generally be dissolved before its addition to the media content in the vessel.

In contrast to glucose, glycerol offers several benefits and is widely utilised as a carbon source in *E. coli* cultivation. For example, glycerol is not easily converted to acetate, therefore higher amounts can be applied, allowing for longer growth periods with less frequent monitoring of nutrient supply. Previous studies have shown that glycerol is preferable to glucose for reduced acetate formation and enhanced RPP, making the reduced maximal growth rate of *E. coli* in glycerol compared to glucose less of an issue in fed-batch processes (Eiteman and Altman, 2006). It also avoids the Maillard reaction process and can be autoclaved together within the growing medium, thus eliminating a step and reducing contamination. Glycerol can also be fed into the culture directly in liquid form, with minimal dilution.

1.4.3.3 Arabinose induction

Using recombinant strains that support tight regulation of induction and that limit the expression leakage has several benefits, just like in the case of tightly regulated promoters. For instance, a strain with a chromosomal copy containing T7 RNA polymerase gene cloned downstream from a regulated promoter, in which T7 RNA polymerase induction is regulated by arabinose, thus the leaking of target protein expression is inhibited until the necessary conditions are met in the culture (Lebendiker and Danieli, 2014). This involves making optimal use of the arabinose inducer concentration. This procedure allows better control of the expression in order to increase the solubility of the RP by reducing the rate of transcription and translation processes, thus allowing for proper folding of the protein to occur.

Previous study has proven that by inducing expression at a lower level utilising weaker promoters or lower inducer concentrations, RPP can be easily balanced and greater biomass concentrations can be obtained (Castiñeiras et al., 2018b). Nevertheless, there are cases whereby these changes are still deemed insufficient for large-scale production of proteins that are susceptible to aggregation. There are instances in which optimal protein production is stimulated without induction, in which low levels of leaky expression is produced at optimised conditions (Jia and Jeon, 2016). However, it is difficult to scale up this strategy due to variations in the media formulation.

1.4.3.4 Temperature

Another strategy to minimise the formation of IBs and increase RP solubility is by reducing the growth cultivation temperature (Baeshen et al., 2015; Mühlmann et al., 2017). Previous studies have demonstrated that the generation of RP is slower when a growth temperature of 25 °C is employed throughout the process as opposed to a growth temperature of 37 °C (Wyre and Overton, 2014b), where effective folding of the protein is improved, leading to greater soluble protein production (Castiñeiras et al., 2018b). Although the rate of recombinant protein synthesis was slowed by these lower growth temperatures, but the physiological stress on the bacteria was reduced, thus plasmid loss was lessened, non-productive bacteria were decreased, and more RP could be folded into a soluble form, resulting in less IB formation.

1.4.4 Fed-batch RPP strategies in *E. coli* and high cell density culture (HCDC) techniques

The increasing demand for RPP and the proposed safety regulations requires the development of appropriate high cell density culture (HCDC) techniques for *E. coli* to achieve a high production level of proteins. Some of the advantages of HCDC include the following; decreased reactor volumes, reduced upstream and downstream processing, enhanced cell separation and higher yield in product recovery. HCDC techniques developed for culturing *E. coli* have also led to significantly improved productivity as well as lowered production costs, reduced wastewater, and reduced investment costs in equipment.

HCDC fermentation of recombinant *E. coli* is an attractive biotechnological method of achieving high space-time yields for the synthesis of heterologous proteins. For instance, cell concentration of more than 100 g/L cell dry mass are obtained using exponential feeding strategies (Wilms et al., 2001), while Faust et al., (2014) reported that cell concentrations of 69-70 g/L cell dry mass are obtained using three different feeding strategies Besides, HCDC are also known to reduce the generation of metabolic by-products such as acetate, ethanol, and lactate that potentially inhibit growth.

In order to obtain HCDC, fed-batch operations are frequently used. This growth culture uses inoculum grown to a maximum specified growth rate, which is sustained by nutrients initially supplemented in the fermenter and subsequent nutrient feed regimes additionally delivered at regular intervals until fermentation

is complete (Kangwa et al., 2015). Studies have discovered that alterations in nutrient composition and its feeding strategy can have effects on cellular stability, production and secretion. High cell densities of *E. coli* have been achieved through the use of relatively straightforward feeding strategies, including increasing feeding, constant rate feeding or exponential feeding (Babu et al., 2000; Faust et al., 2014).

Numerous feedback-controlled nutrient feeding strategies have also been developed. When the main carbon substrate in culture media is used up, the pH or DO shifts, and this is the basis for the pH- or DO-stat approach. When the carbon source is depleted, a concentrated feeding solution is added at a predetermined rate to keep the pH from rising too high and the dissolved oxygen level from falling too low. These feedback-controlled feeding systems are commonly used for *E. coli*'s HCDC due to their ability to avoid the overfeeding of nutrients (Choi et al., 2006; Schaepe et al., 2014).

The development of HCDC for recombinant *E. coli* faces several challenges such as the low solubility of oxygen saturation, carbon dioxide accumulation to levels that can slow growth rate, inefficient mixing in the fermenter, and heat generation. HCDC also are thought to be restricted by substrate inhibitions. In order to overcome heat and mass transfer problems in large-scale fermenter, the optimal HCDC condition must be developed (Faust et al., 2014; Sohoni et al., 2015). To avoid these detrimental effects on RPP, specific strategies are required for the HCDC of recombinant *E. coli*.

1.5 Tumour Necrosis Factor α (TNFα)

The Tumour necrosis factor alpha (TNFα) protein, a member of the cytokine group, is produced by small cells in the nervous and immune systems that affect the behaviour between cells. It is produced mainly by activated macrophages (M1) and a variety of cells including CD4+ lymphocytes, natural killer cells (NK cells), and neurons (Horiuchi et al., 2010).

TNFα has an essential role in controlling the production of immune cells. Uncontrolled TNFα production has been linked to a range of human diseases such as Alzheimer (Swardfager et al., 2010), major depression (Dowlati et al., 2010), apoptotic cell death as well as very serious cases of cancers (Locksley et al., 2001). TNFα was first acknowledged as an endotoxin-induced serum factor in the mid-1970s which led to the necrosis of certain *in vivo* murine tumours (Carswell et al., 1975). Several subsequent studies have shown that this molecule affects many normal and neoplastic cellular processes.

TNFα is a regulator of immune cells and a human protein that plays a critical role in cell signalling during systemic inflammation. To date, TNFα has been accepted as a multi-functional cytokine molecule that mediates major functions in anti-tumour responses, acute and chronic inflammation, and infection. Thus, many studies have been performed to determine their biological action and signal transduction pathways of the receptor.

Human TNFα is first synthesized as a transmembrane precursor protein of 26 kDa uncleaved monomers (Castiñeiras et al., 2018b), that is then proteolytically

processed into its active monomeric form, resulting in soluble TNFα that is a homotrimer of 17-kDa cleaved monomers (Huang et al., 2012).

Currently, recombinant human TNFα (rhTNFα) is commercially available in the biopharmaceutical market. This protein is used as an intravenous infusion for the treatment of soft tissue tumours of the limbs to prevent limb amputation. The non-proprietary name tasonermin is used internationally for rhTNFα. It is produced as a soluble monomer of 17 kDa in *E. coli*. The European Medicines Agency has approved it in 1999 for use in treating soft-tissue sarcoma, and since then, Boehringer Ingelheim has been manufacturing and selling it under the trade name of Beromun®. Beromun® is a recombinant TNFα that is produced in *E. coli* using advanced recombinant DNA technology. Due to its commercial value and availability of reference material, various studies have used it as a model protein for cytoplasmic RPP (Castiñeiras et al., 2018b).

Cimzia® (UCB) and Simponi® (Janssen), on the other hand, are both approved TNFα inhibitors that are commercially available as humanised antibodies. They are predominantly used for the treatment of rheumatoid arthritis (RA) and Crohn's disease (Walsh, 2010). The role of immune-based cytokines such as TNFα as key regulators of the inflammatory process in RA pathogenesis has been well established in the past decade. TNFα, a prototype member of the TNF-related ligands, has various pro-inflammatory and immune-regulatory functions, thus making it a suitable therapeutic target for the selective blockade treatment in antibody therapy. Owing to their significant clinical efficiency, the development of anti-TNFα biologics have significantly revolutionised the treatment of various

autoimmune diseases such as RA, inflammatory bowel disease (IBD), ankylosing spondylitis (AS), and psoriasis.

One of the major limitations of functional and structural TNFα synthesis is its reduced expression in natural sources. However, there have been promising studies that achieved a high volume of bioactive and soluble TNFα due to advancements in recombinant DNA technology. For instance, *E. coli* is a well-characterised prokaryotic host system and a popular choice for heterologous protein expression due to several advantages such as its easy handling, large-scale production and inexpensive cultivation. Although TNFα expression in *E. coli* has been well studied, the reducing environment of the cytoplasm has caused the product to be unstable with the tendency of producing inactive inclusion bodies. In addition, the refolding processes in the cell may not fully restore the native fold of the protein, thus reducing the function of the recombinant protein generated.

1.6 Fumarase B (FumB)

Fumarase B (FumB) is an iron-sulphur (FeS) protein. The FeS centre insertion into FumB is likely to be a limiting factor in the formation of an active enzyme and will require further optimisation procedures. Fumarase (EC 4.2.1.2), also known as fumarate hydratase, participates in an important series of chemical pathways known as the tricarboxylic acid (TCA) cycle (Krebs cycle). Specifically, it catalyses the reversible hydration of fumarate ion to L-malate ion and the dehydration of D-tartrate to oxaloacetate (Barbara et al., 2013).

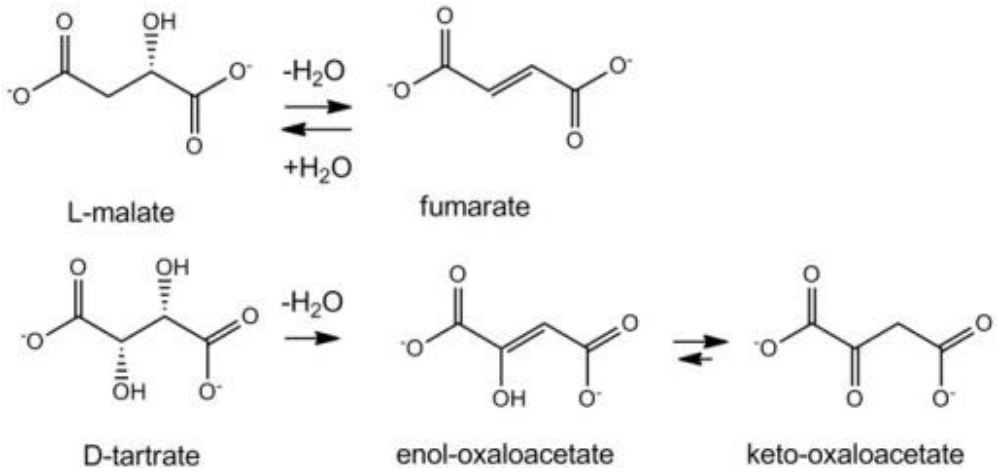


Figure 1.6: Diagram of the fumarase-catalysed reactions of reversible hydration of L-malate and dehydration of D-tartrate (Barbara et al., 2013).

E. coli expresses three distinct fumarases, namely FumA, FumB, and FumC. The comparison of their gene sequences and biochemical characteristics reveal that they consist of two biochemically distinct types of fumarase. Class I fumarases consist of FumA and FumB, and contain an oxygen-sensitive catalytic cluster [4Fe-4S] that can be inactivated upon oxidation to produce a 3Fe-4S centre. They are rendered inactive by exposure to heat or radiation, and very sensitive to superoxide anion as they are iron II (Fe^{2+}) dependent. One of the ions in the iron cluster is not coordinated by amino acid residues and can therefore participate in the catalytic reaction (Barbara et al., 2013). FumC, on the other hand, is a class II enzyme that does not depend on iron for its activity. FumA and FumB yield thermolabile proteins, while the FumC gene product is thermostable (Tseng et al., 2001).

Unlike most of the FumA that is synthesised under aerobic conditions, the FumB enzyme showed a greater abundance under anaerobic conditions and its expression essentially requires FNR (Fumarate and Nitrate reductase Regulatory) protein as a transcriptional activator (Constantinidou et al., 2006; Tseng et al., 2001). In particular, FumB is required for its anaerobic growth on D-tartrate and has been proven to be involved in the formation of biofilm (Herzberg et al., 2006). The *fumB* gene expression is four times higher under anaerobic conditions as compared to aerobic conditions (Tseng, 1997). This is in contrast with the *fumA* and *fumC* gene products that were highly expressed under the same aerobic conditions (Tseng, 1997).

1.7 Flow cytometry (FCM)

Flow cytometry (FCM) is an effective method to analyse several characteristics of individual particles, usually cells, in which thousands of cells move through a laser beam one at a time, and the light emitted by each cell is collected concurrently. Flow cytometry software allows for the statistical analysis of collected data, allowing for the evaluation of cellular characteristics including size and complexity.

As compared to current bulk fluorescence measurement techniques such as OD and fluorimetry, flow cytometry is a powerful method used to analyse multiple characteristics of individual cells within a heterogeneous population of cells. Díaz et al., (2010) stated that there are significant levels of heterogeneity in even the simplest organisms such as microbes. For instance, Davey and Kell, (1996) acknowledged three major causes of culture heterogeneity which are based on the position of cells in the cell cycle, physiological differences, and irregular gene expression. RPP is also known to introduce heterogeneity in the culture due to cell viability, plasmid loss and IBs formation. Hence, the flow cytometry is a good technique in bioprocess monitoring or RPP production, particularly for single-cell analysis within a heterogeneous population (Want et al., 2009; Broger et al., 2011).

1.7.1 FCM components and functions

Flow cytometry is employed in diverse applications, ranging from immunophenotyping of cells to cell counting, green fluorescent protein (GFP) expression analysis and ploidy analysis. **Figure 1.7** summarises the main components of flow cytometry in the standard operation. In general, there are three main systems in the flow cytometer which comprise fluidics, optics and detectors. In the fluidics system, the sample is presented to the examination point and the waste is subsequently removed. The laser beam is the light source used for scattering and fluorescence emission of the sample. The optics system collects and directs the lights, while the detector system collects the lights and converts the signals collected from the detectors into digital data for analysis. The examination point is the key feature of the flow cytometry system as this is the section where the laser and the samples intersect, and the optics system collects the emerging scatter and fluorescence of the sample.

For transportation of samples to the laser, the samples are transported through the examination point which has the laser beam. For the accuracy of data collection, it is essential for particles or cells to move individually through the laser beam. In most flow cytometers, the sample stream containing the cells are injected into a flow of sheath fluids or a solution of saline. The sample stream becomes compressed to approximately one cell diameter in size. This process is known as hydrodynamic focusing. This step is important to enable individual cell analysis and make precise measurements as well as avoid nozzle obstruction (Díaz et al., 2010).

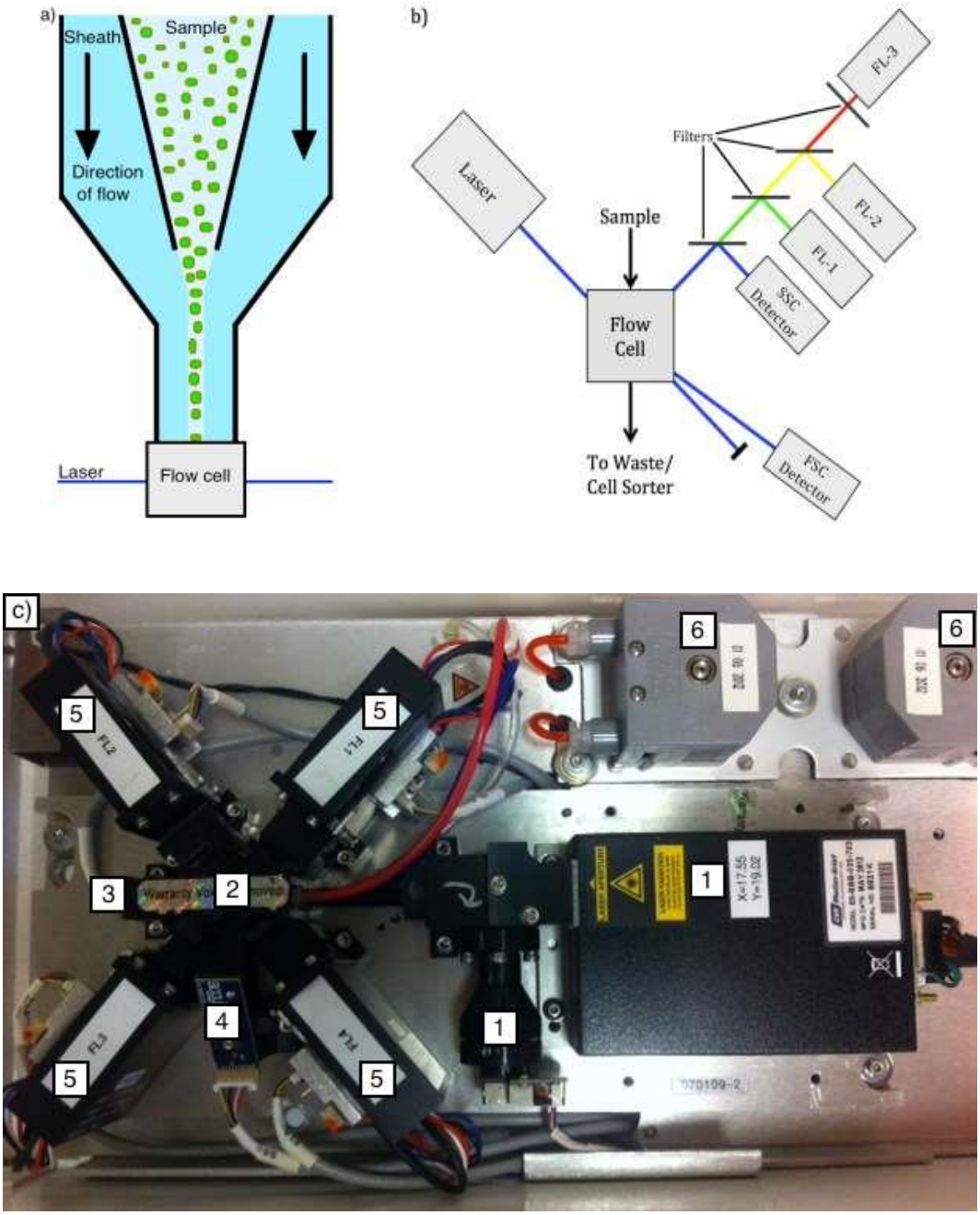


Figure 1.7: Standard operation of flow cytometry (Wyre, 2014).

a) Insertion of a sample into sheath fluids to obtain laminar flow and coordination of the cells into a single line (to approximately one cell diameter) before moving through the flow cell. b) The schematic optical system in flow cytometry: cells intersect with a laser beam in the flow cell to scatter the light either to the forward scatter or the side scatter and emits a fluorescence signal. c) Optical and fluidic systems in BD-Accuri C6. The optical detectors are placed around the flow cell (1 is lasers, 2 is flow cell, 3 is FSC-detector, 4 is SSC-detector, 5 is fluorescence detectors and 6 is peristaltic pumps).

When the cell moves through the laser beam, the light is scattered at all angles. In forward scattering, the amount of light is scattered in the forward direction as soon as the laser beam strikes the cell. The intensity of the forward scattering is approximately proportional to the size of the cell and the data is used to specify this parameter. The amount of light is quantified by an FSC-detector in the system that converts the light intensity into voltage. Once the cells cross the laser beam, light is also scattered around the obstruction bar. The light that is scattered to the side is due to the complexity and granularity of the cells (Marti et al., 2000). This side scatter is directed through a lens system and collected by SSC-detector which is generally positioned 90 degrees from the laser path. The signals collected by both forward and side scatter detectors will be translated into a histogram or dot plot.

Furthermore, flow cytometry is able to identify immediate light from excited fluorescence molecules such as fluorescence labelled antibodies, dyes, or stains. As the cells pass through the laser beam, the fluorescence molecule is excited by the wavelength of the laser light. Subsequently, the light emitted by the fluorescence molecules or fluorophore is directed along the path to the emission filter that allows the further detection of numerous fluorophore emissions in a cell. Hence, flow cytometry is used to identify the existence of different fluorophores and subsequently, quantify the number of fluorophores in the same sample.

1.7.2 Fluorescent dyes and detection

The ability to detect light within cells at the single-cell level is a great advantage of flow cytometry. The most common application of FCM and fluorescent dyes is the determination of cellular viability, which is commonly accomplished through the use of a dual staining system consisting of propidium iodide (PI) and bis-oxonol (BOX).

PI, is a red-fluorescent DNA binding dye, in which the fluorescence of PI is increased by a factor of 20-30 after binding to DNA. Its excitation maximum is shifted to the red by ~30-40 nm, and its emission is shifted to the blue by ~15 nm (excitation 530 nm, emission 625 nm when bound to DNA). Since PI cannot pass through the cell membrane as it is impermeable, PI can only enter a cell and bind to DNA if the cell membrane is damaged, which is therefore PI is employed as a stain for dead cells, which fluoresce red when they are dead (Johnson & Spence, 2010; Fernández-Castané et al., 2017).

Anionic, lipophilic dyes like BOX can penetrate the membrane of a living cell, however they are not allowed inside of cells with a healthy membrane potential. To put it another way, only dead (membrane compromised) or injured cells (membrane depolarised) will appear fluorescent when exposed to these dyes.

Lehtinen et al. (2004) described an alternative viability staining method that makes use of GFP and PI to identify dead cells; in this method, the loss of GFP fluorescence serves as an additional indicator of cell death.

1.7.3 The benefits of FCM

With the introduction of flow cytometry, the study of bacterial physiology and viability as well as their metabolic capabilities and basic functions at the single-cell level can be performed easily (Müller and Nebe-von-Caron, 2010). Flow cytometry can be used to measure and analyse numerous physical characteristics of particles or cells simultaneously.

Furthermore, flow cytometry also enhances the understanding of physiological diversity in closely related cell populations. For instance, the diversity and individuality of cells indicate the complexity and concerted actions of the microbial consortia. Therefore, flow cytometry acts as a mediating technology that offers an increased understanding of the heterogeneity of cell populations and the functions of microbial communities.

1.8 GFP as a reporter protein technology

1.8.1 Introduction to GFP

The term 'GFP' refers to the green fluorescent protein, in which it was initially isolated from *Aequorea Victoria*, a bioluminescent jellyfish. The GFP absorbs blue light with an excitation maximum of 395 nm, while emits green light with an emission maximum of 510 nm (Cormack et al., 1996). Due to the fact that it doesn't need any additional cofactors to work, GFP is a great marker for gene expression and a tag for researching protein localisation in a wide range of

organisms. There are GFP mutations that increase the excitation maximum to around 490 nm, in which excitation at 488 nm resulting proteins that fluoresce more intensely (Heim and Tsien, 1996).

In cell and molecular biology, the *gfp* gene is often used as a reporter molecule in gene expression studies. This gene is attached by researchers to a regulatory sequence of a specific gene of interest and these genes have been expressed in bacteria, cell culture, animals or plants, and mammalian cells including humans (Phillips, 2001). This gene can also be modified to be used as biosensors. These genes can be potentially inserted into the organisms and maintained in their genome through breeding, vector injection, or transformation of cells. These reporter genes can indicate whether the cell or organism has incorporated a certain gene into their genome.

1.8.2 Applications of GFP for RPP in *E. coli*

Reporter genes are selected based on characteristics such as ease of identification, measurable, and suitability as selection markers. However, GFP and other functionally related fluorescent proteins (FPs) have an advantage over other reporters or marker proteins as they are functional molecules in intact cells. They are also characterised as simple proteins that are relatively small in size and can be synthesised by numerous cells ranging from bacteria to mammalian cells. Moreover, as they are fluorescent, they can be easily detected at the single-cell level using non-destructive techniques, both on bulk samples; flow cytometry or fluorescence microscopy (Vizcaino-Caston et al., 2012).

Fluorescent proteins have been utilised in many studies as reporters of recombinant protein (RP) synthesis and folding. When *gfp* gene fused to the C-terminus of recombinant genes of interest, and were translated, the translational fusion amount and folding state of the recombinant proteins produced can be determined by measuring the fluorescence. GFP appears to fold correctly and become fluorescent when a recombinant protein does, however when a recombinant protein misfolds, GFP misfolds and exhibits low fluorescence (Vera et al., 2007). Previous studies employed FCM to improve production of model RP-GFP fusions in both lab-scale (Sevastsyanovich et al., 2009a) and larger scale of bioreactor (Wyre and Overton, 2014b).

Considering the diversity of RPP platforms, the ability to generate RP-GFP fusions in a broad range of organisms is a huge benefit. However, there is considerable evidence that accurate protein folding, solubility and FP fluorescence are not directly related, thus caution is suggested when using fluorescent protein fusions for this purpose. In the first evidence of its kind, García-Fruitós et al. (2005) demonstrated that GFP fused to aggregation-prone proteins can form inclusion bodies in *E. coli* that retain significant fluorescence.

Further study showed that the ratio of GFP+ to GFP- cells can be used as a quick and reliable indicator of plasmid retention in the cells. This technique is much faster than other plasmid retention procedures like replica plating, but it has the potential problem of allowing plasmid-free cells to retain their GFP+ phenotype (Ishii et al., 2009).

Another potential downside of using *Aequorea* GFP in microbial research is that the chromophore maturation process requires oxygen. It is possible to recover GFP from an anaerobic environment since chromophore maturation does occur upon exposure to oxygen (Zhang et al., 2005), however this process is not conducive to real-time monitoring because it delays measurement.

1.9 Study aims and objectives

This study focuses on two areas, involving the generation of fumarase B enzyme and recombinant human tumour necrosis factor alpha (rhTNFα) protein. *Escherichia coli* strains are selected as the host expression system and various strategies are employed to increase RPP in this bacterium. Even though several organisms have been explored as RPP expression hosts in recent years, yet the genetically well-characterised and easily cultivated bacterium *E. coli* remains the popular choice in various applications.

The main focus of this study is to optimise the best conditions for the expression of soluble recombinant proteins in *E. coli*. Specifically, the production of rhTNFα protein, a human cytokine utilised as a biopharmaceutical product (Walsh, 2014), is emphasised in this study. Using flow cytometry, the expression of a plasmid-encoded TNFα attached to a GFP from either an arabinose-induced T7 promoter or an arabinose-inducible paraBAD promoter was analysed. Various effects of growth parameters including carbon source, temperature and inducer concentration on TNFα-GFP cell viability, its folding states and productivity were investigated at the single-cell level. The production of TNFα using a TNFα-GFP

fusion was compared to the GFP expression individually using the same plasmid under similar conditions. The comparison of promoter strength and culture heterogeneity in bacterial cell physiology was investigated, therefore determining the effect of TNFα on gene expression strength and heterogeneity.

Another focus is to aim for the expression of FumB protein in *E. coli* since the protein is needed as material for research projects in the laboratory. This area comprising the cloning of fumarase B (fumB) genes into pET41c plasmid and making fumarase B (FumB) protein in *E. coli*. The challenge, however, is not only due to FumB being an iron-sulphur (FeS) protein, but also the aggregation of inclusion bodies in recombinant bacteria as a consequence of protein misfolding. For instance, FumB is known to readily generate IBs due to protein misfolding. Thus, stress minimisation can be used to improve protein folding.

One of the major advantages of using *E. coli* is that it can be grown to a high cell density using high cell density cultivation (HCDC) techniques in appropriate bioreactors, thus allowing the production of high quantities of heterologous proteins. Thus, this study will include the development of HCDC in intensified fed-batch bioreactor, in which all the developmental conditions can be controlled to regulate the production of IBs in recombinant bacteria.

CHAPTER 2

MATERIALS AND METHODS

This chapter describes in detail all the materials that have been used to conduct the experiments as well as the procedures and methods involved in the research study. This includes all bacterial strains, chemicals, reagents, and equipment. All solutions were prepared and diluted using double distilled water (ddH_2O), unless otherwise stated.

2.1 Materials

All commercially-produced reagents, chemicals, consumables and media components used in the present work were obtained from Thermo Fisher Scientific, Sigma-Aldrich, VWR International, New England BioLabs, Life Technologies (Invitrogen) and Becton, Dickinson and Company (BD), unless indicated otherwise.

2.1.1 Liquid growth media

LB broth (LB) of the following composition: 10 g/L tryptone-peptone, 5 g/L yeast extract and 5 g/L NaCl was prepared by dissolving 20 g of LB powder into 1 L of double distilled water (ddH_2O). The broth was sterilised by autoclaving at 121 °C for 15 minutes for sterilisation.

Terrific broth (TB) of the following composition: 12 g/L tryptone, 24 g/L yeast extract and 4 mL/L glycerol solution (all were obtained from Oxoid™ Thermo Fisher Scientific) were dissolved in ddH_2O and sterilised by autoclaving at 121 °C

for 15 minutes. 10 % (v/v) of potassium phosphate solution was prepared and sterilised separately and mixed in the mixture broth solution before use.

For the preparation of 1 L potassium phosphate solution the final desired concentrations are: 0.17 M monobasic potassium phosphate (23.135 g/L KH_2PO_4, molecular weight = 136.09 g/mole) and 0.72 M dibasic potassium phosphate (125.410 g/L K_2HPO_4, molecular weight = 174.18 g/mole).

If glucose was required in the medium, 0.4 % (w/v) was added in the mixture broth solution; this medium had only less than 1 % glucose in it, and therefore it was possible to autoclave all of the ingredients together without the risk of unwanted Maillard (browning) reactions. If antibiotics (in this case Kanamycin) were essential, these were also added to the medium by filtered sterilisation to prevent thermal inactivation. All solutions were prepared using double distilled water (ddH_2O), unless otherwise mentioned.

2.1.2 Solid growth media

Solid growth media was prepared by dissolving 11.2 g of nutrient agar (NA) (Oxoid™ Thermo Fisher Scientific) into 400 mL ddH_2O and autoclaved for sterilisation at 121 °C for 15 minutes. Preparation of selective agar with antibiotics or other supplements was done when the molten agar has cooled to at least 50 °C to limit thermal degradation of antibiotics upon addition. Filtered-sterilised Kanamycin was added to a final concentration of 50 μg/mL. Molten agar was poured aseptically into sterile petri dishes to an approximate volume of 20 mL per

dish and allowed to set and dried at room temperature to remove excess moisture. To keep its freshness, all the plates were kept in 4 °C refrigerator and those containing antibiotic were used within 2 months.

2.2 Buffers and solutions

2.2.1 Antibiotics and other supplement additions

The 1000x concentrated Kanamycin stock solution (50 mg/mL) was prepared by dissolving Kanamycin sulfate from *Streptomyces kanamyceticus* (purchased from Sigma-Aldrich) in 0.9 % NaCl and filter sterilised with 0.22 μm filter. When needed, Kanamycin was added to culture media at a final concentration of 50 μg/mL. The stock was stored at -20 °C when not in used.

Arabinose, used as inducer for arabinose-inducible expression systems (for GFP and TNFα studies) was diluted in ddH_2O to make a 50 mg/mL stock solution. Two different concentrations of 0.015 % and 0.2 % arabinose were used in shake-flask batch culture and were added at the required time point.

IPTG (Isopropyl β-D-1-thiogalactopyranoside) stock solutions used as inducer for lactose-inducible expression systems (for Fumarase B studies) were required at two different concentrations (8 and 500 μM), depending on batch culture protocol; 1000x working concentration was prepared for both solutions and diluted with ddH_2O to produce the desired concentration for culture media. For safe and

durable storage, both solutions were wrapped in aluminium foil for the protection from light and stored at 4 °C.

Lysozyme from chicken egg white (purchased from Sigma-Aldrich); used for fractionation (cell lysis) protocol, was prepared at 20 mg/mL (stock solution) by dissolving the powder into ddH_2O and stored at -20 °C until use. All solutions were filter sterilised using a sterile 0.22 µm disposable filter unit (Millipore). A summary of antibiotics and supplement additions used in this work is presented in **Table 2.1**.

2.2.2 Phosphate buffered saline (PBS)

PBS was used for cells dilution to maintain specific osmolarity and pH of culture. The commercially produced PBS tablets (Oxoid) with Dulbecco's formulation (8.0 g/L NaCl, 0.2 g/L KCl, 1.15 g/L Na_2HPO_4, 0.2 g/L KH_2PO_4, pH 7.3) was prepared according to the manufacturer's instructions and autoclaved for sterilisation. PBS was also used to prepare serial dilution of culture samples.

2.2.3 Buffers for agarose gel electrophoresis

Preparation of TBE buffers were made as follows:

- A 5x TBE stock solution was made by using 0.44 M Tris-HCl, 0.445 M boric acid and 0.01 M EDTA, and then adjusted to pH 8.0.
- 1x TBE buffer was freshly prepared for agarose gel electrophoresis; 5x TBE stock solution was diluted with ddH_2O to a 1:5 solution.

Table 2.1: Stock and working solution concentrations, solvents used for dilution and storage conditions of supplements and antibiotics addition.

Compound	***Stock solution concentration***	***Working solution concentration***	***Solvent***	***Storage***
Kanamycin	50 mg/mL	50 µg/mL	0.9 % NaCl	-20 °C
Arabinose	50 mg/mL	0.015 %	ddH_2O	4 °C
	50 mg/mL	0.2 %	ddH_2O	4 °C
IPTG	8 mM	8 µM	ddH_2O	-20 °C
	500 mM	500 µM	ddH_2O	-20 °C
Lysozyme	20 mg/mL	200 µg/mL	ddH_2O	-20 °C

2.2.4 Buffers and solutions for protein analysis (SDS-PAGE)

Preparation of SDS-PAGE solution were made as follows:

- 0.75 M Tris/HCl stock running gel buffer was prepared by dissolving 18.172 g of Tris (hydroxymethyl) methylamine into 200 mL ddH_2O. The pH was adjusted to 8.3 with HCl.
- 1.25 M Tris/HCl stock stacking gel buffer was prepared by dissolving 75.73 g of Tris in 500 mL ddH_2O. The pH was adjusted to 6.8 with HCl.
- 20 % Sodium Dodecyl Sulphate (SDS) solution was prepared by dissolving 20 g SDS in 80 mL ddH_2O with gentle heating and stirring. The volume was topped up to 100 mL with ddH_2O.
- A volume of 80 mg/mL of ammonium per sulphate (APS) solution was prepared by dissolving 80 mg of APS in 1 mL ddH_2O. The solution was then stored at -20 °C in 1 mL aliquots.
- Preparation of sample buffer was done by mixing 2 g SDS, 20 mL glycerol and 5 mg bromophenol blue (sulphone form) in 92 mL of 10x diluted stock stacking gel buffer (made by diluting 10 mL of stock stacking gel buffer into 90 mL ddH_2O). A 1 mL aliquot of sample buffer was prepared in 1.5 mL microcentrifuge tubes and stored at ambient temperature. 87 µL of β-mercaptoethanol was immediately added to the tube prior to use.
- Electrode buffer stock solution was prepared by dissolving 150 g of glycine in 600 mL ddH_2O with gentle heating and stirring. Then 30 g Tris was dissolved in and the buffer was made up to 1 L with ddH_2O.

- Electrode buffer working solution was freshly prepared by mixing 5 mL of 20 % SDS into 100 mL of electrode buffer stock solution, then made up to 1 L with ddH_2O.
- 1x Tris-glycine-SDS (TGS) running buffer (25 mM Tris Base, 192 mM glycine, 0.1 % w/v SDS, pH 8.3) was prepared by diluting the commercially available 10x stock solution with ddH_2O.
- The SimplyBlue SafeStain Solution from Invitrogen (Life Technologies) was used for staining of SDS- PAGE.

2.2.5 Dyes for flow cytometry

Stock solution of propidium iodide (PI) (1 mg/mL) was prepared by dissolving 4 mg of PI powder (purchased from Sigma-Aldrich, UK) in 4 mL of ddH_2O. The working concentration of PI (200 µg/mL) was prepared by diluting 1 part of the stock solution to 4 part of ddH_2O. This was divided into Eppendorf tube, each containing 1 mL aliquot. The tubes were wrapped with aluminium foil to protect from light and stored at 4 °C for up to 6 months.

2.3 Bacterial strains

The *Escherichia coli* strains that were used in the present work are *E. coli* BL21(T7) and BL21(A). *E. coli* BL21(T7) and *E. coli* BL21(A) were both obtained from Cobra Biologics (Keele, Staffordshire, UK). Both strains were used for protein expression studies and all subsequent cultures and recombinant protein production of GFP and TNFα-GFP. For work involving Fumarase B studies the strain *E. coli* K-12

MG1655 and *E. coli* BL21(DE3)* were used. *E. coli* BL21(DE3)* was the parental strain used in the expression experiments for several recombinant proteins from pET vectors (Alfasi, 2010). *E. coli* strains used in this work and their genotypes are listed in **Table 2.2**.

Table 2.2: *E. coli* strains used in this work.

E. coli strain	***Description***	***Source***
BL21(A)	*F- ompT gal dcm lon hsdSB (rB-mB-) [malB+]$_{K-12}$(λS) Δ(araBAD)*	Cobra Biologics
BL21(T7)	*F- ompT gal dcm lon hsdSB (rB-mB-) [malB+]$_{K-12}$(λS) araBAD::T7RNAP*	Cobra Biologics
*BL21(DE3)**	*F-ompT gal[dcm][ion]hsdsB rne*131*(rB-mB-;* an *E. coli* B strain); with DE3, a λ prophage carrying the T7 RNA polymerase gene	Invitrogen, GSK
K-12 MG1655	*actB-his6 ΔoxyR::Kan*	*E. coli* MG1655 Δ*oxyR::kan*

2.3.1 Stock culture bacteria for long-term storage (-80 °C)

A single colony of the strain (or 1 % overnight culture of strain growth in LB) was inoculated in 10 mL LB media and grown overnight at 37 °C. After bacterial growth, 500 µL of the overnight culture was transferred into a 2 mL screw top tube or cryovial containing 500 µL of 40 % glycerol and mixed gently. The tube was stored at -80 °C and will remain stable for years at this temperature. Note that the shelf life of culture is reduced upon repeated freeze and thaw cycle.

2.4 Plasmids

Expression vectors used in this study are listed in **Table 2.3**. Plasmids T7 and pBAD encoding the recombinant protein product GFP and TNFα-GFP were used in this study for all RPP experiments. Both expression vectors were obtained from Cobra Biologics (Keele, Staffordshire, UK). Whereas pET41c (Novagen), used in Fumarase B studies, is high copy number plasmid encodes a T7 promoter and the expression of the target gene is under the control of T7 polymerase (Kasem Balasiny, 2016). Plasmid maps are given in **Figure 2.1, 2.2 and 2.3**.

Table 2.3: List of plasmid vectors used in the present research study.

Plasmid name	*Description*	*Promoter used*
Expression of recombinant Green Fluorescence Protein (GFP)		
pORT2_2T72_GFPmut2	T7 expression vector (Kan^R)	T7 promoter
pORT2_2BAD2_GFPmut2	pBAD expression vector (Kan^R)	ParaBAD promoter
Expression of recombinant human Tumour Necrosis Factor α (TNFα) fused to GFP		
pORT2_2T72_TNFα-GFP	T7 expression vector (Kan^R)	T7 promoter
pORT2_2BAD2_TNFα-GFP	pBAD expression vector (Kan^R)	ParaBAD promoter
Vector used for the development of the Fumarase B screening		
pET41c	High copy number vector expressing GST fusion protein and also encoding *lacI* sequence (Kan^R)	T7 promoter

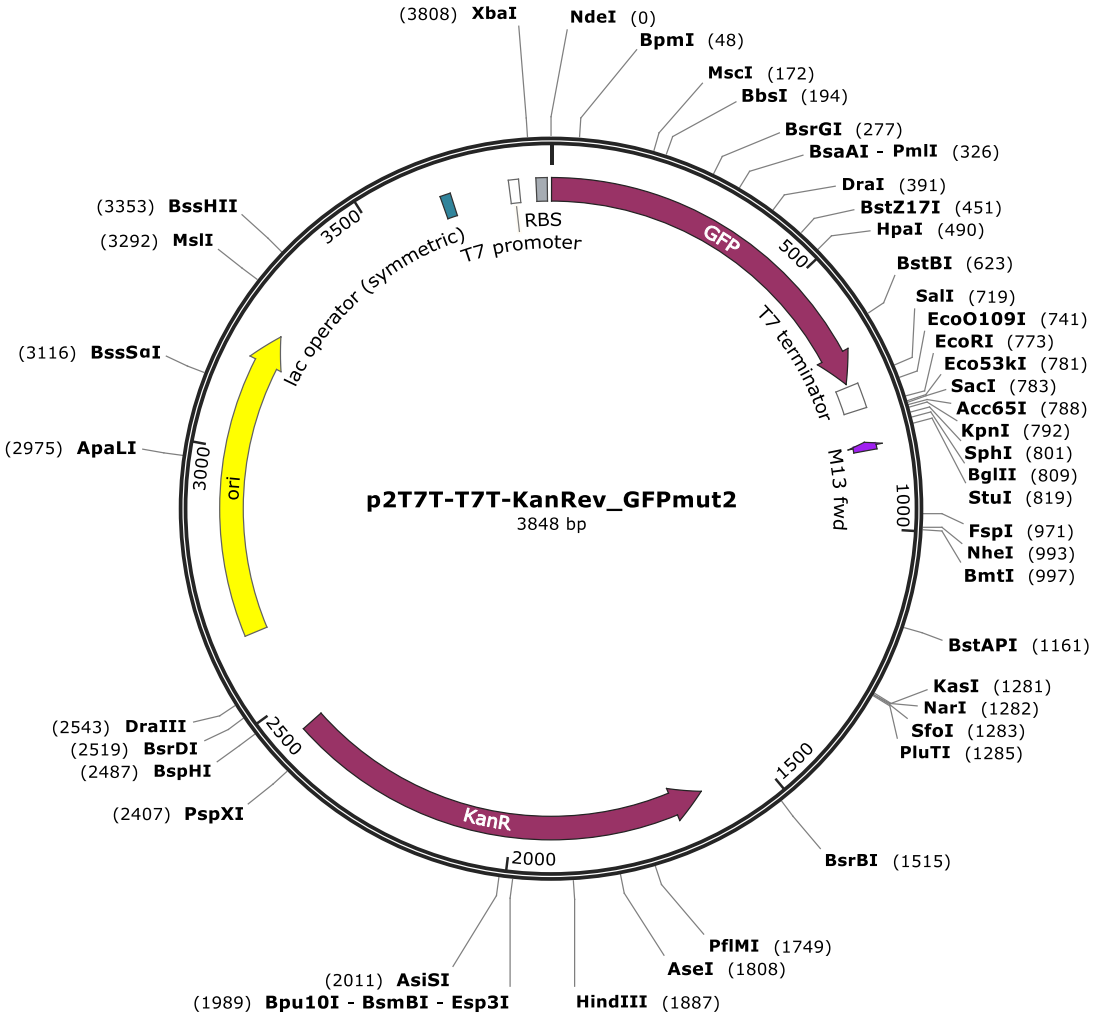


Figure 2.1: T7 expression vector (Kan^R) pORT2_2T72_GFPmut2 for expression of recombinant Green Fluorescence Protein (GFP)

(Courtesy image: Cobra Biologics)

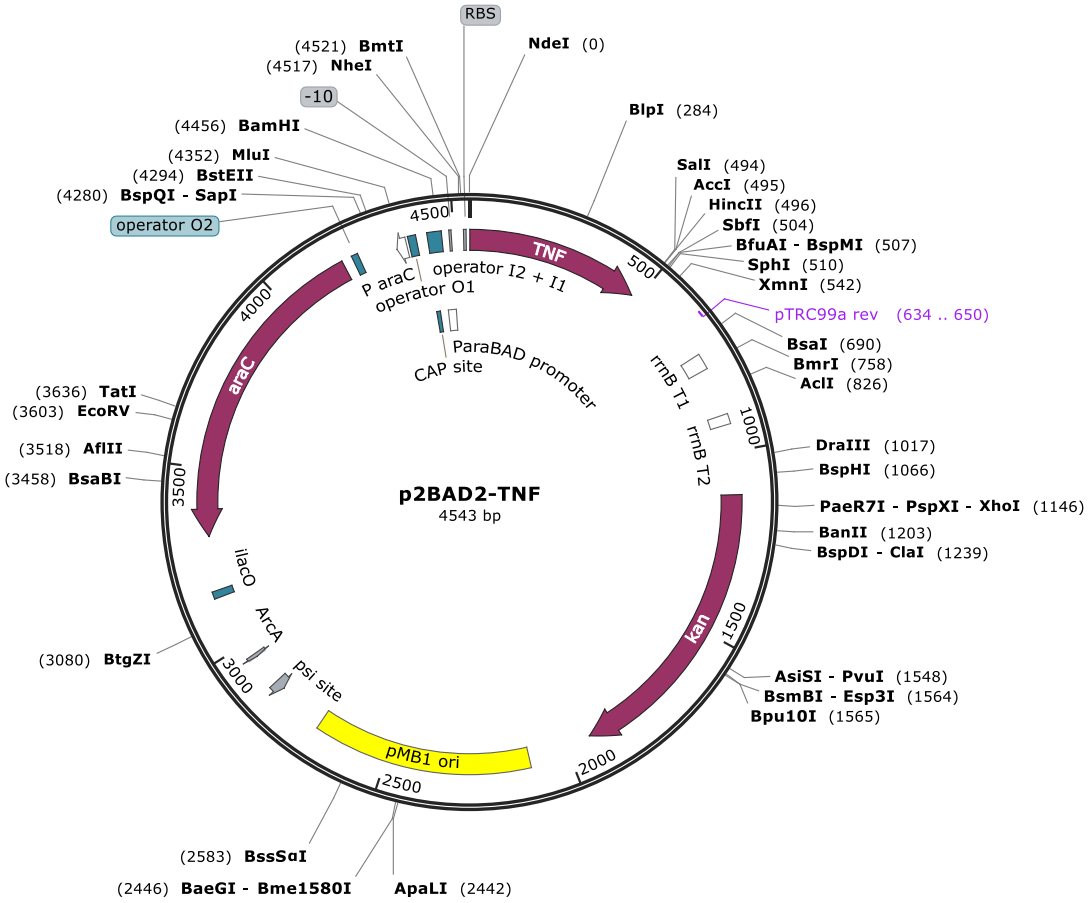


Figure 2.2: pBAD expression vector (Kan^R) pORT2_2BAD2_TNF for expression of recombinant human Tumour Necrosis Factor α (TNFα)

(Courtesy image: Cobra Biologics)

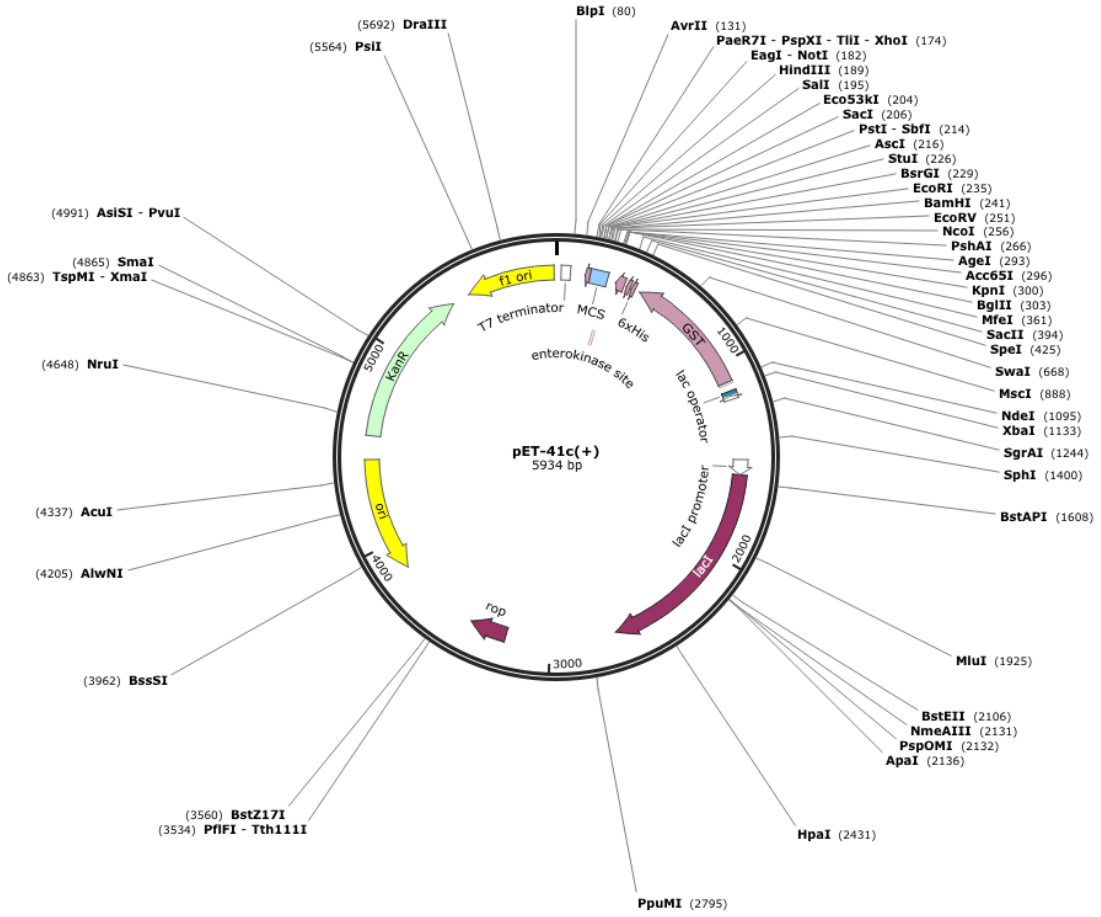


Figure 2.3: pET41c expression vector (Kan^R) for expression of GST fusion protein Fumarase B

(www.snapgene.com)

2.5 Bacterial transformation

2.5.1 Preparation of *E. coli* competent cells

In preparing competent cells, an amount of 5 mL LB growth medium was inoculated with a single colony of the strains and incubated overnight (37 °C, 150 rpm, 18 h). 200 µL (1 %) of overnight culture was inoculated into 20 mL LB growth medium and grown to mid-logarithmic phase (OD_{600} ≈ 0.5) at 37 °C and 180 rpm. Cultures were transferred into sterile, disposable 50 mL falcon tubes (BD-Falcon), and placed on ice bath for 30 minutes. Cells were harvested by centrifugation (3075 g at 4 °C for 5 min) and resuspended in 5 mL $CaCl_2$ (0.1 M) before incubating on ice bath for another 20 minutes. Cells were harvested by centrifugation and resuspended in 5 mL mixture of $CaCl_2$ (0.1 M) and 15 % glycerol. Cells were aliquoted (100 µL) in cryovials or Eppendorf tubes. Competent cells were immediately stored at –80 °C

2.5.2 Transformation of competent *E. coli* with plasmid DNA

An amount of 1 µL plasmid DNA was added into 100 µL aliquots of competent cells (which were first thawed on ice) and incubated on ice for 30 minutes. Cells were heat-shocked using heat block at 42 °C for 90 seconds. Then, the cells were transferred to ice bath for 5 minutes before adding 1 mL of LB broth and incubated at 37 °C for at least 1 hour to allow the cell to recover and express the antibiotic resistance gene. Cells were harvested by centrifugation at 3075 g for 5 minutes at

ambient temperature. The pellet was resuspended in 100 µL LB and spread-plated onto Kanamycin (50 µg/mL) supplemented NA plate. Plates were then incubated at 37 °C overnight to allow for cell growth. Plates with visible colony growth were stored at 4 °C and used within 2 weeks.

2.6 Bacterial growth and RPP in shake-flask for batch culture fermentation

Bacterial strains were aerobically grown in 500 mL shake-flasks containing 100 mL of terrific broth (TB) media supplemented with 50 µg/mL of Kanamycin (**Section 2.1.1**) for plasmid maintenance. The frozen working solution of Kanamycin were thawed at 4 °C overnight before used. To start with, an amount of 10 mL LB growth medium containing 50 µg/mL of Kanamycin was inoculated with a single colony of the strains and incubated overnight (37 °C, 150 rpm, 18 h). 1 mL (1 %) of overnight inoculum were added into the shake-flasks and incubated at 30 °C, 180 rpm. Arabinose was induced for recombinant gene expression at mid-logarithmic phase when OD_{650} reached ≈ 0.5. Cultivation temperature varied between experiments.

2.7 Bacterial growth and RPP in bioreactor for fed-batch fermentation

2.7.1 Vessel and fermentation equipment

A FerMac 310/60 bioreactor (Electrolab) with a 5 L cylindrical vessel was used for all fed-batch fermentations run in this study. The reactor was equipped with 2 vertical six blade Rushton turbine impeller, 4 baffles and an aeration system

located below the lower impeller that capable to sparge the air at a rate of 3 L per minute through an autoclavable, reusable 0.22 μm filter (Sartorius). To prevent possible blockage to the exhaust filter from foamed culture, two 1 L catch pots filled with 1 mL polypropylene glycol (PPG) as antifoam (Sigma-Aldrich) were placed between the condenser and filter. The vessel was autoclave-sterilised at 121°C for 15 min once filled with medium.

2.7.2 Media and growth condition

A semi-defined medium was used for the growth of high cell density fed-batch fermentations in 5 L bioreactor. Initial batch phase fermentation started with 1.5 L medium, with the addition of 0.5 L feeding solution later. 2% (v/v) of overnight culture was inoculated into Kanamycin supplemented medium (50 μg/mL) adjusted to pH 6.3. Overnight inoculum cultures were grown from a sweep of transformant cells in 250 mL conical flasks containing 35 mL Luria Bertani (LB) medium with 50 μg/mL Kanamycin at 25°C, 150 rpm for 18 - 21 hours. Prior to addition to the vessel, a small amount (5 mL) of inoculum was withdrawn for screening and analysis. The medium was also supplemented with trace metal solution and polypropylene glycol was added to the medium as antifoam. Glycerol was utilised as the carbon source. Analysis of OD_{650} and/or DOT/off-gas was done to determine the stage of cell growth. Fermentations were terminated once no more growth was observed. The recipe of semi-defined growth medium, trace metal solution, feed composition and feeding rate equation were summarised in **Tables 2.4** and **Table 2.5.**

Table 2.4: Recipe of semi-defined growth medium and trace metal solution.

	Item	***Stock solution concentration***	***Final volume (1.5 L)***	***Remarks***
Semi-defined medium				
1	$(NH_4)_2SO_4$	14 g/L	21 g	
2	Yeast extract	20 g/L	30 g	
3	KH_2PO_4	2 g/L	3 g	
4	K_2HPO_4	16.5 g/L	24.75 g	
5	Citric acid	7.5 g/L	11.25 g	
6	H_3PO_4	1.5 mL/L	2.25 mL	Concentrated 85%
7	Polypropylene Glycol	0.66 mL/L	0.99 mL	Antifoam
8	Glycerol	35 g/L @ 28mL/L	52.5g @ 42mL	Carbon source
		Top up with ddH_2O	1429.5 mL	
Post-sterilisation additions (All filtered-sterilise)				***Concentration of the added solution***
9	Trace metal solution*	34 mL/L	51 mL	
10	$MgSO_4.7H_2O$	10 mM	15 mL	1 M (24.65 g/100 mL)
11	$CaCl_2.2H_2O$	2 mM	3 mL	1 M (7.35 g in 50 mL)
12	Kanamycin	50 mg/L	1.5 mL	50 µg/mL
		Total	70.5 mL	
**Trace metal solution*				
13	$FeSO_4.7H_2O$	3.36 g/L		
14	$ZnSO_4.7H_2O$	0.84 g/L		
15	$MnSO_4.H_2O$	0.15 g/L		
16	$Na_2MoO_4.2H_2O$	0.25 g/L		
17	$CuSO_4.5H_2O$	0.12 g/L		
18	H_3BO_3	0.36 g/L		
19	H_3PO_4	48 L/L		Concentrated 85%

Table 2.5: Recipe of feeding composition and feeding rate equation.

	Item	***Stock solution concentration***	***Feeding volume (0.5 L)***
Feeding composition			
1	Glycerol	714 g/L	357 g @ 283 mL
2	$MgSO_4.7H_2O$	7.4 g/L	3.7 g
3	Arabinose (*Filtered-sterilise*)	4 mM	5 mL (400mM)

**Feeding started when the initial carbon source has depleted as indicated by an increase in the DOT.*

Feeding rate	
Equation	$$F = \left(\frac{1}{S}\right) \times \left(\frac{\mu}{Y_{XS}} + m\right) \times X_0 \times e^{\mu t}$$
	F = Feeding rate into the bioreactor (L/h), X_0 = Total biomass in bioreactor at start of feed (g DCW), μ = Specific growth rate set at 0.2/h t = Time (h), S = Feed glycerol concentration (714 g/L), Y_{XS} = Cell yield [Amount of biomass (x) produced per mass of substrate (s)] M = Maintenance coefficient [g substrate. g biomass^{-1}. h^{-1}] μ = Set as 0.2

2.8 Cloning of *fumB* gene

2.8.1 Polymerase Chain Reaction (PCR) amplification

The targeted *fumB* gene was first isolated from the genomic DNA of *E. coli*. In running a 1st round PCR, all chemicals (20 µL 5x Hifi buffer, 1 µL 100 mM dNTPs, 72 µL sterile-distilled water, 2 µL 10 mM primer mix (forward and reverse) and 0.5 µL template DNA) were mixed thoroughly. A volume of 2 µL of velocity DNA polymerase was then added, mixed and spun well. Amplification reactions were performed in 2 PCR tubes of 50 µL volume. PCR conditions for targeted gene amplification started with a single denaturation step at 98°C for 5 min. A cycle of denaturation at 98°C for 30 s was done for 35 repetitive cycles followed by annealing at 55°C for 30 s. Elongation steps was carried out at 72°C for 1 min before completing the final elongation step at 72 °C for 10 min. **Table 2.6** summarises the PCR condition used in 1st round PCR.

E. coli K-12 MG1655 genomic DNA was used to amplified *fumB* gene. Generally, primers were about 60 bases long. The forward primer had around 25 bases complementary to the 5' end of the desired plasmid insertion point, a start codon and 20 - 25 bases of the 5' end of the interest gene. The reverse primer annealed to the plasmid with around 25 bases complementary to the 3' end of the point of insertion. The expected sizes of the PCR products of *fumB* gene were 1647 bp (base pairs). 3 µL sample of the 1st round PCR amplification was analysed through gel electrophoresis to confirm the PCR product. Primers used during the Fumarase B studies are listed in **Table 2.7**.

Table 2.6: The PCR condition used in 1st round PCR.

	Cycles	***Temperature (°C)***	***Duration (min)***
Initial denaturation	1	98	5
Denaturation	35	98	0.5
Annealing	35	55	0.5
Elongation	35	72	1
Final elongation	1	72	10
Hold		4	

Table 2.7: Primers used in Fumarase B studies.

Primer name	***Sequence 5' to 3'***	***Source***
Forward primer FumB RF_FRW	TTT AAC TTT AAG AAG GAG ATA TAC ATA TGT CAA ACA AAC CCT TTA TCT ACC AGG CAC	Invitrogen
Reverse primer FumB RF_RVS2	TTA GTG GTG GTG GTG GTG GTG GTG GTG CTC GAG CTT AGT GCA GTT CGC GCA CTG TTT GTT	Invitrogen

2.8.1.1 Agarose gel electrophoresis

To prepare the agarose gel for electrophoresis, 1% (w/v) of agarose powder (Sigma-Aldrich) was dissolved into 100 mL of 1x TBE buffer (**Section 2.2.3**). The solution was heated for 1 to 3 minutes in microwave until nice rolling boil present and the agarose was fully dissolved. The agarose solution was cooled for 10 - 15 min and poured into a gel tray with the well comb in place. Once it set, it was placed into the electrophoresis unit and filled with 1x TBE buffer until the gel was covered. Ethidium bromide (EtBr) was supplemented into the buffer to a final concentration of approximately 0.2 - 0.5 µg/mL (usually about 2 - 3 µL of lab stock solution). Attachment of EtBr to DNA made the visualisation of the DNA on the gel possible under ultraviolet (UV) light to locate the genes.

To estimate the size of DNA samples, each DNA sample (3 µL) was mixed with 1 µL of loading dye (Gel Loading Dye, Blue (6X), New England BioLabs). Samples were carefully pipetted into the gel well with the molecular weight ladder (HyperLadder II, New England BioLabs) was loaded on the first lane. The electrophoresis was operated at 100 V for 30 to 40 minutes. The power was switched off once the dyes has successfully moved through 75 - 80 % of the gel. The gel was removed, and DNA fragments were visualised by a UV transilluminator.

2.8.2 DNA purification

It is very important to have an extremely clean and concentrated PCR products from the 1st round PCR amplification. In this study, PCR products were first

purified using a QIAquick PCR purification kit (Qiagen) prior to use in the subsequent linear amplification reaction (restriction-free cloning) (Van Den Ent and Lowe, 2006) via 2nd round PCR.

2.8.2.1 QIAquick PCR purification kit

The PB buffer were mixed with the PCR reaction solution in a volume ratio of 5:1. A 10 µL of 3 M sodium acetate (pH 5.0) was added to the mix if the colour of the solution was orange/ violet colour. This will turn the colour of solution into yellow.

The solution was transferred to the QIAquick purification column for centrifugation at 20,784 *g* for 60 seconds. This caused the DNA to bind to the column and the flow-through was discarded. Next, the column was rinsed with 750 µL of PE buffer (wash buffer) and centrifuged for 60 s. The flow-through was discarded and the QIAquick column re-centrifuged for another 1 min to remove the remaining wash buffer. The QIAquick column was transferred to a fresh 1.5 mL microcentrifuge tube. 30 µL of sterile distilled water was added (instead of elution buffer) to the centre of the QIAquick membrane and left for 1 min on the bench. The flow-through (concentrated purified DNA) was collected by centrifugation for 1 min.

2.8.3 Restriction-free cloning (2nd round PCR)

A volume of 4 µL purified PCR product *fumB* gene was used as a pair of primers in the linear amplification reaction with following components; 0.5 µL of destination vector pET41c, 1 µL 100 mM dNTPs (25 mM each), 20 µL 5x Hifi buffer, up to a

final volume of 100 µL with sterile distilled water. 2 µL of velocity DNA polymerase was then added last, mixed and spun well. Restriction-free (RF) reactions were performed in 0.2 mL PCR tubes with a final volume of 33.3 µL. The RF reaction started with a single initial denaturation step (95°C, 3 min), followed by 35 repetitive cycles of 1 min denaturation at 95°C, 1.5 min temperature gradient annealing at 50, 55 or 60°C, and 12 min elongation at 68 °C. The process was ended with 10 min of final elongation at 68°C. **Table 2.8** summarises the PCR condition used in 2nd round PCR.

Positive candidates of ligated pET41c-*fumB* gene were checked by gel electrophoresis to confirm the band size. Following completion of the RF reaction in 2nd round PCR, the samples in all 3 tubes were pooled and treated with 1.5 µL DpnI for 2 hours at 37°C to specifically eliminate the methylated parental plasmid DNA. The QIAquick PCR purification kit was used to purify the reaction mix and the DNA was then directly transformed into appropriate competent cells.

2.8.4 Cloning of *fumB* gene into pET41c plasmid

fumB gene were cloned into pET41c plasmid due to its T7 RNA polymerase-dependent promoter and Kanamycin resistance marker. It also has the popular GST (Glutathione S-transferase) fusion tag for enhanced production and solubility.

Table 2.8: The PCR condition used in 2nd round PCR

	Cycles	***Temperature (°C)***	***Duration (min)***
Initial denaturation	1	95	3
Denaturation	35	95	1
Annealing	35	Gradient (50 - 60)	1.5
Elongation	35	68	12
Final elongation	1	68	10
Hold		4	

2.8.5 Transformation of BL21(DE3)* with pET41c-*fumB*

The recombinant pET41c plasmid containing the *fumB* gene was transformed into BL21(DE3)* competent cells and grown on Kanamycin supplemented (50 μg/mL) nutrient agar (NA). In preparing competent cells, cell pellets were resuspended in 30 mL cold TFB1 buffer and incubated for 90 minutes in ice bath. TFB1 was prepared by mixing 15% glycerol, 30 mM potassium acetate, 100 mM RbCl, 50 mM $MnCl_2$ and 10 mM $CaCl_2$. Solution was adjusted to pH 5.8 and filter-sterilised. Cells were then collected by centrifugation and resuspended in 4 mL of ice cold TFB2 buffer (containing 10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2$ and 15 % glycerol; adjusted to pH 6.8 with KOH and filter-sterilised). A volume of 100 μL aliquots were prepared in 1.5 mL microcentrifuge tubes and immediately stored at -80°C. Kanamycin was used to select the transformed cells.

Transformation method used in this study are heat shock transformation. Exposing the cells to heat shock or a sudden increase in temperature will create a pressure difference between inside and outside the cell. This led to pores formation in the cell walls and allows the supercoiled plasmid DNA to enter. The cell walls will self-heal when returned to a more normal temperature. Cells with the plasmid will be able to grow on Kanamycin enriched agar plates.

To start the heat shock transformation, competent cells were thawed on ice. 25 μL of cold plasmid (possibly ligated pET41c-*fumB*) was gently mixed with the competent cells and returned to ice for 30 min. Next, the heat shock was carried out in a 42°C water bath for 30 s and immediately returned to ice for another 5

min. 250 µL of LB (Luria Bertani) broth was added to the cells, and the mixtures were incubated for 90 min at 37°C (200 rpm) for cell recovery. The complete transformation reaction mixtures were plated onto Kanamycin supplemented nutrient agar plates (NA-Kan plates).

2.9 Cell analysis techniques

2.9.1 Spectrophotometer

Optical density (OD) of the cultures were measured in semi-micro disposable polystyrene cuvettes (ThermoFisher Scientific, UK) at the wavelength of 650 nm using an Evolution 200 Spectrophotometer (ThermoFisher Scientific, UK), diluting with PBS where necessary (if the reading was more than 0.8). Another cuvette was used as a blank, zeroing the machine with ddH_2O. Then, the OD readings were recorded for further analysis of the data.

2.9.2 Bacterial count (Colony Forming Unit – CFU)

Cultures were diluted in a serial dilution up to 9-fold dilution times and 900 µL of sterile aliquot phosphate-buffered saline (PBS; Oxoid, UK) solution were used as diluent. To determine the culturability, 100 µL of diluted cultures were spread on non-selective NA plates under aseptic technique. Next, the plates were incubated at 30 °C for 24 h. The number of viable colonies grown on the plates were counted using the colony counter. The plate count was always performed in replicates. The

acceptable counting range for a standard size petri dish was between 30 to 300 CFU per plate.

2.9.3 Plasmid retention

To determine plasmid retention, the colonies were replica-plated on NA supplemented Kanamycin (NA-Kan) plates and incubated for 24 h at 30 °C. Viable colonies were counted using colony counter. The percentage of the plasmid retention expression were determined by using the formula below:

$$Percentage\ of\ Plasmid\ Retention\ (\%) = \frac{Number\ of\ colony\ grown\ on\ NA\ plate}{Number\ of\ colony\ grown\ on\ Na - Kan\ plate} \times 100$$

2.9.4 Flow cytometry

A BD-Accuri C6 flow cytometer (Becton–Dickinson, Oxford, UK) was used to measure the green fluorescence produced by the cells and the red fluorescence from propidium iodide (PI). Flow cytometry is an extremely versatile technique used to analyse individual cells within a population. It was used as an integral part of this project to characterise heterogeneities in bacterial physiology and productivity at the single cell level. This would indicate the quantity and quality of GFP generated by each cell. Propidium iodide (PI; Sigma-Aldrich, Poole, UK) was used to detect dead cells; so, both live/dead and productive/non-productive subpopulations were identified. Unstained sample was analysed by diluting 100 µL

of cultures into 1 mL of filtered-sterilise (0.22 μm) PBS. Stained samples were analysed by incubating samples with 4 μg/mL PI for 5 min. A total of 25,000 events were collected for each sample at a rate of 1000 - 2500 events per second with the threshold set at 8,000 using a forward scatter height (FSC-H) threshold to reduce particulate noise. Detection filter FL1-A (533/30) was used for GFP and FL3-A (670 LP) for PI. Data collection were analysed using CFlow (Becton–Dickinson, Oxford, UK).

The flow cytometry data were analysed in accordance with the workflow presented in the following figure (**Figure 2.4**), similar to previous studies (Sevastsyanovich et al., 2009; Wyre and Overton, 2014; Hothersall et al., 2021; Hothersall et al., 2022a; Hothersall et al., 2022b). Data were collected for each sample using an FSC-H (forward scatter height signal) threshold that significantly filtered out particulate noise. A forward scatter (FSC-A) versus side scatter (SSC-A) plot was constructed to visualise the data and a gate (P1) was utilised to select cells which were used for further investigation (**Figure 2.4a**).

Samples unstained with PI were analysed to determine mean green fluorescence values and the green fluorescence histograms. While live/dead analysis was performed by staining samples with PI (which is taken up by only dead cells). Compensation was not used; rather, the line separating PI+ from PI- cells was drawn diagonally such that highly green fluorescent cells (where there is some spill over into the red fluorescence channel) were still in the PI- sector of the plot. **Figure 2.4b** shows live cells without staining and **Figure 2.4c** with PI staining, demonstrating the gating principle.

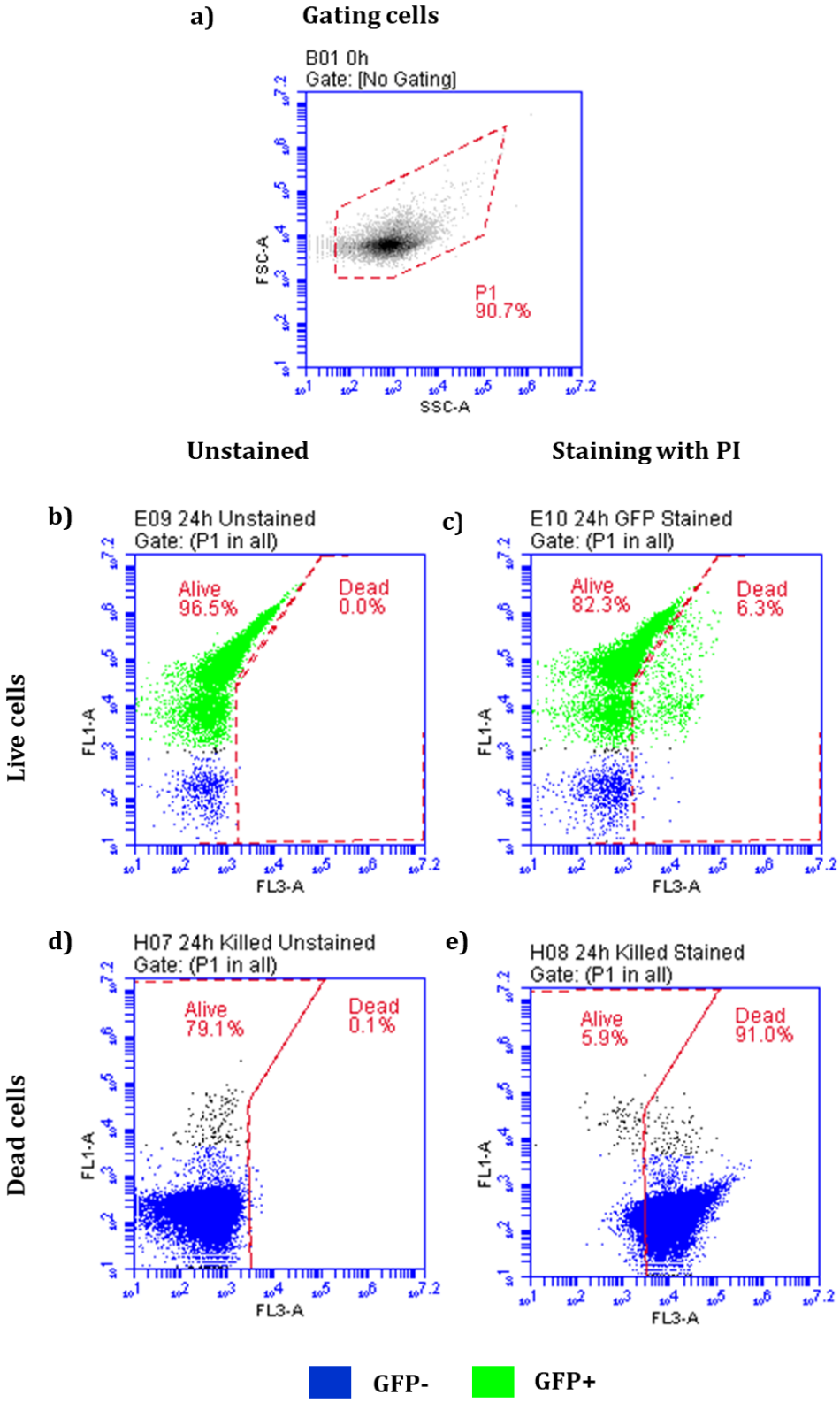

a)
Gating cells
B01 0h
Gate: [No Gating]
P1
90.7%
FSC-A
SSC-A
Unstained
Staining with PI
b)
E09 24h Unstained
Gate: (P1 in all)
Alive
96.5%
Dead
0.0%
c]
E10 24h GFP Stained
Gate: (P1 in all)
Alive
82.3%
Dead
6.3%
Live cells
d)
H07 24h Killed Unstained
Gate: (P1 in all)
Alive
79.1%
Dead
0.1%
e)
H08 24h Killed Stained
Gate: (P1 in all)
Alive
5.9%
Dead
91.0%
Dead cells
FL1-A
FL3-A
GFP-
GFP+

Figure 2.4: Demonstration of flow cytometry workflow.

Flow cytometry data was analysed according to the following workflow. Data were collected for each sample using an FSC-H (forward scatter height signal) threshold that significantly filtered out particulate noise. A forward scatter (FSC-A) versus side scatter (SSC-A) plot was constructed to visualise the data and a gate (P1) was utilised to select cells (a) which were used for further investigation. Live/dead analysis was performed by staining samples with PI (which is taken up by only dead cells). Example data is shown: live cells (b) without staining and (c) with PI staining, demonstrate that PI staining has a negligible effect on the red fluorescence (FL3-A), while figures (d) and (e) show that PI staining has an enormous effect on the red fluorescence of heat-killed cells (PI+).

For the purpose of validating the PI staining, heat-killed cells of an overnight culture were used. Dead cells for PI control stain were prepared by heating the culture for 10 min at 99 °C. Unstained heat-killed cells are depicted in **Figure 2.4d**, whereas **Figure 2.4e** shows heat-killed cells stained with PI, resulting in an increase in red fluorescence. It is also important to note that heat-killing causes GFP fluorescence to decrease because of denaturation.

2.9.5 Fluorescent microscopy

Optical and fluorescent microscopy (Zeiss Axioplan) were used to observe the structure of fluorescence cells with different physical properties. The fluorescence microscope is an optical microscope that excite fluorescence at the specific parts of the cells. The fluorescence cells were placed on a microscope slide and gently covered with a cover slip. The images were acquired under objective lens 100x magnification (oil immersion) with a digital camera system Axiocam ICm1 using a 1.4-megapixel monochrome CCD camera via AxioVision Software (Zeiss). The fluorescence cells were observed at room temperature.

2.10 Protein analysis

2.10.1 Cells fractionation (Freeze-thaw method)

To begin with, a volume of bacterial culture equivalent to a 1 mL at an OD_{650} of 1 was harvested by centrifugation at 12,000 *g* in a microcentrifuge for 10 min and the pellet was resuspended in 250 µL of PBS buffer (pH 8; adjusted with 0.1 M NaOH). Lysozyme (Sigma-Aldrich) were added to a final concentration of 200 µg/mL, followed by 1 µL of benzonase (Sigma-Aldrich). Samples mixture were placed in an ice bath with a gentle shaking platform for 30 - 45 min before freezing in dry ice/ethanol bath and thaw at 37 °C. Freeze-thaw cycle was repeated for at least 3 times. As for precaution, dry ice/ethanol bath were kept in a Dewar jar and strictly handled in a ducted fume cupboard. Forceps were used to hold the samples at all times. Then, the samples were subjected to centrifugation for 30 min at 9000 *g*. The soluble fraction (supernatant) were placed into a different Eppendorf tube. The leftover pellet (insoluble fraction) were resuspended in 250 µL of PBS buffer (pH 8).

2.10.2 SDS-PAGE

SDS-PAGE is the procedure used to separate extracted protein fractions by gel, using two different apparatus; an Omni-PAGE mini (Cleaver Scientific, UK) with gel-pouring preparation and a Mini Gel Tank (ThermoFisher Scientific, UK) with

Tris-Glycine SDS precast 12% (w/v) polyacrylamide gels (Novex WedgeWell, ThermoFisher Scientific, UK).

2.10.2.1 Omni-PAGE mini with gel-pouring preparation

The following gels were prepared in a 50 mL falcon tube and mixed well. Gels were prepared using a 15 % (w/v) resolving gel and a 6 % (w/v) stacking gel in a vertical electrophoresis tank in the fume hood with butyl gloves as the precaution. Ethanol-polished gel plates (Cleaver Scientific, UK; glass notched 20 x 10 cm, 2 mm thick) were separated by 0.5 mm spacers. Buffer preparation details for Omni-PAGE mini were explained in **Section 2.2.4**.

The resolving gel contained 10 mL stock resolving gel (or stock running gel) buffer, 10 mL acrylamide solution and 100 µL 20 % (w/v) SDS. This was polymerized by the addition of 200 µL APS (80 mg/mL) and 10 µL N, N, N', N'-tetramethylethylenediamine (TEMED), mixed well and poured between the assembled glass plates immediately. An overlay of 0.1 % SDS was applied very gently once poured the separating gel to exclude air. The resolving gel was allowed to polymerise for approximately 1 hour, after which the 0.1 % SDS overlay was poured off and replaced with the stacking gel.

The stacking gel consisted of 1.5 mL stock stacking gel buffer, 3 mL acrylamide solution, 10.5 mL deionised water, 75 µL 20 % SDS, and was polymerised with 150 µL APS and 7.5 µL TEMED. A comb was inserted into the stacking gel to create wells for ladder and samples loading.

For sample preparation, 87 µL of β-mercaptoethanol was mixed with 1 mL aliquot of sample buffer, thus the samples were resuspended with 60 µL of the mixture.

Before loading, the samples were boiled in heat block at 99 °C for 10 min, with Eppendorf caps on lids. Both ladder (5 µL) and processed samples (10 µL) were loaded into the gel wells and run at 110 V for approximately 2 h at room temperature.

2.10.2.2 Mini Gel Tank with Tris-Glycine SDS precast polyacrylamide gels

While for Mini Gel Tank with Tris-Glycine SDS precast 12 % (w/v) polyacrylamide gels (Novex™ WedgeWell™), samples were prepared by mixing the fractionated protein sample with 60 µL of mixture solution containing 40 µL Tris-Glycine SDS sample buffer (Life Technologies), 8 µL NuPAGE reducing agent and 12 µL ddH_2O. The samples were heated in heat block at 85 °C for 2 min (500 rpm) before loading. Both ladder (5 µL) and samples (10 µL) were loaded on the gel and SDS-PAGE gels run at 200 V for approximately 40 min at room temperature.

Calibration of molecular weight for all gels was done using 5 µL of Blue Protein Standard (BPS) ladder (New England BioLabs, USA) of known concentration 0.2 mg/mL, with size range from 11 to 190 kDa. To visualise protein bands, SDS-PAGE gels were first stained with SimplyBlue™ Safe Stain solution from Invitrogen (Life Technologies), then destained with water, and imaged on a flatbed scanner CanoScan 9000F (Canon, UK) at 4800 dpi resolution. The images were analysed densitometrically (**Figure 2.5**) using ImageJ software (Schneider et al. 2012).

Firstly, images were subjected to background subtraction (default settings) followed by lane description. Histogram of band colour intensity were produced from defined lanes. GFP/TNFα-GFP protein peaks were defined and the areas

underneath were calculated. To estimate the concentration of total GFP/TNFα-GFP proteins, the band intensity of the peak was calculated as the percentage of the intensity of the total lane. GFP/TNFα-GFP peak areas for soluble and insoluble samples were compared to estimate the percentage ratio of both proteins.

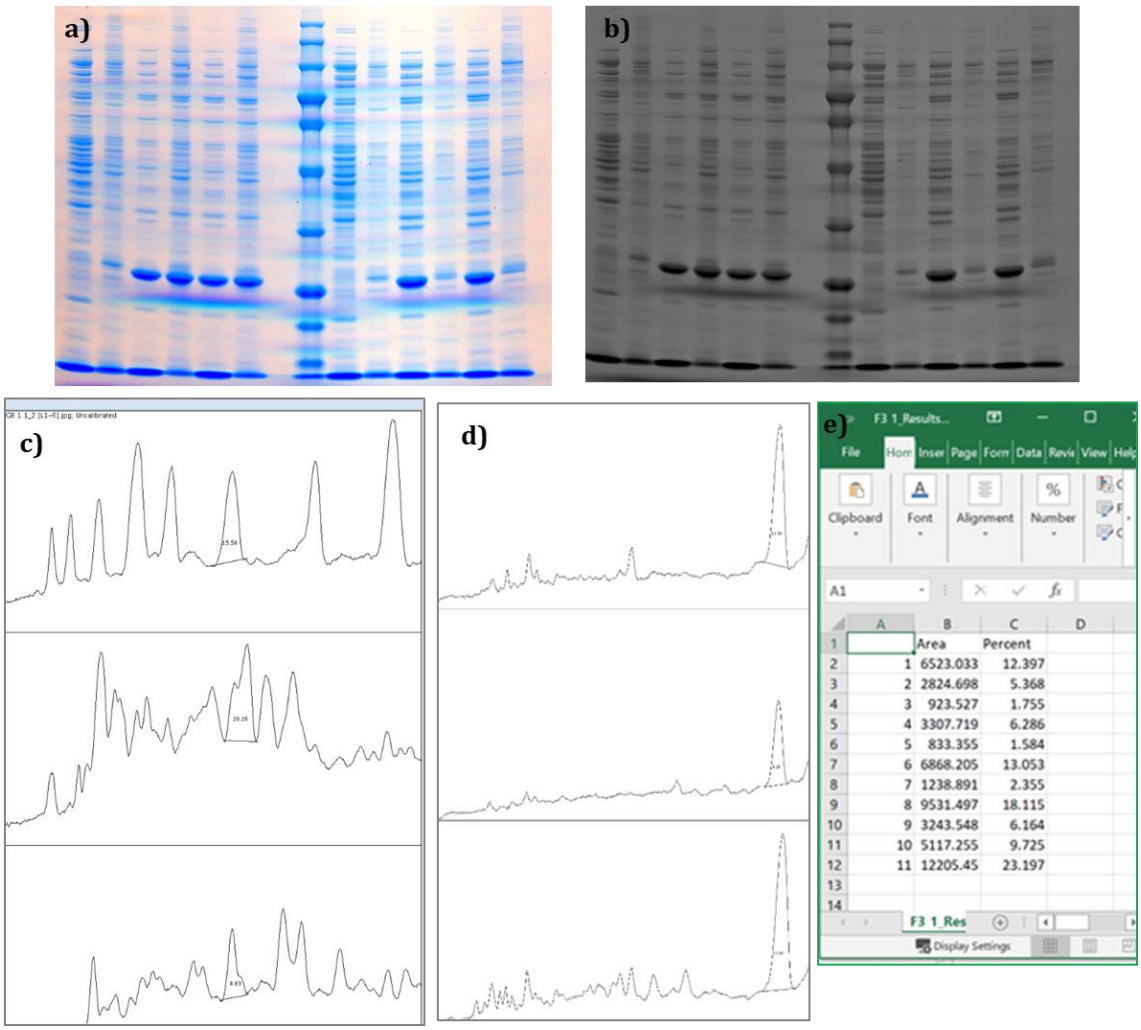


Figure 2.5: Densitometric analysis steps of SDS-PAGE gels

Analysis of an SDS-PAGE gel sample by ImageJ. a) Gel was scanned at 4800 dpi and opened in ImageJ. b) Background was removed by default settings and converted to be a gray-scale image. c) Lanes were defined and lane histograms were plotted. d) GFP/TNFα-GFP peaks were defined and selected e) Area under the peak automatically calculated and used for analysis (for total protein concentration and solubility analysis).

CHAPTER 3

BATCH CULTURE OF BL21(T7)-pORT(T7)-GFP/TNFα-GFP

The first part of the project's objective was to investigate the strength of the promoter and culture heterogeneity in terms of physiology and protein production during the generation of Tumour necrosis factor alpha (TNFα). This chapter investigates the heterogeneity of the cultures and the potency of pORT(T7) plasmid promoter to produce GFP and TNFα-GFP proteins. The production of GFP (using GFP only) and TNFα proteins by TNFα-GFP fusion using the same plasmid pORT(T7) will be investigated by few analyses including flow cytometry for comparison at the single cell level.

3.1 Introduction

The investigation was initiated by conducting several series of experiments to determine the best optimised condition for recombinant protein production (RPP), which in this study involved production of two recombinant proteins, green fluorescent protein (GFP) and human Tumour necrosis factor alpha fused to GFP (TNFα-GFP).

An arabinose-induced T7 expression system was used in these initial studies, supplied by Cobra Biologics Ltd. (Keele, UK). It is comparable to the widely-used commercial pET/DE3 system (Studier, 1991), with the exception that arabinose is added to the BL21(T7) strain to induce expression of the chromosomally-encoded T7 RNA polymerase (Castiñeiras and Overton, 2018).

For this study, two different plasmids used were p2T7-TNFαGFP; henceforth referred to as pORT(T7)-TNFα-GFP, and p2T7-GFP; referred to as pORT(T7)-GFP.

The gene for TNFα is translated in pORT(T7)-TNFα-GFP, with its C-terminus is fused to GFPmut2 (Cormack et al., 1996), while in pORT(T7)-GFP, the gene for TNFα is deleted, leading to the synthesis of GFPmut2 alone. Thus, this allow the comparison of both GFP and TNFα-GFP synthesis.

GFP is relatively easy for *E. coli* to produce as it readily folds correctly into a functional form (Sørensen and Mortensen, 2005). GFP is also easy to measure by fluorescence measurement. Various excellent reviews can be obtained on the function, structure and evaluation of GFP molecules in terms of folding efficiency, maturation speed and fluorescent spectrum (Shaner et al., 2007; Remington, 2006). In this work, we compared production of GFP to that of TNFα-GFP under the same growth conditions. TNFα is not as easy to produce in *E. coli* as GFP as it is prone to misfolding (Castiñeiras et al., 2018b). Previous work showed that an RP-GFP fusion could be used to measure both the quantity and folding state of the RP (Sevastsyanovich et al., 2009b; Wyre and Overton, 2014a). In the present study, TNFα-GFP production was compared to GFP production to determine promoter strength, heterogeneity and the effect of TNFα on expression level.

The production of GFP versus TNFα-GFP was studied in several sets of experiments; this protocol has been designed to grow the cells in shake flasks in batch culture. The objective was to compare the production of TNFα-GFP fusion and GFP alone, improve the solubility of TNFα-GFP, decrease formation of insoluble inclusion bodies, and therefore maximise the overall yield of functional TNFα. The culture conditions have been established for the experimental design of TNFα (**Table 3.1**), with several different measurable parameters for comparison

purposes, to evaluate the optimum yield of soluble and active TNFα-GFP. The parameters investigated in this study were the carbon source and its concentration, the growth temperature, and the concentration of inducer (arabinose). The optimum conditions obtained in these experiments will be used for the production of TNFα-GFP in fed-batch fermentation at 5 L scale, and this is elaborated upon in Chapter 5.

Table 3.1: TNFα experimental design for cell growth in shake flasks in batch culture

Parameter	*Culture Conditions*
Medium used	Terrific Broth supplemented with Kanamycin (50 µg/mL)
Carbon source	Glycerol or glucose (0.4% or 0.8%)
Temperature	30°C or 37°C
Inducer concentration	Arabinose (None, 0.015% or 0.2%)
Induction point	Mid-logarithmic phase (OD_{650} ≈ 0.5)

Based on **Table 3.1**, the medium used for cell growth in shake flasks is Terrific Broth (TB) supplemented with Kanamycin (50 μg/mL) to maintain plasmid retention. Carbon sources used for comparison were glycerol and glucose. These two carbon sources were selected based on previous work on RPP in *E. coli* (Wyre and Overton, 2014b). The use of glycerol as a carbon source does not repress the expression system. Glucose, in contrast to glycerol, negatively regulates the arabinose-inducible promoter, thus inhibits T7 RNA polymerase synthesis. Moreover, acid production due to overflow metabolism is often not an issue when glycerol is used as a feed, unlike when utilising glucose as the carbon source.

The temperature used was either 30°C or 37°C. The temperature of 30°C was chosen as an approach to increase the folding of recombinant proteins by minimising the stress, achieved by reducing the production rate of recombinant proteins. Wyre and Overton (2014b) believed that lower temperature and low inducer concentrations can slow down the transcription rate of recombinant protein-encoding genes, thus slowing down translation and folding rates. Reduction of translational rate allows cells to have sufficient time to properly fold and produce functional proteins. In recombinant protein processing, *E. coli* is grown below an optimum temperature of 37°C. However, the high-level expression of heterologous proteins in *E. coli* often fails to achieve its native conformation, thus stimulates the formation of IBs, unless lower temperature is used to grow the cells (Kim et al., 1996; Wong et al., 1990).

Arabinose was used as an inducer in this study, at concentrations of 0.015% or 0.2%. Arabinose induces expression of the T7 RNA polymerase from the

chromosome of *E. coli* BL21(T7) strain used; the T7 RNA polymerase activates the T7 promoter on the plasmid. In all experiments, expression of GFP was compared to that of the TNFα-GFP fusion.

Cultures were sampled at regular intervals. Growth was determined by measuring OD_{650} and colony-forming unit (CFU). Plasmid retention was measured using replica plating. Flow cytometry (FCM) was used to determine levels of GFP fluorescence per cell and viability using propidium iodide (PI). Cell samples were also analysed by SDS-PAGE and subcellular fractionation to determine the quantity and solubility of recombinant protein generated.

3.2 Batch culture with 0.4% glycerol as carbon source at 30°C

The initial batch culture designed were expected to allow the generation of soluble recombinant protein. This first experiment was the growth of cells in 100 mL of TB with 0.4% glycerol supplemented with 50 µg/mL Kanamycin. A temperature of 30°C was used to grow the cells. Recombinant protein production was induced when the OD_{650} reached around 0.5 (mid-logarithmic phase), with either 0.015% arabinose or 0.2% arabinose. An uninduced control with no arabinose was also included. Samples were taken at regular intervals and the OD_{650} measured. Some samples were also analysed for CFU on nutrient agar (NA) plates to indicate the culturability of the cultures, and were replica-plated onto nutrient agar plates supplemented with 50 µg/mL Kanamycin to evaluate the retention of plasmids. All experiments in shake flask were executed in duplicate.

In general, measurements of OD_{650} as showed in **Figure 3.1a** revealed a slightly difference in growth between cultures expressing GFP and TNFα-GFP. Cultures of pORT(T7)-GFP growing with uninduced of arabinose showed the higher cell densities than those induced, reaching the final OD_{650} of 16.5.

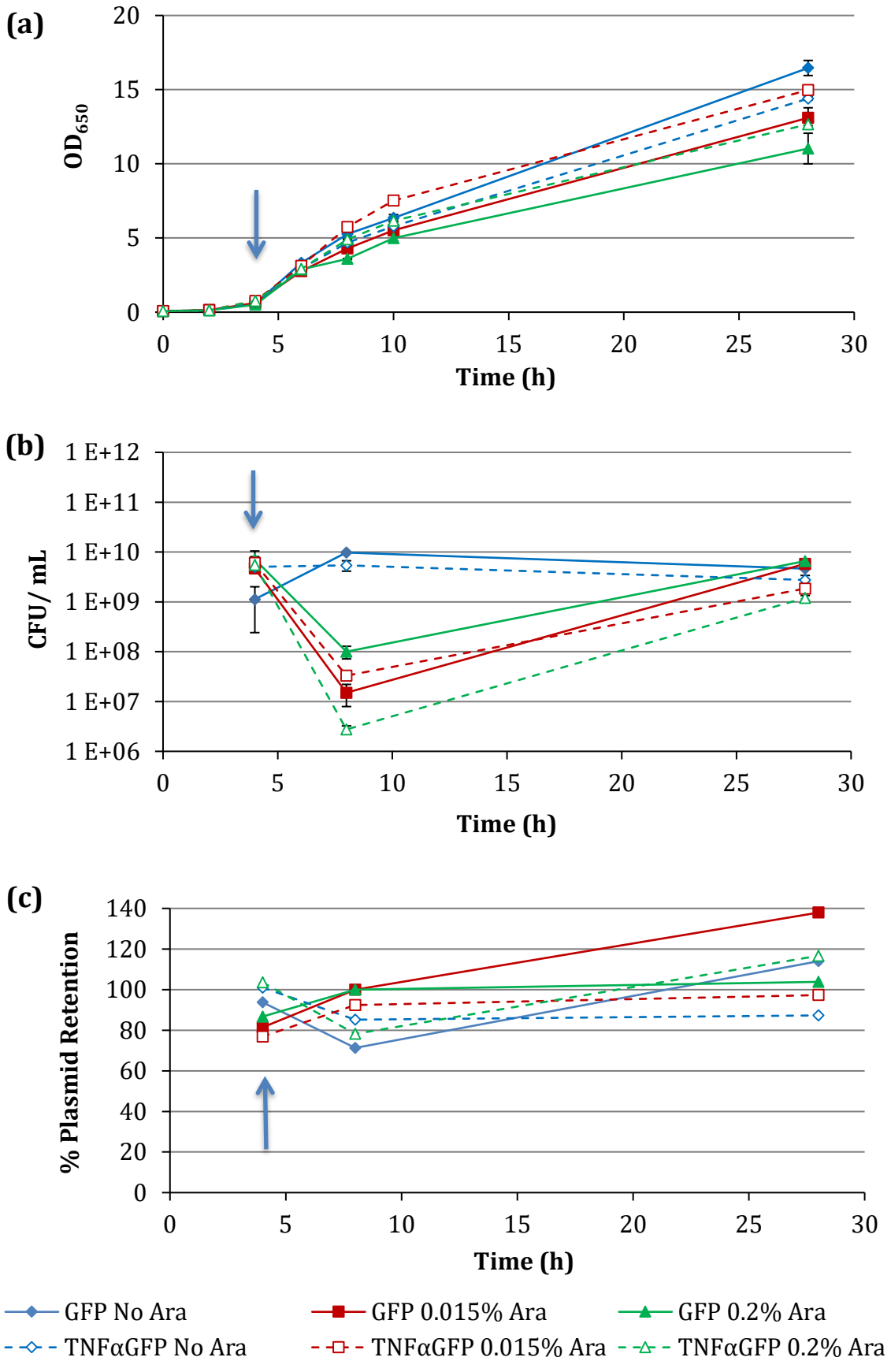

(a)
OD650
Time (h)
(b)
CFU/ mL
1 E+12
1 E+11
1 E+10
1 E+09
1 E+08
1 E+07
1 E+06
Time (h)
(c)
% Plasmid Retention
Time (h)
GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara
TNFαGFP No Ara
TNFαGFP 0.015% Ara
TNFαGFP 0.2% Ara

Figure 3.1: Correlation of *E. coli* BL21(T7) growth for the study of optimisation for GFP and TNFα-GFP production, grown in TB medium at 30°C with 0.4% glycerol

E. coli BL21(T7) cultures carrying the plasmid pORT(T7)-GFP or pORT(T7)-TNFα-GFP were induced with three different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at OD_{650} ≈ 0.5 (approximately after 4 h of growth; indicated by blue arrow), evaluated by (a) OD_{650} measurement, (b) CFU/mL counting, and (c) plasmid retention. (a) The OD_{650} was measured at intervals. (b) Serial dilution of culture samples at 0 h, 4 h and 24 h post-induction was plated onto NA agar for indication of cell viability and, (c) were replica-plated onto NA agar supplemented with Kanamycin (50 μg/mL) to reveal plasmid retention. All data shown are mean values from replica flasks, error bars are ±1 standard deviation.

But for pORT(T7)-TNFα-GFP, the final OD of the uninduced culture is lower (14.4) than that 0.015% arabinose culture (15.0). In contrast, both cultures of pORT(T7)-GFP and pORT(T7)-TNFα-GFP induced with the higher concentration of 0.2% arabinose reached less final OD_{650} of 11.0 and 12.7, respectively, after 28 hours of growth. Another induced cultures with a lower concentration of 0.015% arabinose generally reached an intermediate level of OD_{650} reading. It was shown that growth was suppressed by inducing cells with high concentrations of arabinose.

The slower growth of cultures with arabinose induction might be explained as the result of cultures experiencing stress when forced to express recombinant proteins (both GFP and TNFα-GFP), where the induction of T7 RNA polymerase was controlled by arabinose (Overton, 2014). Both an increased metabolic burden and incorrect folding of recombinant proteins might contribute to this kind of stress. However, this negative effect was not caused by the arabinose itself, as there have been no reports so far of its toxicity on cell growth (Narayanan, 2009).

The observation of the colony forming unit (CFU) showed a significant reduction in all induced cultures, both for pORT(T7)-GFP and pORT(T7)-TNFα-GFP. The culturability dropped almost immediately after induction (**Figure 3.1b**), but eventually increased over time. The CFU recovered 24 hours after induction, reaching 10^9 to 10^{10} CFU/mL. The non-induced cultures showed no reduction of cell culturability throughout the growth.

Plasmid retention was the lowest for non-induced cultures of pORT(T7)-GFP at 4 h post-induction, which was 71.3% (**Figure 3.1c**), led to some ambiguity in this situation. Non-induced cultures were known to be less pressured to produce RP, thus they were the least likely to lose the plasmid. All induced cultures of pORT(T7)-TNFα-GFP also showed plasmid loss of 10-20% after 4 h post-induction. If they lost the plasmid inside the cells, they were not able to survive and grow on the plate with Kanamycin, thus will die (not culturable). The plasmid-free cells overgrowth could describe the unexpectedly high cell density observed in OD measurement for pORT(T7)-GFP non-induced cultures and most of pORT(T7)-TNFα-GFP induced cultures, as the result of low plasmid retention. Cultures of pORT(T7)-TNFα-GFP induced with 0.2% arabinose had the least OD growth and plasmid retention, as well as the lowest CFU/mL, compared to other pORT(T7)-TNFα-GFP cultures.

3.2.1 Flow cytometry measurement at single cell level

The green fluorescence of each cell, which provides an indication of the amount of functional GFP, was measured using flow cytometry (FCM), represented by the mean value of FL1-A fluorescence unit. GFP is a very good reporter in interpreting the folding status for proteins that are fused to it, so the FCM can be significantly used to determine the quantity and quality of a recombinant protein (Wyre and Overton, 2014a). FCM was also used to measure the average of GFP production per cell at time points post-induction.

From **Figure 3.2**, mean green fluorescence showed significant differences between induced and non-induced cultures of both pORT(T7)-GFP and pORT(T7)-TNFα-GFP, observed in data plotted at 0 h, 2 h, 4 h, 6 h and 24 h post-induction. The mean green fluorescence of cells that produced GFP alone was higher than their corresponding cells that produced TNFα-GFP at all timepoints, despite equivalent arabinose concentrations being used. It was observed that induction with arabinose greatly increased the green fluorescence. The green fluorescence measured was the highest observed at 6 h post-induction for pORT(T7)-GFP induced cultures with 357,954 for 0.2% arabinose-induced cultures, and 318,383 for 0.015% arabinose-induced cultures (**Figure 3.2a**). Of these figures, pORT(T7)-TNFα-GFP induced cultures showed a lower green fluorescence producing that was almost half lower than those of pORT(T7)-GFP. pORT(T7)-TNFα-GFP cultures with 0.015% arabinose-induced recorded the green fluorescence of 208,047, while 204,666 was measured for pORT(T7)-TNFα-GFP with induced cultures of 0.2% arabinose. The dynamics of recombinant protein production can be easily observed and rapidly studied using FCM. The data in **Figure 3.2a** also showed that cells producing TNFα-GFP were less fluorescent than GFP-generating cells alone. This illustrates the fact that TNFα-GFP production and its folding is less straightforward than that of GFP.

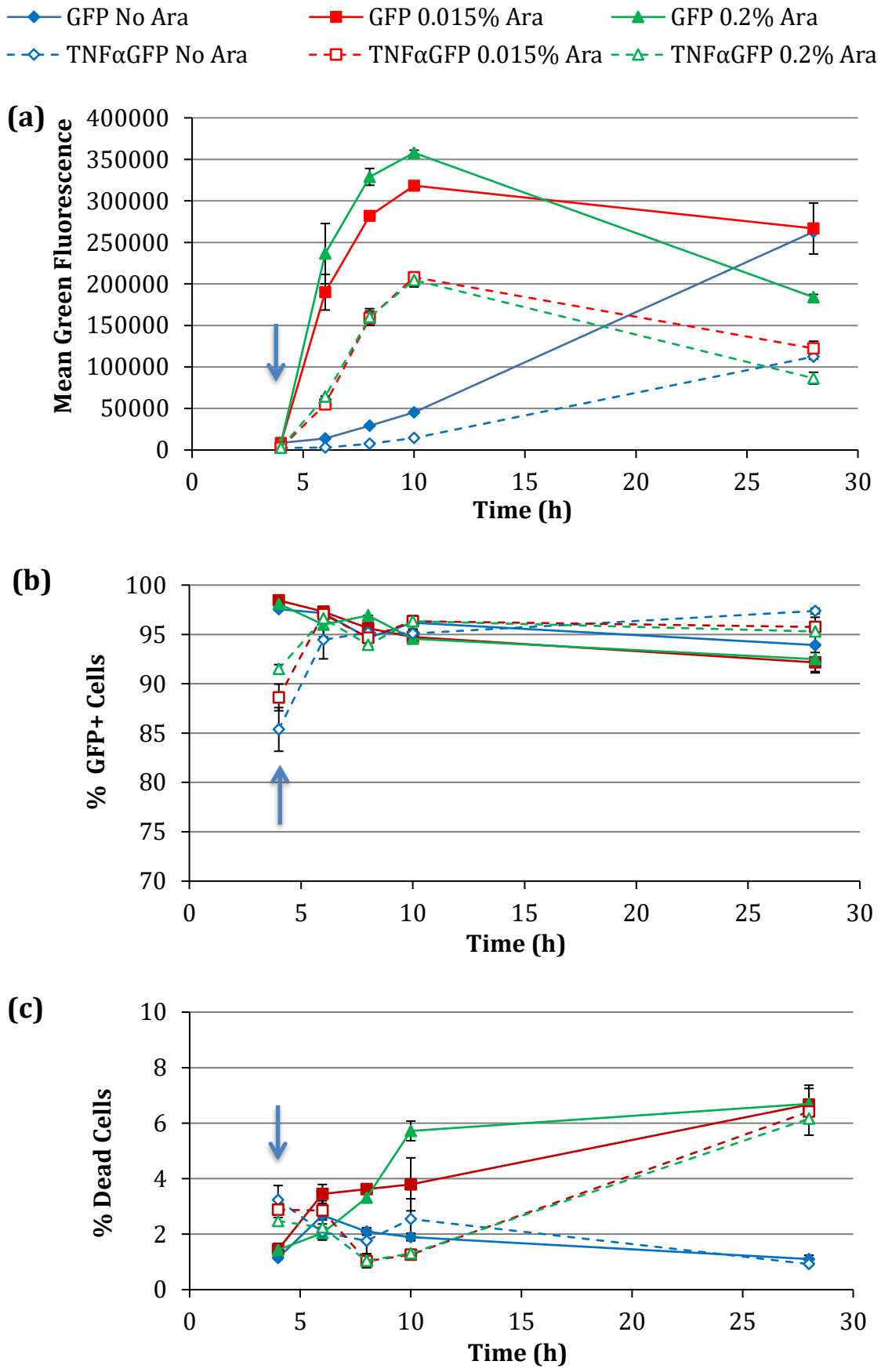

GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara
TNFαGFP No Ara
TNFαGFP 0.015% Ara
TNFαGFP 0.2% Ara
(a)
Mean Green Fluorescence
400000
350000
300000
250000
200000
150000
100000
50000
0
0
5
10
15
20
25
30
Time (h)
(b)
% GFP+ Cells
100
95
90
85
80
75
70
0
5
10
15
20
25
30
Time (h)
(c)
% Dead Cells
10
8
6
4
2
0
0
5
10
15
20
25
30
Time (h)

Figure 3.2: Mean green fluorescence measured by flow cytometry for GFP and TNFα-GFP in batch culture with 0.4% glycerol at 30°C

Recombinant human TNFα fused to GFP (or also known as rhTNFα-GFP) was expressed from an arabinose-induced plasmid system with different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow), and compared to the production of GFP from the same plasmid pORT(T7). (a) FL1-A (green fluorescence) plots of cells taken from culture samples at interval time points post-induction. (b) Percentage of cells that producing GFP (GFP+ cells) versus time, and (c) Percentage of dead cells at interval 0 h to 24 h post-induction. Data are mean values of duplicate samples, error bars are ±1 standard deviation.

Increasing the concentration of arabinose with a higher concentration of 0.2% did not further increase fluorescence from both GFP and TNFα-GFP cultures significantly, indicating that the concentration of 0.015% arabinose is adequate to completely activate the promoter. While all induced cultures reached their maximum green fluorescence after 10 hours of growth (in which 6 h post-induction), non-induced cultures nevertheless also exhibited a very high fluorescence when measured at the end of 28 hours growth. This is due to the promoter system's leakiness.

Preliminary observations of FL1-A (fluorescence units) histogram data showed a high percentage of more than 90% cell populations producing GFP (GFP+ cells) in all induced and non-induced cultures of both pORT(T7)-GFP and pORT(T7)-TNFα-GFP (**Figure 3.2b**) starting from the point of induction at mid-log phase up to 24 h post-induction, even though there were a slight decrease observed at the end of the incubation. Mid-log phase ($OD_{650} \approx 0.5$) is the optimized induction point that is frequently recorded (Kasli et al., 2019; Su et al., 2019). To monitor precise events that occurred at the level of single cells, a more thorough analysis is necessary. Therefore, propidium iodide (PI) staining was used to determine the proportion of dead cells in both BL21(T7)-pORT(T7)-GFP and BL21(T7)-pORT(T7)-TNFα-GFP cultures. The results are shown in **Figure 3.2c**.

The percentage of dead cells for both induced GFP cultures were higher (between 4 to 6%) compared to other cultures as early as 6 h post-induction. This might happen because the cultures had lack of energy and nutrients in the medium, since they were seen actively producing GFP especially at 6 h post-induction (**Figure**

3.2a), thus eventually they died after producing a huge amount of GFP. The percentage of dead cells were increased above 6% for these two cultures after 24 h post-induction. While for induced cultures of TNFα-GFP, the percentage of dead cells observed was lower up to 6 h post-induction, as they also produced less green fluorescence compared to GFP cultures, reflecting less metabolic stress on the cells. However, the dead percentage of these TNFα-GFP cultures were also increased to more than 6% after 24 h post-induction. PI staining also demonstrated a larger number of dead cells in induced cultures compared to uninduced. Dead cells were not significantly higher in TNFα-GFP expressing cultures compared to GFP expressing cultures. The majority of dead cells in all cultures were GFP+.

The flow cytometry measurement of FL1-A histograms

The measurement of green fluorescence obtained by flow cytometry can be observed in the histograms of mean FL1-A values as demonstrated in **Figure 3.3 and Figure 3.4**. Histogram plotted were based on samples taken at 0 h, 2 h, 4 h, 6 h and 24 h post-induction. The FL1-A histogram obtained for *E. coli* BL21(T7)-pORT(T7)-GFP and BL21(T7)-pORT(T7)-TNFα-GFP cultures was compared to the FL1-A histogram of *E. coli* BL21(T7) only without any recombinant plasmid attached to it. This untransformed *E. coli* BL21(T7) acted as a control figure for comparative analysis of FL1-A histograms since they were plasmid-free cells and therefore, were unable to produce green fluorescence (GFP- cells). Thus, this FL1-A plasmid-free cells histogram was compared to the FL1-A histogram representing GFP-producing cells and TNFα-GFP-producing cells.

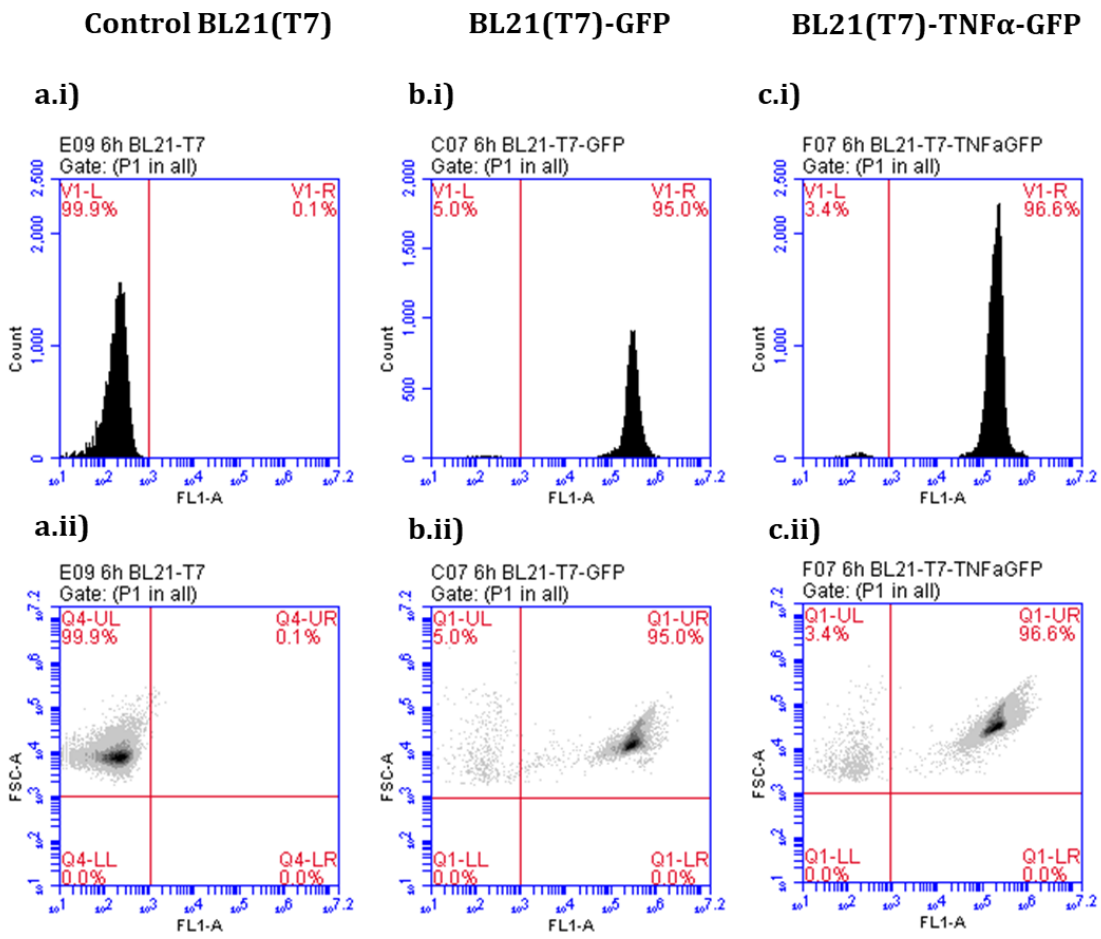


Figure 3.3: Green fluorescence histograms of *E. coli* BL21(T7) producing GFP or TNFα-GFP at 6 h post-induction

Histogram of mean FL1-A (fluorescence units) showing typical gating for GFP+ cells (V1-R) and GFP- cells (V1-L). (a.i) The histogram of FL1-A for untransformed *E. coli* BL21(T7) cells, which did not produce green fluorescence, where the histogram peak is below 10^3 fluorescence units. (b.i) *E. coli* BL21(T7)-pORT(T7)-GFP cultures and (c.i) *E. coli* BL21(T7)-pORT (T7)-TNFα-GFP cultures showed green fluorescence peaks inside the V1-R gate (GFP+ cells), specifically at 10^5 to 10^6 fluorescence units. (a.ii) The same data as (a.i) were plotted on an FSC-A versus FL1-A (forward scatter versus fluorescence unit) plot, revealing no green fluorescent populations. (b.ii) and (c.ii) corresponds to the same data as (b.i) and (c.i) respectively, showing the existence of green fluorescent populations. Duplicated data for all cultures showed almost similar figures and gate P1 included in all analysis to eliminate noise.

It was observed that histogram peaks of the untransformed cultures were located around 2 x 10^2 fluorescence units (**Figure 3.3a**), reflecting the autofluorescence of *E. coli*. Analysis of the heterogeneity of green fluorescence on a single cell level revealed that non-induced cultures of GFP and TNFα-GFP had a small non-productive population around 2×10^2 fluorescence units, likely plasmid-free cells as above (untransformed cells; **Figure 3.4**). Therefore, we can certainly state that these populations with very low fluorescence are plasmid-free cells. Aside from that, there was a considerably larger single population of GFP+ cells in non-induced cultures at all timepoints, including 4 hours (pre-induction), which once again reflects the leakiness of the promoter.

Histogram peaks representing green fluorescence-producing GFP or TNF-GFP cells were located specifically at 10^4 to 10^6 of fluorescence units (**Figure 3.3b, 3.3c**). These induced cultures displayed a similar pattern for the first 10 hours of growth. At 28 hours, however, cultures expressing both GFP and TNFα-GFP comprised multiple populations with different green fluorescence, likely indicating subpopulations with different RP accumulation and folding state (Wyre and Overton, 2014a).

Since only 7-10% of cells were GFP-, this indicates that plasmid selection was efficient enough to prevent the emergence of a sizable plasmid-free population. The addition of Kanamycin to the culture medium also results in tighter plasmid selection; in contrast to the work with plasmids conferring ampicillin resistance which the selection was much looser and led to a greater loss of plasmids (Sevastsyanovich et al., 2009a; Wyre and Overton, 2014b).

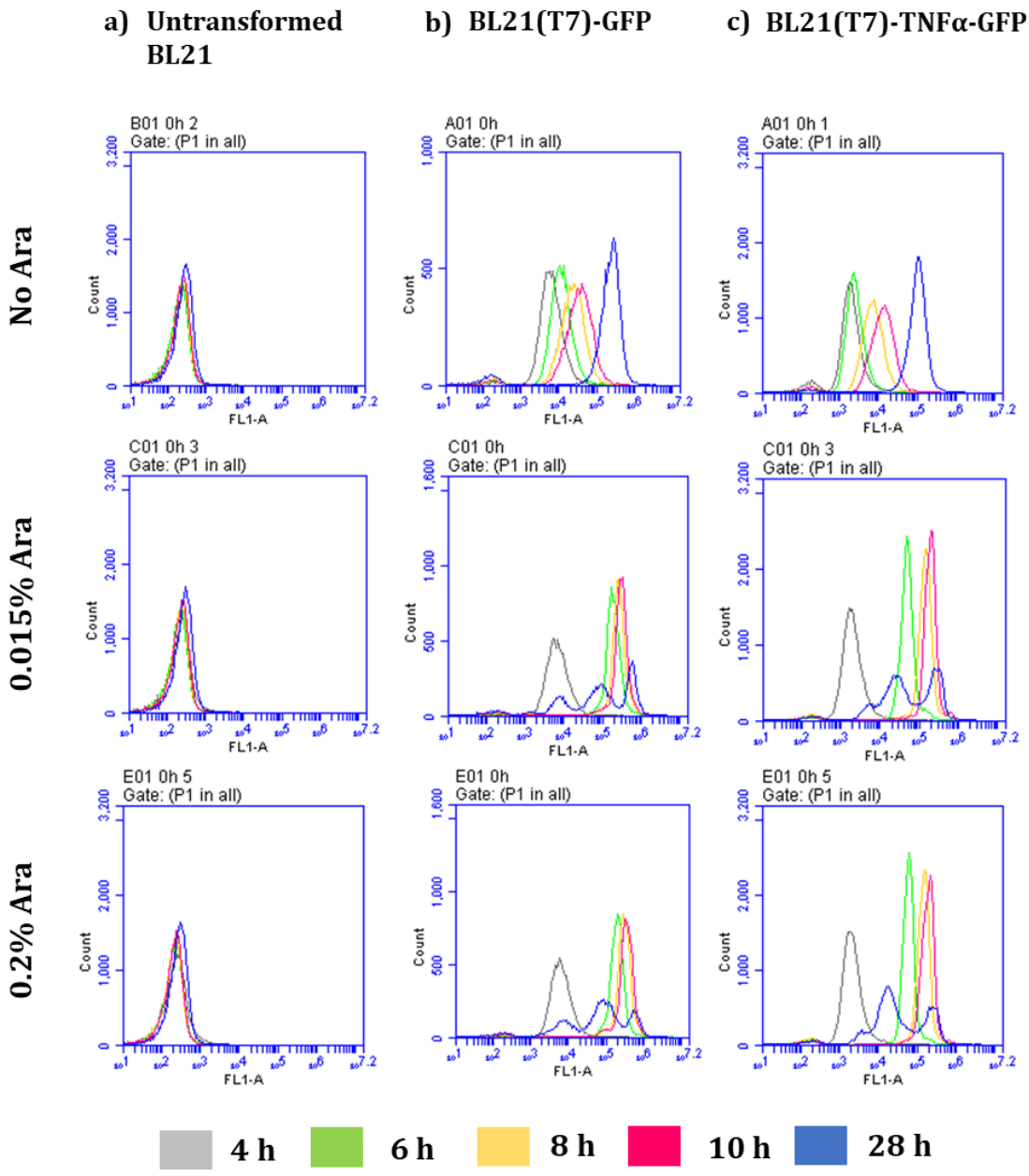


Figure 3.4: Single cell analysis of both T7 expression system cultures (GFP and TNFα-GFP) and untransformed *E. coli* with 0.4% glycerol at 30°C

(a) Untransformed *E. coli* BL21-A cultures, (b) *E. coli* BL21-T7 cultures carrying p2T7-GFP (GFP) and (c) p2T7-TNFα-GFP (TNFα-GFP) were grown in Terrific Broth with 0.4% glycerol as a carbon source at 30°C. At an OD_{650} of 0.5 (approximately after 4 h growth), arabinose (0.015% or 0.2%) was added. At regular intervals, samples were taken and FCM was used to measure green fluorescence of individual cells. Timepoints relate to the time the cultures were grown; 4 h refers to pre-induction with arabinose. Duplicate cultures of each plot showed a similar pattern throughout the growth.

Previous studies have shown that cells producing higher green fluorescence contained correctly-folded GFP or RP-GFP fusions (Wyre and Overton, 2014b; Vera et al., 2007). While plasmid-free cells mean cells that were no longer able to maintain in the host cell the plasmid containing target gene (in this case GFP and TNFα-GFP genes), thus cannot produce the protein. Through the observation in **Figure 3.3**, we can clearly see for cultures at 6 h post-induction that there were two major populations that dominate, it was either the lower fluorescence of plasmid-free cells (GFP- or TNFαGFP-) or higher green fluorescence-producing cells (GFP+ or TNFαGFP+).

Inspection of green fluorescence histograms at 0 h, 6 h and 24 h post-induction revealed increasing heterogeneity over time (**Figure 3.5**). At 0 h and 6 h, a few plasmid-free cells were present but the productive cells were in a single population. There was a shoulder present on the left of the peaks in induced cultures especially at 6 h post-induction, suggesting incomplete induction. At 24 h, there were multiple populations in terms of green fluorescence in induced cultures. Uninduced cultures are homogenous in terms of green fluorescence throughout. In observations of **Figure 3.5**, it was obviously seen that there were multiple populations of varying green fluorescence intensity at 24 h post-induction, in which, as cells aged, some of them gradually transitioned from the higher fluorescence population to the intermediate. This intermediate fluorescence had been tracked to be located in between histogram peaks of lower fluorescence plasmid-free cells population and higher green fluorescence-producing cells, specifically at 10^3 - 10^5 fluorescence units.

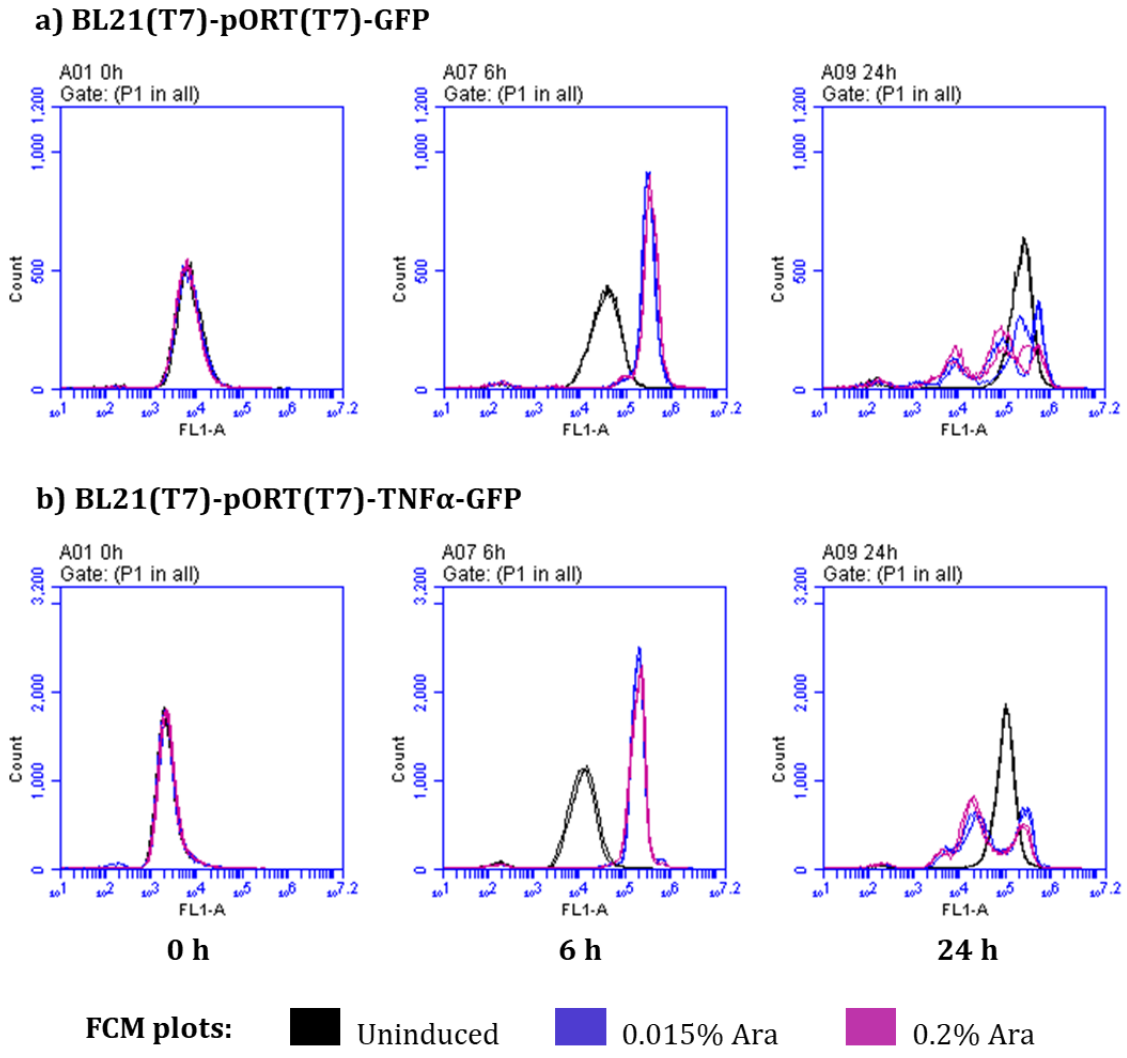


Figure 3.5: FCM green fluorescence histograms of individual time points

Inspection of green fluorescence histograms at 0 h, 6 h and 24 h post-induction revealed increasing heterogeneity over time. (a) pORT(T7)-GFP and (b) pORT(T7)-TNFα-GFP induced cultures comprised a single population with respect to green fluorescence at timepoint 6 h post-induction. At 24 h post-induction, there were multiple green fluorescence populations in cultures producing both GFP and TNFα-GFP. Data are mean values derived from duplicate cultures; error bars are ±1 standard deviation. Duplicate cultures of each plot showed a similar pattern throughout the growth.

Intermediate fluorescence was believed to be produced by cells population that produced poor quality of green fluorescence or incorrectly folding protein (Wyre and Overton, 2014a). There were so many reasons that contribute to the production of intermediate fluorescence. The increased metabolic burden brought on the cells as a result of being forced to produce RP following induction increased the physiological stress of the cells, leads to protein misfolding and arise in the amount of non-functional RP generated (Castiñeiras and Overton, 2018), thus lowering the production of green fluorescence. The phenomenon of protein misfolding has the potential to contribute to the formation of inclusion bodies (IBs), as well as to a reduction in the generation of functional enzymes (Bhatwa et al., 2021; Su et al., 2019).

In most of the cases, intermediate fluorescence is often associated with the production of IBs. IBs is frequently produced by *E. coli* in the form of densely particles of denatured protein molecules, where rapid recombinant protein expression in *E. coli* had resulted in its in vivo accumulation as insoluble IBs (Fahnert et al., 2004). Most cells such as *E. coli* tend to form IBs rather than recombinant proteins due to several reasons including the differences in physicochemical conditions and folding pathways between bacteria and eukaryotic cells (Wyre and Overton, 2014b). IBs are relatively pure source of recombinant proteins and can be treated to produce functional proteins through chemical treatment, but these refolding steps promise different success. Some recombinant proteins in bacteria cannot be produced directly in soluble form, and in most cases, their attempts resulted in the formation of IBs, thus reducing host cell viability.

A more-in depth study was conducted to obtain more detailed information and observations related to each of the multiple populations that exist in the heterogeneity. It is obvious from the FL1-A histogram data presented in **Figure 3.6**, the existence of cells population that produce intermediate fluorescence at 24 h post-induction in both GFP and TNFα-GFP cultures. This intermediate fluorescence was still giving some value to FL1-A and considerable amounts of green fluorescence was also measured, but the quality of the GF-producing might be lower. Wyre and Overton (2014b) in their research on recombinant protein production of CheY::GFP believed that the formation of intermediate fluorescence detected at the location of 10^3 to 10^5 fluorescence unit was the result of that recombinant protein aggregation into inclusion bodies.

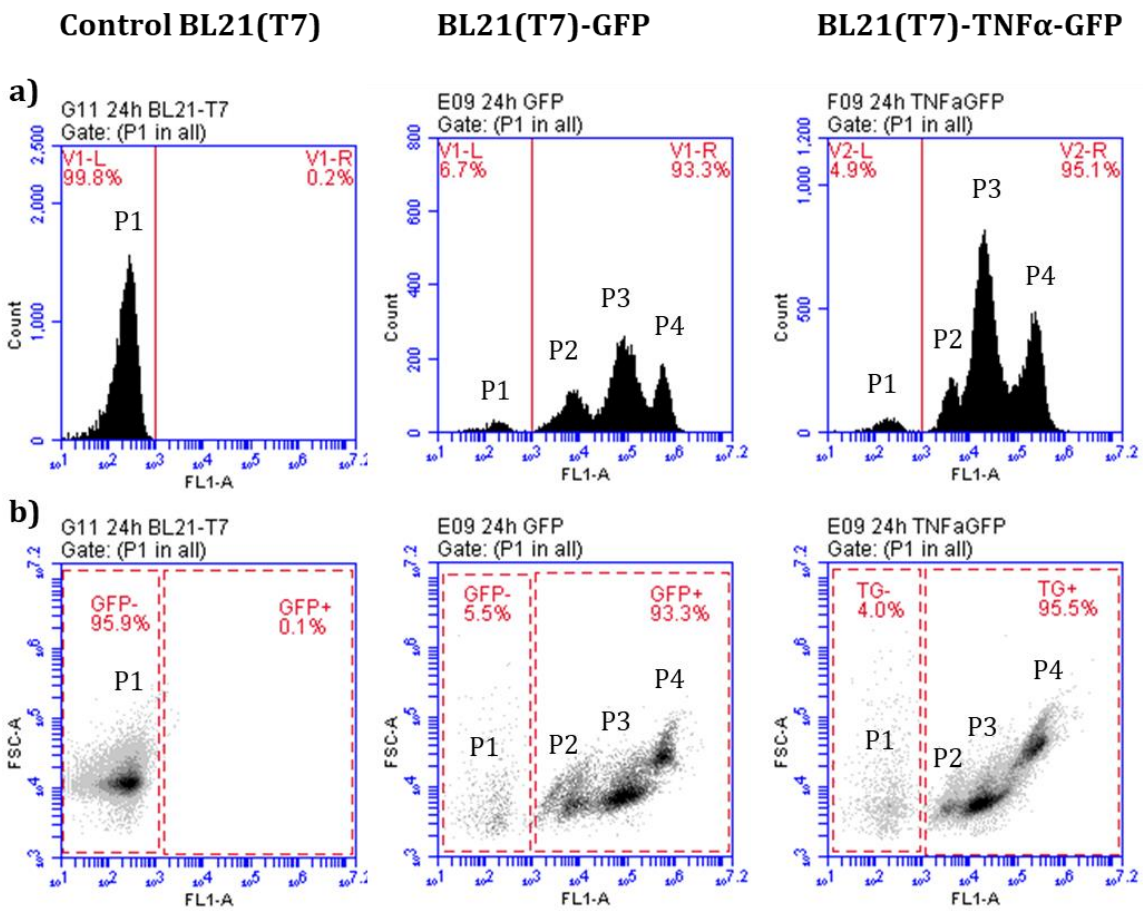


Figure 3.6: FCM detection of cultural heterogeneity after 24 h post-induction

(a) Population P1 (below 10^3 fluorescence units) comprises plasmid-free cells, as confirmed by plasmid-free control cultures (See **Figure 3.3**). Population P4 (on 10^5 fluorescence units) represent cells displaying higher green fluorescence measured, suggesting cells producing functional recombinant protein GFP or TNFα-GFP. The populations P2 and P3 as intermediate fluorescence (from 10^3 to 10^4 fluorescence units) had lower green fluorescence than population P4, likely reflecting subpopulations with different RP accumulation and folding state. (b) Mean forward scatter (FSC-A) versus FL1-A plots of data (a); note that the population with higher FL1-A value has higher FSC-A value, suggesting cells with GFP production are bigger in size.

It is known that incorrectly-folded fluorescent proteins that make up IBs can still retain the appropriate amounts of fluorescence, although not as much as the equivalent amounts of soluble protein (García-Fruitós et al., 2005), thus it is reasonable to assume this intermediate fluorescence becomes less fluorescent and can prone to aggregation. They have a huge probability of incorrectly folded and became non-functional.

Based on **Figure 3.6a**, population P1 consists of plasmid-free cells, as confirmed by the cultures of plasmid-free control (See **Figure 3.3**). Population P4 represents cells expressing higher green fluorescence, in which based on batch culture experiments performed in this condition for pORT(T7)-GFP, 0.2% arabinose-induced cultures exhibited higher mean green fluorescence at 6 h post-induction (**Figure 3.2a**) with FL1-A histogram data represented by a single population at 10^5 to 10^6 fluorescence units (**Figure 3.5a**). The best green fluorescence produced was shown by the histogram at the 10^5 green fluorescence unit, detected via FL1-A channel (533/30 BP). The populations of P2 and P3 have lower green fluorescence (intermediate fluorescence) than population P4 and were thought to contain insoluble inclusion bodies and have lower levels of recombinant protein production, based on Wyre and Overton (2014b).

Mean FSC-A value can be used as an estimation of cell size and indication of cell size changes in each single cell. Observation on **Figure 3.6b**; the mean forward scatter (FSC-A) versus FL1-A plots of data, gives the impression that when the cells population produced green fluorescence (higher in mean FL1-A value), mean

forward scatter appeared to increase, suggesting that cells size elongate and became bigger in size when producing GFP.

Staining with propidium iodide (PI)

In addition to measuring GFP, further analysis of FCM was run to determine cell viability by using propidium iodide (PI) staining (more details in **Figure 2.4**). Therefore, all subpopulations; live or dead, and productive or non-productive GFP-producing, can also be identified using flow cytometry. Gating strategy is designed to set boundaries for all different populations that exist in the cultures, especially the one that was obviously seen at 24 h post-induction. This gating strategy can help to identify the type of populations exist and can also facilitate the enumeration of percentages for each population in the cultures. For ease of quantification and analysis, FCM data were displayed as mean green fluorescence (FL1-A) versus red fluorescence (FL3-A) plots of cells. Based on the predetermined gating as **Figure 3.7**, all populations that were present in the cultures after 24 h post-induction can be easily identified. Among the four identified populations were live populations that do not produce GFP (GFP- PI-), live populations producing either intermediate or higher green fluorescence (GFP+ PI-), populations that initially produced green fluorescence but then died (GFP+ PI+) and populations of non-productive but also died (GFP- PI+). The detection of multi populations within cultures using FCM revealed the heterogeneity of cultures after 28 h of growth, reflecting the variations in the accumulation and folding state of GFP/ TNFα-GFP.

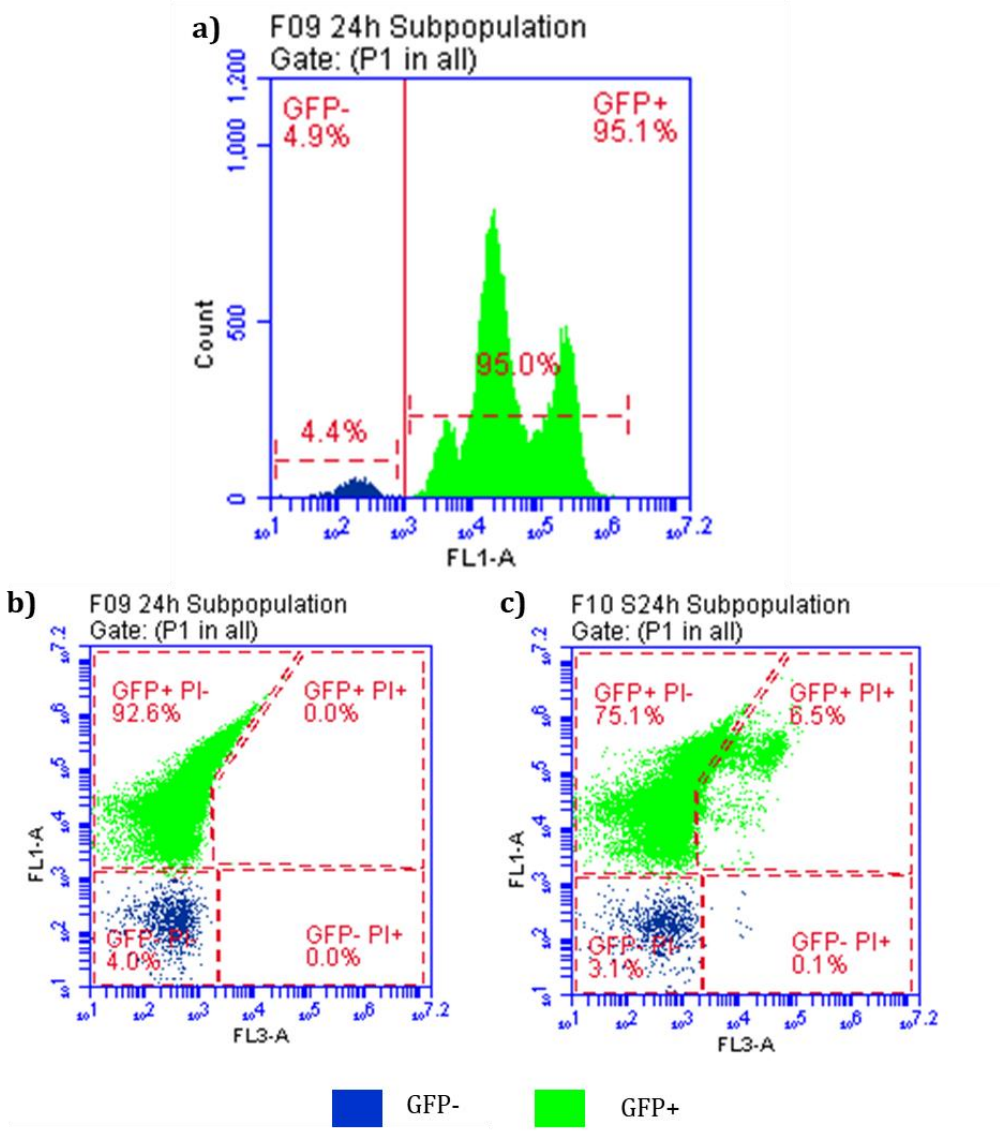


Figure 3.7: Gating strategy to identify live/dead cells and productive/non-productive GFP-producing cells

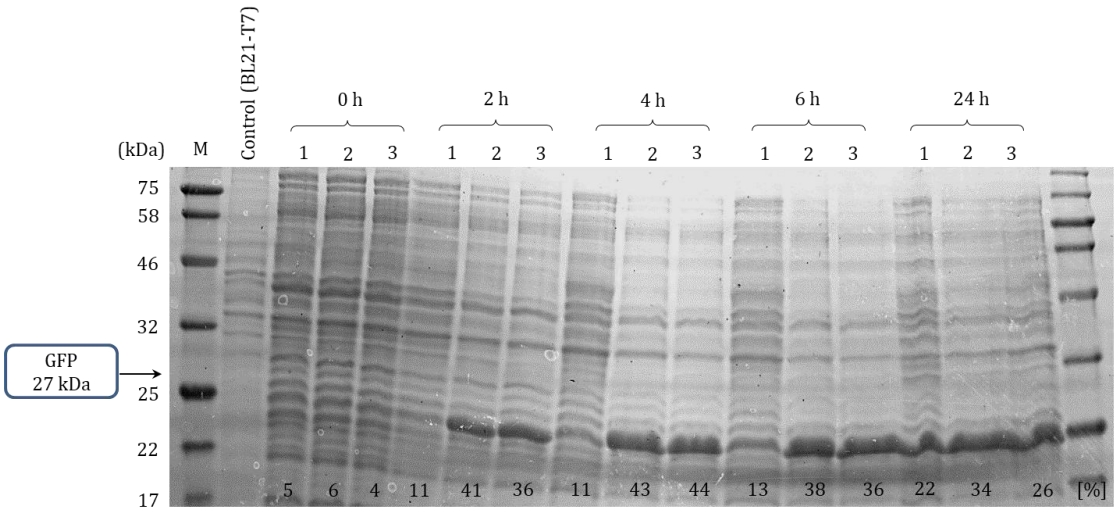
(a) FL1-A histogram data showed subpopulations in the cultures that represents heterogeneity after 24 h post-induction of BL21(T7)-pORT(T7)-GFP. (b) is data (a) when plotted as mean green fluorescence FL1-A versus red fluorescence (FL3-A) to determine subpopulation of live plasmid-free cells (GFP- PI-) and live intermediate and higher green fluorescence-producing cells (GFP+ PI-). (c) is (b) when stained with propidium iodide (PI) to identify population percentage of dead green fluorescence-producing cells (GFP+ PI+) and dead cells that did not produce GFP (GFP- PI+).

3.2.2 SDS-PAGE and fractionation

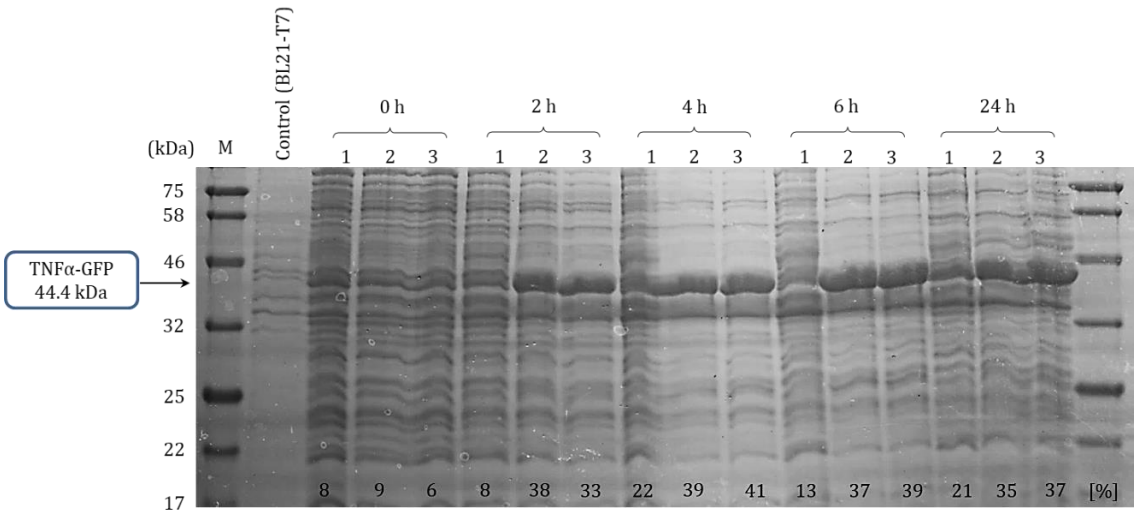
Further analysis had been done to investigate the production of GFP and TNFα-GFP. Samples were collected at intervals time point of 0 h, 2 h, 4 h, 6 h and 24 h post-induction and has been evaluated by using SDS-PAGE. SDS-PAGE was used in order to measure the amount of recombinant protein (RP) synthesised. The results of the analysis were expressed as the percentage of total cellular protein (TCP) that was the RP (GFP or TNFα-GFP). Cell protein bands were compared to the Blue Protein Standard (BPS) ladder band. GFP has a molecular weight of 27 kDa while TNFα-GFP is 44.4 kDa. Densitometry data for the percentage of cellular protein that was GFP or TNFα-GFP is shown on the gel pictures in **Figure 3.8** and expressed as estimated protein concentrations in cell extracts (**Figure 3.9**).

From **Figure 3.9**, the quantity of GFP and TNFα-GFP was observed to increase over the time up to 8 h growth (4 h post-induction) and RP content was comparable in all induced cultures up to 10 h growth (6 h post-induction), regardless of inducer concentration was being added. It was found that the highest amount of GFP was produced at 4 h post-induction for 0.015% arabinose-induced pORT(T7)-GFP cultures with 43% of TCP (0.43 mg/mL), while the highest amount of TNF-GFP was produced at the same timepoint for 0.2% arabinose-induced pORT(T7)-TNF-GFP cultures, reaching 41% of TCP (0.55 mg/mL).

(a)



(b)



① No Ara ② 0.015% Ara ③ 0.2% Ara

Figure 3.8: The accumulation of GFP and TNFα-GFP in cell extracts expressed by SDS-PAGE in batch culture with 0.4% glycerol at 30°C

SDS-PAGE gel picture with arrow showing (a) GFP (27 kDa) and (b) TNFα-GFP (44.4 kDa) with the percentage of total GFP/TNFα-GFP proteins expressed inside the cell extracts. Sample bands at times post-induction were compared to BPS ladder band (**M:** Marker used; Blue Protein Standard (BPS), broad range of 11-190 kDa).

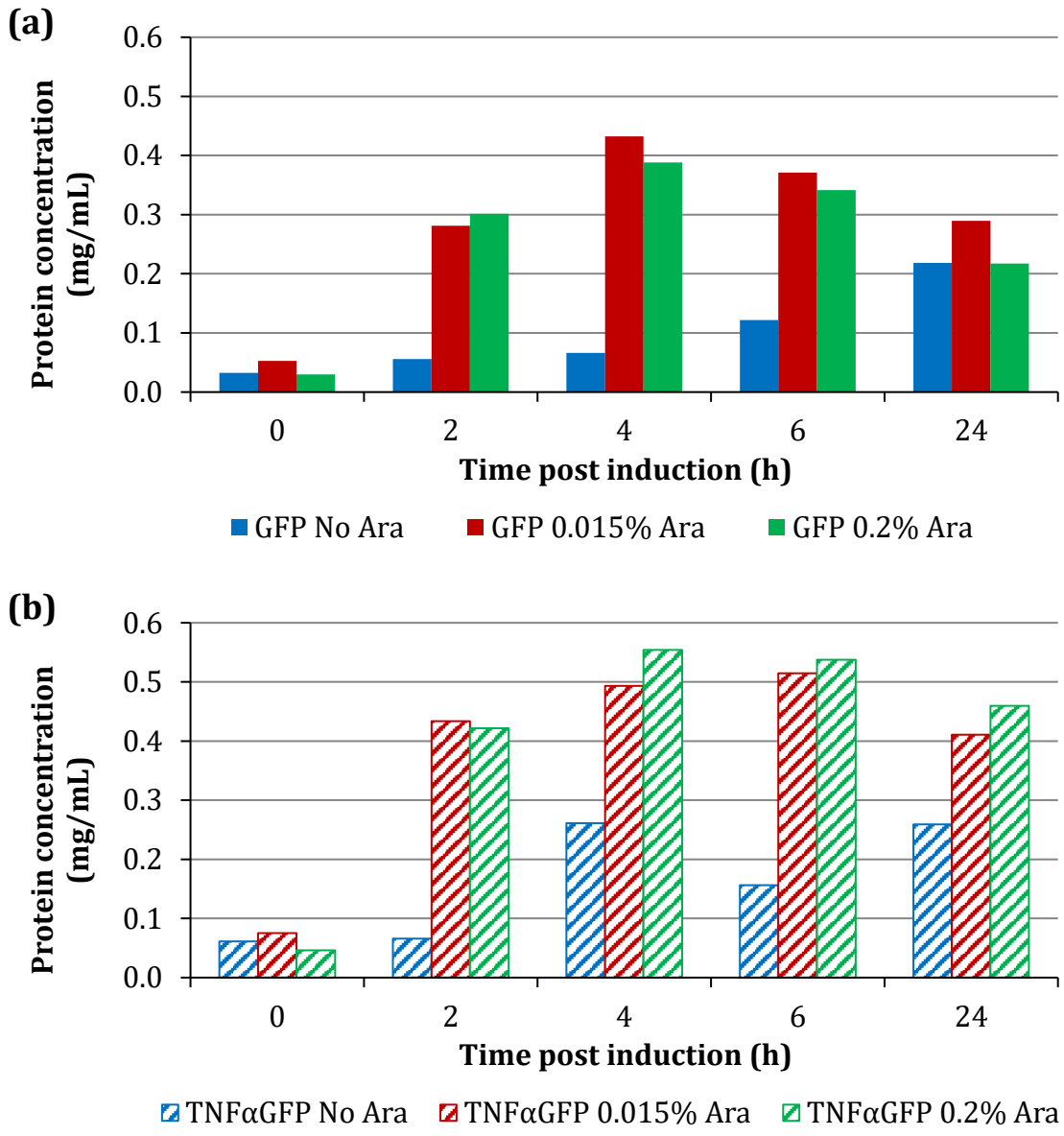


Figure 3.9: The expression of total proteins in cell extracts measured by SDS-PAGE for GFP and TNFα-GFP in batch culture with 0.4% glycerol at 30°C

Predicted total protein concentration in cell extracts of (a) GFP and (b) TNFα-GFP measured in mg/mL at times post-induction.

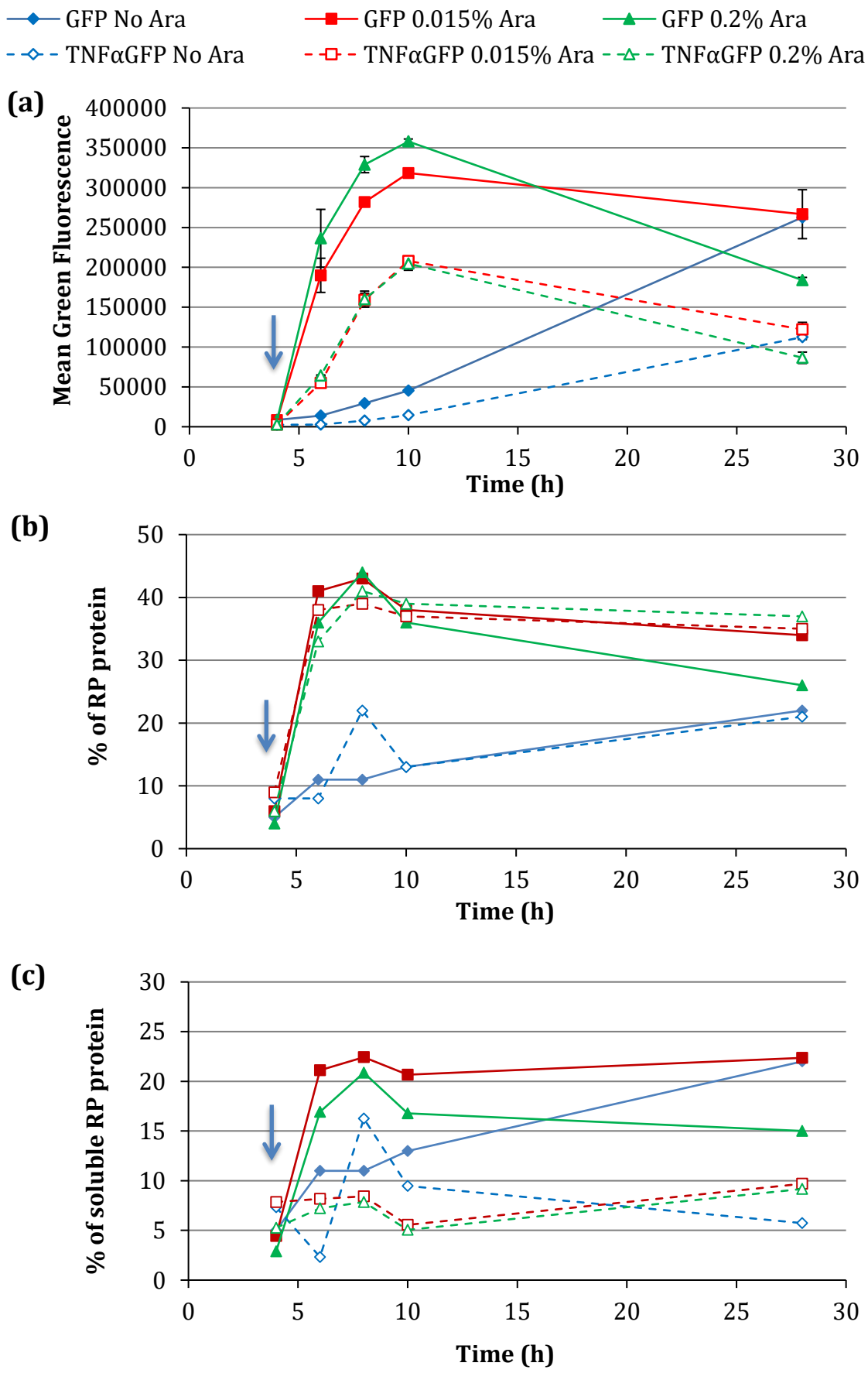

GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara
TNFαGFP No Ara
TNFαGFP 0.015% Ara
TNFαGFP 0.2% Ara
(a)
Mean Green Fluorescence
400000
350000
300000
250000
200000
150000
100000
50000
0
0
5
10
15
20
25
30
Time (h)
(b)
% of RP protein
50
40
30
20
10
0
0
5
10
15
20
25
30
Time (h)
(c)
% of soluble RP protein
30
25
20
15
10
5
0
0
5
10
15
20
25
30
Time (h)

Figure 3.10: Comparison of mean green fluorescence as determined by FCM with soluble RP content for GFP and TNFα-GFP in batch culture with 0.4% glycerol at 30°C

Recombinant human TNFα fused to GFP (or also known as rhTNFα-GFP) was expressed from an arabinose-induced plasmid system with different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow), and compared to the production of GFP from the same plasmid pORT(T7).

(a) FL1-A (mean green fluorescence) plots of cells taken from culture samples at interval time points post-induction. Samples were analysed by SDS-PAGE for separation, followed by subcellular fractionation to separate soluble and insoluble proteins. The percentage of cellular protein, that is, recombinant protein (b) and the percentage of cellular protein, that is, soluble recombinant protein (c) were determined. Data points represent the mean values of duplicate cultures ± standard deviation.

At 28 h of growth, the synthesized RP was lower in GFP cultures induced with 0.2% arabinose compared to the other three induced conditions (**Figure 3.10b**). Meanwhile uninduced cultures contained comparable proportions of RP at this timepoint. The percentage of GFP and TNFα-GFP content in cell extracts were comparable, despite the fact that the fluorescence of cultures expressing GFP was higher than that of those expressing TNFα-GFP (**Figure 3.10a**). This is probably due to misfolding of TNFα-GFP. Thus, fractionation was performed in order to determine RP solubility to correlate fluorescence data from FCM with protein data generated by SDS-PAGE.

Subcellular fractionation was done using the freeze-thaw method on bacterial pellets collected at intervals time point to separate soluble and insoluble protein fractions. The fractionation of soluble and insoluble proteins allowed us to determine the proportion of RP present in the cell that was soluble (**Figure 3.10c**).

The highest proportion of soluble RP was observed on 0.015% arabinose-induced GFP cultures at 24 h post-induction with 0.26 mg/mL (**Figure 3.11a**), reaching 66% solubility of the total GFP generated (**Figure 3.12**). While the percentage of RP content from TCP in cell extracts of TNF-GFP induced cells was equivalent to that of GFP, subcellular fraction revealed that TNF-GFP produced substantially lower soluble protein (**Figure 3.10c**). At 6 h post-induction, TNFα-GFP was accumulated the most in insoluble form for 0.2% arabinose-induced pORT(T7)-TNFα-GFP cultures with 0.61 mg/mL (**Figure 3.11b**), accounting for 87% of the total TNFα-GFP that was produced (**Figure 3.13**). The overall results showed that

GFP was more soluble than TNFα-GFP, and that increasing concentrations of arabinose reduced GFP solubility.

When compared side-by-side, the soluble RP content (Figure 3.10c) and the mean green fluorescence as measured by flow cytometry (Figure 3.10a) reveal that the two measurements are broadly comparable to one another. Multiplying the mean green fluorescence value obtained from FCM by the number of cells per unit volume yields an estimation of the amount of soluble RP present in the culture.

The data as a whole reveal that although the levels of production of GFP and TNFα-GFP are very similar to one another (**Figure 3.10b**), the difference in their solubility (**Figure 3.10c**) essentially explains the difference in their fluorescence (**Figure 3.10a**). The synthesis of TNFα-GFP in a functional (fluorescent) form poses greater challenges for *E. coli* compared to the production of GFP alone.

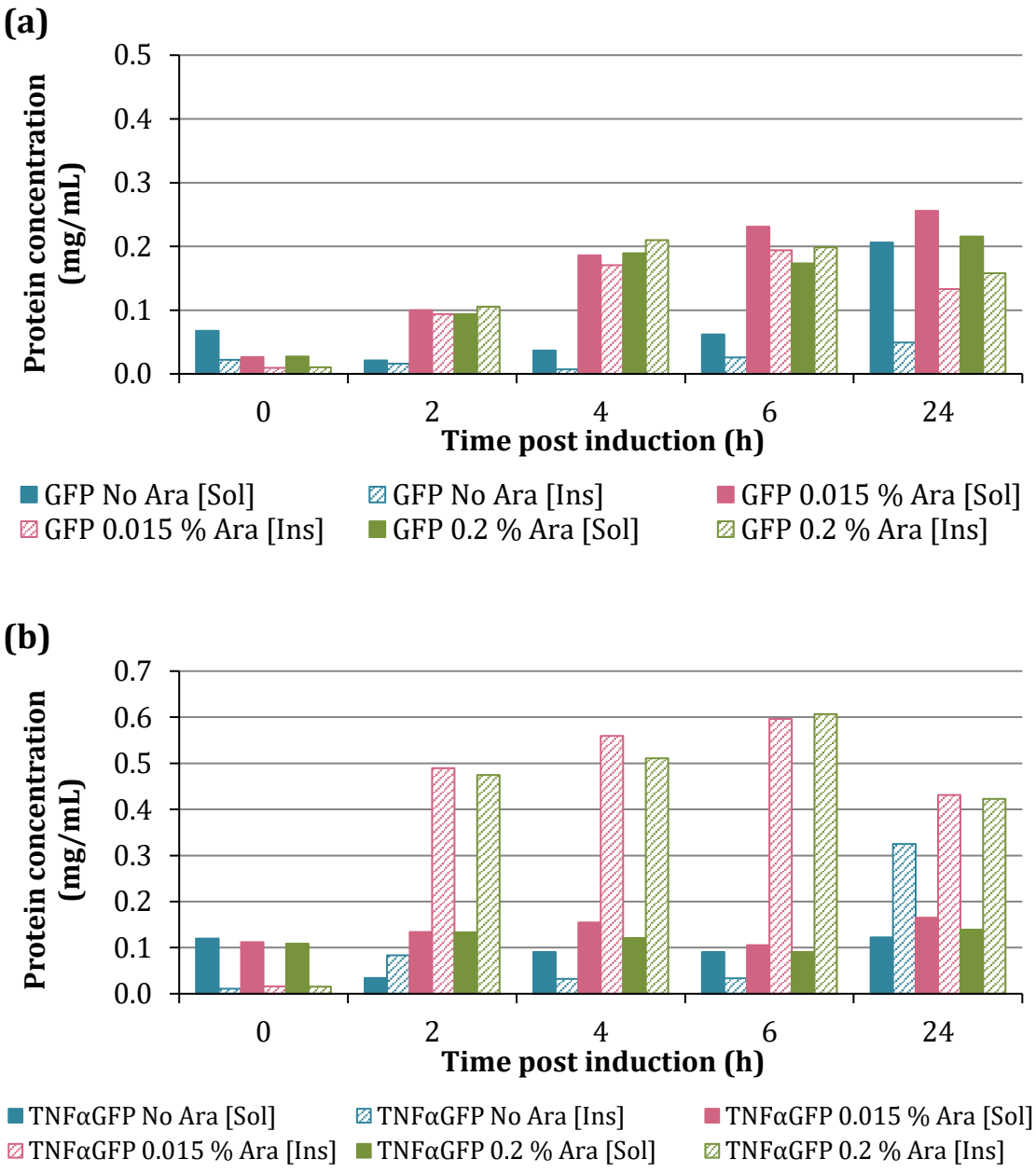
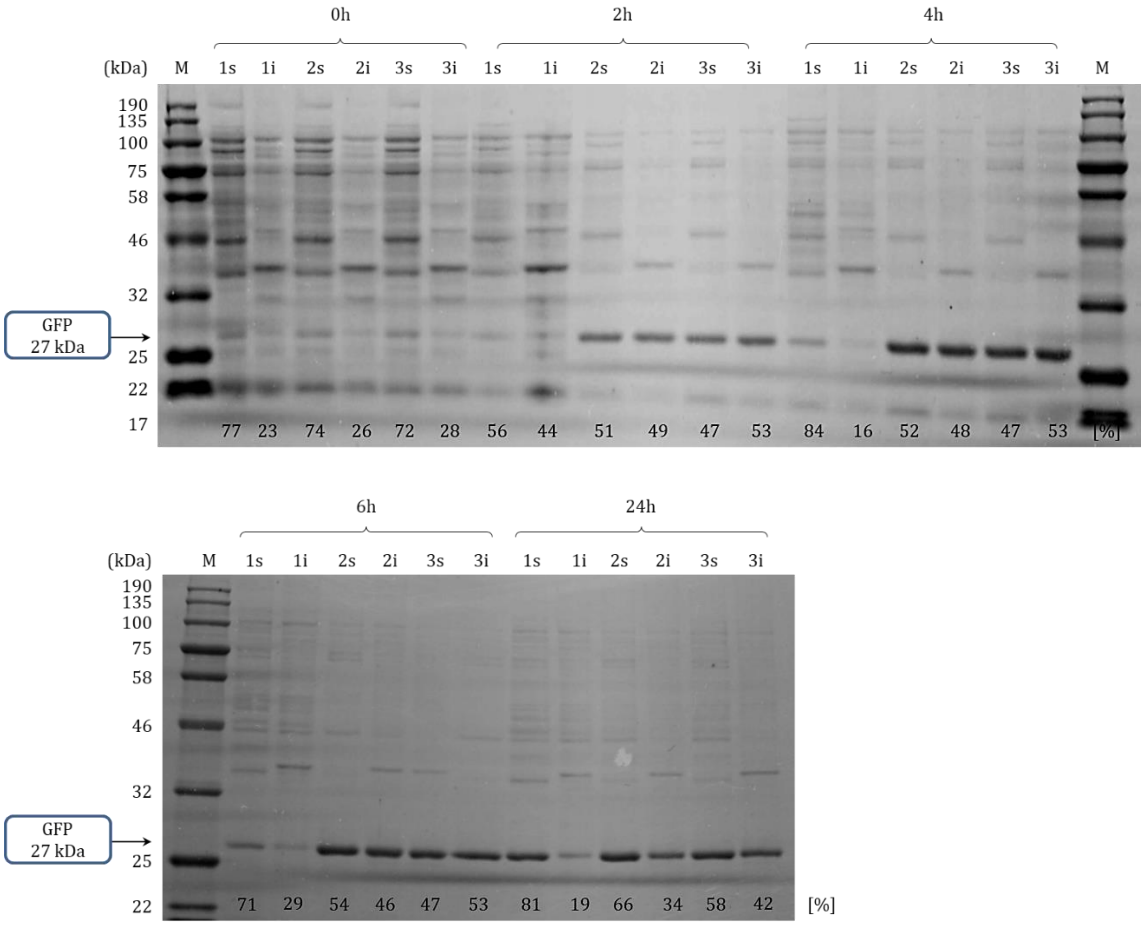


Figure 3.11: The expression of protein solubility in cell extracts measured by SDS-PAGE for GFP and TNFα-GFP in batch culture with 0.4 % glycerol at 30 °C

Estimated protein solubility of soluble and insoluble fractions in cell extracts of (a) GFP and (b) TNFα-GFP at concentration of mg/mL at times post induction.



M: Marker used; Blue Protein Standard (BPS), broad range of 11-190 kDa;

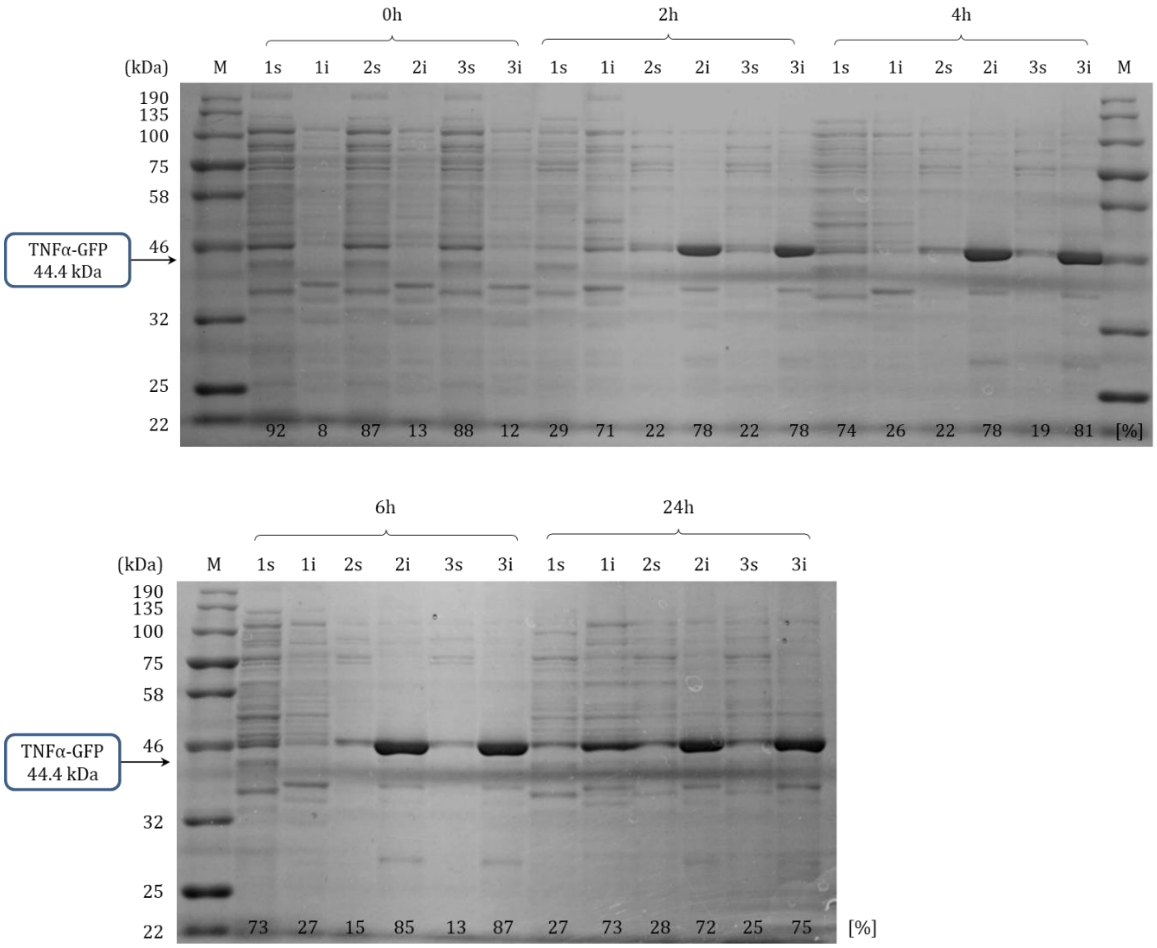
1s: GFP No Ara (Soluble); **1i**: GFP No Ara (Insoluble);

2s: GFP 0.015 % Ara (Soluble); **2i**: GFP 0.015 % Ara (Insoluble);

3s: GFP 0.2 % Ara (Soluble); **3i**: GFP 0.2 % Ara (Insoluble)

Figure 3.12: The accumulation of GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in batch culture with 0.4% glycerol at 30°C

SDS-PAGE gel picture with arrow showing GFP (27 kDa) and the ratio percentage of soluble versus insoluble GFP proteins accumulated inside the cell extracts of pORT(T7)-GFP, as shown at the bottom of the gel. All sample bands at times post-induction were compared to BPS ladder band.



M: Marker used; Blue Protein Standard (BPS), broad range of 11-190 kDa;

1s: TNFα-GFP No Ara (Soluble); **1i**: TNFα-GFP No Ara (Insoluble);

2s: TNFα-GFP 0.015 % Ara (Soluble); **2i**: TNFα-GFP 0.015 % Ara (Insoluble);

3s: TNFα-GFP 0.2 % Ara (Soluble); **3i**: TNFα-GFP 0.2 % Ara (Insoluble)

Figure 3.13: The accumulation of TNFα-GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in batch culture with 0.4 % glycerol at 30 °C

SDS-PAGE gel picture with arrow showing TNFα-GFP (44.4 kDa) and the ratio percentage of soluble versus insoluble TNFα-GFP proteins accumulated inside the cell extracts of pORT(T7)-TNFα-GFP, as shown at the bottom of the gel. All sample bands at times post-induction were compared to BPS ladder band.

3.2.3 Discussion

Previous research conducted by Huang et al. (2012) on the investigation of various carbon sources and their influence on the transcription profile of *E. coli* had revealed that, in numerous cases, the use of glycerol led to even superior production of RP in *E. coli*. The introduction of glycerol into the cell leads to a decrease in the rate of carbon utilisation during glycolysis, which lowers the concentration of acetate produced significantly (Lee, 1996). In an alternative perspective, glycerol exhibits a comparatively slower rate of cellular uptake by facilitated diffusion. Consequently, this resulted in a reduced carbon flow through glycolysis and a reduction in the metabolic stress on *E. coli* while producing RP.

The findings of this study demonstrated that using 0.4% glycerol as the carbon source in the growing medium is sufficient for the production of improved and increased solubility RP. Small-scale fermentations often use glycerol in concentrations between 0.2% and 0.4% (Wyre and Overton, 2014b). A previous study conducted by Sivashanmugam et al. (2009) revealed that medium containing 0.2% and 0.3% glycerol resulted in a limited production of the desired protein, whereas medium containing 0.4% and 0.5% glycerol exhibited a yield comparable to that achieved with 0.75% glycerol. Following the procedure as Kasli et al. (2019), a carbon source of 0.4% glycerol had been used for the growing of all cultures in 100 mL of TB medium contained within 250 mL conical flasks, as it was deemed appropriate for the growth conditions of the *E. coli* strains under investigation.

Besides, according to Lebendiker and Danieli (2014), the ability to keep low concentrations of an inducer is advantageous for the efficient synthesis of RPP, hence potentially avoiding dramatic decrease in protein productivity. In 0.4% glycerol cultures, the concentration of 0.015% arabinose, which was more than tenfold lower than 0.2% arabinose-induced, resulted in a similar level of green fluorescence production (**Figure 3.2a**). Nevertheless, the results of subcellular fractionation exhibited that GFP solubility was decreased when arabinose concentrations increased (**Figure 3.10c**). Higher inducer concentrations, as previously mentioned, did not result in an increase in the rate of particular protein expression or the cellular productivity (Lecina et al., 2013). This lower concentration of 0.015% arabinose cultures, in addition, had fewer dead cells than cultures with 0.2% arabinose-induced (**Figure 3.2c**).

The comparability between the amount of soluble RP measured by SDS-PAGE and the mean green fluorescence obtained by FCM was observed to be extensive. One notable benefit of utilising FCM is that it generates data much more quickly, often within 5–10 min on average, than SDS-PAGE which is in contrast demands several hours to complete. Therefore, FCM can serve as a real-time tool for in-process analysis, allowing for the dynamic adjustment of processes in response to the obtained results at that particular instant.

The utilisation of glucose as a feed may provide challenges associated with the accumulation of acids resulting from overflow metabolism. This becomes particularly problematic when growth rates fluctuate, leading to a drop in the pH of the growth medium, thereby inhibiting the growth. The utilisation of glycerol as

a carbon source often does not pose this issue (Overton, 2014; Want et al., 2009). Nevertheless, a further investigation was conducted in order to explore the extent to which the finding is reliable. The subsequent experiment thus was developed for batch culture, using TB medium with glycerol substituted by glucose, to determine whether the use of glucose may potentially enhance both growth and RP production.

3.3 Replacing glycerol with glucose as a carbon source

The next experiment was the growth of cells in 100 mL of TB medium supplemented with 50 µg/mL of Kanamycin, and 0.4% glucose as the carbon source. All other experimental conditions remained the same as in the previous experiment; a temperature of 30°C was applied, and either 0.015% or 0.2% of arabinose concentration was used to induce RPP when the OD_{650} reached around 0.5, corresponding to the mid-logarithmic phase. An uninduced control without arabinose was also included. Samples were collected at regular intervals and the OD_{650} were measured. Further, some samples were analysed for CFU on NA plates to determine the culturability of the cultures and were subsequently subjected to replica-plating onto NA plates with 50 µg/mL Kanamycin to assess plasmid retention. All experiments conducted in shake flasks were executed in duplicate.

The use of glucose as the carbon source has been compared to the use of glycerol in the growth media. Glucose is well recognised as the favoured carbon source for bacteria, especially *E. coli* due to its ability to facilitate faster growth rate in comparison to other types of sugars. Furthermore, glucose provides a significant

amount of energy to the cells as glucose is converted in glycolysis into glucose 6-phosphate and consumed through the tricarboxylic acid (TCA) cycle (Deutscher et al., 2006). However, glucose represses the arabinose-inducible promoter, which is responsible for regulating the production of T7 RNA polymerase (T7 RNAP).

The growth of all pORT(T7)-GFP cultures were slower when the carbon source used was glucose, in comparison to pORT(T7)-TNFα-GFP cultures (**Figure 3.14a**). Cultures expressing TNFα-GFP with 0.015% arabinose-induced were observed to show the highest cell densities than those induced, reaching the final OD_{650} of 17.02. The cultures generating GFP and were both induced with 0.015% and 0.2% arabinose had the lowest OD_{650} growth of 10.5 and 9.93, respectively, following 24 h post-induction.

CFU data was comparable to the data obtained when using 0.4% glycerol as carbon source. All induced cultures of GFP and TNFα-GFP showing a drop in culturability at 4 h post-induction then recovered 24 h after induction (**Figure 3.14b**). Non-induced cultures showed no reduction in culturability throughout the growth. Plasmid retention data showed very uncertain results, but the most severely affected were both GFP and TNFα-GFP cultures induced with 0.2% arabinose (**Figure 3.14c**).

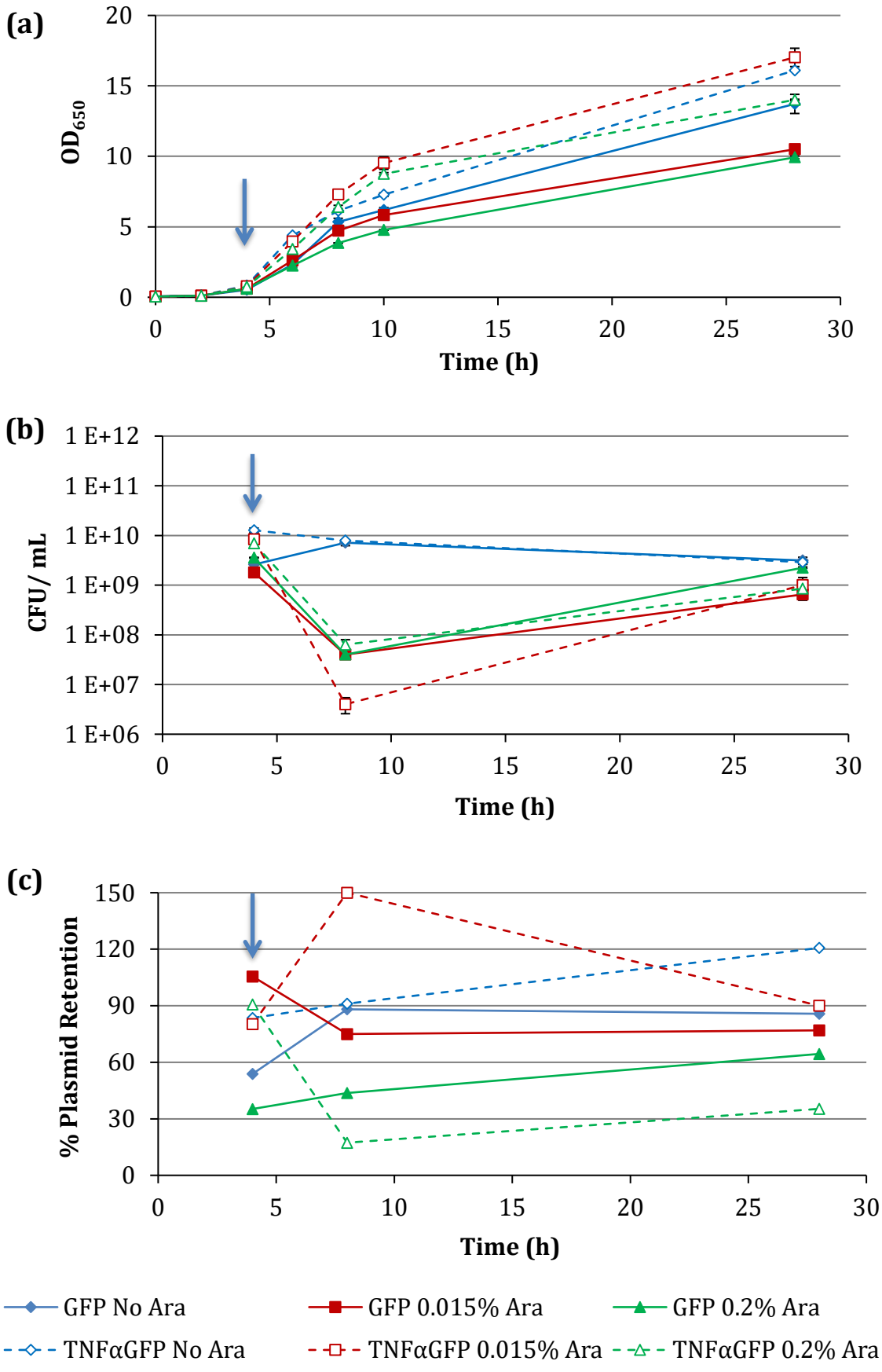

(a)
OD650
Time (h)
(b)
CFU/ mL
Time (h)
(c)
% Plasmid Retention
Time (h)
GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara
TNFαGFP No Ara
TNFαGFP 0.015% Ara
TNFαGFP 0.2% Ara

Figure 3.14: Correlation of *E. coli* BL-21(T7) growth for the study of optimisation for GFP and TNFα-GFP production, grown in TB medium at 30°C with 0.4% glucose

E. coli BL-21(T7) cultures carrying the plasmid pORT(T7)-GFP or pORT(T7)-TNFα-GFP were induced with three different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at OD_{650} ≈ 0.5 (approximately after 4 h of growth; indicated by blue arrow), evaluated by (a) OD_{650} measurement, (b) CFU/mL counting, and (c) plasmid retention. (a) The OD_{650} was measured at intervals. (b) Serial dilution of culture samples at 0 h, 4 h and 24 h post-induction was plated onto NA agar for indication of cell viability and, (c) were replica-plated onto NA agar supplemented with Kanamycin (50 µg/mL) to reveal plasmid retention. All data shown are mean values from replica flasks, error bars are ±1 standard deviation.

3.3.1 Flow cytometry measurement

Consistent with the findings obtained in batch culture using 0.4% glycerol, the mean green fluorescence of cultures that produced GFP was observed to be higher compared to cultures producing TNFα-GFP, despite the induction of equal arabinose concentrations. Both induced TNFα-GFP cultures revealed roughly half reduced green fluorescent (GF) production than cultures induced with GFP alone. While the initial phase of GF production in non-induced cultures was notably low, there was a subsequent increase in GF levels observed towards the end of the 28-h growth, particularly in non-induced GFP cultures.

It was demonstrated that the induction with arabinose greatly enhanced the green fluorescence. A comparison of FCM data showed that glucose decreased the production of GFP or TNFα-GFP in cultures induced with the lower concentration (0.015%) of arabinose (**Figure 3.15a**), in contrast to cultures grown with glycerol (**Figure 3.2a**), where the green fluorescence of bacteria was not significantly affected by different arabinose concentrations. The phenomenon of 'dampening' has been previously reported, whereby the *araBAD* promoter is repressed by glucose and activated by arabinose, allowing for 'fine tuning' production of recombinant proteins (Castiñeiras et al., 2018b).

Individual green fluorescence histograms (**Figure 3.16**) exhibited partial induction by 0.015% arabinose, observed in both GFP and TNFα-GFP cultures, as indicated by a shoulder on the left side of the fluorescence peak, which appeared most noticeable at 6 h (2 h post-induction) but remained visible until 10 h of

growth. There was no such partial induction found in either cultures with 0.2% arabinose.

Observations on FL1-A histogram data showed all GFP cultures produced a nearly 100% of cell populations producing GFP (**Figure 3.15b**). All induced cultures exhibited multiple peaks at 28 h of growth. Specifically, TNFα-GFP cultures induced with 0.2% arabinose displayed a significant GFP- peak (around 2×10^2 fluorescence units) corresponding to a population of non-productive cells or plasmid-free cells (**Figure 3.16**). Equivalent to the data shown by the proportion of GFP+ cells within these cultures that showed a notable reduction of GFP+ cells, with a decline to 62% at 24 h post-induction (**Figure 3.15b**). This could potentially be related to the huge loss of plasmids in these cultures as early as 4 h post-induction (**Figure 3.14c**).

The results obtained from PI staining indicated that GFP-expressing cultures had a higher proportion of dead cells than TNFα-GFP cultures (**Figure 3.15c**). Both induced GFP cultures had extremely high percentages of dead cells, which reached 10% at 24 h post-induction. This likely explains for the comparatively lower recorded OD_{650} values in the GFP cultures (**Figure 3.14a**).

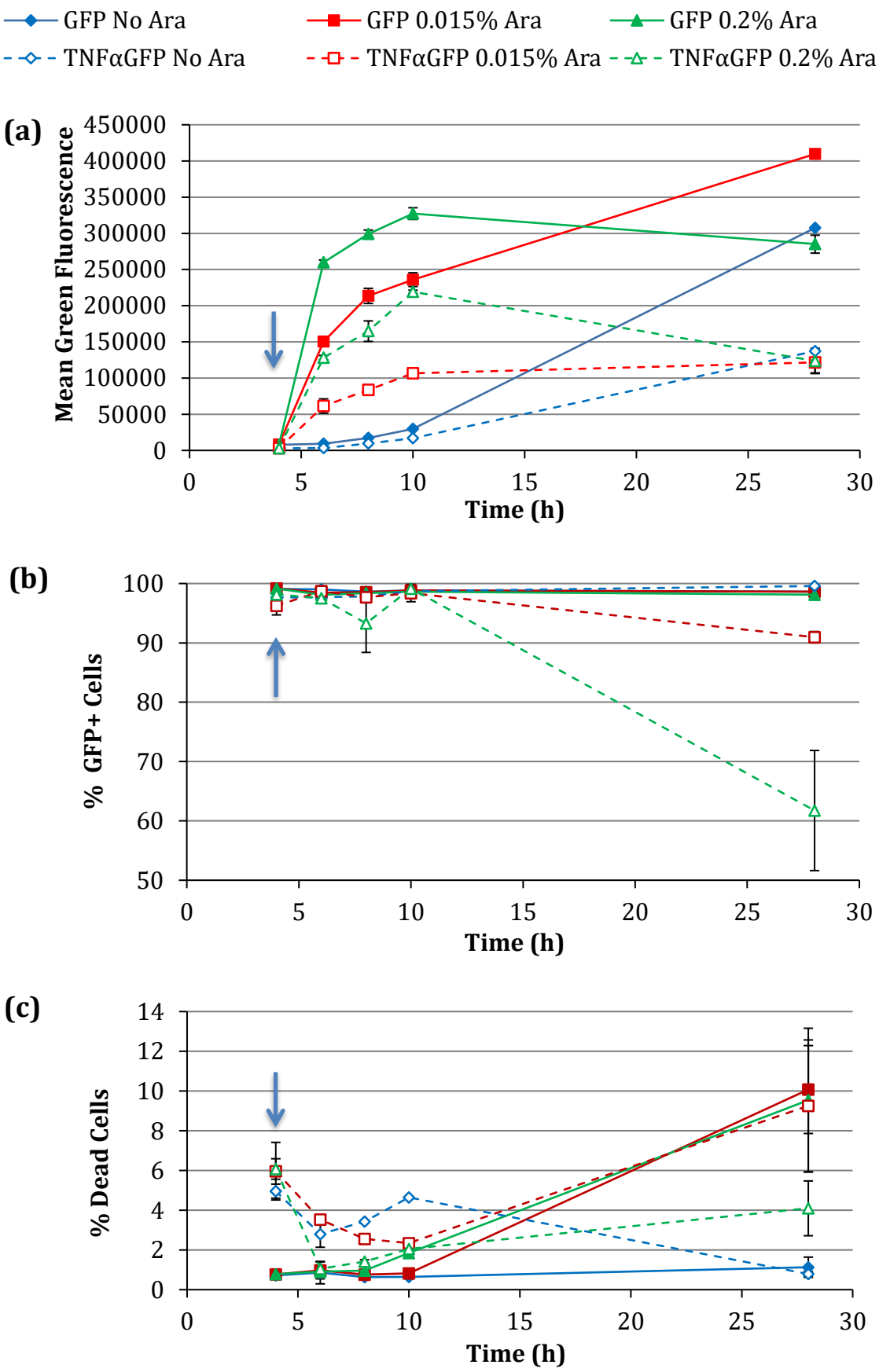

GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara
TNFαGFP No Ara
TNFαGFP 0.015% Ara
TNFαGFP 0.2% Ara
(a)
Mean Green Fluorescence
450000
400000
350000
300000
250000
200000
150000
100000
50000
0
0
5
10
15
20
25
30
Time (h)
(b)
% GFP+ Cells
100
90
80
70
60
50
0
5
10
15
20
25
30
Time (h)
(c)
% Dead Cells
14
12
10
8
6
4
2
0
0
5
10
15
20
25
30
Time (h)

Figure 3.15: Mean green fluorescence measured by flow cytometry for GFP and TNFα-GFP in batch culture with 0.4% glucose at 30°C

Recombinant human TNFα fused to GFP (or also known as rhTNFα-GFP) was expressed from an arabinose-induced plasmid system with different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow), and compared to the production of GFP from the same plasmid pORT(T7). (a) FL1-A (green fluorescence) plots of cells taken from culture samples at interval time points post-induction. (b) Percentage of cells that producing GFP (GFP+ cells) versus time, and (c) Percentage of dead cells at interval 0 h to 24 h post-induction. Data are mean values of duplicate samples, error bars are ±1 standard deviation.

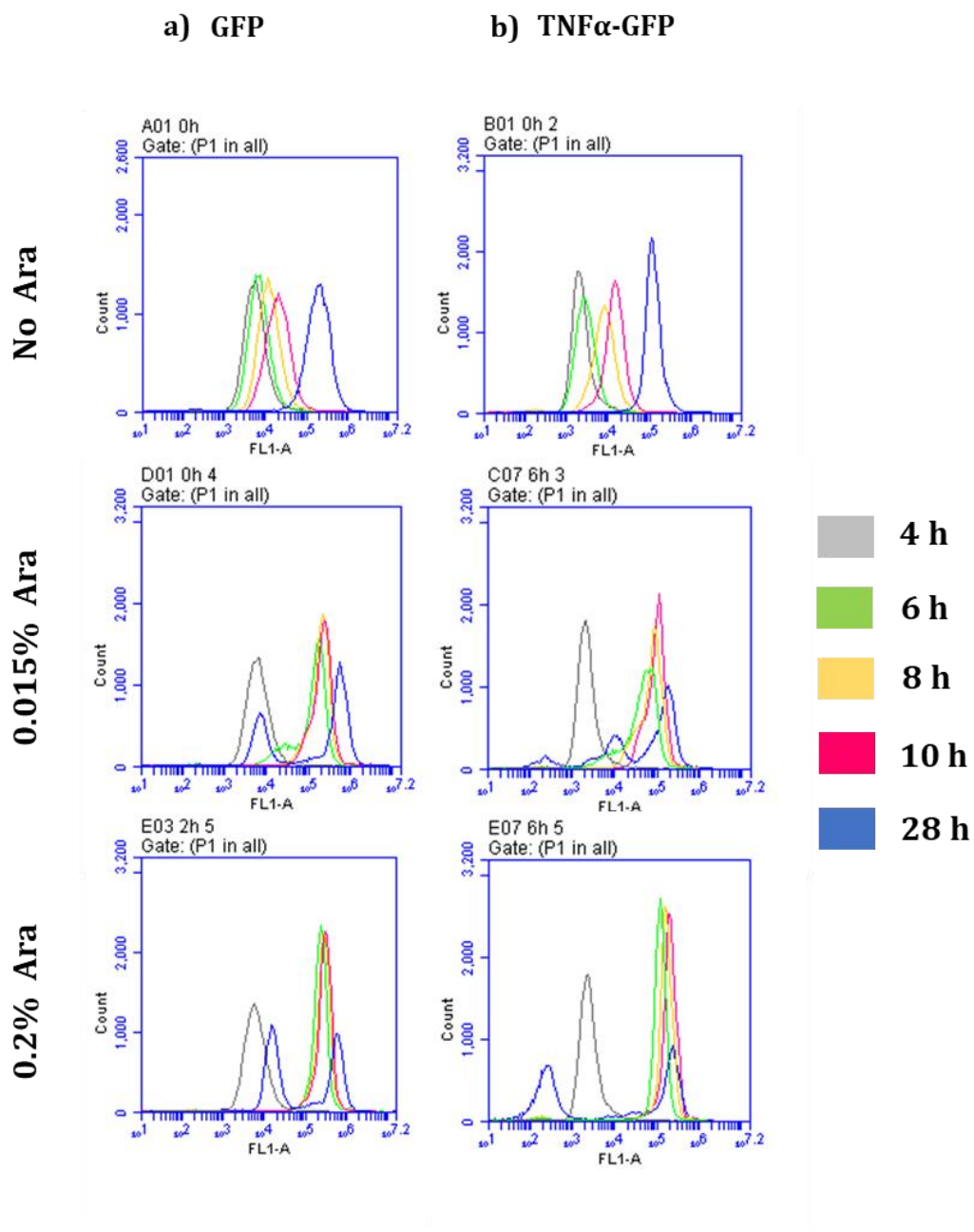


Figure 3.16: Single cell analysis of T7 expression system cultures for GFP and TNFα-GFP with 0.4% glucose at 30°C

(a) *E. coli* BL21-T7 cultures carrying p2T7-GFP (GFP) and (b) p2T7-TNFα-GFP (TNFα-GFP) were grown in Terrific Broth with 0.4% glucose as a carbon source at 30°C. At an OD_{650} of 0.5 (approximately after 4 h growth), arabinose (0.015% or 0.2%) was added. At regular intervals, samples were taken and FCM was used to measure green fluorescence of individual cells. Timepoints relate to the time the cultures were grown; 4 h refers to pre-induction with arabinose. Duplicate cultures of each plot showed a similar pattern throughout the growth.

3.3.2 Discussion

Previous studies have indicated that the utilisation of glycerol is more advantageous compared to glucose for better soluble RP synthesis, particularly when employed at an optimal concentration, which typically tends to be lower (Wang et al., 2015; Zhang et al., 2010). This is supported by Huang et al. (2012) who strongly suggested that the use of glycerol provide better production of RP in *E. coli* compared to glucose. The integration of glycerol as carbon source in the cells lead to a reduction in carbon flow of glycolysis. Hence, reduce acetate accumulation (Peng and Shimizu, 2003; Lee, 1996; Korz et al., 1995).

Several approaches have been established for controlling the specific growth rate in the purpose of regulating the metabolic activity of *E. coli* in glucose cultures. These methods have demonstrated the potential of growing the cultures at higher cell densities without the build-up of acetic acid, achieved by reducing the glucose concentration in the growth medium. Nevertheless, it had been observed that the metabolism of these *E. coli* was reorganized again under induction circumstances, resulting in the accumulation of acetic acid. This accumulation subsequently hampers cell growth and diminishes the effectiveness of protein expression, even at low concentrations (Lecina et al., 2013).

In glycerol-based cultures, the presence of different concentrations of arabinose, either 0.015% or 0.2%, did not yield any significant effects on both cell growth and green fluorescence produced (**Figure 3.1a**). However, this was not the case in glucose-based cultures. The repression of the arabinose-promoter, which regulates

T7RNAp, by glucose, leads to the inhibition of RPP. Nevertheless, a higher dose of arabinose was able to boost RPP because a "dampening effect" that took place in the cells where suppression by glucose had been tailored by arabinose for RPP (Castiñeiras et al., 2018b). Put simply, there is a positive correlation between the amount of glucose used as a carbon source and the amount of arabinose required to finely adjust the RPP.

3.4 Increasing glucose concentration from 0.4% to 0.8%

The subsequent batch culture was designed for the growth of cells in 100 mL of TB medium with increasing 0.8% glucose supplemented with 50 μg/mL Kanamycin and temperature of 30°C. This procedure was carried out to investigate whether increasing the glucose concentration might affect the growth profile of bacterial culture or influence any physiological effects on bacterial cells. Further, to explore any potential inhibition of cell growth that might be triggered by different glucose concentrations, or how they may impact the regulation of RPP. All experiments in shake flask were executed in duplicate.

Remarkably similar growth profiles were obtained with the increasing 0.8% glucose concentration compared to the 0.4% glucose cultures. Overall, the growth of all pORT (T7)-TNFα-GFP cultures were higher compared to pORT (T7)-GFP cultures (**Figure 3.17a**). Cultures of 0.015% arabinose-induced pORT(T7)-TNFα-GFP showed the highest cell densities, reaching the final OD_{650} of 16.32 at 24 h post-induction. The least OD_{650} growth was cultures of 0.2% arabinose-induced pORT(T7)-GFP with 10.99 at 24 h post-induction.

Similar to the data obtained when utilising 0.4% glucose as carbon source, all GFP and TNFα-GFP induced cultures exhibited a reduction in cellular culturability at 4 h post-induction (**Figure 3.17b**). The non-induced cultures exhibited a consistent level of cell culturability throughout the growth period, with no observed decrease. All cultures retained more than 75% of their plasmids throughout 28 h of growth (**Figure 3.17c**).

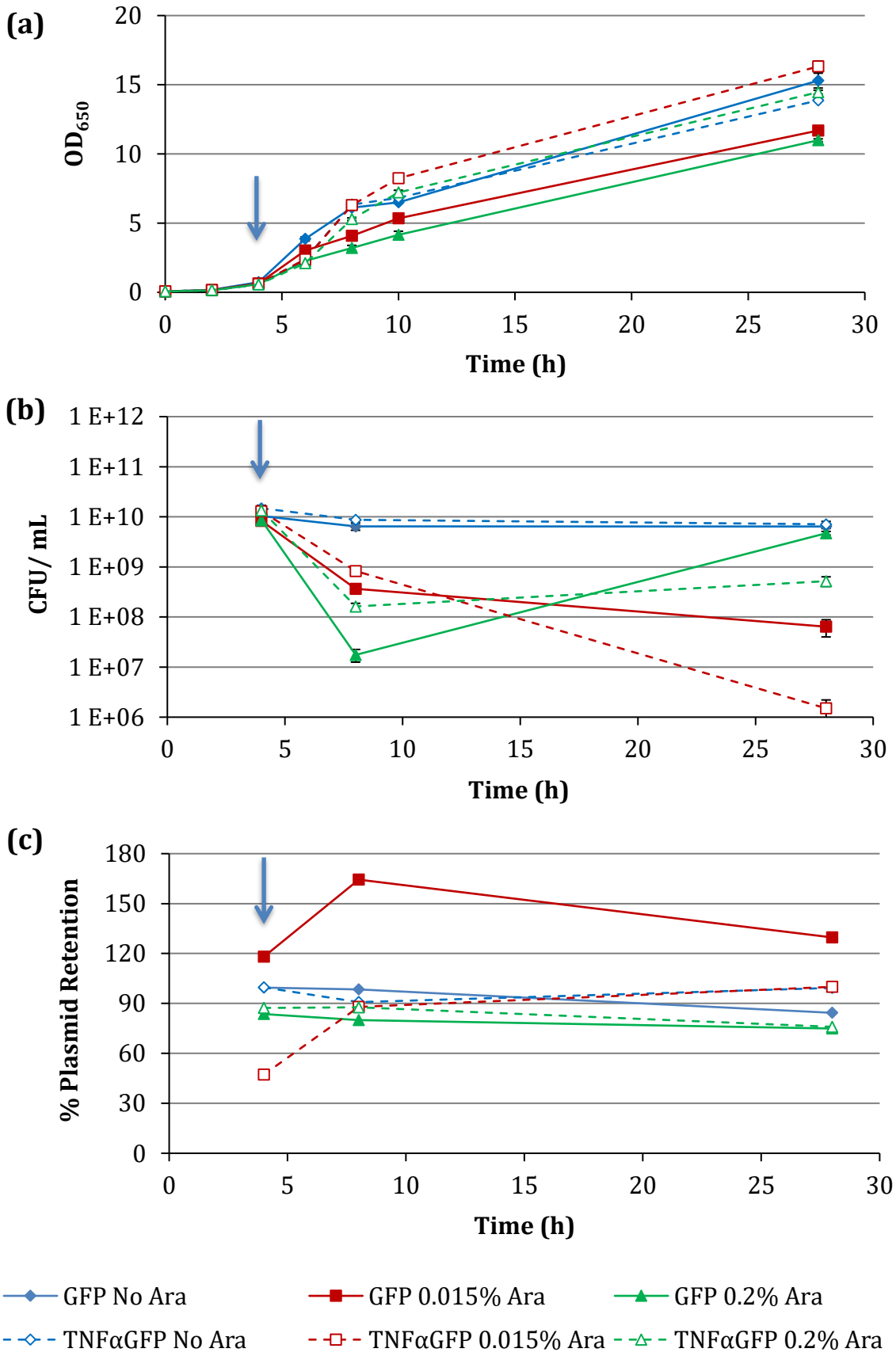

(a)
OD650
Time (h)
(b)
CFU/ mL
Time (h)
(c)
% Plasmid Retention
Time (h)
GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara
TNFαGFP No Ara
TNFαGFP 0.015% Ara
TNFαGFP 0.2% Ara

Figure 3.17: Correlation of *E. coli* BL-21(T7) growth for the study of optimisation for GFP and TNFα-GFP production, grown in TB medium at 30°C with 0.8% glucose

E. coli BL-21(T7) cultures carrying the plasmid pORT(T7)-GFP or pORT(T7)-TNFα-GFP were induced with three different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at OD_{650} ≈ 0.5 (approximately after 4 h of growth; indicated by blue arrow), evaluated by (a) OD_{650} measurement, (b) CFU/mL counting, and (c) plasmid retention. (a) The OD_{650} was measured at intervals. (b) Serial dilution of culture samples at 0 h, 4 h and 24 h post-induction was plated onto NA agar for indication of cell viability and, (c) were replica-plated onto NA agar supplemented with Kanamycin (50 µg/mL) to reveal plasmid retention. All data shown are mean values from replica flasks, error bars are ±1 standard deviation.

3.4.1 Flow cytometry measurement

The mean green fluorescence displayed a comparable pattern over time as the 0.4% glucose cultures (**Figure 3.18a**) although several cultures demonstrated higher fluorescence values at 28 h. These cultures included the uninduced and 0.015% arabinose-induced GFP cultures, as well as the 0.2% arabinose-induced TNFα-GFP cultures. The highest amount of recorded green fluorescence was measured at 28 h of growth for GFP induced with 0.015% arabinose, with a value of 530,304 green fluorescence units, which was equivalent to the green fluorescence observed in non-induced GFP cultures. This might reflect to lower metabolic burden in these cultures due to the higher concentration of carbon source. Despite the fact that biomass concentrations in cultures generated with 0.8% glucose and 0.4% glycerol were comparable, it has been suggested that a greater proportion of the carbon source could have been utilised for the synthesis of recombinant protein (Retamal et al., 2018).

Histograms of green fluorescence exhibited a more notable partial induction of TNFα-GFP expression with 0.015% arabinose concentration. This was evident from the presence of a broad left shoulder on the green fluorescence peaks (**Figure 3.19**). Due to the established knowledge that glucose represses the arabinose-promoter that is responsible for regulating T7RNAp, increasing the glucose concentration resulted in repressing RP synthesis. This was expected for glucose cultures, especially those with a higher concentration. Hence, as previously mentioned, additional arabinose is needed to 'fine-tune' the process for RP expression.

The TNFα-GFP cultures that were induced with 0.2% arabinose did not exhibit such a shoulder but rather a single population, demonstrating that a higher concentration of arabinose was required to induce RPP in cultures containing 0.8% glucose. A lower number of GFP- cells was also observed compared to the cultures with 0.4% glucose (**Figure 3.18b**).

According to the results of PI staining, GFP-induced cultures had more dead cells than the other cultures (**Figure 3.18c**). The dead cell percentages in both induced GFP cultures were incredibly high at 28 h post-induction. Once again, this presumably provides an explanation for the comparatively lower recorded OD_{650} values in the GFP cultures (**Figure 3.17a**). Non-induced cultures had the least impact, with fewer than 3% of the cells being dead throughout the growth.

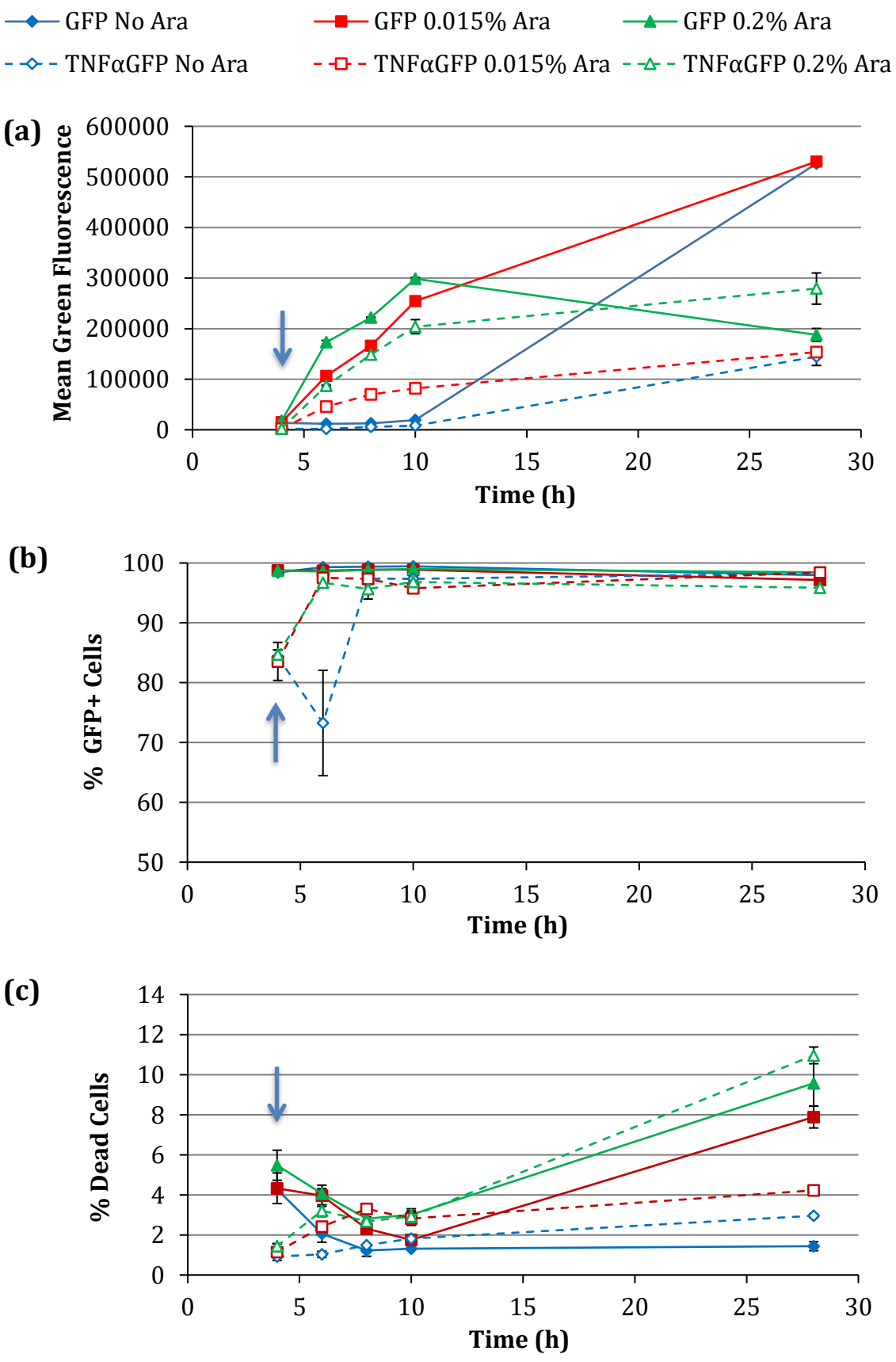

GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara
TNFαGFP No Ara
TNFαGFP 0.015% Ara
TNFαGFP 0.2% Ara
(a)
Mean Green Fluorescence
600000
500000
400000
300000
200000
100000
0
0
5
10
15
20
25
30
Time (h)
(b)
% GFP+ Cells
100
90
80
70
60
50
0
5
10
15
20
25
30
Time (h)
(c)
% Dead Cells
14
12
10
8
6
4
2
0
0
5
10
15
20
25
30
Time (h)

Figure 3.18: Mean green fluorescence measured by flow cytometry for GFP and TNFα-GFP in batch culture with 0.8% glucose at 30°C

Recombinant human TNFα fused to GFP (or also known as rhTNFα-GFP) was expressed from an arabinose-induced plasmid system with different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow), and compared to the production of GFP from the same plasmid pORT(T7). (a) FL1-A (green fluorescence) plots of cells taken from culture samples at interval time points post-induction. (b) Percentage of cells that producing GFP (GFP+ cells) versus time, and (c) Percentage of dead cells at interval 0 h to 24 h post-induction. Data are mean values of duplicate samples, error bars are ±1 standard deviation.

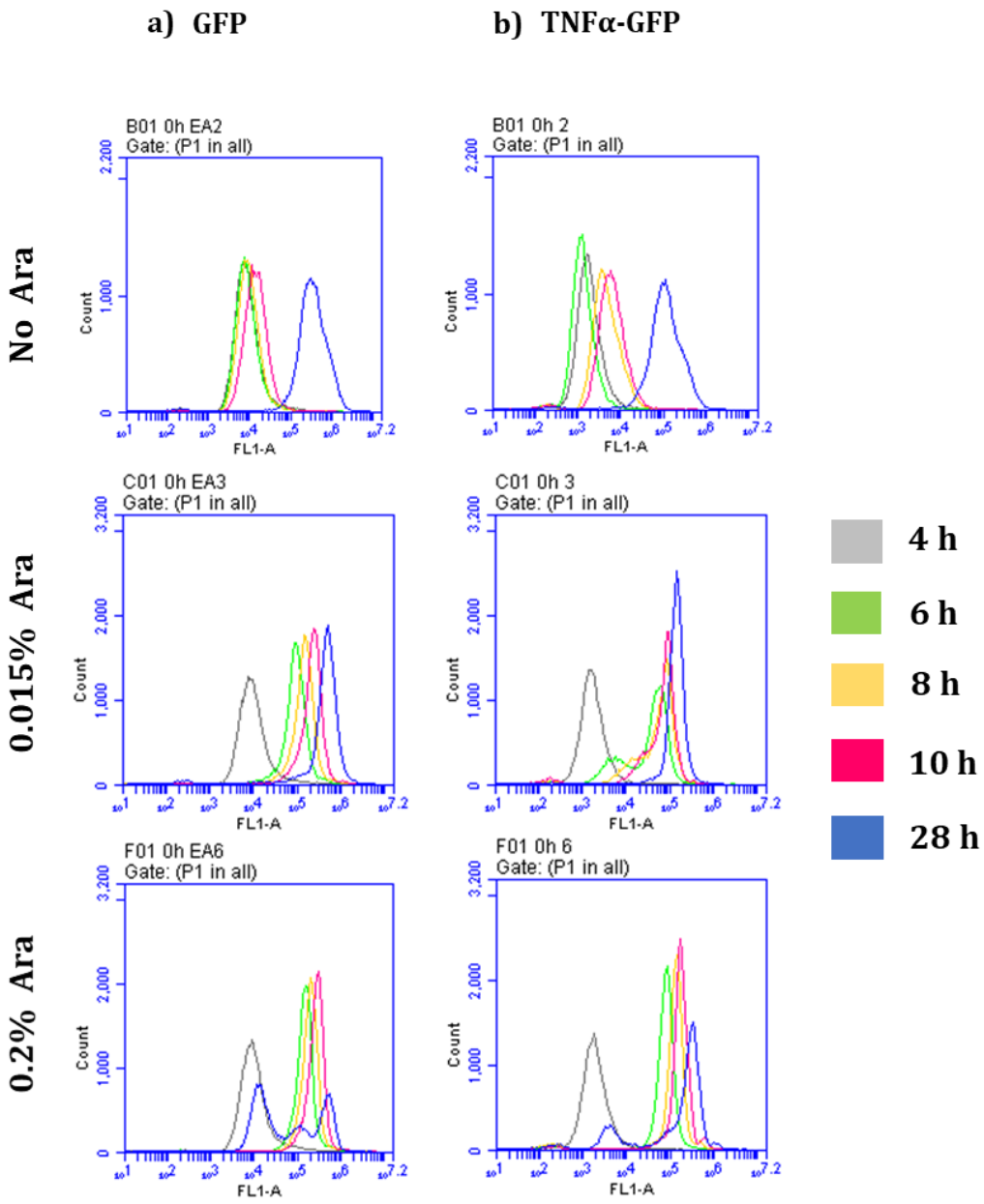


Figure 3.19: Single cell analysis of T7 expression system cultures for GFP and TNFα-GFP with 0.8% glucose at 30°C

(a) *E. coli* BL21-T7 cultures carrying p2T7-GFP (GFP) and (b) p2T7-TNFα-GFP (TNFα-GFP) were grown in Terrific Broth with 0.8% glucose as a carbon source at 30°C. At an OD_{650} of 0.5 (approximately after 4 h growth), arabinose (0.015% or 0.2%) was added. At regular intervals, samples were taken and FCM was used to measure green fluorescence of individual cells. Timepoints relate to the time the cultures were grown; 4 h refers to pre-induction with arabinose. Duplicate cultures of each plot showed a similar pattern throughout the growth.

3.4.2 SDS-PAGE and fractionation

The SDS-PAGE analysis yielded results that were consistent with those obtained from 0.4% glycerol cultures. Specifically, the amounts produced of GFP or TNFα-GFP were found to be similar when using both concentrations of 0.015% and 0.2% arabinose. The quantity of GFP was observed to increase over the time. Based on the data shown in **Figure 3.20a**, the highest amount of GFP was produced at 24 h post-induction for 0.015% arabinose-induced GFP cultures, with 0.39 mg/mL, reaching 46% of TCP (**Figure 3.21b, Appendix 3.1a**).

However, TNFα-GFP exhibited greater production compared to GFP in the presence of 0.8% glucose (**Figure 3.20b**). The highest amount of TNFα-GFP was seen as early as 4 h post-induction for cultures induced by 0.2% arabinose, reaching 41% of TCP (**Appendix 3.1b**) with a protein concentration of 0.87 mg/mL. This finding shows a contradiction to the FL1-A results, as it indicates that the mean green fluorescence seen in this particular culture was lower (**Figure 3.21a**).

Subcellular fractionation data revealed the state level of RP solubility. Figure **3.22a** demonstrates that the production of GFP in a soluble form was significantly higher in comparison to TNFα-GFP, which predominantly existed in an insoluble state (**Figure 3.22b**). The highest soluble GFP content was observed for 0.015% arabinose induction (0.96 mg/mL) at 24 h post-induction, accounting for 74% of the total GFP generated (**Appendix 3.2**). This finding provides evidence that the

mean FL1-A green fluorescence data aligns with the observation that this GFP cultures exhibit the highest levels of soluble GFP production (**Figure 3.21a, c**).

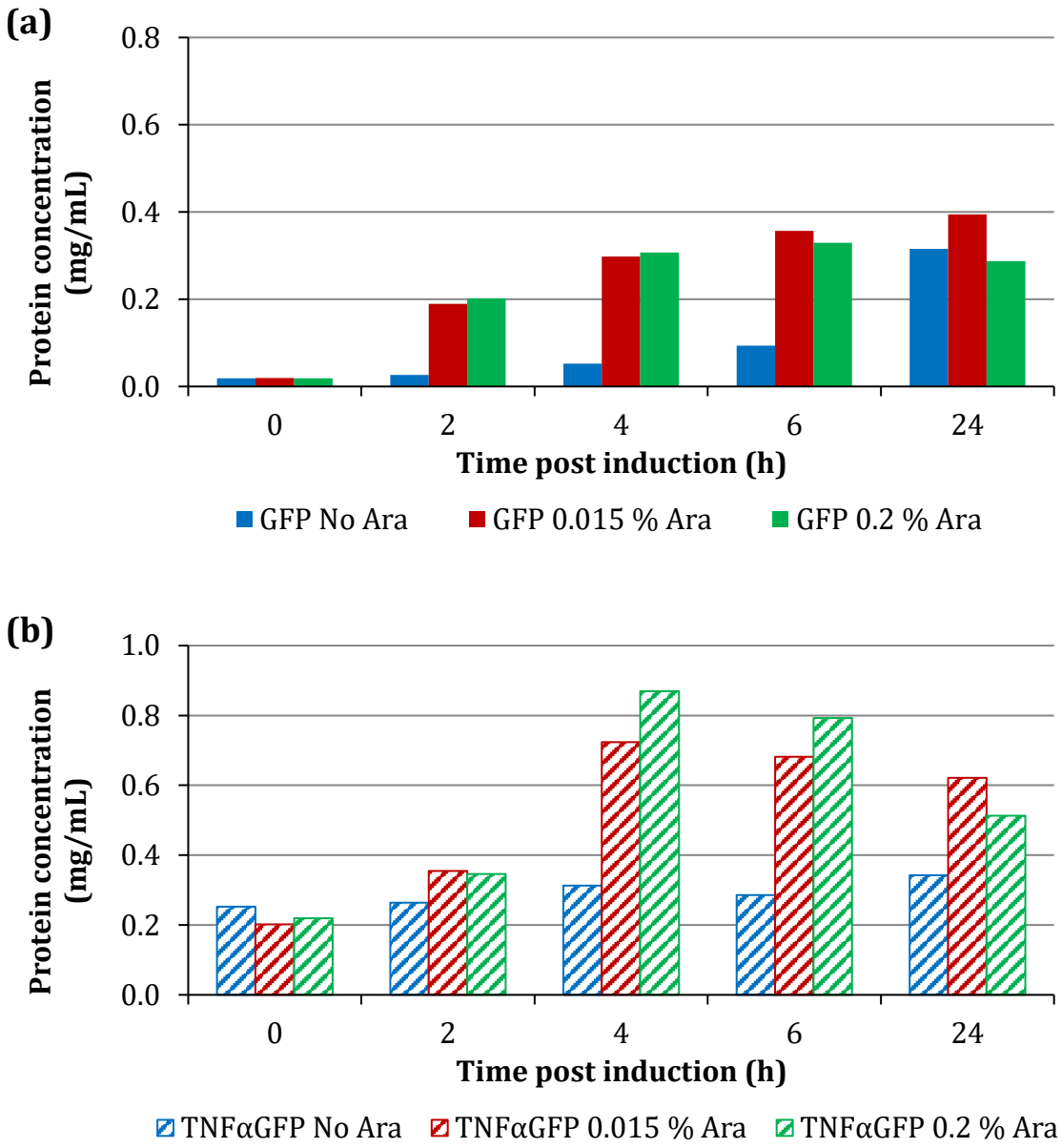
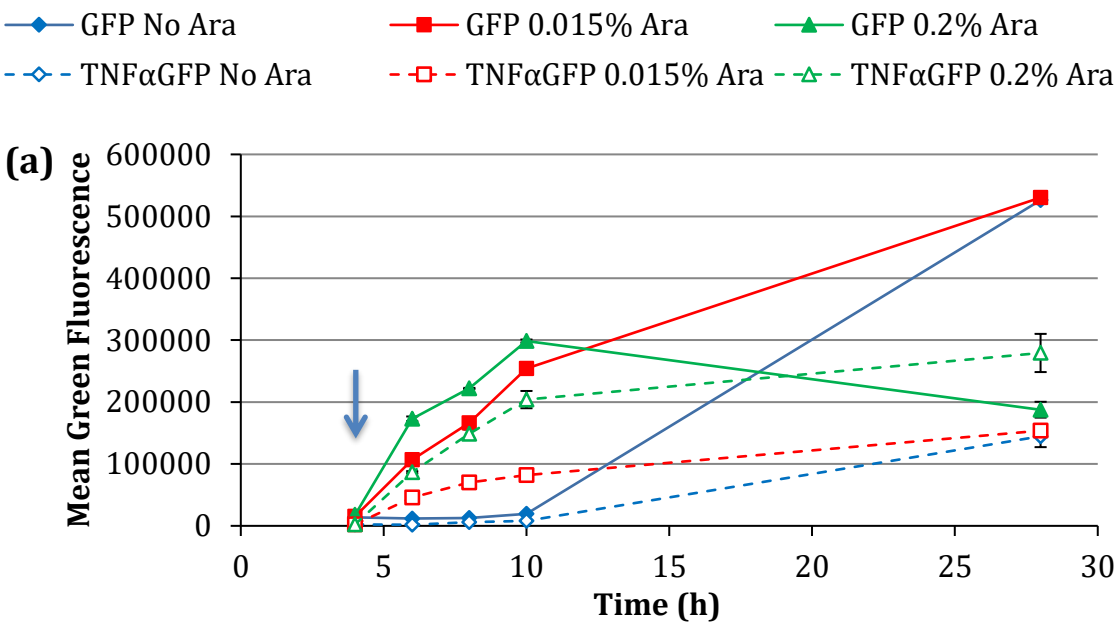
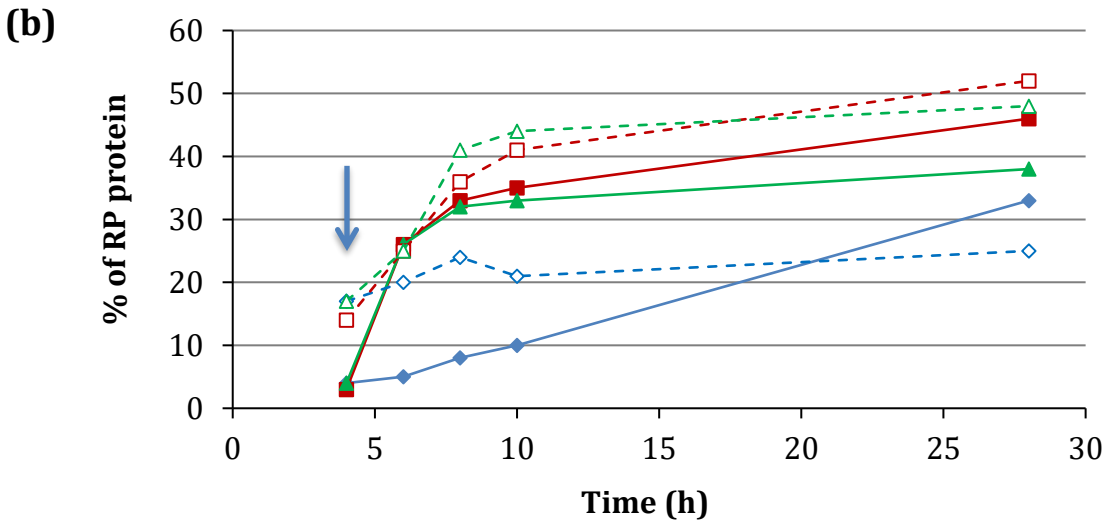


Figure 3.20: The expression of total proteins in cell extracts measured by SDS-PAGE for GFP and TNFα-GFP in batch culture with 0.8% glucose at 30°C

Predicted total protein concentration in cell extracts of (a) GFP and (b) TNFα-GFP measured in mg/mL at times post-induction.


GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara
TNFαGFP No Ara
TNFαGFP 0.015% Ara
TNFαGFP 0.2% Ara
(a)
Mean Green Fluorescence
600000
500000
400000
300000
200000
100000
0
0
5
10
15
20
25
30
Time (h)


(b)
% of RP protein
60
50
40
30
20
10
0
0
5
10
15
20
25
30
Time (h)

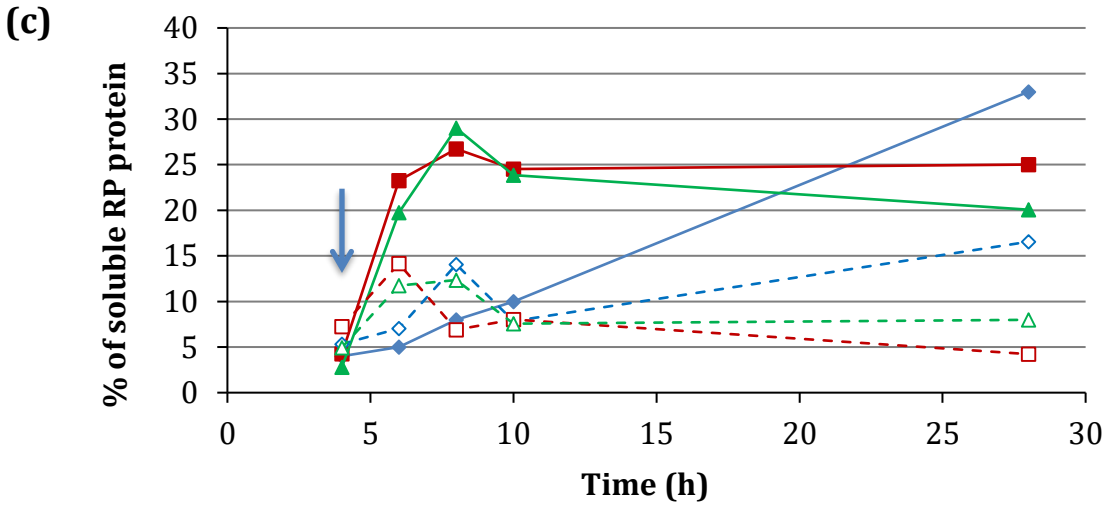

(c)
% of soluble RP protein
40
35
30
25
20
15
10
5
0
0
5
10
15
20
25
30
Time (h)

Figure 3.21: Comparison of mean green fluorescence as determined by FCM with soluble RP content for GFP and TNFα-GFP in batch culture with 0.8% glucose at 30°C

Recombinant human TNFα fused to GFP (or also known as rhTNFα-GFP) was expressed from an arabinose-induced plasmid system with different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow), and compared to the production of GFP from the same plasmid pORT(T7).

(a) FL1-A (mean green fluorescence) plots of cells taken from culture samples at interval time points post-induction. Samples were analysed by SDS-PAGE for separation, followed by subcellular fractionation to separate soluble and insoluble proteins. The percentage of cellular protein, that is, recombinant protein (b) and the percentage of cellular protein, that is, soluble recombinant protein (c) were determined. Data points represent the mean values of duplicate cultures ± standard deviation.

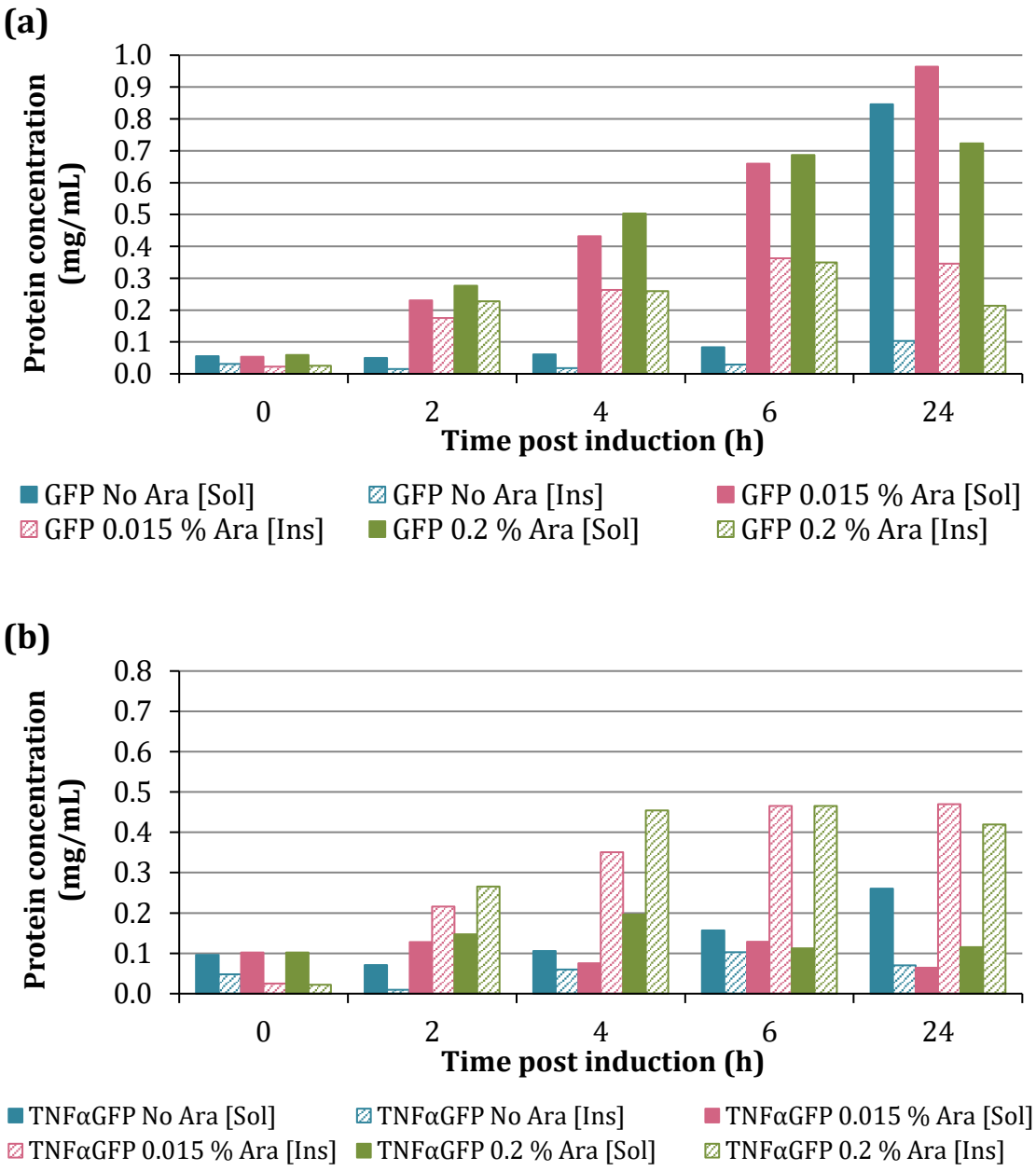


Figure 3.22: The expression of protein solubility in cell extracts measured by SDS-PAGE for GFP and TNFα-GFP in batch culture with 0.8 % glucose at 30 °C

Estimated protein solubility of soluble and insoluble fractions in cell extracts of (a) GFP and (b) TNFα-GFP at concentration of mg/mL at times post induction.

In general, our findings indicate that the utilisation of glucose as the carbon source in TB medium did not result in higher-quality TNFα-GFP. Despite the substantial synthesis of TNFα-GFP, its solubility was shown to be significantly poor. This protein was highly accumulated in insoluble form, potentially resulting from the misfolding of TNFα-GFP and subsequent aggregation as IBs (**Figure 3.21c**). This phenomenon offers a logical rationale for the significantly low levels of green fluorescence detected in the culture (**Figure 3.21a**). Again, this finding strongly suggests that the green fluorescence data obtained from FCM were broadly reflective to the solubility of the recombinant protein.

3.4.3 Discussion

According to our study findings, there is no apparent difference in cell growth when the amount of glucose used as a carbon source in the growth medium was increased from 0.4% to 0.8%. While 0.4% glycerol was preferable than 0.4% glucose, increasing the concentration of glucose in the growth medium only has detrimental impacts on the solubility of RP, hence impeding its successful execution (Lecina et al., 2013). Therefore, it is strongly recommended to culture *E. coli* in glycerol to ensure that minimal or no acetate is produced. Using glycerol as the carbon source in growth medium can reduce the rate of protein transcription and translation, reducing the metabolic burden of *E. coli* cells while producing RP, thus resulting better RPP (Shih et al., 2002).

Insufficient induction was indicated by the appearance of the left shoulder on the FL1-A histogram from FCM data for 0.4% glucose cultures with 0.015% arabinose-

induced (**Figure 3.16**). Increasing the glucose concentration only broadening the shoulder, in which its existence can be more clearly seen, indicating partial induction. As previously stated by Castiñeiras et al. (2018) that increasing the arabinose was able to tailor the protein expression that was inhibited by the presence of glucose in the T7 expression system, induction with a higher concentration of 0.2% arabinose in 0.8% glucose cultures might reduce this issue (**Figure 3.19**).

3.5 Increasing the growth temperature to 37°C

The final experiment designed for BL21(T7)-pORT(T7) in batch culture involved growing cells in 100 mL of TB medium, wherein 0.4% glycerol was present as the carbon source, supplemented with 50 μg/mL Kanamycin, and an increase in temperature from 30°C to 37°C.

In general, the growth was slower at temperature of 37°C with 0.4% glycerol as carbon source than it was at 30°C. A decrease in the final OD_{650} was seen in induced cultures than uninduced (**Figure 3.23a**). Specifically, both GFP cultures induced with 0.015% and 0.2% arabinose had lower final OD_{650} values of 11.04 and 9.12, respectively, after 28 h of growth. This was presumably due to the additional physiological stress triggered by RPP at 37°C (Wyre and Overton 2014a; Sevastsyanovich et al. 2009; Vera et al. 2007).

This is reflected by a greater decrease in CFU (**Figure 3.23b**) and an increase in dead PI^+ cells (**Figure 3.24c**) than at 30°C. The observation of CFU showed a significant reduction in all induced cultures after 4 h of induction, especially in GFP cultures, although the CFU recovered 24 h post-induction. GFP cultures with 0.2% arabinose-induced also had the least percentage of plasmid retained at 4 h post-induction (**Figure 3.23c**).

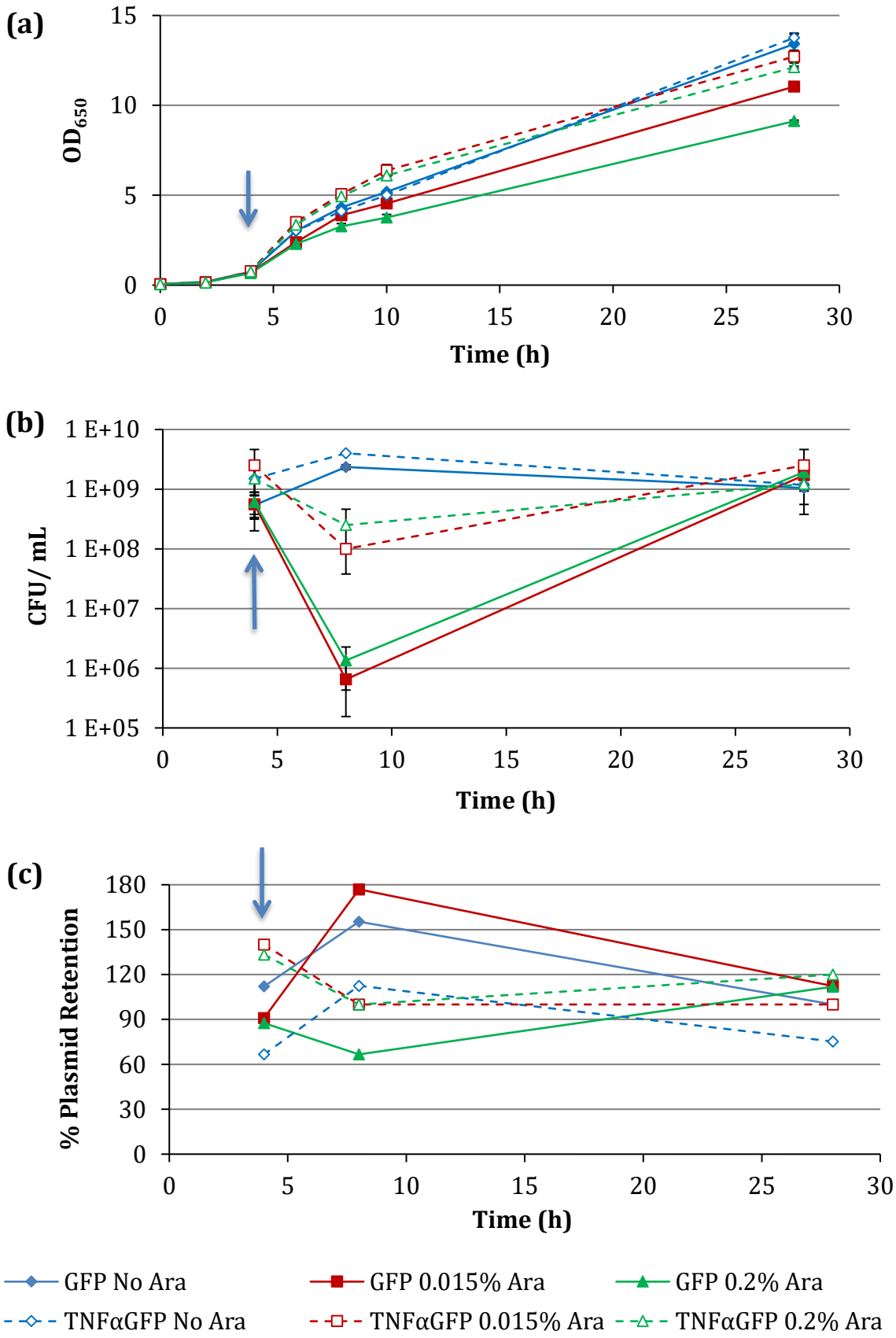

(a)
OD650
Time (h)
(b)
CFU/ mL
Time (h)
(c)
% Plasmid Retention
Time (h)
GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara
TNFαGFP No Ara
TNFαGFP 0.015% Ara
TNFαGFP 0.2% Ara

Figure 3.23: Correlation of *E. coli* BL-21(T7) growth for the study of optimisation for GFP and TNFα-GFP production, grown at 37°C with 0.4% glycerol

E. coli BL-21(T7) cultures carrying the plasmid pORT(T7)-GFP or pORT(T7)-TNFα-GFP were induced with three different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at OD_{650} ≈ 0.5 (approximately after 4 h of growth; indicated by blue arrow), evaluated by (a) OD_{650} measurement, (b) CFU/mL counting, and (c) plasmid retention. (a) The OD_{650} was measured at intervals. (b) Serial dilution of culture samples at 0 h, 4 h and 24 h post-induction was plated onto NA agar for indication of cell viability and, (c) were replica-plated onto NA agar supplemented with Kanamycin (50 µg/mL) to reveal plasmid retention. All data shown are mean values from replica flasks, error bars are ±1 standard deviation.

Non-induced cultures grown at 37°C was slightly lower in final OD_{650} measured than that at 30°C, but neither the dead cells nor the culturability were significantly affected. Both uninduced cultures of GFP and TNFα-GFP displayed higher cell densities than those induced (**Figure 3.23a**), with the final OD_{650} of nearly 14.0. These cultures also showed no reduction of cell culturability throughout the growth. This is presumably because the delayed accumulation of RP in these cultures results in lower level of stress compared to the rapid generation occurred in induced cultures.

3.5.1 Flow cytometry measurement

Green fluorescence of GFP-producing cells appears to be higher at 37°C compared to 30°C, suggesting that GFP production is enhanced at this higher temperature (**Figure 3.24a**), and both arabinose concentration showed comparable GFP production. The fluorescence of TNFα-GFP at 37°C exhibited a decrease compared to that at 30°C, regardless of whether the cultures were induced or uninduced. Significant difference in the mean green fluorescence of FL1-A was detected between the induced cultures of GFP and TNFα-GFP. The green fluorescence measured was the highest for GFP induced cultures at 6 h post-induction, with roughly 620,000 green fluorescence units, and was similar for both arabinose-induced cultures. In contrast, green fluorescence production in TNFα-GFP induced cultures was about four times lower than in GFP induced cultures at the same timepoint, with approximately 160, 000 green fluorescence units being recorded. Non-induced GFP cultures demonstrated much higher fluorescence values than all

other cultures at 28 h post-induction, with nearly 500, 000 green fluorescence units.

Observation on single cells analysis of cultures at 37°C revealed a closely similar histogram pattern to cultures grown at 30°C (**Figure 3.25**). Following induction, all induced cultures had a single GFP^+ population, and after 28 h of growth, these cultures displayed the emergence of multiple GFP^+ populations. Furthermore, there were barely any GFP^- cells visible in any of the cultures. According to FL1-A histogram data, all induced and non-induced cultures of GFP and TNFα-GFP exhibited a cell population expressing GFP (GFP^+ cells) above 95% throughout the growth (**Figure 3.24b**).

PI staining revealed that the percentage of dead cells appears to be higher at 37°C compared to 30°C cultures (**Figure 3.24c**). Based on the results from PI staining, all induced cultures had more dead cells than uninduced cultures, with 0.2% arabinose-induced GFP cultures having up to 14% dead cells at 6 h post-induction. This might suggest that the cultures were forced to produce GFP/TNFα-GFP following induction, thus increased cellular pressure and eventually died after producing GFP/ TNFα-GFP. This provides an explanation for the lower measured OD_{650} values in induced cultures at 37°C than those at 30°C (**Figure 3.23a**). Non-induced cultures had minimal effects, as less than 3% of the cells were dead throughout the growth.

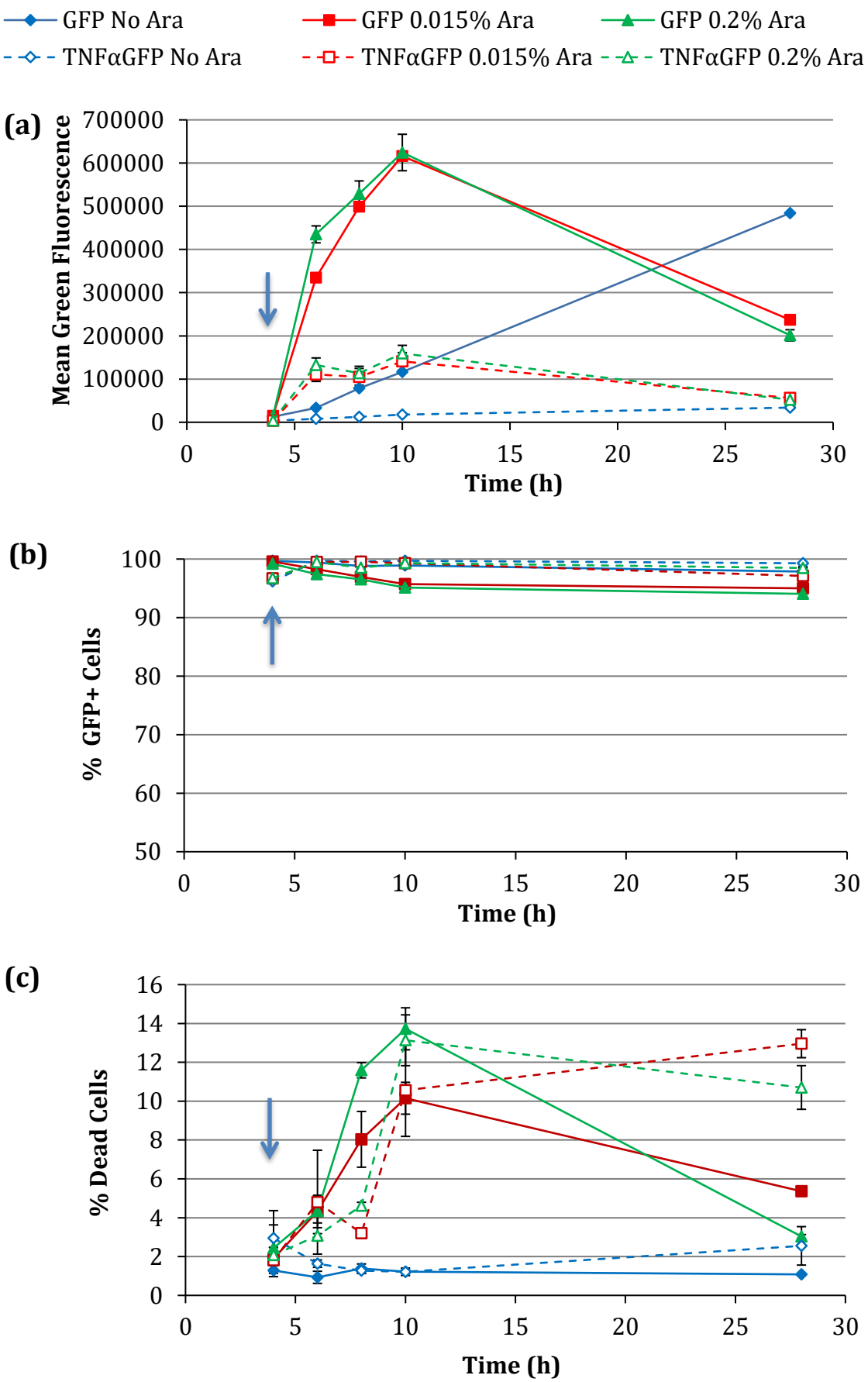

GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara
TNFαGFP No Ara
TNFαGFP 0.015% Ara
TNFαGFP 0.2% Ara
(a)
Mean Green Fluorescence
700000
600000
500000
400000
300000
200000
100000
0
0
5
10
15
20
25
30
Time (h)
(b)
% GFP+ Cells
100
90
80
70
60
50
0
5
10
15
20
25
30
Time (h)
(c)
% Dead Cells
16
14
12
10
8
6
4
2
0
0
5
10
15
20
25
30
Time (h)

Figure 3.24: Mean green fluorescence measured by flow cytometry for GFP and TNFα-GFP in batch culture with 0.4% glycerol at 37°C

Recombinant human TNFα fused to GFP (or also known as rhTNFα-GFP) was expressed from an arabinose-induced plasmid system with different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow), and compared to the production of GFP from the same plasmid pORT(T7). (a) FL1-A (green fluorescence) plots of cells taken from culture samples at interval time points post-induction. (b) Percentage of cells that producing GFP (GFP+ cells) versus time, and (c) Percentage of dead cells at interval 0 h to 24 h post-induction. Data are mean values of duplicate samples, error bars are ±1 standard deviation.

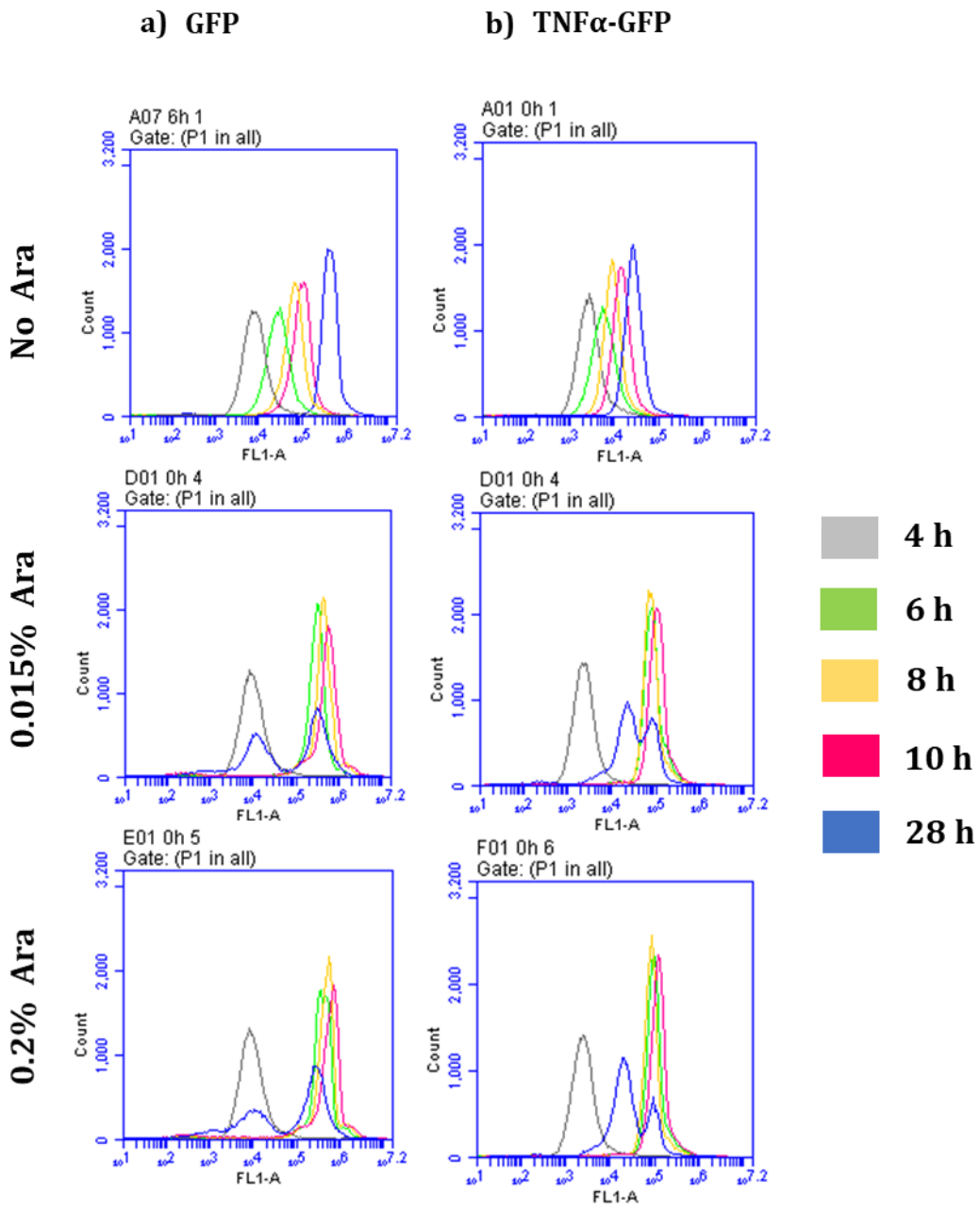


Figure 3.25: Single cell analysis of T7 expression system cultures for GFP and TNFα-GFP with 0.4% glycerol at 37°C

(a) *E. coli* BL21-T7 cultures carrying p2T7-GFP (GFP) and (b) p2T7-TNFα-GFP (TNFα-GFP) were grown in Terrific Broth with 0.4% glycerol as a carbon source at 37°C. At an OD_{650} of 0.5 (approximately after 4 h growth), arabinose (0.015% or 0.2%) was added. At regular intervals, samples were taken and FCM was used to measure green fluorescence of individual cells. Timepoints relate to the time the cultures were grown; 4 h refers to pre-induction with arabinose. Duplicate cultures of each plot showed a similar pattern throughout the growth.

3.5.2 SDS-PAGE and fractionation

SDS-PAGE analysis demonstrated that RP production was comparable for all induced cultures (**Figure 3.27b**), and the amount of RP produced at 37°C cultures was quite similar to that at 30°C cultures (**Figure 3.26a, 3.26b**). However, the solubility was different.

Subcellular fractionation revealed that around 30% of GFP made in induced cultures was soluble 2 h after induction, and that proportion rose to approximately 50% over time (**Figure 3.27c, 3.28a**). By contrast, almost 50% of GFP was soluble starting 2 h after induction in cultures at 30°C (**Figure 3.10c**), suggesting better folding at the lower growth temperature.

Even though TNFα-GFP production was comparable with the production of GFP, but their solubility was very low (**Figure 3.27c, 3.28b**), reflected by the low green fluorescence recorded for these cultures by FCM (**Figure 3.27a**). Non-induced GFP cultures showed significantly high GFP production at 28 h post-induction, and also a higher proportion of soluble GFP produced compared to other cultures at the same timepoint, again reflected by green fluorescence measured in FCM.

Densitometry data for the percentage of cellular protein that was GFP/ TNFα-GFP was also shown on the SDS-PAGE gel pictures in **Appendix 3.4**, while the percentage ratio of soluble versus insoluble GFP/ TNFα-GFP accumulated inside the cell extracts were also shown on the SDS-PAGE gel pictures in **Appendix 3.5** and **Appendix 3.6**.

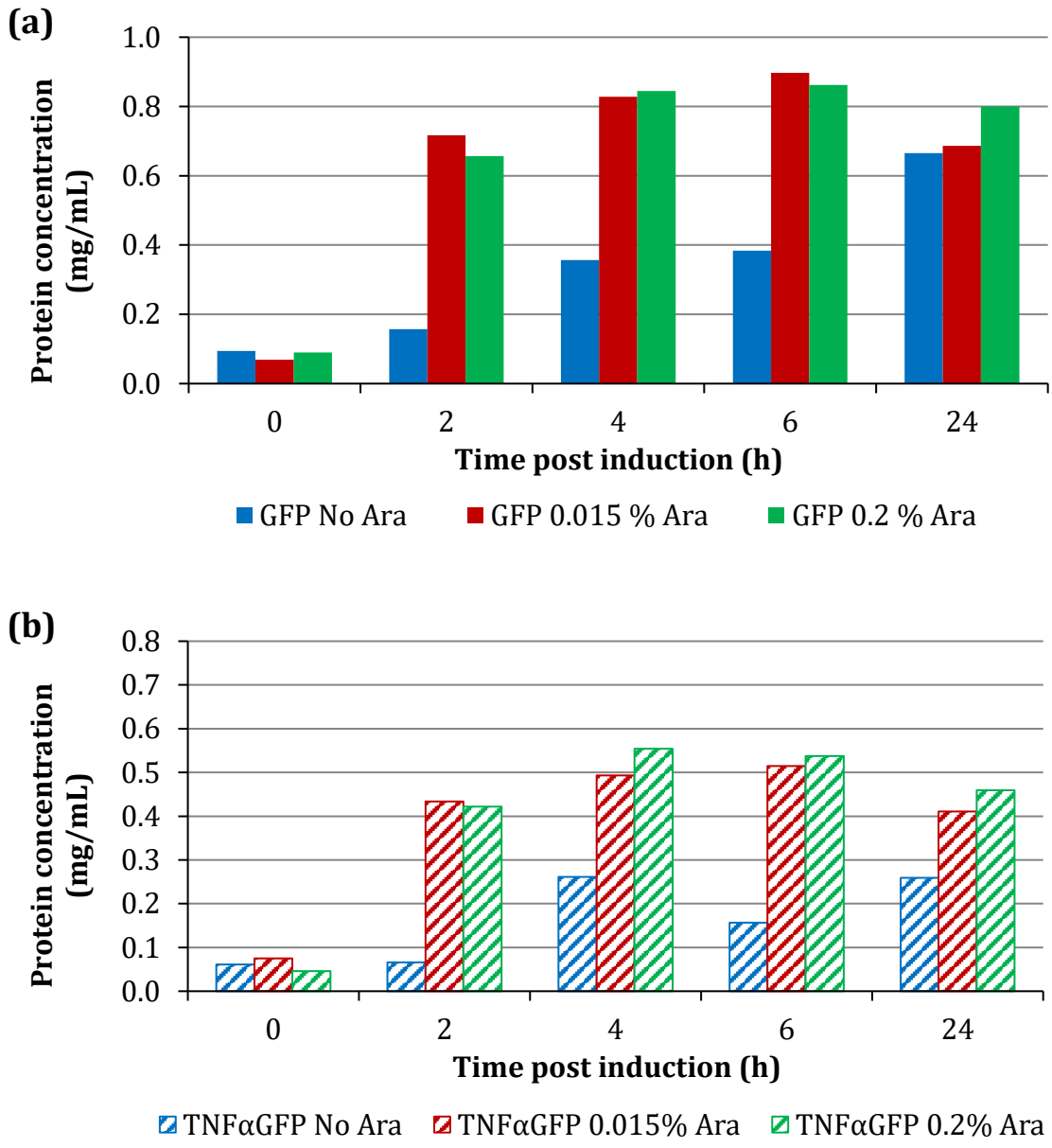


Figure 3.26: The expression of total proteins in cell extracts measured by SDS-PAGE for GFP and TNFα-GFP in batch culture with 0.4% glycerol at 37°C

Predicted total protein concentration in cell extracts of (a) GFP and (b) TNFα-GFP measured in mg/mL at times post-induction.

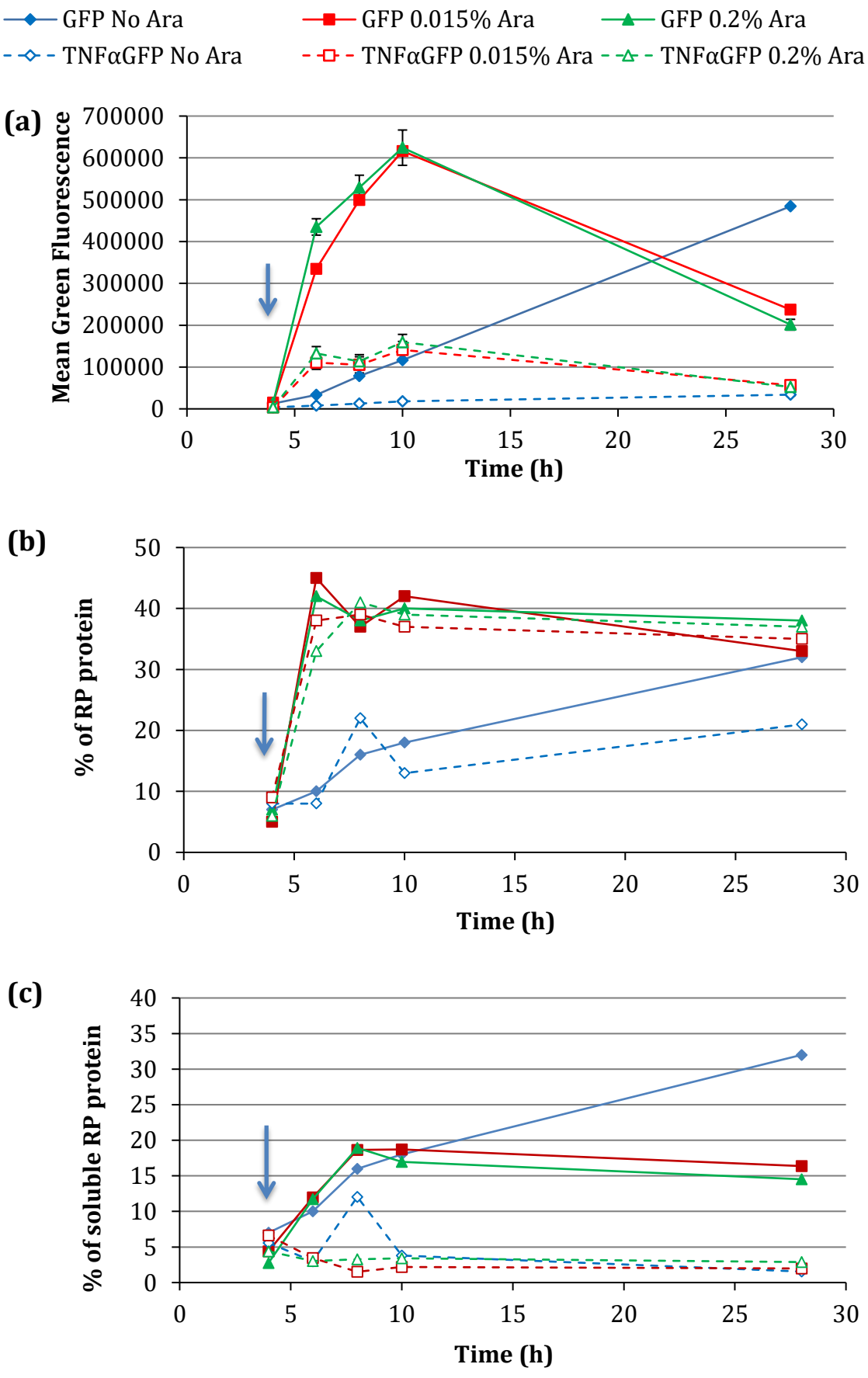

GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara
TNFαGFP No Ara
TNFαGFP 0.015% Ara
TNFαGFP 0.2% Ara
(a)
Mean Green Fluorescence
700000
600000
500000
400000
300000
200000
100000
0
0
5
10
15
20
25
30
Time (h)
(b)
% of RP protein
50
40
30
20
10
0
0
5
10
15
20
25
30
Time (h)
(c)
% of soluble RP protein
40
35
30
25
20
15
10
5
0
0
5
10
15
20
25
30
Time (h)

Figure 3.27: Comparison of mean green fluorescence as determined by FCM with soluble RP content for GFP and TNFα-GFP in batch culture with 0.4% glycerol at 37°C

Recombinant human TNFα fused to GFP (or also known as rhTNFα-GFP) was expressed from an arabinose-induced plasmid system with different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow), and compared to the production of GFP from the same plasmid pORT(T7).

(a) FL1-A (mean green fluorescence) plots of cells taken from culture samples at interval time points post-induction. Samples were analysed by SDS-PAGE for separation, followed by subcellular fractionation to separate soluble and insoluble proteins. The percentage of cellular protein, that is, recombinant protein (b) and the percentage of cellular protein, that is, soluble recombinant protein (c) were determined. Data points represent the mean values of duplicate cultures ± standard deviation.

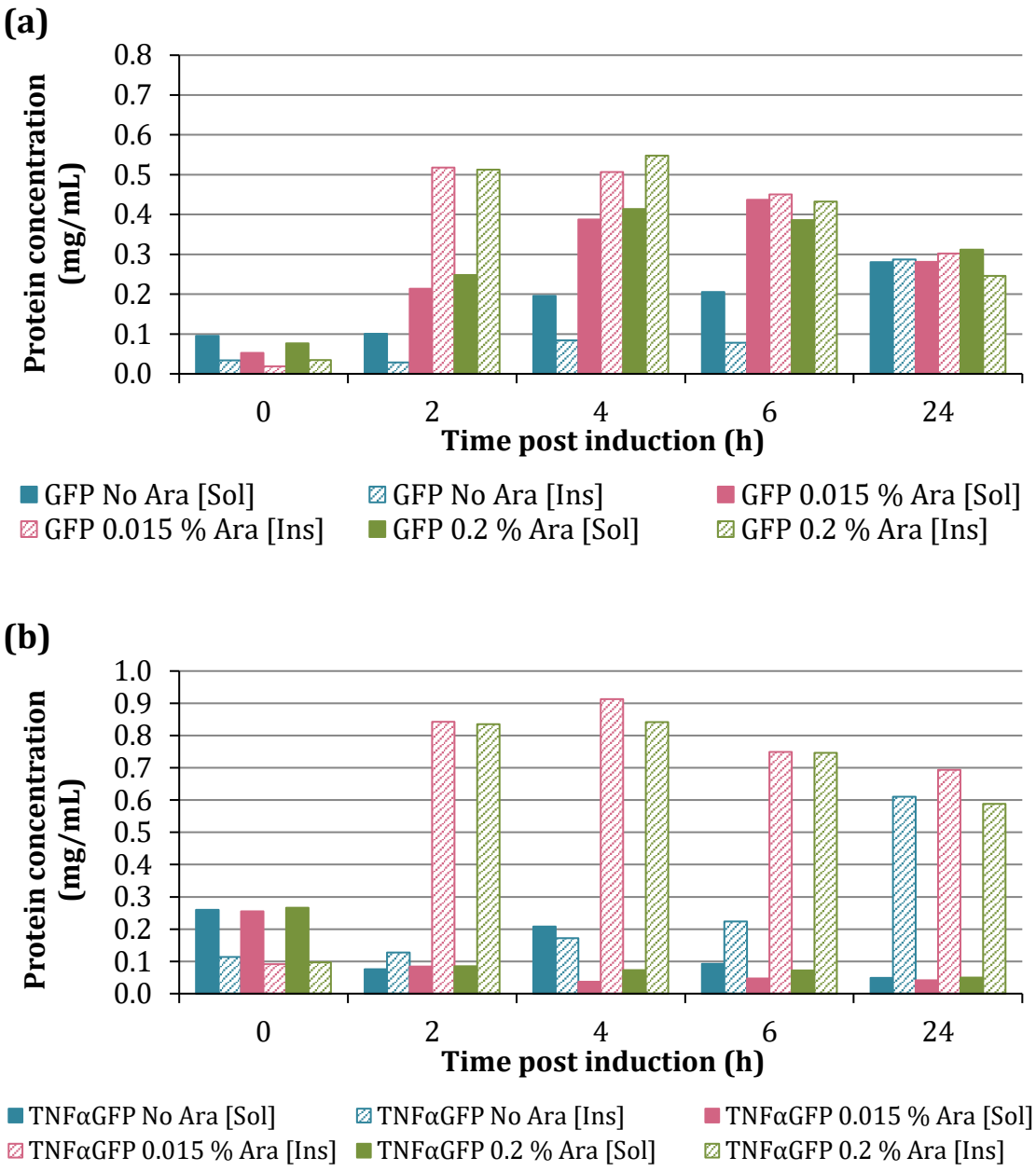


Figure 3.28: The expression of protein solubility in cell extracts measured by SDS-PAGE for GFP and TNFα-GFP in batch culture with 0.4 % glycerol at 37 °C

Estimated protein solubility of soluble and insoluble fractions in cell extracts of (a) GFP and (b) TNFα-GFP at concentration of mg/mL at times post induction.

3.5.3 Discussion

Two different growth temperatures were evaluated for the production of GFP/ TNFα-GFP, with the lower temperature yielded the highest levels of soluble RP. An increase in solubility of GFP was observed upon reduction of the growing temperature. At the temperature of 30°C, almost 50% of GFP was soluble after 2 h of induction (**Figure 3.10c**). In contrast, at the temperature of 37°C, only 30% of the GFP produced in induced cultures was soluble at 2 h post-induction, and that proportion rose to approximately 50% over time (**Figure 3.27c**). These findings suggest that the folding of GFP is enhanced at the lower growth temperature.

Reducing the growth temperature and inducer concentration is often the most obvious approach to slow down the rate of RP synthesis in order to prevent the accumulation of unfolded proteins, favouring the accumulation of soluble and correctly folded protein (Sevastsyanovich et al., 2009b). When a low growth temperature (30°C) and a low concentration of inducer (0.015% arabinose) were used, both GFP and TNFα-GFP exhibited enhanced solubility. Increasing concentrations of arabinose in 30°C cultures reduced GFP solubility. Yet, the solubility of GFP at 37°C was not significantly affected by the differences in arabinose concentration.

In contrast, as indicated by SDS-PAGE analysis, TNFα-GFP solubility was significantly decreased at 37°C than it was at 30°C. As expected, the total TNFα-GFP produced under higher growth temperature was predominantly insoluble (**Figure 3.27c**). These findings, in addition to the observed reduction in

fluorescence as measured by FCM, provide more evidence in support of previous research that have demonstrated that TNFα solubility was significantly higher in *E. coli* at 30°C than 37°C (Castiñeiras et al., 2018b). Glycerol-based cultures at 37°C often grow faster than at 30°C. This promotes fast folding of newly synthesized polypeptides, increases stress and leads to misfolding. Consequently, this may cause a corresponding reduction in the amount of functional proteins generated, thereby lowering the overall yield of soluble RP (Castiñeiras et al., 2018b). The occurrence of folding errors can potentially result in the formation of IBs and also a restriction in the synthesis of active enzymes (Bhatwa et al., 2021; Su et al., 2019).

The generation of TNFα-GFP in *E. coli* found to be challenging in comparison to GFP, particularly when growing at a higher temperature of 37°C. This difficulty arose due to the burden of RPP on the bacteria, thus increased their physiological stress. This resulted in more dead cells (**Figure 3.24c**) and a low solubility of final TNFα-GFP yielded. According to existing literature, it has been shown that the optimal temperature for RPP is 30°C. This is based on the understanding that lower temperatures can impede the rate of transcription and decrease the rate of protein translation. In this way, RP is translated slower, thus allowing more effective folding (Wyre and Overton, 2014b; Alfasi et al., 2011). It is essential that we minimise cellular stress levels during the production of RP in order to achieve a greater quantity of soluble and functional TNFα-GFP. The utilisation of stress minimization approaches has been extensively employed to enhance the synthesis of cytoplasmic model proteins (Sevastsyanovich et al., 2009b).

Yet, it is important to note that GFP retained its ability to undergo proper folding and exhibit fluorescence at 37°C. Moreover, it was observed that cells grown at this higher temperature exhibited an increased amount of fluorescence compared to those produced at 30°C (**Figure 3.24a**). However, even though cultures at 37°C increased green fluorescence for GFP, but they only exhibited low solubility, making the growth temperature of 30°C preferable than 37°C for improved folding and soluble RP. Upon considering both GFP and TNFα-GFP, it can be determined that GFP demonstrates a greater degree of tolerance towards production at 37°C**.** Utilising a higher temperature for RPP did not seem to be a more favourable alternative for TNFα-GFP as it generated RP that were incredibly insoluble.

3.6 General discussion

Optimising the conditions for RPP is often challenging due to the fact that *E. coli* has not been naturally tuned as a protein factory for the pharmaceutical industry (Hoffmann et al., 2004). Indeed, despite over three decades of industrial biopharmaceutical production, no existing method has proven universally effective in the manufacturing of all protein products. The host strain, the expression vector, the medium composition (including the carbon source used), the induction strategy and antibiotic resistance, as well as cultivation temperature, are only a few of the various variables that are frequently assessed (Sevastsyanovich et al., 2010).

The physiological balance of microbial host is often disturbed by growth conditions, gene expression and plasmid retention, leading to a heavy metabolic burden on the microbial host. According to Rahmen et al. (2015), the term "metabolic burden" refers to the draining of essential materials, nutrients as well as energy from the physiological microbial metabolism in order to facilitate the expression of foreign DNA or proteins. Numerous studies have addressed the primary obstacles encountered in RPP and potential influences on the metabolic burden. Additionally, such reviews have discussed alternative strategies for reducing the metabolic burden and optimising protein synthesis (Kasli et al., 2019; Castiñeiras et al., 2018; Hsu et al., 2016; Wyre and Overton, 2014b). Cells might encounter varying degrees of toxicity and metabolic stress when expressing recombinant products, potentially resulting in a decline in biomass production, cellular viability, and productivity. Hence, it is crucial to measure the metabolic

burden caused by the RPP and its significant influence on the overall metabolism of *E. coli* while optimising recombinant bioprocesses (Song et al., 2013).

Ukkonen et al. (2013a) posit that the amounts of the RPP and the metabolic state of the cells will both influence the magnitude of the metabolic burden. As such, it is necessary to evaluate various conditions through a trial-and-error procedure with the aim to identify the ideal growing parameters. This approach strives to reduce metabolic burden and cellular stress, thus encouraging the successful RPP. The primary stages of the experiments conducted in this study aimed to optimise the parameters for the synthesis of GFP and TNFα-GFP utilising the T7 expression system. Several crucial variables were taken into consideration, including the selection of carbon source (glycerol or glucose), growth temperature (30°C or 37°C) and induction concentration (0.015% arabinose or a higher concentration of 0.2% arabinose).

The T7 expression system demonstrates an impressive ability for efficient and rapid production of RPP, exhibiting a high degree of solubility (Bhatwa et al., 2021). However, it is important to note that the T7 expression system is limited by the issue of promoter leakiness. This clarifies the rationale behind the observation of the emergence of comparable amounts of green fluorescence in uninduced cultures at the end of fermentation (**Figures 3.2a, 3.15a, 3.18a, 3.24a**). In addition to the above observation, it was found that most cultures that were non-productive (GFP- cells) and exhibited minimal fluorescence experienced growth inhibition, decreased ability to be cultured and loss of plasmid. This highlights the significance of achieving ideal environmental conditions and regulating gene

expression to lessen undesirable metabolic burden on microbial host, which could negatively impact them and hinder RPP.

An excessive production of RP may cause a detrimental impact on cell division, resulting in a decreased ability to form colonies on agar plates. This phenomenon, also known as "viable but non-culturable" (or VBNC), is characterised by cells that retain metabolic activity and continue to produce RP, however, are unable to undergo division when cultivated on agar plates (Wyre and Overton, 2014a). The presence of VBNC bacteria exist within most microbial cell populations (Khan et al., 2010) and is frequently observed in RPP cultures (Sundström et al., 2004; Nebe-Von-Caron et al., 2000).

Expression instability might occur as a consequence of the leaky expression driven by T7 RNA polymerase. Inducible configurations of T7 RNA polymerase are advantageous for the expression of a diverse range of RPP. Nevertheless, the polymerase exhibits such high activity that even a minimal basal level within the uninduced cell can prevent the establishment or maintenance of relatively harmful target genes. Furthermore, it should be noted that even in cases when the target gene is harmless, the increased level of transcriptional activity of the target gene reaches a threshold that ultimately stop the cells from growing and dividing (Studier, 1991).

Following the induction for GFP and TNFα-GFP production in the T7 expression system, a drastic drop in the CFU associated with the increase in biomass and green fluorescence produced was frequently observed (**Figures 3.1b, 3.23b**). The fact that FCM does not rely on microbial growth for cell viability measurement is

an essential benefit. The limitations of growth-based methods in accurately determining total viable cell counts, particularly in detecting the VBNC phenotype and hence underestimating total viable cell concentrations, are not relevant to FCM. Thus, the findings obtained from FCM can be considered more reliable (Fernández-Castané et al., 2017).

Favourable carbon source for GFP/ TNFα-GFP production in T7 expression system

The selection of medium composition possesses an ability to exert an influence on RPP through modulating growth rates and plasmid content (Hoffmann et al., 2004). In general, the utilisation of glycerol as a carbon source in the growth medium for the synthesis of GFP and TNFα-GFP demonstrates advantages over glucose due to its ability to produce rapid and soluble RPP, while avoiding the formation of acetate, a common occurrence in glucose-based cultures. Acetate has been proven to have negative effects on protein synthesis and cell growth. However, by using alternate carbon sources in place of glucose, this issue can be reduced (Luo et al., 2006). Thus, glycerol has been chosen as an alternate source of carbon to lower metabolic pressure during expression in *E. coli* since *E. coli* exhibits limited acetate production during growth on glycerol (Zhang et al., 2010).

Induction of arabinose for GFP/ TNFα-GFP production in T7 expression system

The utilisation of arabinose induction was of crucial significance in facilitating the efficient induction of RPP by stimulating the lac-derived promoter systems of T7 (Jia and Jeon, 2016). This induction mechanism operated by means of the binding

between arabinose and the Lac repressor protein. It was reported that arabinose induction increased RP solubility, decreases the number of inclusion bodies and had a good effect on RP activity (Narayanan, 2009).

However, careful monitoring is required due to the potential for cellular stress in the microbial host as a result of the high pressure to make RPP during induction. Castiñeiras (2017) stated that, after the diversion of metabolic resources from biosynthesis, the force that generated RPP upon induction was the second major factor of cell stress in microbial host. When RPP is induced, excessive cell stress not only inhibits growth and reduces the viability of cells, but it also causes a build-up of misfolded or unfolded proteins. This build-up occurs due to the inability of the cellular machinery to keep up with the rapid rate of protein synthesis (Sevastsyanovich et al., 2010). The findings of our study align with the statement, that the introduction of arabinose to the cultures for GFP and TNFα-GFP production in T7 expression system resulted in sharp decline in CFU, illustrating the substantial stress experienced by the cells, leading to a reduction in their ability to form colonies (**Figures 3.14b, 3.17b**), and most of them, especially TNFα-GFP cultures yielded insoluble RP.

Insufficient induction was indicated by the appearance of the left shoulder on the FL1-A histogram from FCM data for 0.4% glucose cultures with 0.015% arabinose-induced (**Figure 3.16**). The presence of this shoulder was not exhibited in 0.4% glycerol cultures. Increasing the glucose concentration only broadening the shoulder, and its existence can be more clearly seen, indicating partial induction.

Induction with a higher concentration of 0.2% arabinose in 0.8% glucose cultures can reduce this problem (**Figure 3.19**).

Introducing induction at an early stage imposes an increased metabolic burden on cells, leading to a decrease in protein production and yield. Conversely, delaying induction can limit the efficacy of the induction process and subsequent protein production. (Su et al., 2019). The optimal induction point, that often recorded in lab-scale growth, is during the mid-logarithmic phase (Kasli et al., 2019; Su et al., 2019).

There are situations where the best results can be achieved without the need for induction, with minimal leaky expression for as when using the T7 expression system, to accommodate soluble RPP. However, scaling up this process may present occasional challenges due to variations in medium formulations (Lebendiker and Danieli, 2014).

FCM and population heterogeneities in T7 expression system

FCM was a great choice for the development of the analytical methods used in this study. FCM is highly desirable due to its rapid analysis, minimal sample volume requirements and ability to provide data on cellular parameters without requiring further cell growth, thus offering a "snapshot" of the physiological state of the cells at that particular instant (Fernández-Castané et al., 2017).

Previous studies have proven the significant benefits of utilising FCM for the rapid analysis of microbial physiology (Geng et al., 2014) and the assessment of auto-

fluorescent protein expression (Lagendijk et al., 2010). Furthermore, the utilisation of FCM has been extensively employed to monitor RPP (Wyre and Overton, 2014a; Sevastsyanovich et al., 2009b) and investigate population heterogeneities within cultures (Fernández-Castané et al., 2017). This study employed the use of FCM to quantify the RPP in fermentation cultures of pORT(T7)-GFP and pORT(T7)-TNFα-GFP at the single cell level. In addition to measuring the green fluorescence produced, the FCM was capable of visualising the relative size and granularity (internal complexity) of the cells.

FCM allows determination of fluorescence properties of individual cells, thereby facilitating the identification of population heterogeneities in GFP/TNFα-GFP cultures during fermentation. This was accomplished by reflecting subpopulations that exhibit different levels of RP accumulation and folding states. These subpopulations include a highly productive GFP-producing population, an intermediate GFP-producing population, a non-productive GFP-producing (GFP- cells) population, as well as a population of plasmid-free cells (**Figure 3.6**). The FL1-A histogram peak generated from the FCM analysis was utilised to determine the presence of either a single population or multiple populations within the cultures (**Figure 3.4**), Additionally, it was used to evaluate the sufficiency of the induction supplied to the cultures (**Figure 3.19**). The quantification of dead cells in the cultures was also achieved through PI staining (**Figure 3.7**).

The expression of TNFα-GFP in *E. coli* was monitored using GFP that had been fused to the C-terminus of TNFα genes. This fluorescent GFP served as a marker for quantifying the efficiency of TNFα synthesis. This finding demonstrated that GFP

indicated TNFα and had a correlation to RP's proper folding. It is widely acknowledged that GFP has been utilised as a library-screening tool in the optimisation of RPP, and that the fusion of FP is extremely useful in observing and assessing protein folding (Vizcaino-Caston et al., 2012; Alfasi et al., 2011).

FCM outperformed SDS-PAGE by a wide margin

In many occasions, despite observing considerably lower levels of green fluorescence recorded in FCM for TNFα-GFP cultures, it was frequently turned up that the data from SDS-PAGE showed comparable amounts of RP generated between GFP and TNFα-GFP cultures. Up until subcellular fractionation revealed that the level of RP solubility produced in the two cultures were clearly different. The solubility of RP in TNFα-GFP cultures was significantly lower than in GFP cultures, consistent with the findings recorded by FCM. Again, this finding strongly suggests that the green fluorescence data obtained from FCM were broadly reflective to the solubility of the RP (**Figures 3.10, 3.21**).

Extensive comparability was established between the quantification of soluble GFP/TNFα-GFP via SDS-PAGE and the determination of mean green fluorescence via FCM. Utilising FCM offers several important advantages, including a fairly simple procedure and substantially more rapid data generation, typically within 5–10 min on average. In contrast to SDS-PAGE, which requires several hours to complete and involves numerous complicated procedures. Therefore, FCM can serve as a real-time tool for in-process analysis, permitting the process of dynamically modification in response to the outcomes that have been generated.

3.7 Conclusions and future work

The generation of RP keeps being a challenge, demanding different strategies for ensuring the production of different RP. The strategy approached for the production of GFP/TNFα-GFP involved the utilisation of glycerol or glucose as carbon source, different concentrations of arabinose to induce RPP and low growth temperature to slow down protein synthesis rate, thus preventing protein misfolding. In general, the optimisation of the conditions for GFP/TNFα-GFP production in T7 expression system was successful. Specifically, the use of glycerol instead of glucose, a lower temperature of 30°C instead of 37°C and a lower concentration of arabinose (0.015%) were shown to be sufficient to boost RPP.

In conclusion, the T7 expression system appeared to generate rapid RPP; the highest fluorescence values, which correspond to high amounts of RPP, were achieved 4-6 hours following induction in the presence of glycerol, whereas growth in the presence of glucose appeared to extend the duration of productivity. This goal was successfully achieved as the T7 expression system itself was a construct of an expression vector with a strong promoter (Terpe, 2006).

While uninduced cultures had equal mean green fluorescence to induced cultures at the end of the fermentation, indicating the T7 expression system is also far leakier. A crucial consideration when selecting bacterial expression vectors is the ability to regulate expression levels. Since the T7 expression system exhibits leaky expression, these promoters may potentially be problematic for protein expression, especially in the higher cell densities fermentation. Hence, it is

essential to evaluate alternative expression systems featuring stricter regulatory features in order to determine their effectiveness in producing GFP as well as TNFα-GFP, that generally referred to as "difficult-to-express" proteins in *E. coli*.

Further investigation was conducted on an alternative strategy utilising the paraBAD expression system for RPP. This paraBAD system is recognised for its tighter regulation of promoter, and akin to T7 system, it is also arabinose-inducible promoter. The purpose of this investigation was to evaluate the potential advantages of the paraBAD expression system, specifically its tightly regulated promoter, in enhancing the performance of the RPP. Additionally, the study aimed to analyse the effects of utilising this system on cellular physiology and viability.

CHAPTER 4

BATCH CULTURE OF BL21(A)-p*ara*(*BAD*)-GFP

The second part of the project's objective was to investigate the strength of the promoter and culture heterogeneity in terms of cellular physiology and protein expression during the generation of RPP using a system based on the *E. coli* arabinose promoter. This p*araBAD* system was sourced from Cobra Biologics (Castiñeiras et al., 2018a) and is comparable to the commercial pBAD system (Guzman et al., 1995; Invitrogen, 2010).

As well as using a T7 polymerase-based expression system, this protocol has been developed to investigate the potential of p*araBAD* expression system in generating GFP. The *araBAD* promoter is reported to be much more tightly regulated (Guzman et al., 1995), so there is much less leakiness than the T7 expression system. The production of GFP in p*araBAD* expression system will be compared by few analyses including FCM for comparison at the single cell level.

4.1 Introduction

The production of GFP was studied in several sets of experiments; the design of the protocols was very similar to those for T7 expression system in Chapter 3. The experimental conditions that yield the best results will be used for the generation of GFP in p*araBAD* expression system using fed-batch fermentation at 5 L scale. Further details regarding this process is elaborated upon in Chapter 5.

Due to the fact that *araBAD* is an arabinose-inducible promoter, arabinose induction was crucial for serving as an effective inducer for RPP in the *araBAD* promoter system (Overton, 2014). In fact, RP accumulation may not occur in a

p*araBAD* expression system lacking arabinose induction. Arabinose give positive effect on protein activity, resulting in a decrease in IBs production and an increase in protein solubility (Narayanan, 2009).

As stated in **Table 4.1**, the medium used for cell growth in shake flasks was Terrific Broth (TB) with Kanamycin (50 μg/mL) as an addition in order to sustain plasmid retention. Glycerol and glucose were used for comparison of carbon source. The temperature used was either 30°C or 37°C. In this study, arabinose was added as an inducer at concentrations of 0.015% or 0.2%. Cultures were sampled at regular intervals. Growth was determined by measuring OD_{650} and CFU. Plasmid retention was measured using replica plating. FCM was used to determine levels of GFP fluorescence per cell and viability using PI. SDS-PAGE and subcellular fractionation were further utilised to analyse cell samples in order to determine the production of RP and evaluate their solubility.

Table 4.1: GFP experimental design for cell growth in shake flasks in batch culture using p*araBAD* expression system

Parameter	***Culture Conditions***
Medium used	Terrific Broth supplemented with Kanamycin (50 µg/mL)
Carbon source	Glycerol or glucose (0.4% or 0.8%)
Temperature	30°C or 37°C
Inducer concentration	Arabinose (None, 0.015% or 0.2%)
Induction point	Mid-logarithmic phase (OD_{650} ≈ 0.5)

4.2 Batch culture with 0.4% glycerol as carbon source at 30°C

In the first experiment, cells were grown in 100 mL of TB medium with 0.4% glycerol and 50 µg/mL Kanamycin. The cells were grown at a growth temperature of 30°C. Similar to the earlier work using the T7 expression system, RPP was induced by the addition of arabinose when the OD_{650} reached around 0.5 (mid-logarithmic phase). Additionally, an uninduced control lacking arabinose was also included. At regular intervals, samples were collected and the OD_{650} measured. In order to determine the viability of the cultures, some samples were analysed for CFU on NA plates, and were replica-plated onto NA plates supplemented with 50 µg/mL Kanamycin to evaluate the retention of plasmids. All shake flask experiments were conducted in duplicate.

In general, the results in **Figure 4.1a** showed a slightly difference in growth of *araBAD* promoter cultures expressing GFP. Cultures with 0.015% arabinose-induced showed the highest cell densities than other cultures, reaching the final OD_{650} of 12.7. Cultures induced with 0.2% arabinose exhibited the lowest final OD_{650} with a value of 10.28, comparable to uninduced cultures, which had a final OD_{650} of 10.72.

The observation of the CFU showed no reduction in all induced and non-induced cultures, before and after induction (**Figure 4.1b**). All cultures maintained their culturability, reaching 10^9 to 10^{10} CFU/mL throughout the growth. There was no drop in CFU following induction as had been observed previously in T7 promoter cultures. Plasmid retention was likewise not much affected by induction in *araBAD*

promoter cultures, only cultures with 0.2% arabinose-induced showed plasmid loss of roughly 15% after 4 h of induction (**Figure 4.1c**). This observation suggests that, compared to the T7 promoter system, the *araBAD* promoter system exhibited more stability and tolerance, and has a significantly lesser detrimental effect on bacterial physiology.

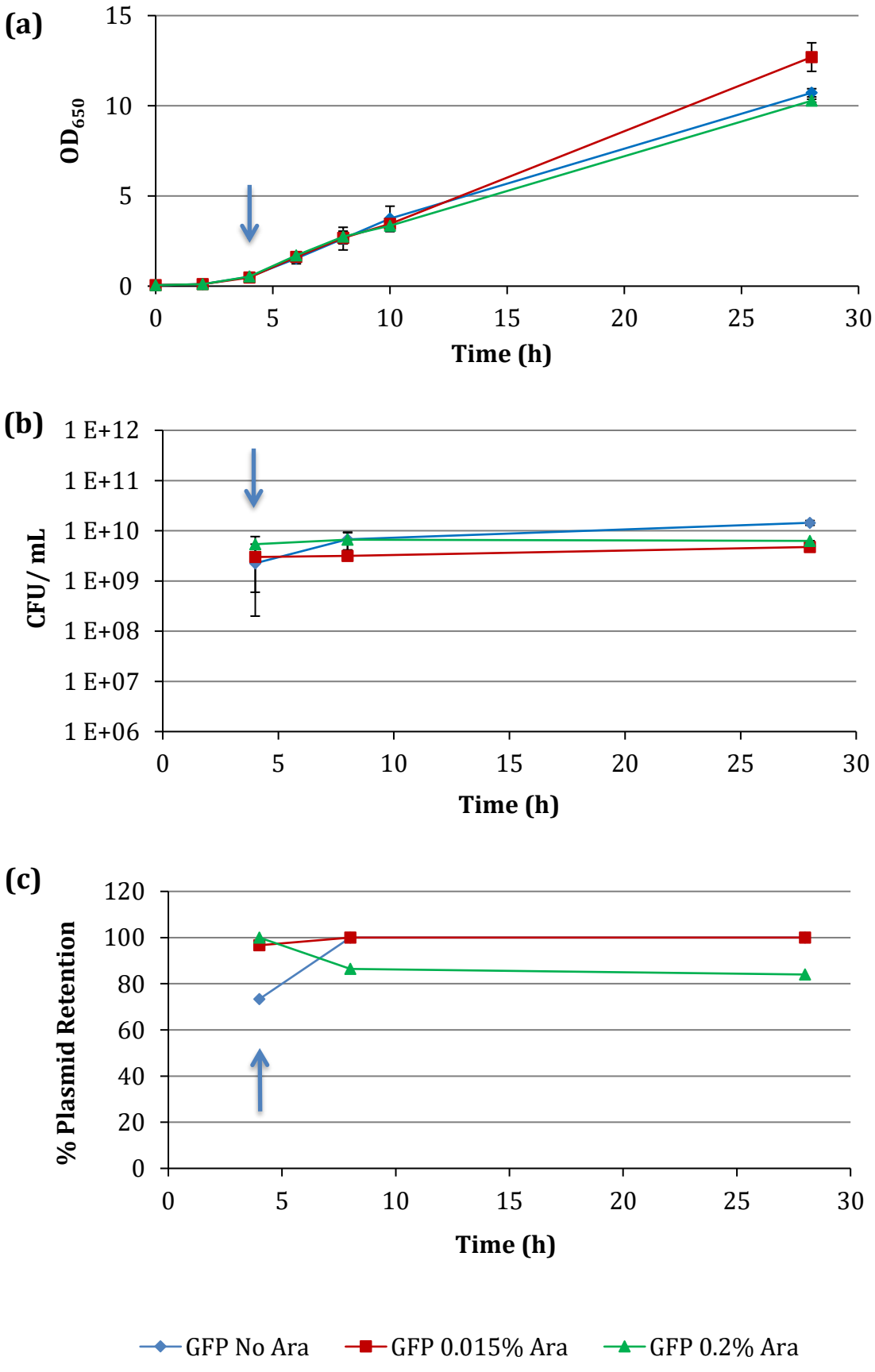

(a)
OD650
15
10
5
0
0
5
10
15
20
25
30
Time (h)
(b)
CFU/ mL
1 E+12
1 E+11
1 E+10
1 E+09
1 E+08
1 E+07
1 E+06
0
5
10
15
20
25
30
Time (h)
(c)
% Plasmid Retention
120
100
80
60
40
20
0
0
5
10
15
20
25
30
Time (h)
GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara

Figure 4.1: Correlation of *E. coli* BL21(A) growth for the study of optimisation for GFP production, grown in TB medium at 30°C with 0.4% glycerol

E. coli BL21(A) cultures carrying the plasmid p*araBAD*-GFP were induced with three different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at OD_{650} ≈ 0.5 (approximately after 4 h of growth; indicated by blue arrow), evaluated by (a) OD_{650} measurement, (b) CFU/mL counting, and (c) plasmid retention. (a) The OD_{650} was measured at intervals. (b) Serial dilution of culture samples at 0 h, 4 h and 24 h post-induction was plated onto NA agar for indication of cell viability and, (c) were replica-plated onto NA agar supplemented with Kanamycin (50 μg/mL) to reveal plasmid retention. All data shown are mean values from replica flasks, error bars are ±1 standard deviation.

4.2.1 Flow cytometry measurement at single cell level

The flow cytometry analysis demonstrated that the production of green fluorescence increased more gradually over time following induction, with the highest green fluorescence values observed at 24 h post-induction with 676,117 fluorescence unit for cultures of p*araBAD*-GFP induced by 0.015% arabinose (**Figure 4.2a**). In contrast to the findings observed in the T7 promoter system, it was noted that the onset of green fluorescence production occurred as early as 2 h following induction, with the highest level of green fluorescence achieved after 4 to 6 h of induction (**Figure 3.2a**). In p*araBAD*-GFP cultures, increasing the arabinose concentration to 0.2% did not yield in an increase in the amount of green fluorescence generated; rather, it was observed to decrease.

Significant observations were also made in GFP non-induced cultures, where uninduced p*araBAD*-GFP cultures were hardly exhibiting any green fluorescence, indicating a lack of RP production. The expression of RP is strictly regulated by the *araBAD* promoter (Overton, 2014). Since the *araBAD* promoter is far more tightly regulated than the T7 promoter system, cultures grown in the absence of arabinose generated relatively little GFP.

According to preliminary observations of FL1-A (fluorescence units) histogram data, all induced cultures of p*araBAD*-GFP showed a sudden increase of GFP-producing cells (GFP+ cells) following induction, reaching nearly 100% (**Figure 4.2b**). The non-induced cultures exhibited a complete absence of cells that

produced GFP, with less than 5% of the cells were GFP+ cells throughout the growth.

Generally, there were low proportions of dead cells in *araBAD* promoter cultures as measured by PI staining. It was clearly seen that the percentage of dead cells were a bit higher before induction, but were decreased upon arabinose induction (**Figure 4.2c**). The percentage of dead cells observed in all cultures were less than 3% at 4 h post-induction until the end of growth, reflecting less metabolic stress on the cells due to gradual production of GFP. Non-induced cultures barely produced any green fluorescence. This indicates that *araBAD* promoter system has a far less negative impact on bacterial physiology than the T7 system,

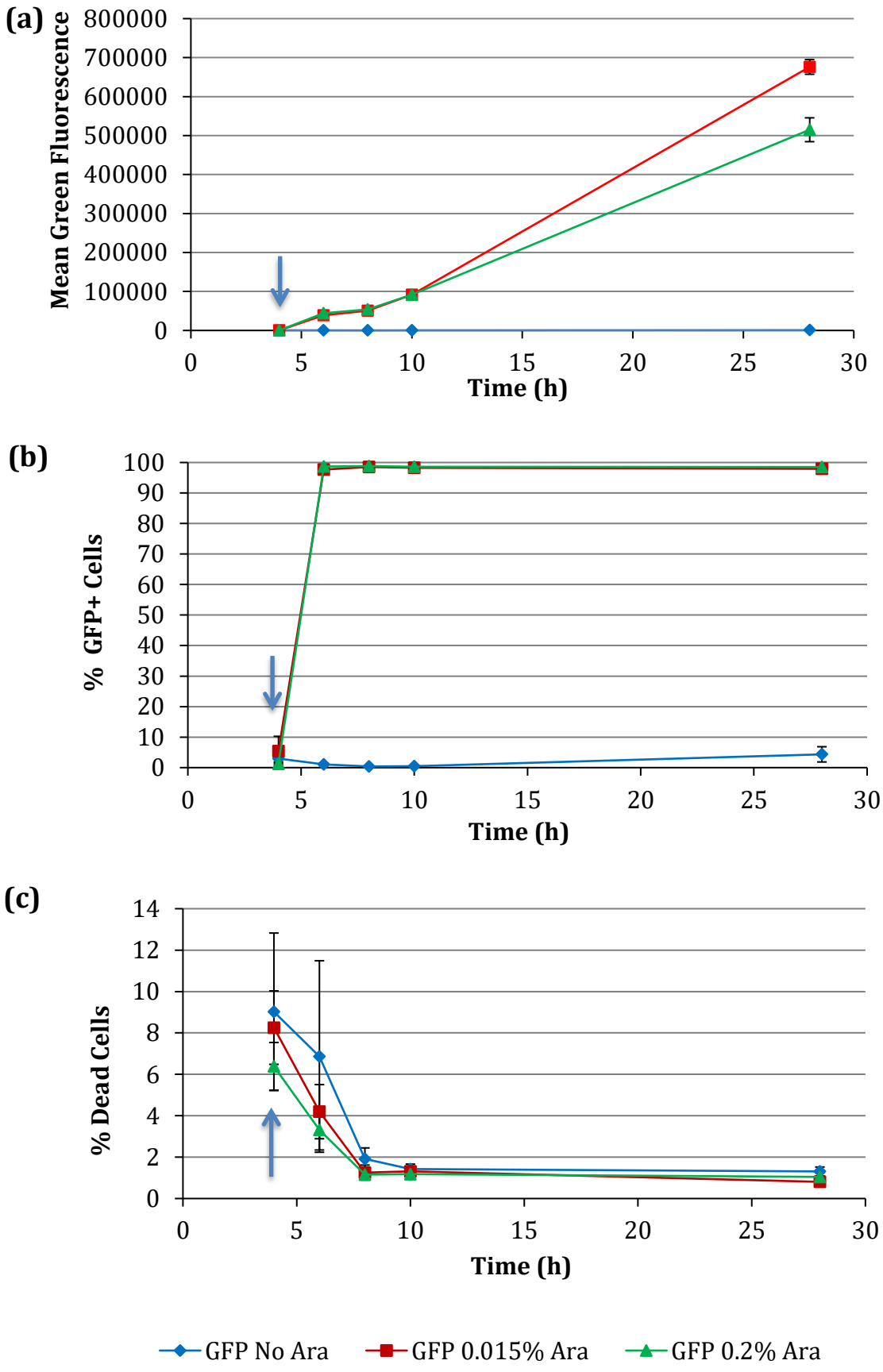

(a)
Mean Green Fluorescence
800000
700000
600000
500000
400000
300000
200000
100000
0
0
5
10
15
20
25
30
Time (h)
(b)
% GFP+ Cells
100
90
80
70
60
50
40
30
20
10
0
0
5
10
15
20
25
30
Time (h)
(c)
% Dead Cells
14
12
10
8
6
4
2
0
0
5
10
15
20
25
30
Time (h)
GFP No Ara
GFP 0.015% Ara
GFP 0.2% Ara

Figure 4.2: Mean green fluorescence measured by flow cytometry for GFP production in batch culture with 0.4% glycerol at 30°C

Green fluorescent protein (GFP) was expressed from an arabinose-induced plasmid p*araBAD* system with different concentrations of arabinose (Uninduced, 0.015% or 0.2%) at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow). (a) FL1-A (green fluorescence) plots of cells taken from culture samples at interval time points post-induction. (b) Percentage of cells that producing GFP (GFP+ cells) versus time, and (c) Percentage of dead cells at interval 0 h to 24 h post-induction. Data are mean values of duplicate samples, error bars are ±1 standard deviation.

The flow cytometry measurement of FL1-A histograms

From the observation of FL1-A histogram peaks, there was a single population of GFP+ cells exhibited in all p*araBAD*-GFP induced cultures following induction at all timepoints, reflecting complete induction (**Figure 4.3b**). Pre-induction with arabinose (4 h of growth) in induced cultures consistently exhibited non-induced subpopulation (GFP- cells), similar to that histogram peaks of the p*araBAD*-GFP uninduced cultures and untransformed BL21(A) cultures that were located around 2 x 10^2 fluorescence units (reflecting the autofluorescence of *E. coli*; **Figure 4.3a**). This once again suggests that the *araBAD* promoter exhibited tight regulation, as RP expression will not occur in the cultures in the absence of arabinose. Induction with 0.015% arabinose concentration was shown to be sufficient in completely boosting RPP in cultures of p*araBAD*-GFP with 0.4% glycerol. Additionally, it was also observed that very few GFP- cells exhibited in all cultures until the end of 28 h growth.

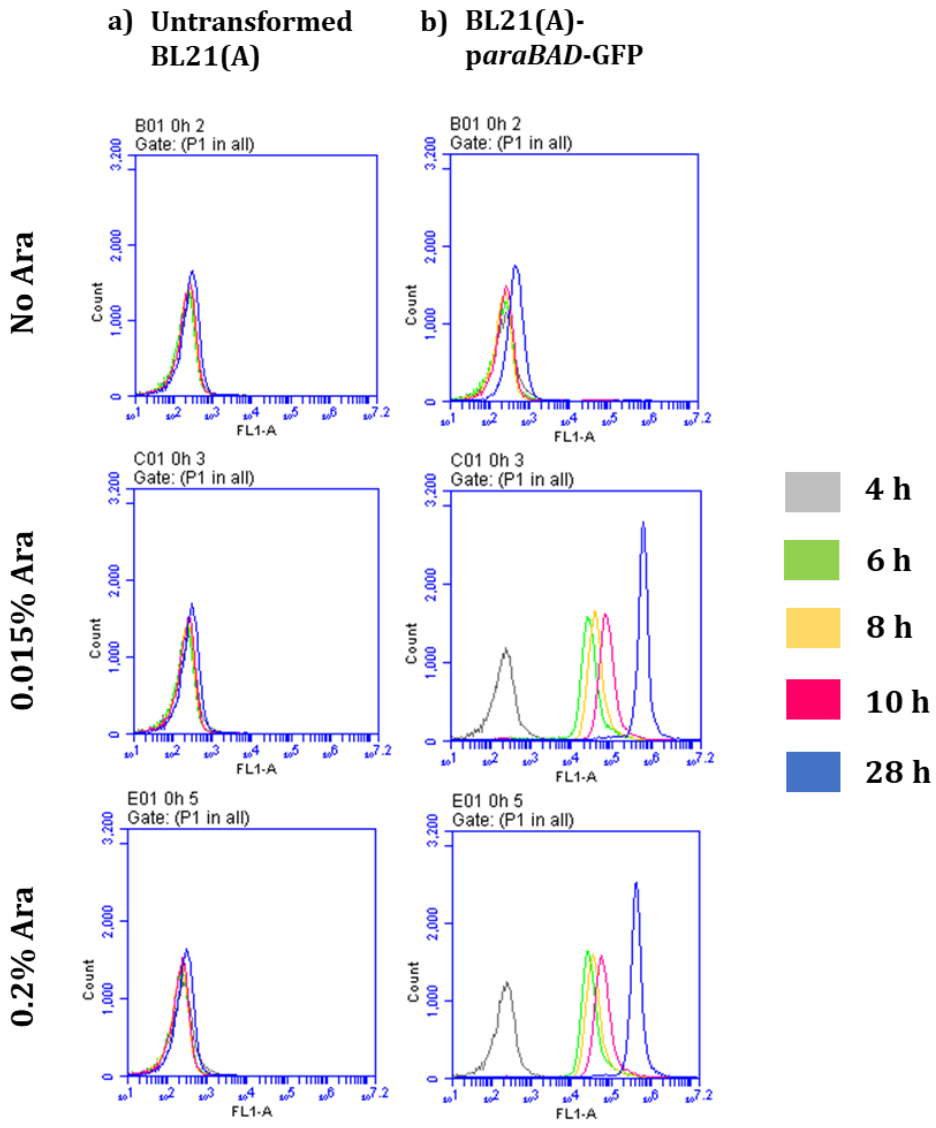


Figure 4.3: Single cell analysis of p*araBAD* expression system for GFP cultures and untransformed *E. coli* BL21(A) with 0.4% glycerol at 30°C

(a) Untransformed *E. coli* BL21(A) cultures and (b) *E. coli* BL21(A) cultures carrying p*araBAD*-GFP were grown in Terrific Broth with 0.4% glycerol as a carbon source at 30°C. At an OD_{650} of 0.5 (approximately after 4 h growth), arabinose (0.015% or 0.2%) was added. At regular intervals, samples were taken and FCM was used to measure green fluorescence of individual cells. Timepoints relate to the time the cultures were grown; 4 h refers to pre-induction with arabinose. Duplicate cultures of each plot showed a similar pattern throughout the growth.

Inspection of green fluorescence histograms at 0 h, 6 h and 24 h post-induction revealed a single population over time (**Figure 4.4**). At 0 h (pre-induction; approximately after 4 h of growth), there was only one subpopulation of non-productive GF-producing (GFP-) cells in all p*araBAD*-GFP cultures. At timepoint 6 h post-induction, one subpopulation of intermediate GF-producing cells, which were positioned between 10^4 and 10^5 fluorescence units, was seen in both p*araBAD*-GFP induced cultures after induction and was present in both cultures with arabinose concentrations of 0.015% and 0.2%. At 24 h post-induction, both of these induced cultures showed the presence of a single subpopulation of significantly high productive GF-producing (GFP+) cells at 10^5 fluorescence units. This observation corresponds to a very high generation of green fluorescence that was recorded, particularly for cultures induced by 0.015% arabinose, with a significant high value of nearly 700, 000 fluorescence units (**Figure 4.2a**). At all timepoints, non-induced cultures comprised non-productive GF-producing cells. Overall, this finding demonstrated that all cultures were non-productive GF-producing cells in the absence of arabinose and were highly productive GF-producing cells following induction, clearly demonstrating tight regulation in the *araBAD* promoter.

FL1-A histogram data in **Figure 4.5** showed populations in cultures at 24 h post-induction; in which p*araBAD*-GFP displayed only a single population in all cultures, consisting of either productive (GFP+) or non-productive (GFP-) cells (**Figure 4.5a**). As opposed to the T7 promoter system, which exhibited multiple populations in both T7-GFP induced cultures, reflecting heterogeneities at 24 h post-induction (**Figure 4.5b**). These T7 promoter cultures likewise showed a

single population of productive GF-producing (GFP+) cells in uninduced cultures as an indication of T7 promoter leakiness. There were also a higher number of GFP- cells observed in all induced T7-GFP cultures compared to p*araBAD*-GFP induced cultures. In general, the p*araBAD* expression system indicated the generation of single population, reduced heterogeneity within the cultures and fewer GFP- cells at 28 h of growth. This finding proved the strength of *araBAD* promoter in effectively controlling protein expression, hence enhancing the efficiency of RP generation. Indeed, the p*araBAD* expression system is advantageous because of its strictly regulation (Guzman et al., 1995).

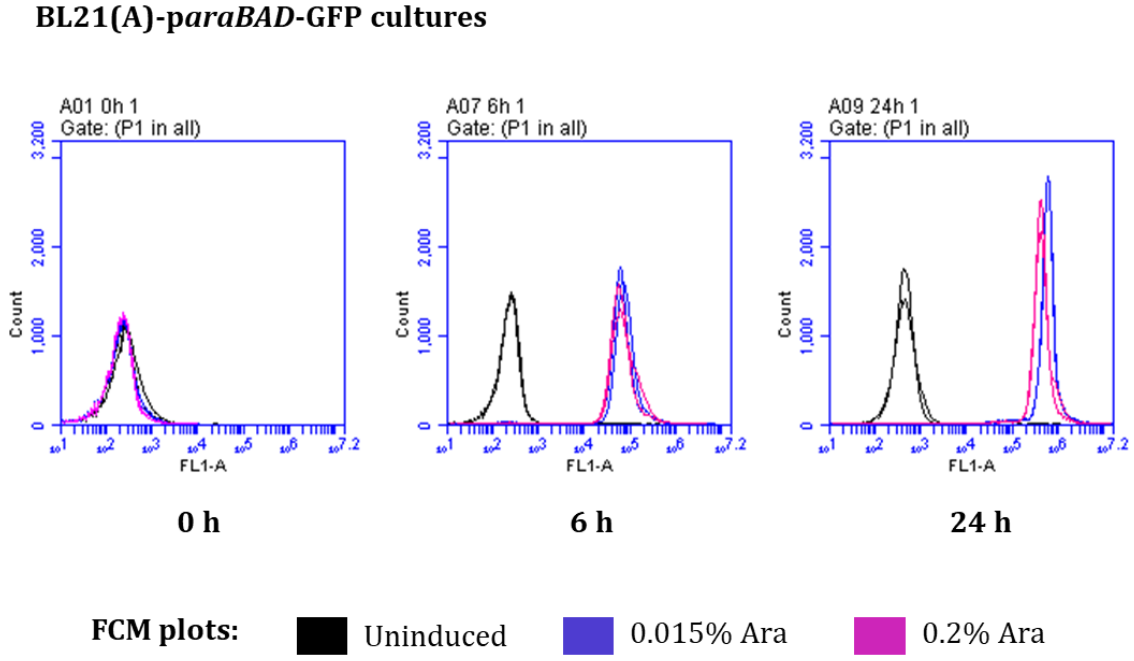


Figure 4.4: FCM green fluorescence histograms of individual time points for paraBAD-GFP cultures in batch culture with 0.4% glycerol at 30°C

Inspection of green fluorescence histograms at 0 h, 6 h and 24 h post-induction revealed a single population over time. p*araBAD*-GFP induced cultures comprised a single population with respect to green fluorescence at timepoint 6 h and 24 h post-induction. Non-productive GFP-producing cells was observed in non-induced cultures at all timepoints. Data are mean values derived from duplicate cultures; error bars are ±1 standard deviation. Duplicate cultures of each plot showed a similar pattern throughout the growth.

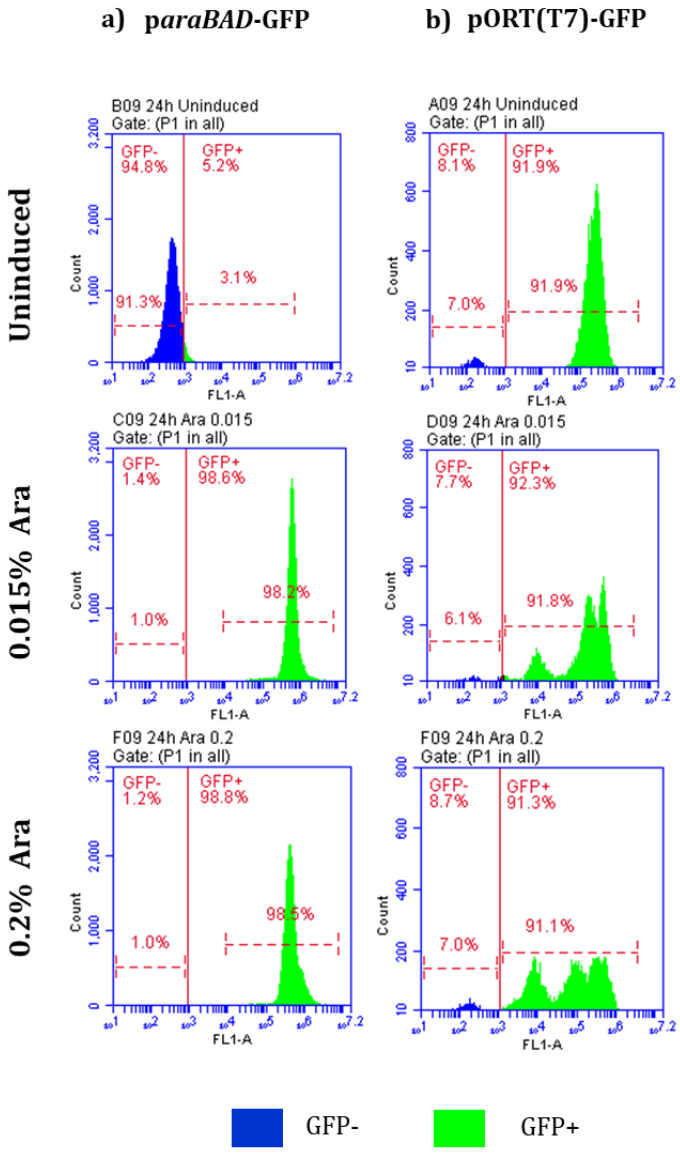


Figure 4.5: Comparison of single cell analysis between p*araBAD* vs pORT(T7) expression system for GFP cultures with 0.4% glycerol at 30°C

FL1-A histogram data showed populations in cultures at 24 h post-induction. (a) BL21(A)-p*araBAD*-GFP and (b) BL21(T7)-pORT(T7)-GFP were grown in TB medium with 0.4% glycerol as a carbon source at 30°C. At an OD_{650} of 0.5 (approximately after 4 h growth), arabinose (0.015% or 0.2%) was added. At regular intervals, samples were taken and FCM was used to measure green fluorescence of individual cells. (a) displayed single populations in the cultures that represents complete induction, while (b) demonstrated multipopulations that represents heterogeneity at 24 h post-induction. Blue represents non-productive GF-producing (GFP-) cells, while green represents productive GF-producing (GFP+) cells. Duplicate cultures of each plot showed a similar pattern throughout the growth.

4.2.2 SDS-PAGE and fractionation

Further analysis on SDS-PAGE and subcellular fractionation was done to investigate protein solubility, both soluble and insoluble fractions. The highest RP accumulation was at 24 h post-induction, with the highest soluble GFP was observed in both induced p*araBAD*-GFP cultures (**Figure 4.6**). It was found that the amount of soluble and insoluble GFP measured were comparable, and different concentration of arabinose-induced cultures gave similar amount of GFP produced, with around 0.3 mg/mL of protein concentration. Non-induced cultures produced significantly lower GFP accumulation with around 0.05 mg/mL, and they were mostly insoluble. Data obtained from batch culture of BL21(A)-p*araBAD* using 0.4% glycerol have successfully proved that production of GFP is significantly high in induced cultures, while non-induced cultures barely producing any GFP. Densitometry data for the proportion of protein solubility that was soluble or insoluble GFP was shown on the gels in **Figure 4.7**.

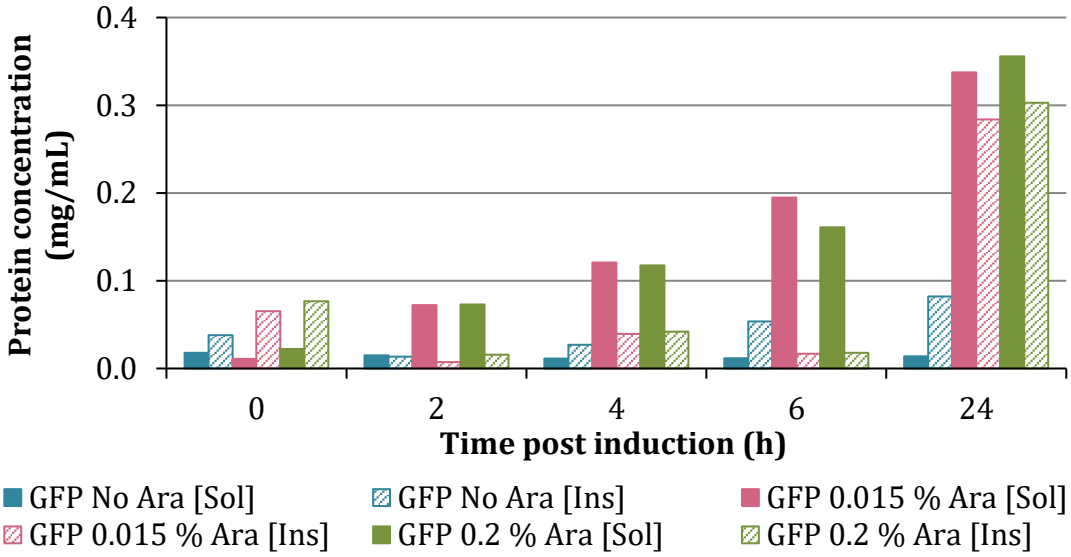
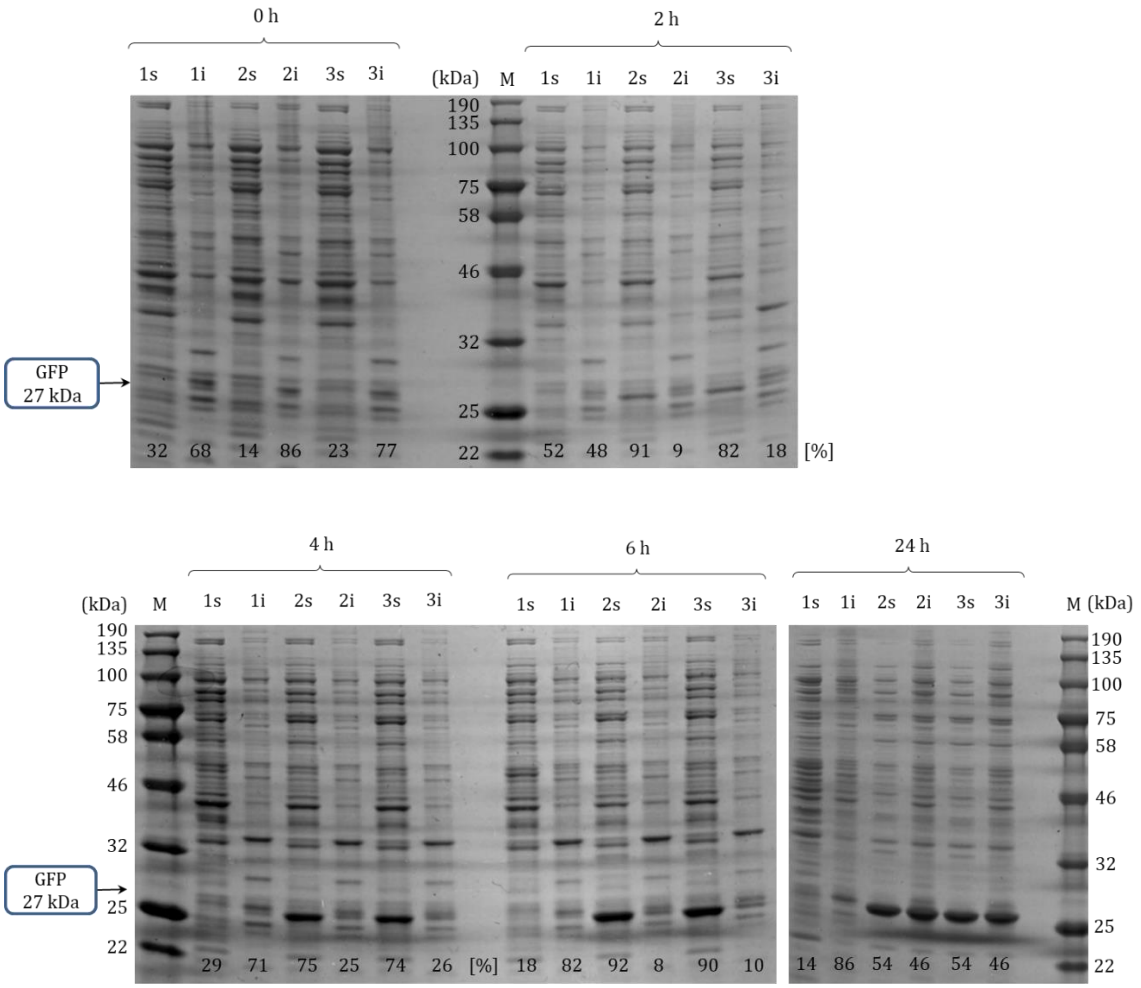


Figure 4.6: The expression of protein solubility in cell extracts measured by SDS-PAGE for p*araBAD*-GFP in batch culture with 0.4% glycerol at 30°C

Estimated protein solubility of soluble and insoluble fractions in cell extracts of p*araBAD*-GFP cultures at concentration of mg/mL at times post induction.



M: Marker used; Blue Protein Standard (BPS), broad range of 11-190 kDa;

1s: GFP No Ara (Soluble); **1i**: GFP No Ara (Insoluble);

2s: GFP 0.015 % Ara (Soluble); **2i**: GFP 0.015 % Ara (Insoluble);

3s: GFP 0.2 % Ara (Soluble); **3i**: GFP 0.2 % Ara (Insoluble)

Figure 4.7: The accumulation of GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in batch culture with 0.4% glycerol at 30°C

SDS-PAGE gel picture with arrow showing GFP (27 kDa) and the ratio percentage of soluble versus insoluble GFP proteins accumulated inside the cell extracts of p*araBAD*-GFP, as shown at the bottom of the gel. All sample bands at times post-induction were compared to BPS ladder band.

4.2.3 Discussion

The findings of this study demonstrated that p*araBAD*-GFP cultures grew well in the medium with 0.4% glycerol, produced green fluorescence gradually after being induced and reached the highest green fluorescence value at 28 h of growth (**Figure 4.2a**). In contrast to T7-GFP cultures that could rapidly generate green fluorescence as early as 2 h following induction, achieved the highest green fluorescence value at 4-6 h post-induction (**Figure 3.2a**).

Similar to T7 promoter cultures, lower concentration of 0.015% arabinose was sufficient to boost the RPP in *araBAD* promoter cultures. Notably, the findings showed that increasing the arabinose concentration to 0.2% did not yield any significant impact to the RP generated. Higher inducer concentrations, as previously stated, did not result in an increase in the rate of particular protein expression or the cellular productivity(Lecina et al., 2013).

In fact, *araBAD* promoter system favoured 0.015% arabinose-induced cultures as growing in cultures induced with 0.2% arabinose slower the cell growth and decreased the green fluorescence produced, but overall were comparable. Additionally, they exhibited less CFU drop and fewer plasmid loss following induction. While T7 promoter system also exhibited some reduction in terms of cell growth and green fluorescence produced in 0.2% arabinose-induced cultures, they also had a dramatic drop in CFU and some plasmid loss following induction in all induced-cultures. This might be due to the increase of metabolic stress within the cells caused by the rapid synthesis of RPP. Therefore, the *araBAD* promoter

system was believed to be more tolerated and have a lesser negative impact on bacterial physiology.

The tight regulation of the *araBAD* promoter inhibited the production of any green fluorescence in the cultures lacking arabinose induction. Unlike the T7 promoter, which exhibited leaky expression, uninduced cultures had the ability to generate green fluorescence and achieve levels of green fluorescence that were comparable to those observed in induced cultures, following 28 h of growth.

Further investigation was conducted in order to explore the efficacy of *araBAD* promoter for RPP when growing in a medium with glucose as the carbon source and to determine whether the utilisation of glucose could potentially enhance both cell growth and RP accumulation in p*araBAD* expression system.

4.3 Use of glycerol vs glucose as carbon source in p*araBAD* expression system

The next experiment was designed to grow p*araBAD*-GFP cultures in 100 mL of TB medium with utilising glucose as carbon source instead of glycerol, at growth temperature of 30°C. The concentration of glucose used was 0.4% and 0.8%. All experiments in shake flask were executed in duplicate. The use of glucose as the carbon source has been compared to the use of glycerol in the growth medium. Glucose is known to be the preferred carbon source for bacteria especially *E. coli*, and it provide faster growth rate in comparison to other type of sugars (Mayer et al., 2014).

Overall, the growth curves were comparable in all p*araBAD*-GFP cultures even though the growth were a bit slower with 0.2% arabinose-induced (**Figures 4.8a, c, e**). The *paraBAD*-GFP cultures of 0.4% glycerol with 0.015% arabinose-induced showed the highest cell densities, compared to all glucose cultures, with the final OD_{650} of 12.7. The least OD_{650} growth was p*araBAD*-GFP 0.2% arabinose-induced of 0.4% glucose cultures, reaching the final OD_{650} of 8.81 after 24 h post-induction. It was very likely that acetate would be produced when glucose was used in the growth medium. Even when present in low concentrations, the accumulation of acetate has the potential to inhibit cells from growing and decrease the effectiveness of protein expression (Lecina et al., 2013).

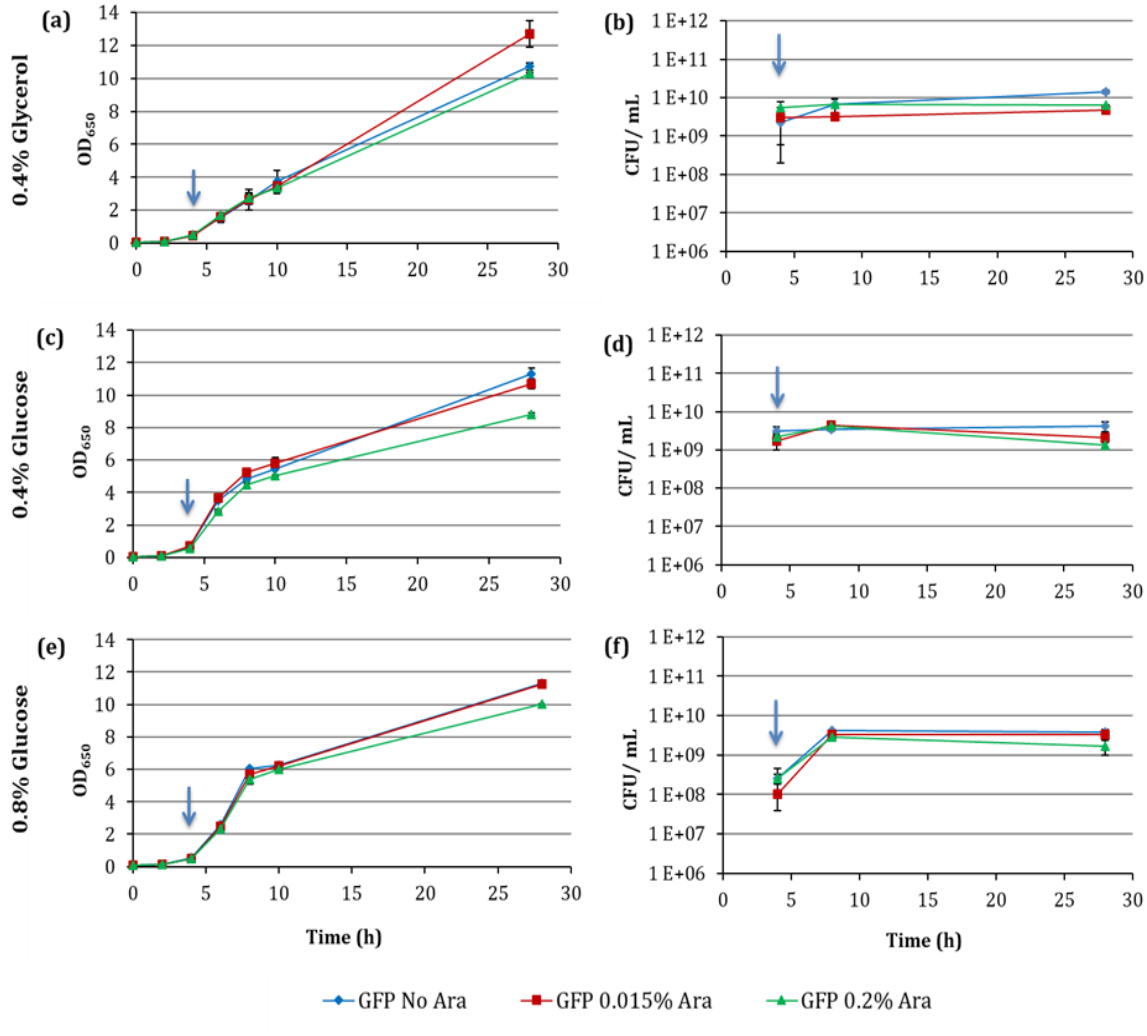


Figure 4.8: Comparison of OD and CFU in p*araBAD* expression system for the production of GFP

E. coli BL21(A) cultures carrying p*araBAD*-GFP were grown in batch culture at 30°C in TB medium with 0.4% glycerol (a-b); 0.4% glucose (c-d); or 0.8% glucose (e-f). Different concentrations of arabinose (Uninduced, 0.015% or 0.2%) were induced at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow), evaluated by OD_{650} measurement and CFU/mL counting. The OD_{650} was measured at intervals (a, c, e). Serial dilution of culture samples at 0 h, 4 h and 24 h post-induction was plated onto NA agar for indication of cell viability (b, d, f). All data shown are mean values from replica flasks, error bars are ±1 standard deviation.

Similar to the data obtained when using 0.4% glycerol as carbon source, there was no drop in CFU following induction in all glucose cultures, both of 0.4% and 0.8% glucose concentration (**Figures 4.8b, d, f**). The T7 promoter system, as previously noted, exhibited a sharp decline in CFU after induction as a result of metabolic stress on cellular physiology brought on by the rapid generation of RP. This shows that the p*araBAD* promoter causes less stress on cells than the T7 promoter, resulting in no reduction in cell culturability throughout incubation. The results from CFU, however, revealed a slight decrease in all glucose cultures after 28 h of growth. This could be caused by the overflow metabolism that produces acetate, which is harmful for cell growth and viability (Retamal et al., 2018).

4.3.1 Flow cytometry measurement

The results obtained from flow cytometry analysis demonstrated that the production of green fluorescence in all *araBAD* promoter cultures occurred in a more gradual manner over time after induction, in contrast to the T7 cultures. *araBAD* promoter cultures exhibited the highest levels of fluorescence after 28 h of growth. The expression of p*araBAD*-GFP in glucose cultures resulted in a much lower level of green fluorescence compared to glycerol cultures, with only about 250, 000 fluorescence units (**Figures 4.9c, e**), which is roughly one-third of the fluorescence generated in glycerol cultures (**Figure 4.9a**; nearly 700, 000 fluorescence units). Increasing glucose concentration did not yield a statistically significant rise in the production of green fluorescence.

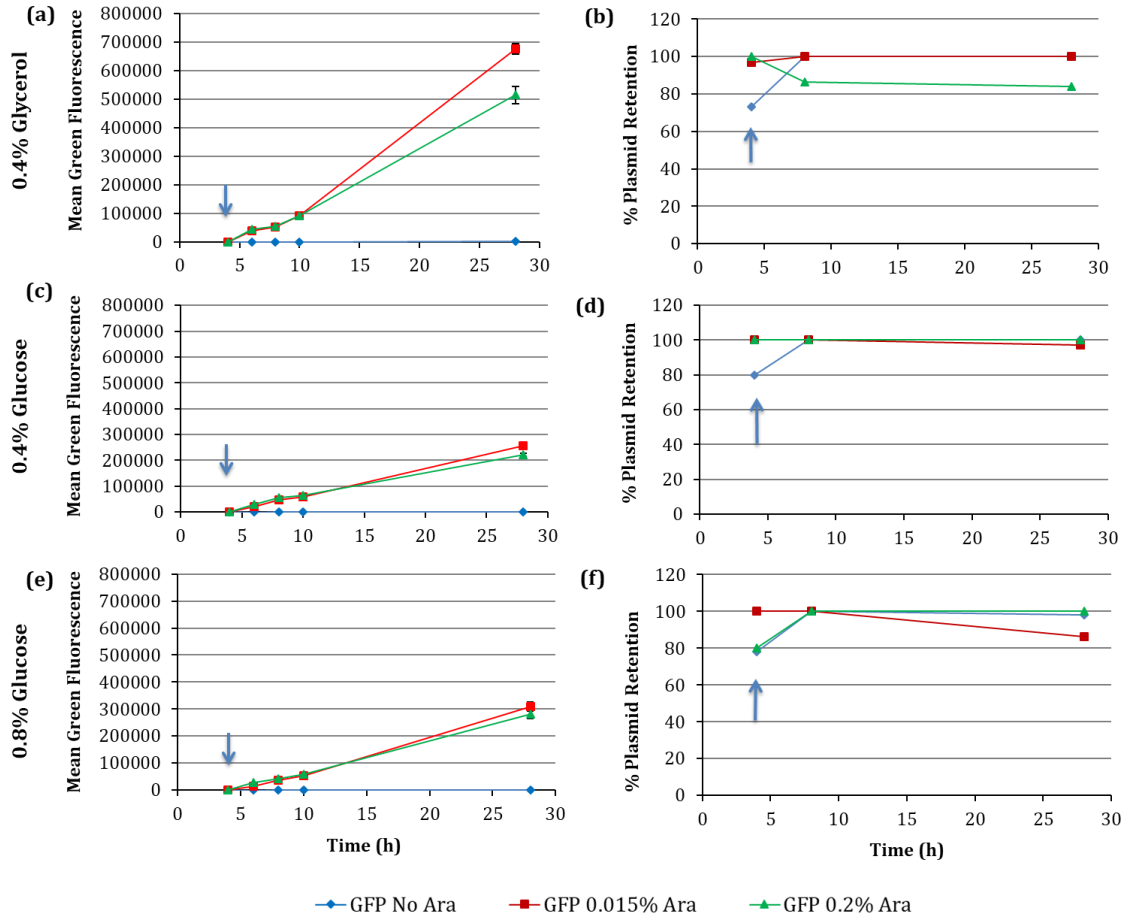


Figure 4.9: Comparison of mean green fluorescence and plasmid retention in p*araBAD* expression system for the production of GFP

E. coli BL21(A) cultures carrying p*araBAD*-GFP were grown in batch culture at 30°C in TB medium with 0.4% glycerol (a-b); 0.4% glucose (c-d); or 0.8% glucose (e-f). Different concentrations of arabinose (Uninduced, 0.015% or 0.2%) were induced at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow). FL1-A (green fluorescence) plots of cells taken from culture samples at interval time points post-induction (a, c, e). Serial dilution of culture samples at 0 h, 4 h and 24 h post-induction was plated onto NA agar and were replica-plated onto NA agar supplemented with Kanamycin (50 µg/mL) to reveal plasmid retention (b, d, f). All data shown are mean values from replica flasks, error bars are ±1 standard deviation.

It is also noteworthy to note that p*araBAD*-GFP non-induced cells barely produced any green fluorescence in the absence of arabinose induction. This demonstrated how tightly the p*araBAD* system was regulated. This observation was very different from uninduced cultures in T7 system, which had produced green fluorescence even without arabinose induction due to leakiness of T7 promoter.

In general, the p*araBAD*-GFP cultures exhibited similar levels of green fluorescence production within the same cultures (either 0.4% glycerol or 0.4% glucose or 0.8% glucose) despite different arabinose concentration, in which increasing the concentration of arabinose did not result in a significant increase in the production of green fluorescence. The plasmid retention of all p*araBAD*-GFP glucose cultures remained above 90% throughout the growth (**Figures 4.9d, f**). The only difference observed was a little decrease in plasmid retention in 0.4% glycerol cultures supplemented with 0.2% arabinose, as compared to 0.015% arabinose cultures (**Figure 4.9b**). This finding correlates with the observed lower levels of green fluorescence in those cultures.

Preliminary observations of FL1-A histogram data on all p*araBAD*-GFP induced cultures revealed a high proportion of productive GF-producing (GFP+) cells that sharply increased after induction, whereas non-induced cultures remained non-productive GF-producing (GFP-) cells until the end of growth (**Figures 4.10a, c, e**), once again illustrating the tightly regulation of *araBAD* promoter. The p*araBAD*-GFP cultures exhibited low percentages of dead cells, as determined by PI staining (**Figures 4.10b, d, f**). This percentage of dead cells was much lower compared to T7-GFP cultures.

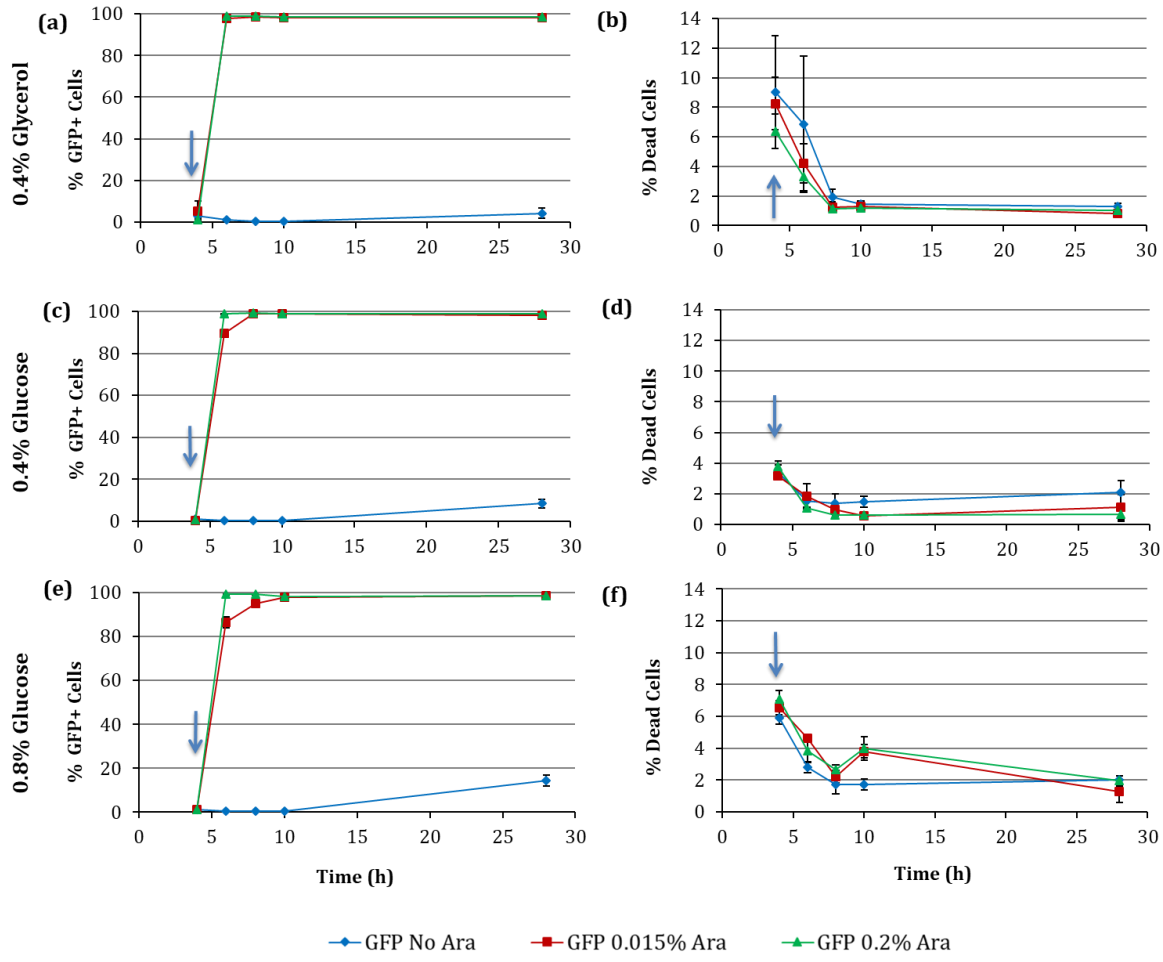


Figure 4.10: Comparison of productive GFP population and dead cells in p*araBAD* expression system for the production of GFP

E. coli BL21(A) cultures carrying p*araBAD*-GFP were grown in batch culture at 30°C in TB medium with 0.4% glycerol (a-b); 0.4% glucose (c-d); or 0.8% glucose (e-f). Different concentrations of arabinose (Uninduced, 0.015% or 0.2%) were induced at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow). Percentage of cells that producing GFP (GFP+ cells) versus time (a, c, e), and percentage of dead cells at interval 0 h to 24 h post-induction (b, d, f). All data shown are mean values from replica flasks, error bars are ±1 standard deviation.

Investigating single cell fluorescence data demonstrated that the p*araBAD* system encountered partial induction under certain growth conditions (**Figure 4.11**). In T7 promoter cultures, the presence of the left shoulder served as a cue for partial induction, whereas in *araBAD* promoter cultures, the appearance of 'on and off' states represented partial induction.

The presence of this phenomenon was observed in cultures of 0.4% glucose that had been induced with 0.015% arabinose; it resulted in a non-induced subpopulation 2 h after induction (after 6 h of growth), and complete induction was achieved after a further 2 hours. The occurrence of that 'on and off' states was more obvious in 0.8% glucose cultures, where the non-induced subpopulation is also visible at 4 h post-induction (after 8 h of growth). This finding is very similar to T7 promoter cultures, where it had previously been observed that increasing the glucose concentration in the medium resulted in more partial induction; indicated by the appearance of a broader shoulder in 0.8% glucose cultures, suggesting that the arabinose induction was not sufficient.

This was expected considering the presence of glucose had a repressive effect on the *araBAD* promoter (Lichenstein et al., 1987; Guzman et al., 1995). This explained the lower amount of green fluorescence produced in p*araBAD*-GFP cultures with glucose (**Figures 4.9c, e**) compared to glycerol cultures, as glucose repressed *araBAD* promoter, thus inhibit GFP expression. Therefore, the more the glucose added, the more the arabinose needed to induce RPP. Because of this, cultures of 0.015% arabinose-induced glucose had non-induced subpopulation (GFP-) cells; the quantity rose with the addition of 0.8% glucose.

Cultures with higher concentration of 0.2% arabinose did not show such things, suggesting arabinose induction was sufficient. The results of our study demonstrated that glucose had a significant inhibitory effect on the p*araBAD* system in comparison to the T7 system.

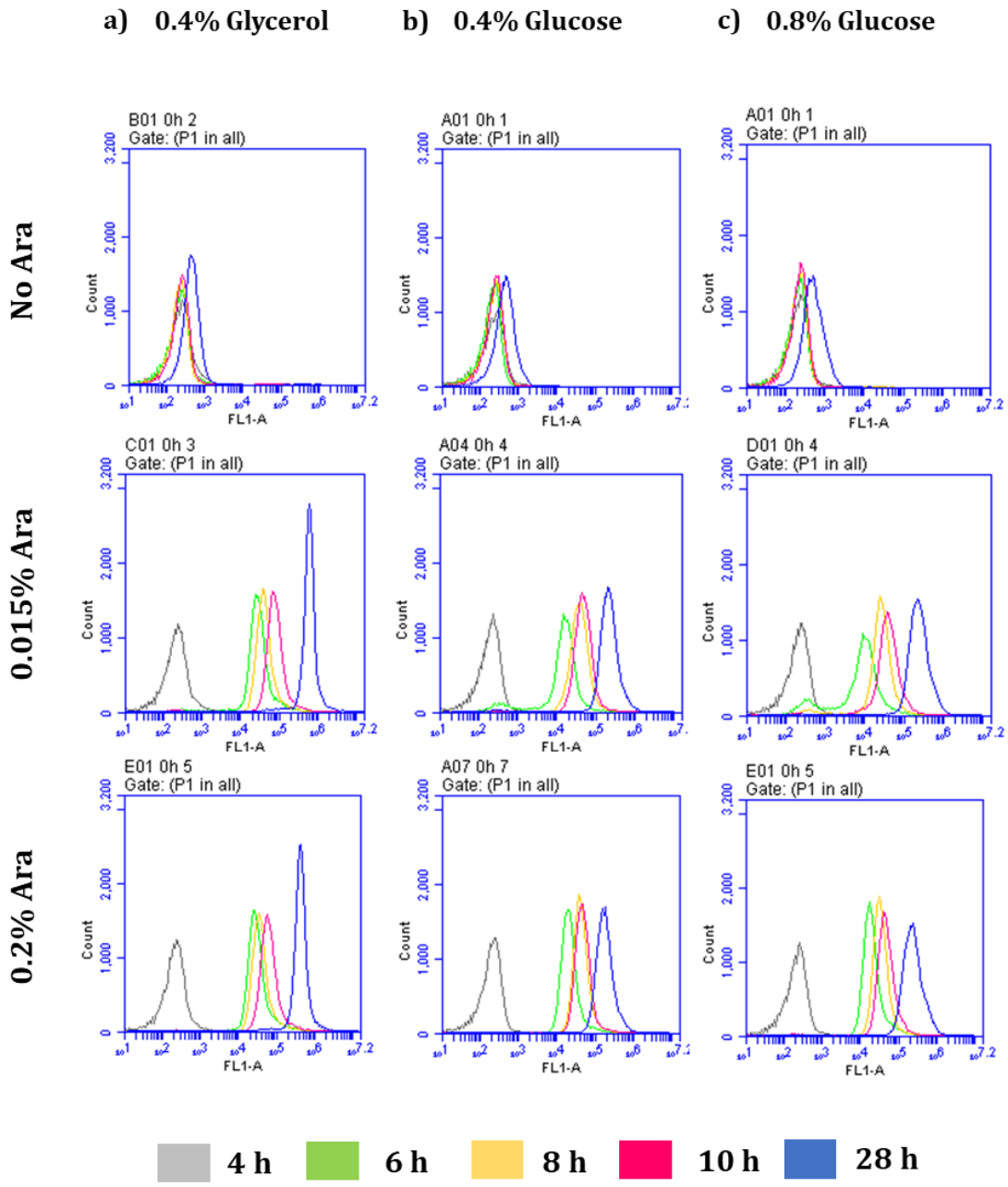


Figure 4.11: Single cell analysis of *araBAD* promoter cultures for the production of GFP

E. coli BL21(A) cultures carrying p*araBAD*-GFP were grown in Terrific Broth with (a) 0.4% glycerol, (b) 0.4% glucose, and (c) 0.8% glucose as a carbon source at 30°C. At an OD_{650} of 0.5 (approximately after 4 h growth), arabinose (0.015% or 0.2%) was added. At regular intervals, samples were taken and FCM was used to measure green fluorescence of individual cells. Timepoints relate to the time the cultures were grown; 4 h refers to pre-induction with arabinose. Duplicate cultures of each plot showed a similar pattern throughout the growth.

4.3.2 SDS-PAGE and fractionation

The protein solubility fractions were obtained from bacterial pellets that underwent freeze-thaw cycles during subcellular fractionation procedure. The highest GFP accumulation were observed at 24 h post-induction in all p*araBAD*-GFP cultures (**Figures 4.12a, b, c**). Cultures with 0.4% glycerol demonstrated comparable amounts of both soluble and insoluble GFP, while glucose cultures, surprisingly, exhibited a higher proportion of soluble GFP compared to insoluble GFP, especially in cultures with 0.4% glucose.

Although it appeared that 0.4% glucose cultures produced a higher estimated proportion of soluble GFP concentration (0.5 mg/mL) than 0.4% glycerol cultures (0.35 mg/mL), the estimated proportion referred to the amount of GFP accumulation in the cultures (**Figures 4.12a, b**). Since the amount of GFP that accumulated in the cultures was mirrored by the green fluorescence detected in FCM, multiplying the green fluorescence by the amount of estimated protein produced in each culture would allow for a direct comparison. In contrast to 0.4% glucose cultures, 0.4% glycerol cultures generated a significantly higher yield of soluble GFP. Densitometry data for the percentage ratio of soluble versus insoluble GFP accumulated inside the cell extracts of p*araBAD*-GFP with glucose cultures were also shown on the SDS-PAGE gel pictures in **Appendix 4.1** and **Appendix 4.2.**

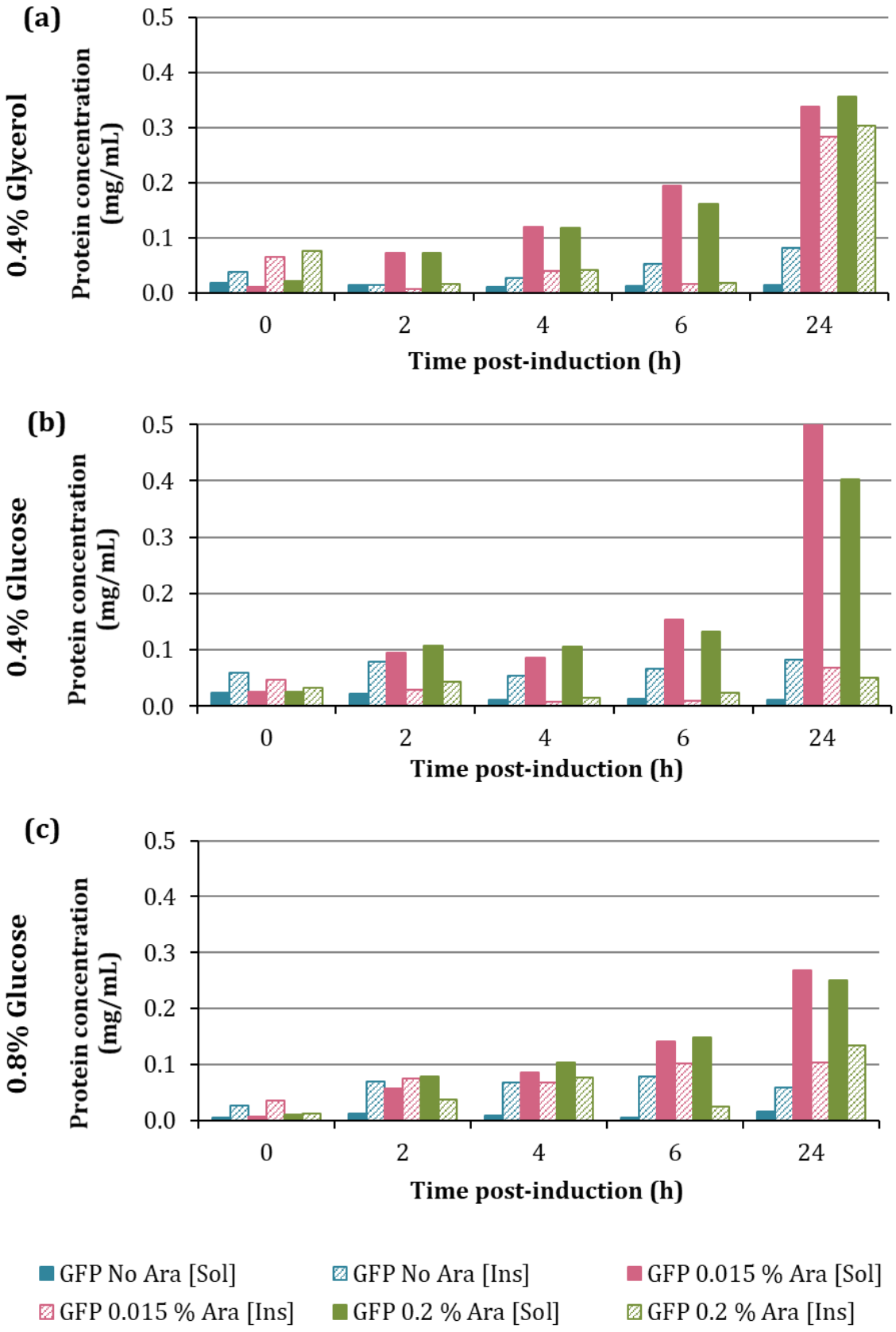

(a)
0.4% Glycerol
Protein concentration (mg/mL)
0.5
0.4
0.3
0.2
0.1
0.0
0
2
4
6
24
Time post-induction (h)
(b)
0.4% Glucose
Protein concentration (mg/mL)
0.5
0.4
0.3
0.2
0.1
0.0
0
2
4
6
24
Time post-induction (h)
(c)
0.8% Glucose
Protein concentration (mg/mL)
0.5
0.4
0.3
0.2
0.1
0.0
0
2
4
6
24
Time post-induction (h)
GFP No Ara [Sol]
GFP No Ara [Ins]
GFP 0.015 % Ara [Sol]
GFP 0.015 % Ara [Ins]
GFP 0.2 % Ara [Sol]
GFP 0.2 % Ara [Ins]

Figure 4.12: The expression of protein solubility in cell extracts measured by SDS-PAGE for p*araBAD*-GFP in batch culture at 30°C

Estimated protein solubility of soluble and insoluble fractions in cell extracts of p*araBAD*-GFP cultures with (a) 0.4% glycerol, (b) 0.4% glucose and (c) 0.8% glucose at concentration of mg/mL at times post-induction.

4.3.3 Discussion

The utilisation of glucose within the growth medium exhibited a notable inclination towards the production of acetate, leads to cell growth inhibition and thus limiting the efficiency of protein expression. As previously noted, acetate synthesis can inhibit cell growth even when it accumulates at a low concentration (Lecina et al., 2013), and increasing the glucose concentration to the medium would undoubtedly prevent cell growth because of overflow metabolism (Retamal et al., 2018). Flow cytometry revealed that the green fluorescence in all *araBAD* promoter cultures were produced more gradually over time following induction compared to T7 promoter cultures, and according to our findings, glucose exhibited a significantly greater repression on the p*araBAD* system compared to the T7 system.

The subsequent experiment was designed for batch culture of *paraBAD*-GFP, using TB medium with 0.4% glycerol as carbon source, yet, the temperature was raised from 30°C to 37°C to determine whether the use of higher temperature could possibly enhance both cell growth and RP accumulation.

4.4 Increasing growth temperature in p*araBAD* expression system

The growth of *araBAD* promoter cultures with 0.2% arabinose-induced was slightly higher compared to 0.015% induced-cultures at 37°C (**Figure 4.13b**). The cultures grown at 30°C, however, had the reverse phenomena (**Figure 4.13a**). The finding from CFU showed a slight reduction in CFU counts after 24 h post-induction in cultures incubated at 37°C, notably, cultures induced with 0.2% arabinose exhibited more considerable decrease in CFU (**Figure 4.13d**). In contrast to the cultures at 30°C, which exhibited no decrease in growth over time (**Figure 4.13c**). This could be explained to the reduction in the cellular culturability of p*araBAD*-GFP cultures when grown at 37°C compared to 30°C. Plasmid retention analysis revealed that, similar to glycerol cultures at 30°C (**Figure 4.13e**), growing the cultures at 37°C also increased plasmid loss in 0.2% arabinose-induced cultures (**Figure 4.13f**).

4.4.1 Flow cytometry measurement

In a little difference from the previously described conditions, p*araBAD*-GFP cultures grown at 37°C produced green fluorescence quicker after induction compared to 30°C, although it reached at a lower level of green fluorescence produced after 28 h of growth compared to 30°C (**Figures 4.14a, b**).

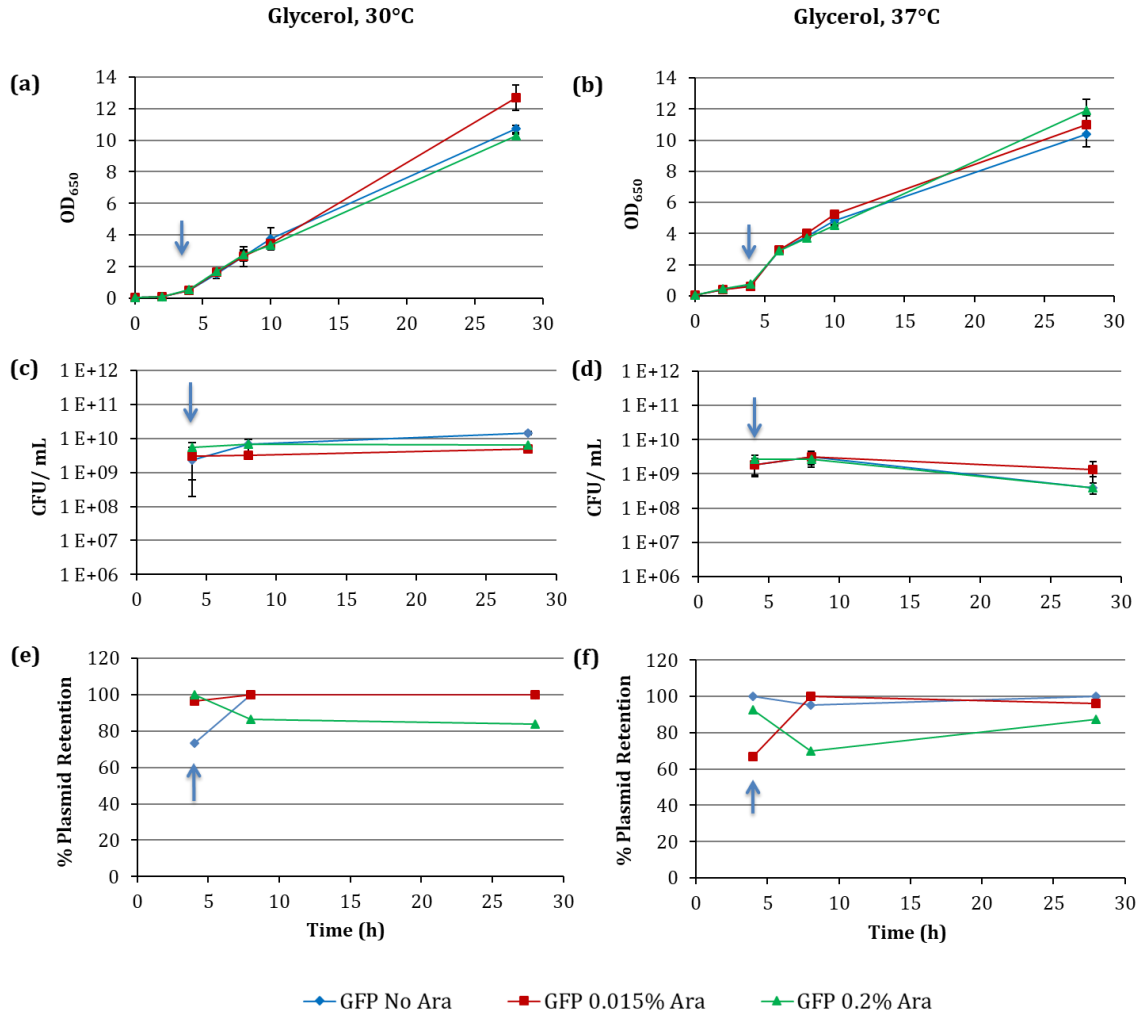


Figure 4.13: Comparison of OD, CFU and plasmid retention in p*araBAD* expression system for the production of GFP

E. coli BL21(A) cultures carrying p*araBAD*-GFP were grown in TB medium with 0.4% glycerol at 30°C (a, c, e); or 37°C (b, d, f). Different concentrations of arabinose (Uninduced, 0.015% or 0.2%) were induced at mid logarithmic phase at OD_{650} ≈ 0.5 (approximately after 4 h of growth; indicated by blue arrow), evaluated by OD_{650} measurement, CFU/mL counting and plasmid retention. The OD_{650} was measured at intervals (a, b). Serial dilution of culture samples at 0 h, 4 h and 24 h post-induction was plated onto NA agar for indication of cell viability (c, d) and were replica-plated onto NA agar supplemented with Kanamycin (50 µg/mL) to reveal plasmid retention (e, f). All data shown are mean values from replica flasks, error bars are ±1 standard deviation.

This might be due to the higher temperature used in the medium, which initially enhanced the rate of protein transcription and translation after induction but ultimately leads to protein misfolding. A higher growing temperature can prevent proper protein folding (Overton, 2014).

Compared to cultures induced with 0.2% arabinose, cultures induced with 0.015% arabinose exhibited less green fluorescence at 37°C. Due to the absence of arabinose, uninduced cultures of p*araBAD*-GFP hardly produced any green fluorescence, indicating the tight regulation of the *araBAD* promoter. This is consistent for all prior p*araBAD*-GFP cultures that were discussed.

In both p*araBAD*-GFP induced cultures growing at 30°C and 37°C, preliminary observations of FL1-A histogram data revealed a sharp increase of productive GF-producing (GFP+) cells after induction (**Figures 4.14c, d**), but there were then a reduced in GFP+ cells at the end of growth in cultures grown at 37°C. PI staining revealed low percentages of dead cells in both p*araBAD*-GFP cultures (**Figures 4.14e, f**).

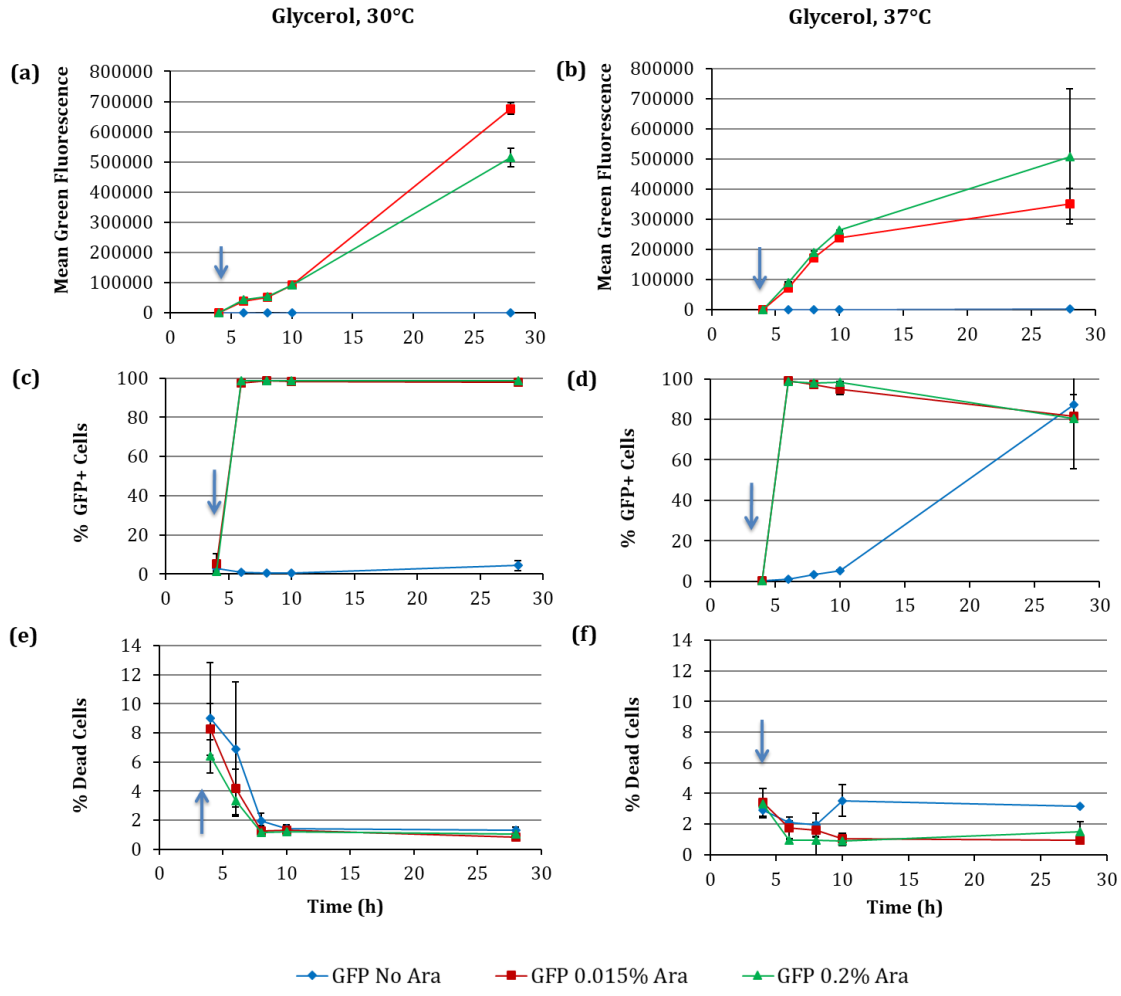


Figure 4.14: Comparison of mean green fluorescence, productive GFP population and dead cells in p*araBAD* expression system for GFP production

E. coli BL21(A) cultures carrying p*araBAD*-GFP were grown in TB medium with 0.4% glycerol at 30°C (a, c, e); or 37°C (b, d, f). Different concentrations of arabinose (Uninduced, 0.015% or 0.2%) were induced at mid logarithmic phase at $OD_{650} \approx 0.5$ (approximately after 4 h of growth; indicated by blue arrow). FL1-A (green fluorescence) plots of cells taken from culture samples at interval time points post-induction (a, b). Percentage of cells that producing GFP (GFP+ cells) versus time (c, d) and percentage of dead cells at interval 0 h to 24 h post-induction (e, f). All data shown are mean values from replica flasks, error bars are ±1 standard deviation.

The analysis of single cell fluorescence data revealed that all induced p*araBAD*-GFP cultures at 37°C were primarily single populations that productively produced green fluorescence after induction, representing the excitement of cultures undergoing protein expression at a higher temperature but then producing more sizable GFP- population after 24 h post-induction, as was evident at the end of 28 h growth (**Figure 4.15b**). These explained the reduction in culturability of these cultures, as was demonstrated by CFU (**F4.15d**) and the lowered plasmid retention (**F4.13f**) at the end of the growth. Host cells frequently experience higher stress levels or a greater metabolic burden when they undergo growth at higher temperatures, and this could be demonstrated through the occurrence of plasmid loss (Overton, 2014). In cultures grown at 30°C, only few GFP- populations were seen (**Figure 4.15a**).

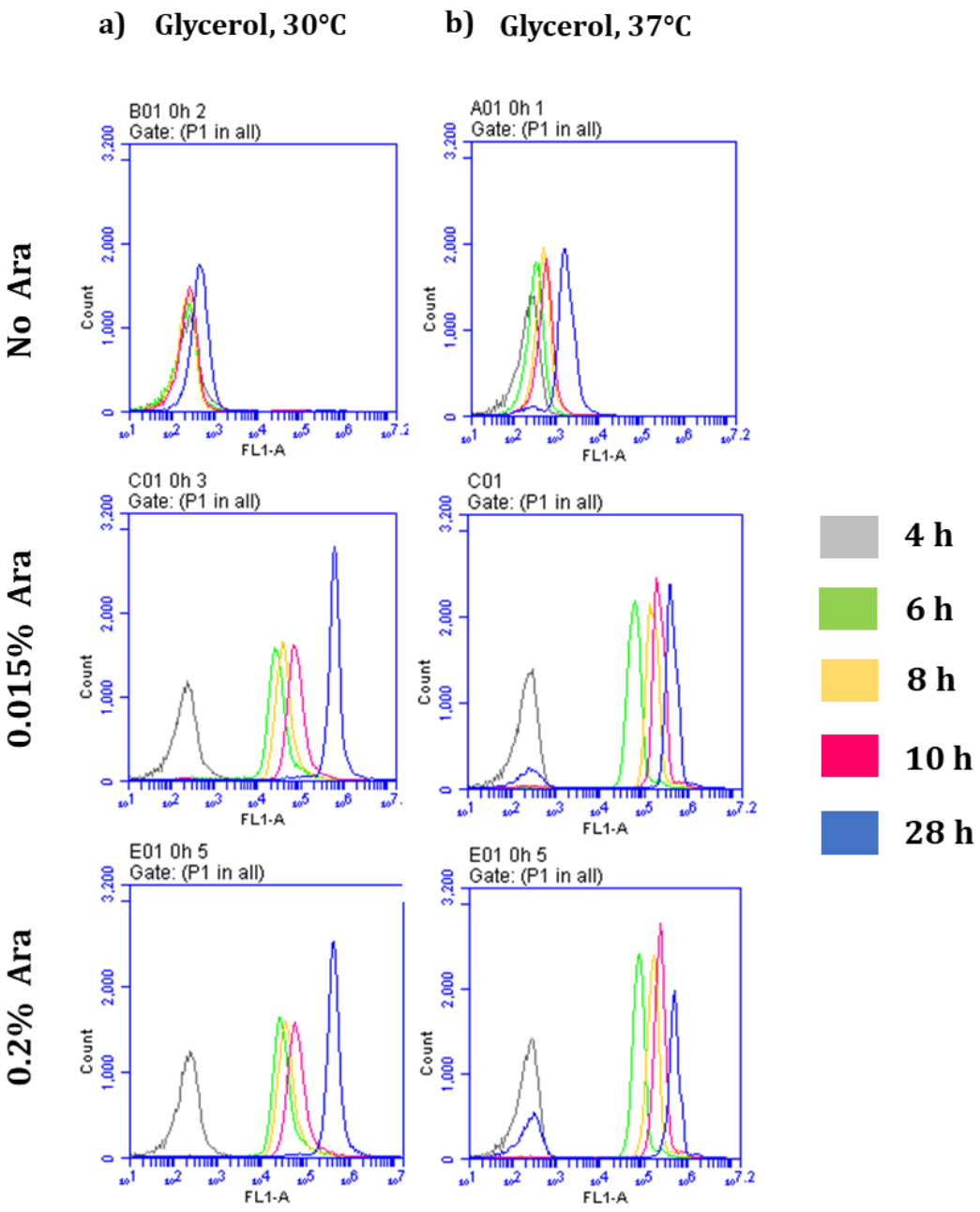


Figure 4.15: Single cell analysis of *araBAD* promoter cultures for the production of GFP

E. coli BL21(A) cultures carrying p*araBAD*-GFP were grown in Terrific Broth with 0.4% glycerol as a carbon source at temperature growth of (a) 30°C and (b) 37°C. At an OD_{650} of 0.5 (approximately after 4 h growth), arabinose (0.015% or 0.2%) was added. At regular intervals, samples were taken and FCM was used to measure green fluorescence of individual cells. Timepoints relate to the time the cultures were grown; 4 h refers to pre-induction with arabinose. Duplicate cultures of each plot showed a similar pattern throughout the growth.

4.4.2 SDS-PAGE and fractionation

The *araBAD* promoter induced cultures grown at 37°C displayed GFP accumulation as early as 2 h after induction (**Figure 4.16b**), indicating that raising the temperature might accelerate the rate at which the *gfp* genes were being transcribed and translated. However, there was a decrease in GFP accumulation at 24 h post-induction, which was clearly seen in cultures induced with 0.015% arabinose. This parallels to the decreased levels of green fluorescence observed in FCM for these cultures, demonstrating once again how effective FCM is in identifying the states of protein solubility. Cultures grown at 30°C exhibited the maximum level of green fluorescence produce at 24 h post-induction (**Figure 4.16a**).

Densitometry data for the percentage ratio of soluble versus insoluble GFP accumulated inside the cell extracts of p*araBAD*-GFP cultures with 0.4% glycerol at 37°C was shown on the SDS-PAGE gel pictures in **Appendix 4.3.**

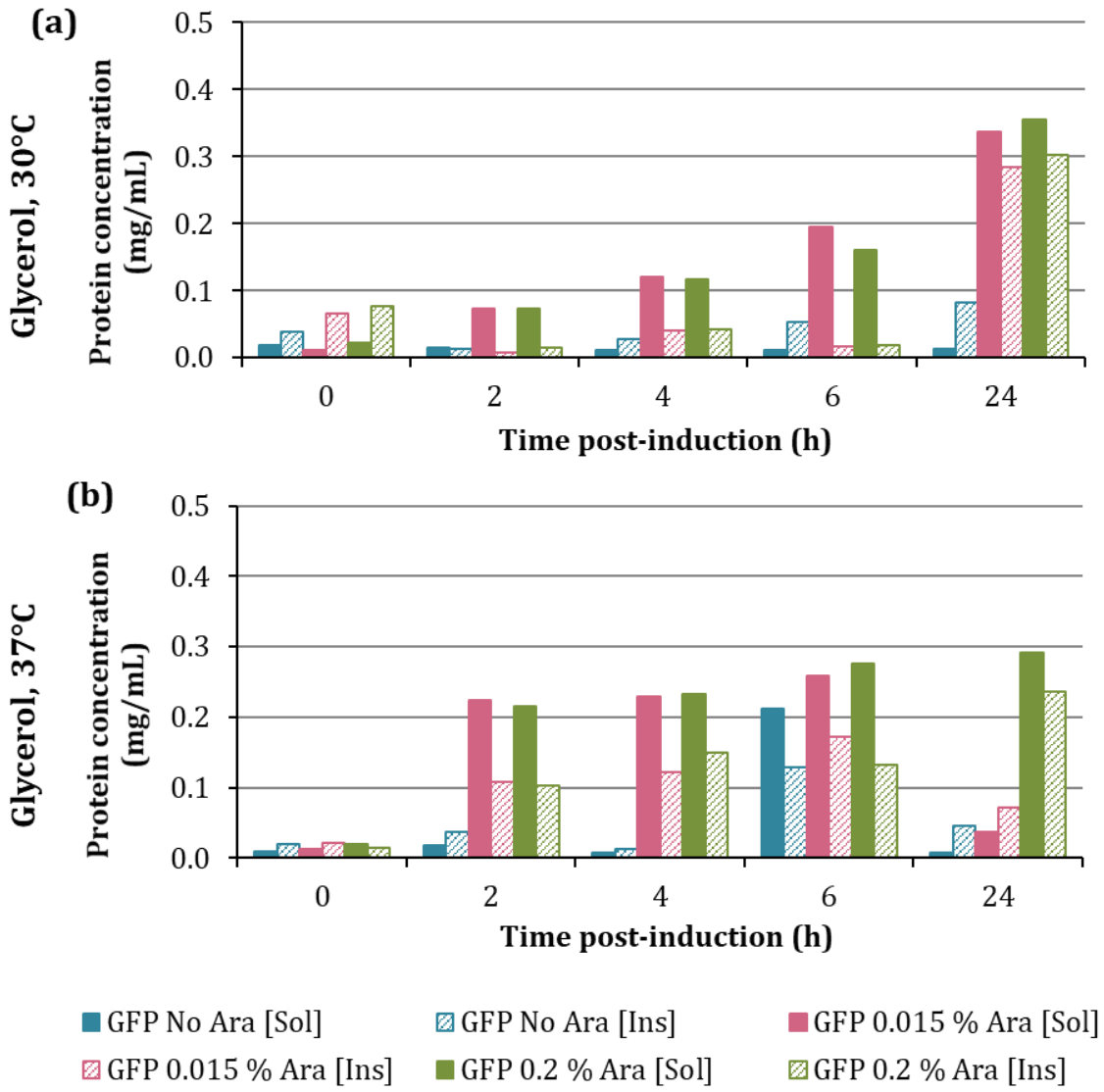


Figure 4.16: The expression of protein solubility in cell extracts measured by SDS-PAGE for p*araBAD*-GFP in batch culture of TB medium with 0.4% glycerol

Estimated protein solubility of soluble and insoluble fractions in cell extracts of p*araBAD*-GFP cultures with 0.4% glycerol as the carbon source at growth temperature of (a) 30°C and (b) 37°C, at concentration of mg/mL at times post-induction.

4.5 General discussion

The p*araBAD* system demonstrated more sustained production of recombinant proteins. The peak fluorescence levels, indicative of high levels of GFP production, were achieved after 28 h of growth in all induced cultures supplemented with either glycerol or glucose as the carbon source. However, it is worth noting that the growth of *araBAD* promoter cultures in the presence of glucose was more significantly repressed.

Similar to the T7 promoter, the *araBAD* promoter is also an arabinose-inducible-promoter, thus arabinose was used as an inducer in this study to boost the expression of GFP in p*araBAD*-GFP cultures. In the presence of glucose as carbon source, glucose repressed *araBAD* promoter (Cronan, 2006). Hence, within cultures containing glucose, an increased amount of arabinose was needed to generate GFP. This was previously addressed in Chapter 3: **Section 3.3.1**, specifically discussing on the phenomenon of 'dampening' as previously reported by (Castiñeiras et al., 2018a). This phenomenon was related to the repression of the *araBAD* promoter by glucose and its activation by arabinose, thereby enabling precise regulation of RPP.

Arabinose was used in this study at concentrations of 0.015% or 0.2%, and p*araBAD*-GFP cultures demonstrated that 0.015% arabinose was sufficient to increase RPP in glycerol-based cultures. Increasing the arabinose concentration did not significantly increase fluorescence, and in certain conditions, it may have even decreased cell growth and protein expression (Lecina et al., 2013).

According to FL1-A histogram single cell fluorescence analysis, insufficient arabinose induction was observed in p*araBAD*-GFP cultures in glucose-based growth medium, as shown by the appearance of 'on and off' states in 0.015% induced cultures (**Figure 4.11b**). The effect was more notably observed in cultures that had a higher glucose concentration of 0.8%, which is indicative of a greater degree of partial induction (**Figure 4.11c**). This finding was very similar to the finding obtained for T7 promoter cultures, which indicated partial induction by the existence of a shoulder in the FL1-A histogram peak. This shoulder became wider by increasing the concentration of glucose in the medium. It was also found that glucose repressed *araBAD* promoter more than T7 system.

Our findings suggested that the *araBAD* promoter provides tighter control of RPP in the p*araBAD* system, with green fluorescence being observed exclusively following induction. Despite the higher level of leakage in the T7 system, it was observed that uninduced cultures might exhibit mean fluorescence values comparable to that of induced cultures after 28 h of growth.

According to this study's observations, p*araBAD*-GFP cultures typically contained a single population of productive GF-producing (GFP+) cells following induction, which also demonstrated the *araBAD* promoter's strict regulation. As opposed to T7 promoter cultures, which demonstrated multi population, which illustrates the heterogeneity of the cultures, particularly at 24 h post-induction and reflects the promoter's leakiness. Since T7 system showed leaky expression, this promoter can be problematic for protein expression (Jia and Jeon, 2016). This demonstrates the *araBAD* promoter's potency in tightly controlling protein expression, which results

in greater control over RP synthesis. Indeed, the strict regulation of the p*araBAD* expression system is favourable (Guzman et al., 1995). The determination of the presence of both single population and multi population, which indicates the heterogeneity of cultures, can potentially be efficiently achieved by the utilisation of FCM (Wyre and Overton, 2014; Fernández-Castané et al., 2017).

Furthermore, compared to the T7 system, the p*araBAD* system exhibits a reduced impact on cellular physiology, leading in a far fewer dead cells. The p*araBAD* system exhibited superior tolerance and stability on cellular viability compared to the T7 system. Based on our study findings, the utilisation of the p*araBAD* system has also been observed to improve plasmid retention, but with a slight reduction observed at the temperature of 37°C. The reduction in culturability of *araBAD* promoter cultures when grown at the temperature of 37°C as shown by CFU (**Figure 4.14d**), the increase in plasmid loss (**Figure 4.14f**) and the larger non-productive GFP- population observed at the end of the growth (**Figure 4.16b**), was indicative of higher stress levels and a greater metabolic burden when these cultures underwent growth at higher temperatures (Castiñeiras and Overton, 2018; Overton, 2014).

The findings of our study revealed that, in the majority of the conditions investigated, the amounts of protein concentrations accumulated in the cultures, either soluble or insoluble, as indicated by subcellular fractionation and SDS-PAGE, accurately corresponded to the green fluorescence levels measured by FCM. This finding further demonstrates the efficacy of FCM in precisely determining the different stages of protein solubility.

4.6 Conclusions and future work

In general, the optimisation of the conditions for GFP production in p*araBAD* expression system was successful. Specifically, the use of glycerol instead of glucose, a lower temperature of 30°C instead of 37°C and a lower concentration of arabinose (0.015%) were shown to be sufficient to boost GFP in glycerol-based medium. These optimal conditions obtained for the GFP production in p*araBAD* expression system is very similar to the optimal conditions obtained for the production of GFP and TNFα-GFP in T7 expression system. In conclusion, the p*araBAD* system affords more sustained recombinant protein production, tighter control, greater glucose repression and a lesser impact on the cellular physiology, resulting in far fewer dead cells.

Further investigation was conducted to grow cultures expressing GFP and TNFα-GFP in fed-batch fermentation at 5 L scale following shake-flask growth. This approach that utilising the T7 and p*araBAD* expression system, were intensified in fed-batch bioreactors for higher cell densities for RPP. The optimal conditions obtained in the previously discussed shake-flask growth experiments will be used for the production of GFP and TNFα-GFP in fed-batch fermentation. The purpose of this investigation was to evaluate the potential advantages of the T7 and p*araBAD* expression system in boosting the RPP performance. Additionally, the study aimed to analyse the effects of utilising these systems at a higher cell density on cellular physiology and viability.

CHAPTER 5

FED-BATCH FERMENTATION

For the scale-up production of GFP and TNFα-GFP cultures in intensified fed-batch fermentation at a 5 L scale of bioreactor, we constructed high cell density fermentations in this study applying the optimal conditions identified from the previously discussed shake-flask growth studies. RPP optimisation studies are frequently performed at a larger pilot scale after being conducted at the bench scale in the lab for industrial purposes. The aim of our study on fed-batch fermentations were to obtain optimally high biomass production, a high proportion of productive bacteria, a high yield of RP, and improved solubility of the produced RP.

5.1 Introduction

The investigation included the conducting of several fed-batch series of experiments utilising either the T7 expression system or the p*araBAD* expression system, aimed for the production of two recombinant proteins; GFP and human Tumour necrosis factor alpha fused to GFP (TNFα-GFP). Protocols were developed following shake flask growth; cultures that expressing GFP and TNFα-GFP were intensified in fed-batch bioreactors with the use of a semi-defined medium and glycerol feed. The chosen protocol (Wyre and Overton, 2014b) was adapted an industrially derived-improved protocol that has been utilised successfully on numerous times in the past (Want et al., 2009) . A few adjustments of parameters were made to fermentation conditions according to the optimal conditions that we acquired from batch culture optimisation.

Glycerol was chosen as a carbon source rather than glucose as it normally minimises overflow metabolism and therefore acid production, allowing it to be used in high quantities (Retamal et al., 2018). The optimal glycerol concentration identified in the batch culture studies was deemed to be equivalent to the amount of glycerol utilised in fed-batch fermentation. Glycerol and $MgSO_4$ were the only ingredients in the feed composition, and arabinose induction was carried out concurrently with feeding at mid-logarithmic (mid-exponential) phase (OD_{650} ≈ 40-50). When the DOT increased, suggesting the depletion of initial carbon source, feeding began at a rate of 67.5 mL/h. Induction with arabinose at concentration of 400 mM was comparable to the lower concentration of arabinose induced in shake flask growth. Trace metal solution was introduced in order to meet the nutritional needs of denser cultures. A temperature of 30°C was chosen because it was thought that a lower temperature could reduce cellular stress by slowing the rate of transcription and translation, as evidenced by our previous investigated studies conducted in batch culture.

The cultures were grown in a FerMac 310/60 bioreactor (Electrolab) 5 L (Details in **Section 2.7.1**) containing 1.5 L initial batch medium of semi-defined medium supplemented with 35 g/L glycerol and 50 µg/mL Kanamycin for plasmid retention, inoculated with 2% (v/v) overnight culture of transformed *E. coli* BL21(T7) or BL21(A) grown at 30°C throughout. Fermentations were terminated once no more growth was observed at 24-26 h post induction. Considering this was a pilot study, a single bioreactor growth experiment was carried out for each of GFP and TNFα-GFP cultures.

Dissolved oxygen tension (DOT) was measured in situ using D150 Oxyprobe, maintained above the set point of 30% by increasing agitation to maximum 800 rpm from a minimum of 300 rpm. pH was measured by Ferm Probe, controlled at a set point 6.3 with automated addition of sterile 10% (v/v) ammonia (NH_3) or 5% (v/v) hydrochloric acid (HCl). Analysis of OD_{650} and/or DOT/off-gas was done to determine the stage of cell growth. The recipe of semi-defined growth medium, trace metal solution, feed composition and feeding rate equation were summarised in **Tables 2.4** and **Table 2.5** (**Section 2.7.2**)**.** The utilisation of FCM was applied as an approach for single-cell analysis in order to optimise the fermentations in context of bacterial physiology and productivity.

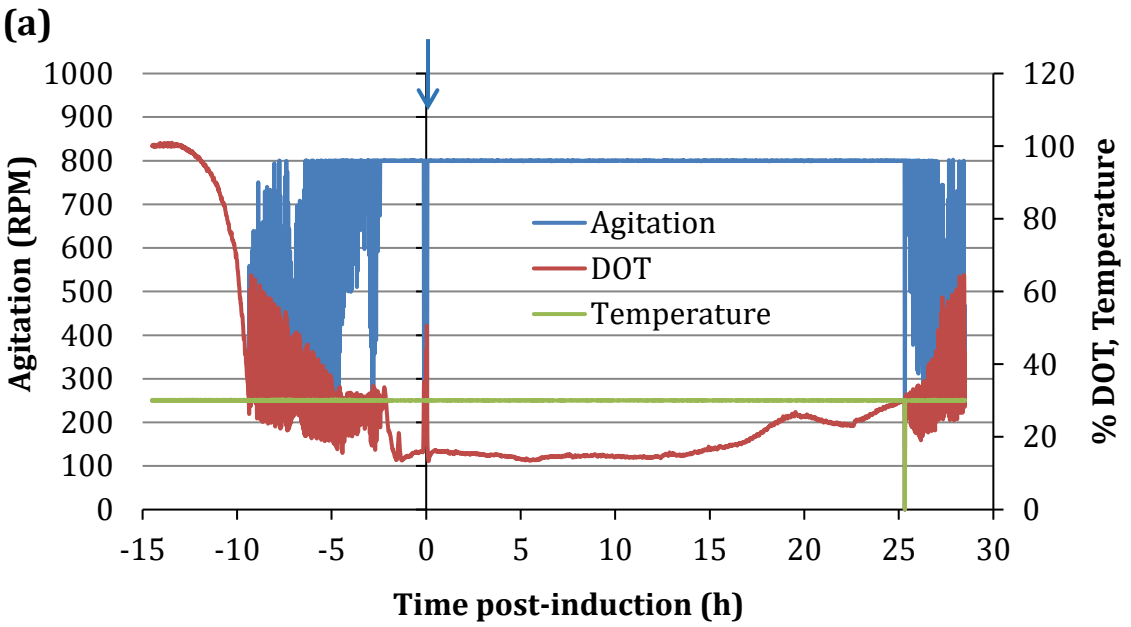
5.2 Fed-batch of BL21(T7)-pORT(T7)-GFP fermentation

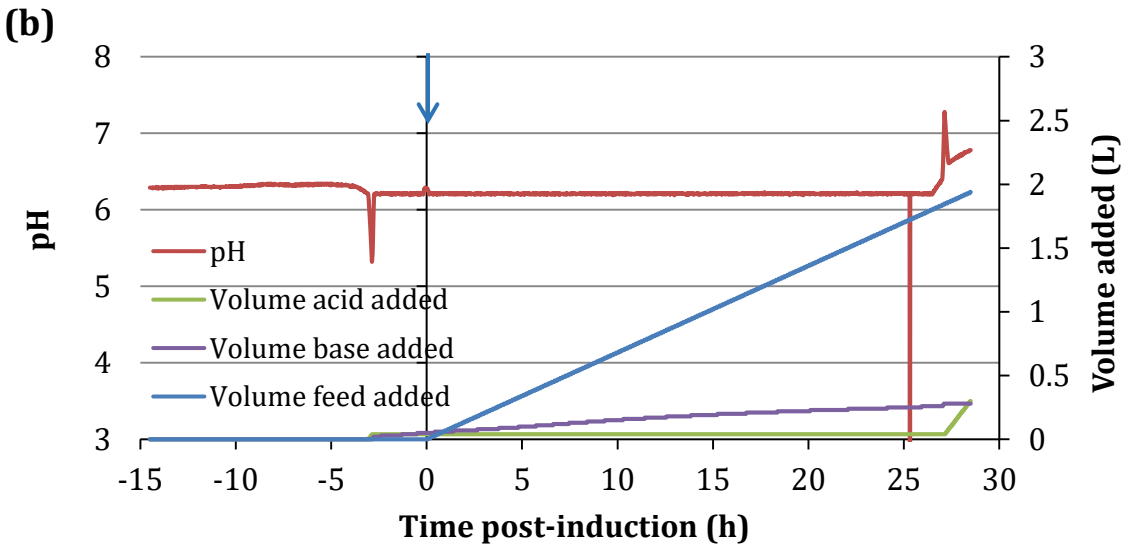
The first fed-batch experiment was the growth of BL21(T7)-pORT(T7)-GFP cultures in a 5 L bioreactor. Feeding began when there was an increase in the DOT, reflecting oxygen demand drop, indicating depletion of initial carbon source (approximately 14 h post-inoculation). Cultures were induced with arabinose at the same time as feeding started (mid logarithmic induction), at an OD_{650} of around 40 (**Figure 5.1c**) when the initial glycerol in the batch medium was exhausted. The DOT remained high at the start of growth and then gradually decreased over time as biomass and thereby oxygen demand increased.

The agitation increased when the DOT reached 25% to increase DOT. The DOT was maintained above 15% by the agitation at 800 rpm. After 25 h post-induction, the DOT increased and the agitation rate decreased as oxygen demand fell, indicating lower metabolic activity (**Figure 5.1a**).

Data obtained from pH probe and pump volume showed that pH remained at the set point of 6.3±1 at the beginning of fermentation. Approximately 3 hours before to induction, there was an unexpected drop in pH to 5.32, resulting in the adjustment of the initial point for introducing bases into the vessel medium (**Figure 5.1b**).

Corroborating to online measurements of FerMac controller data, the growth of pORT(T7)-GFP cultures seemed to be slower at the beginning of incubation, followed by a steady increase upon feeding (**Figure 5.2a**), reaching the final OD_{650} of 133. The observation of the CFU showed a reduction throughout the times post-induction, but slightly increased at the termination, reaching 9.0x108 CFU/mL. The gradual decrease in CFU measurement indicated that the cultures within the bioreactor had a negative impact on its ability to grow on agar plates. This phenomenon had previously been observed and was believed to be a result of the bacteria adapting to the growth conditions in the bioreactor, leading to the emergence of a VBNC phenotype (Wyre and Overton, 2014a). The pORT(T7)-GFP cultures in the bioreactor showed plasmid loss of more than 20% after 4 h of induction (**Figure 5.2b**).


(a)
Agitation (RPM)
% DOT, Temperature
Agitation
DOT
Temperature
Time post-induction (h)


(b)
pH
Volume added (L)
pH
Volume acid added
Volume base added
Volume feed added
Time post-induction (h)

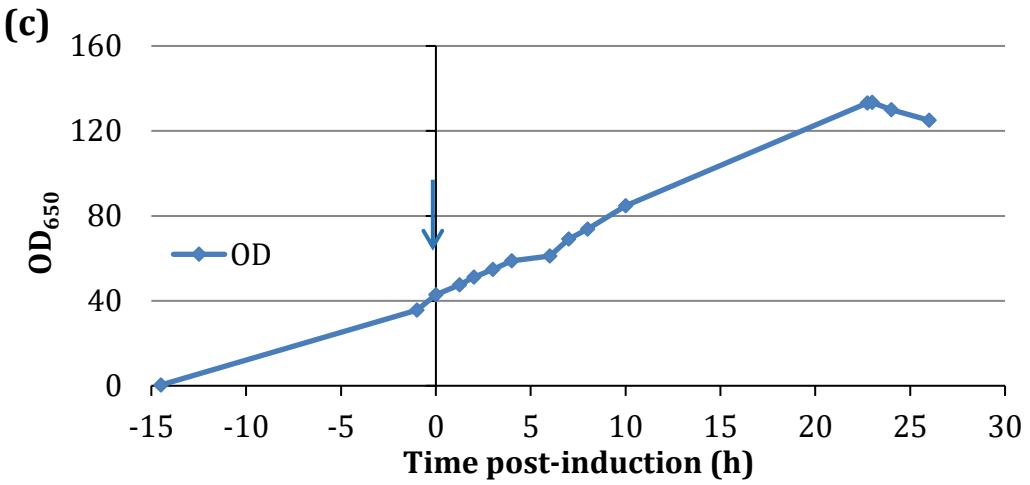

(c)
OD650
OD
Time post-induction (h)

Figure 5.1: FerMac controller data of *E. coli* BL21(T7)-pORT(T7)-GFP growth in fed-batch fermentation for GFP production

E. coli BL21(T7) cultures carrying the plasmid pORT(T7)-GFP were grown at 30°C in 5 L vessel and induced with arabinose at mid exponential phase at OD_{650} ≈ 40-50 (approximately after 14 h post-inoculation; indicated by blue arrow), concurrently with feeding at a rate of 67.5 mL/h.

(a) The data of agitation (blue) and percentage of DOT (red), with the temperature of 30°C (green) throughout the fermentation. (b) pH (red), volumes of acid (green), base (purple) and feed (blue) added to the vessel. (c) OD_{650} (blue) measurement at time points intervals.

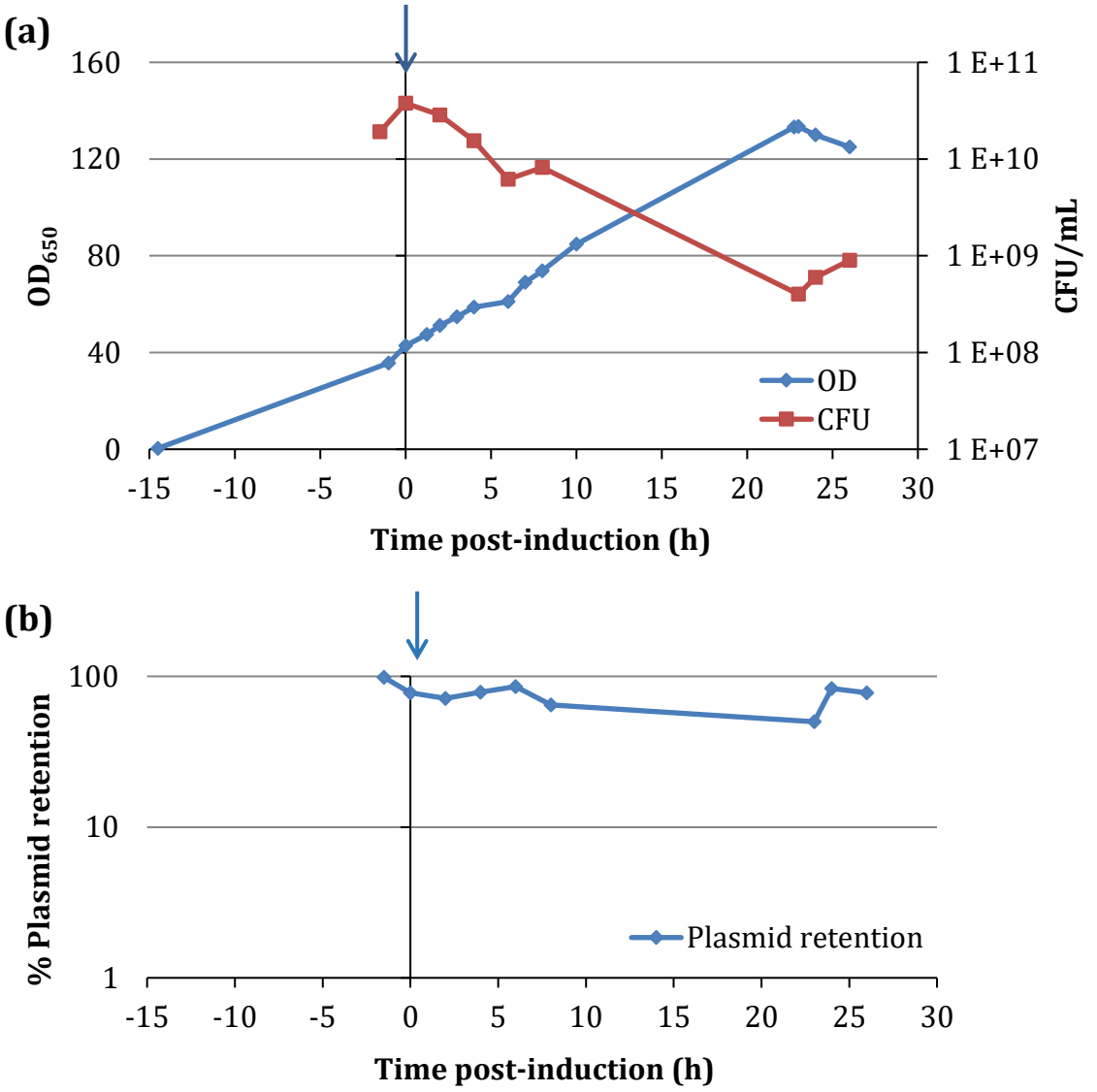


Figure 5.2: Correlation of *E. coli* BL21(T7)-pORT(T7)-GFP growth, viability and plasmid retention in fed-batch culture for GFP production

E. coli BL21(T7) cultures carrying the plasmid pORT(T7)-GFP were grown at 30°C in 5 L vessel and induced with arabinose at mid exponential phase at OD_{650} ≈ 40-50 (approximately after 14 h post-inoculation; indicated by blue arrow), concurrently with feeding at a rate of 67.5 mL/h. (a) The OD_{650} was measured at intervals and serial dilution of culture samples was plated onto NA agar for indication of cell viability (CFU/mL counting), and (b) were replica-plated onto NA agar supplemented with Kanamycin (50 µg/mL) to reveal plasmid retention.

5.2.1 FCM for fed-batch culture of pORT(T7)-GFP

FCM was used to measure the average amount of GFP produced by each cell over time following induction. According to FCM analysis, the mean green fluorescence per cell steadily increased after induction (**Figure 5.3a**), reaching a maximum of 400, 000 fluorescence units at 24 h post-induction. As previously reported (Wyre and Overton, 2014a), the mean forward scatter FSC-A, which served as a measure of cell size, increased following induction. Throughout the fermentation, there was just a small number of non-productive (GFP-) cells (**Figure 5.3b**). The proportion of cells that exhibited GFP expression (GFP+) remained consistently over 98% following induction. The results obtained from PI staining indicated that the proportion of dead cells remained constantly below 6%, with the exception of the 23 h-time point following induction. The absence of non-productive cells that produced GFP, along with the low presence of dead cells, indicates that the bacteria were active and physiologically 'healthy'.

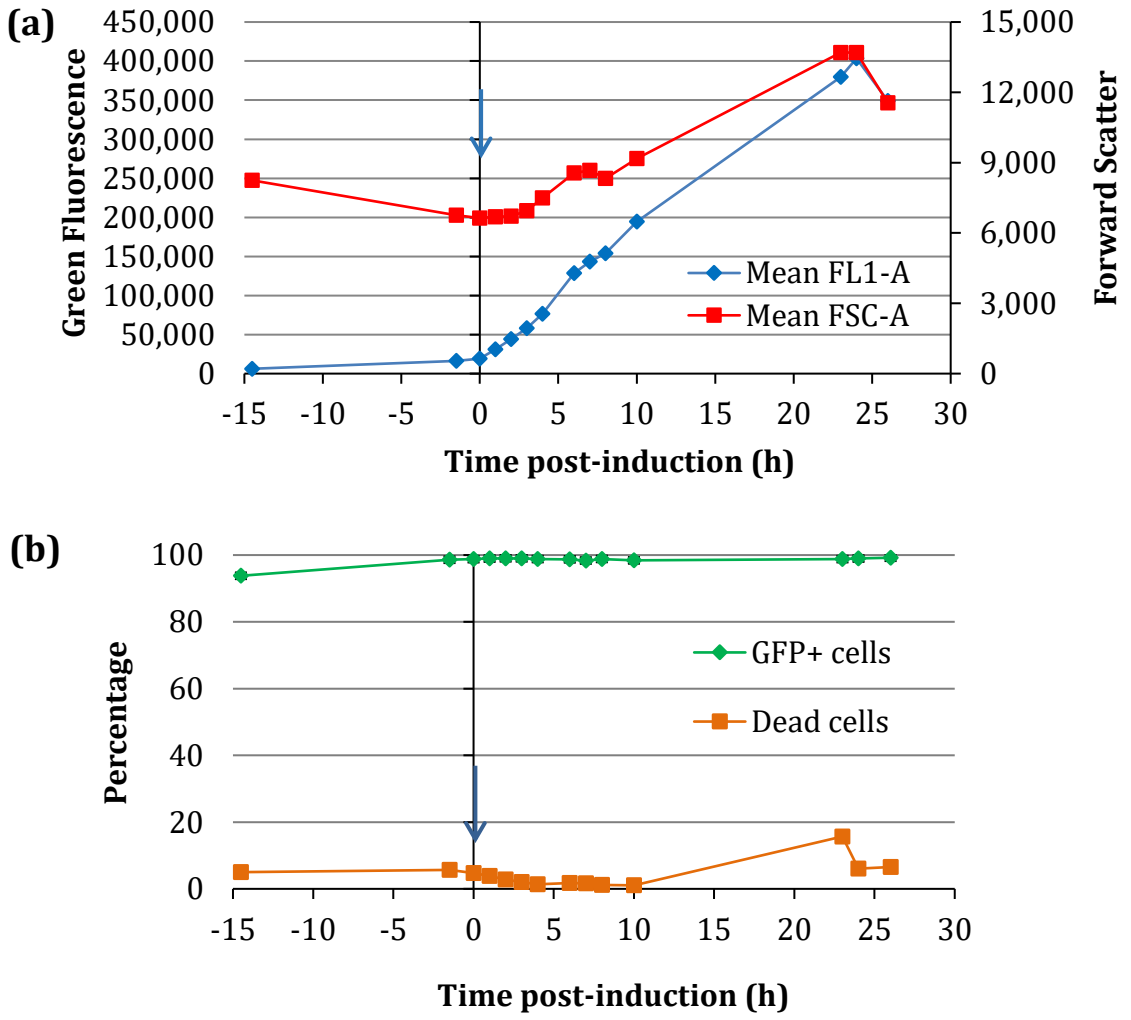


Figure 5.3: FCM data measured for *E. coli* BL21(T7)-pORT(T7)-GFP growth in fed-batch culture for GFP production

Green fluorescent protein (GFP) was expressed from an arabinose-induced T7 promoter system; induction at mid exponential phase at OD_{650} ≈ 40-50 (approximately after 14 h post-inoculation; indicated by blue arrow), concurrently with feeding at a rate of 67.5 mL/h. (a) Mean green fluorescence (Mean FL1-A) versus mean forward scatter (Mean FSC-A) plots of cells taken from culture samples at interval time points of incubation. (b) Percentage of cells that producing GFP (GFP+ cells) and percentage of dead cells (PI+ cells) versus time.

5.2.2 SDS-PAGE for fed-batch culture of pORT(T7)-GFP

Densitometry data for the percentage of GFP that was in the soluble or insoluble fraction was shown in **Figure 5.4** and expressed as estimated protein concentrations (mg/mL) in cell extracts (**Figure 5.5**). Overall, the soluble GFP produced was increased over the time. The higher the concentration of soluble GFP were produced throughout the fermentation, the lesser the insoluble form. The soluble GFP was highest measured at the end of fermentation with 0.86 mg/mL, reaching a final solubility of 96% of total GFP produced. Yield of GFP accumulation was lower at the initial stages of feeding, but then gradually increased over time.

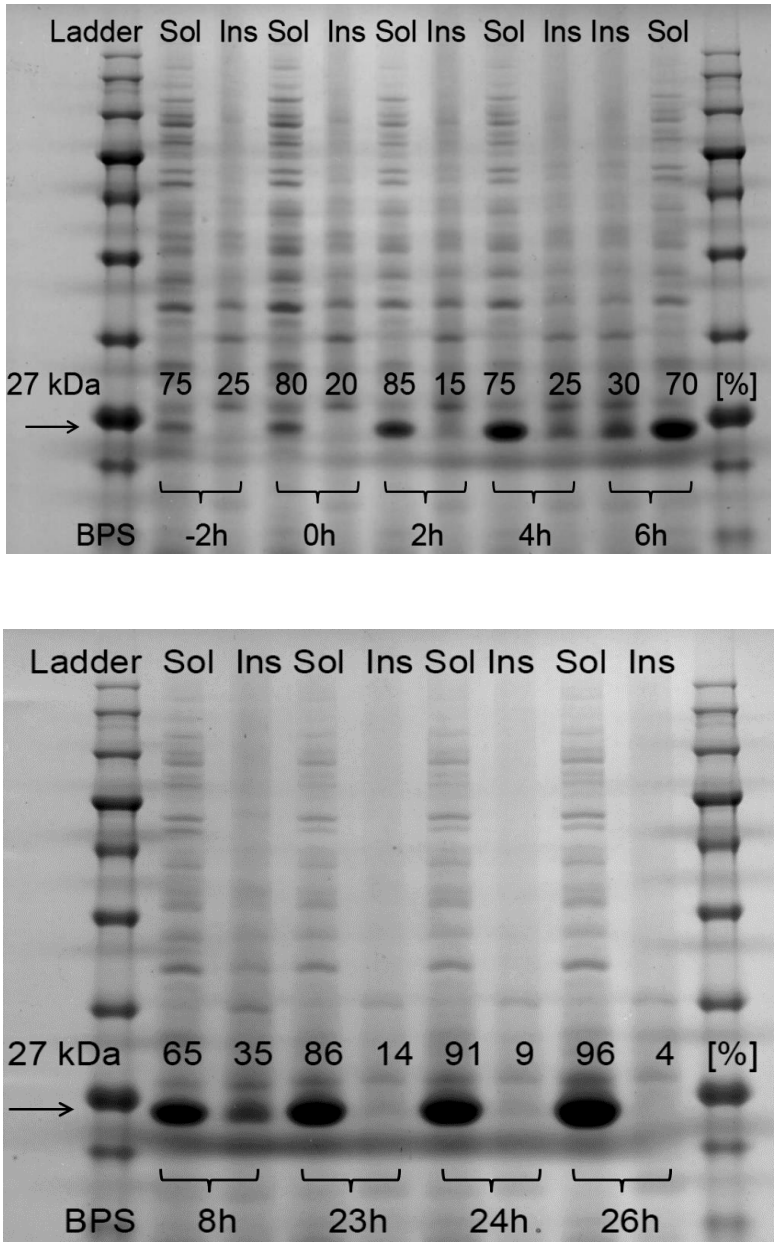


Figure 5.4: The accumulation of GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in fed-batch of pORT(T7)-GFP cultures

SDS-PAGE gel picture with arrow showing GFP (27 kDa) and the ratio percentage of soluble versus insoluble GFP proteins accumulated inside the cell extracts of pORT(T7)-GFP, as shown above the band. All sample bands at times post-induction were compared to BPS ladder band. [**BPS**: Marker used (Blue Protein Standard), broad range of 11-190 kDa; **Sol**: Soluble GFP; **Ins**: Insoluble GFP]

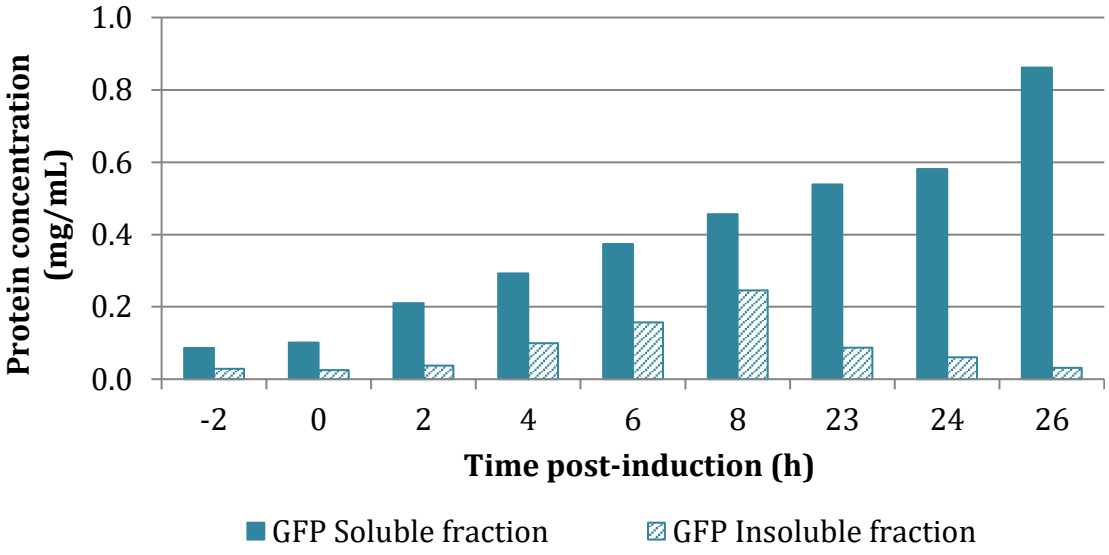


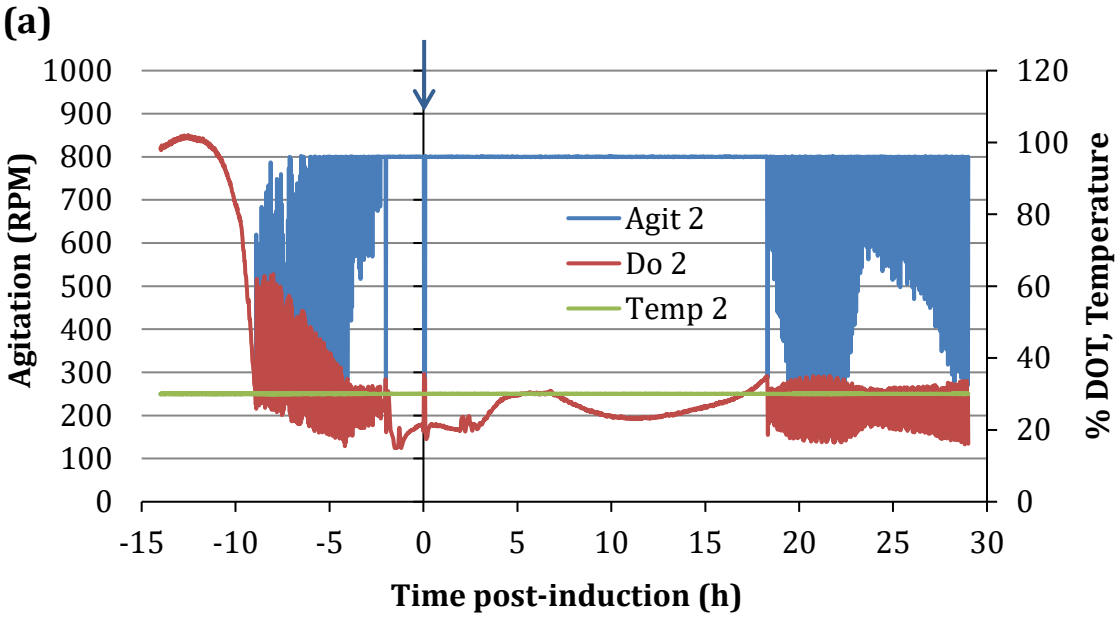
Figure 5.5: The expression of protein solubility in cell extracts measured by SDS-PAGE for GFP production in fed-batch cultures of pORT(T7)-GFP

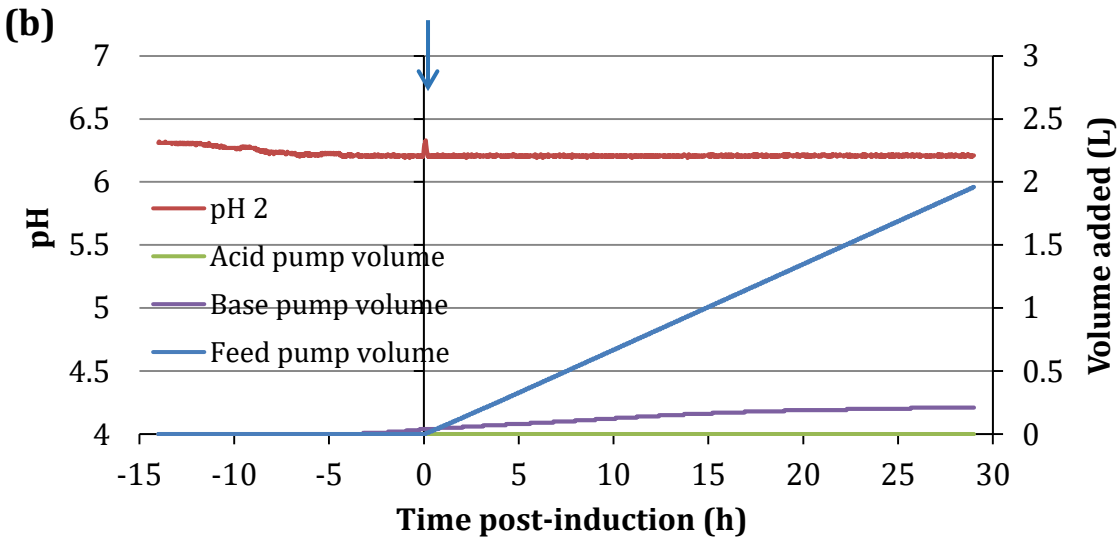
Estimated protein solubility of soluble and insoluble fractions in cell extracts of GFP at concentration of mg/mL at times post-induction.

5.3 Fed-batch of BL21(T7)-pORT(T7)-TNFα-GFP fermentation

It was essential to find out whether the same fed-batch environment might boost the productivity of TNFα-GFP in pORT(T7)-TNFα-GFP cultures after encouraging findings for the production of greater soluble GFP were obtained. Feeding began when the DOT increased, indicating depletion of glycerol in the initial batch medium at 11-12 h post-inoculation (**Figure 5.6a**). The agitation and DOT data showed increased oxygen consumption levels in the period between 2 h before induction and 18 h post-induction. It was presumed to correlate with an increase in cell metabolic activity under batch growth and feeding, and in turn the deficiency of that metabolic activity at termination due to carbon source exhaustion and transition of cells into stationary phase.

At 0 h pre-induction the pH was slightly increased, indicating glycerol supplied was not enough for the cell concentration presented (**Figure 5.6b**). Cultures were induced with arabinose at the same time as feeding at an OD_{650} of around 44 (**Figure 5.6c**). Following induction, the pH remained under the pH 6.3 until termination, suggesting acid/ base addition for pH regulation. The growth of pORT(T7)-TNFα-GFP cultures increased well in fed-batch fermentation up to 24 h post-induction, reaching a higher final OD_{650} (>200; **Figure 5.7a)** than the GFP cultures. CFU counts decreased at the early stages of feeding, then increased to $7.5x10^{10}$ CFU/mL nearly termination (**Figure 5.7a**). These cultures also showed plasmid loss of more than 50% after 6 h of induction (**Figure 5.7b**). This plasmid loss could disturb the physiological balance of the cells, placing an excessive metabolic burden on the cultures.


(a)
Agitation (RPM)
% DOT, Temperature
Agit 2
Do 2
Temp 2
Time post-induction (h)


(b)
pH
Volume added (L)
pH 2
Acid pump volume
Base pump volume
Feed pump volume
Time post-induction (h)

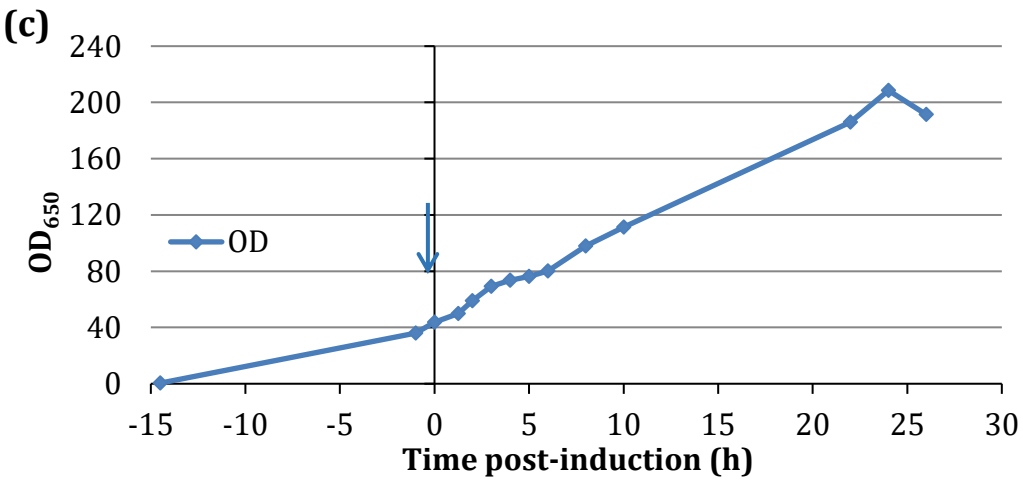

(c)
OD650
OD
Time post-induction (h)

Figure 5.6: FerMac controller data of *E. coli* BL21(T7)-pORT(T7)-TNFα-GFP growth in fed-batch fermentation for TNFα-GFP production

E. coli BL21(T7) cultures carrying the plasmid pORT(T7)-TNFα-GFP were grown at 30°C in 5 L vessel and induced with arabinose at mid exponential phase at OD_{650} ≈ 40-50 (approximately after 14 h post-inoculation; indicated by blue arrow), concurrently with feeding at a rate of 67.5 mL/h.

(a) The data of agitation (blue) and percentage of DOT (red), with the temperature of 30°C (green) throughout the fermentation. (b) pH (red), volumes of acid (green), base (purple) and feed (blue) added to the vessel. (c) OD_{650} (blue) measurement at time points intervals.

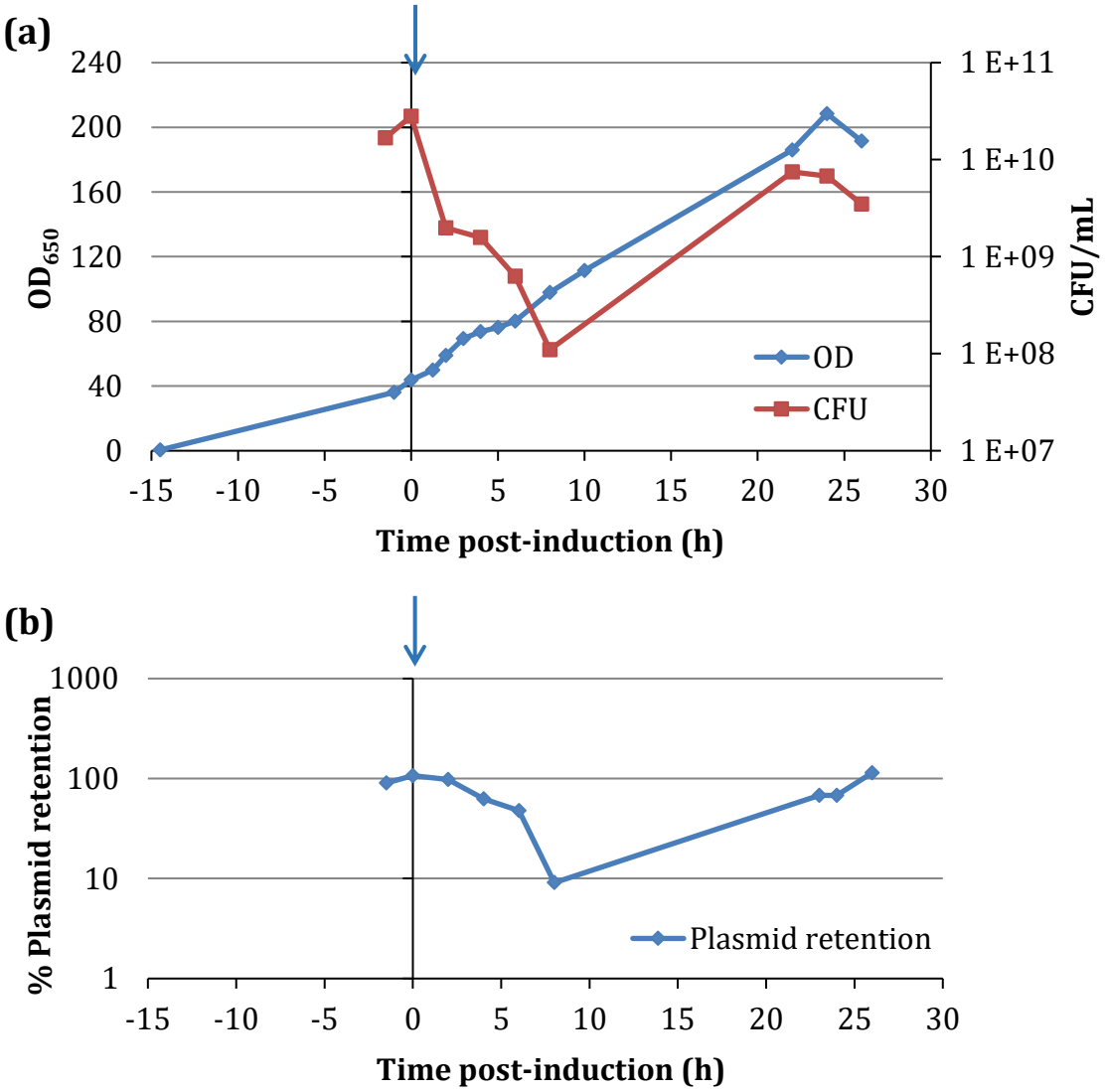


Figure 5.7: Correlation of *E. coli* BL21(T7)-pORT(T7)-TNFα-GFP growth, viability and plasmid retention in fed-batch culture for TNFα-GFP production

E. coli BL21(T7) cultures carrying the plasmid pORT(T7)-TNFα-GFP were grown at 30°C in 5 L vessel and induced with arabinose at mid exponential phase at OD_{650} ≈ 40-50 (approximately after 14 h post-inoculation; indicated by blue arrow), concurrently with feeding at a rate of 67.5 mL/h. (a) The OD_{650} was measured at intervals and serial dilution of culture samples was plated onto NA agar for indication of cell viability (CFU/mL counting), and (b) were replica-plated onto NA agar supplemented with Kanamycin (50 µg/mL) to reveal plasmid retention.

5.3.1 FCM for fed-batch culture of pORT(T7)-TNFα-GFP

The culture of T7 promoter expressing TNFα-GFP attained a lower mean green fluorescence (150, 000 fluorescence unit; **Figure 5.8a**) than the GFP culture (400, 000 fluorescence unit) although the final OD_{650} was higher. Forward scatter FSC-A substantially increased during simultaneous induction and feeding, parallel to the mean green fluorescence generated, indicating an increase in cellular size of the culture when expressing TNFα-GFP. Although considerably less green fluorescence was produced, the data mean green fluorescence versus forward scatter displayed a pattern similar to that observed in pORT(T7)-GFP cultures.

Consistent with the data obtained when generating GFP, a limited proportion of non-productive (GFP-) cells were visible in pORT(T7)-TNFα-GFP culture. The proportion of cells that exhibited GFP expression (GFP+) remained consistently over 98% following induction, with the exception of the time point 2 h post-induction, with 93% were GFP+ cells (**Figure 5.8b**). This indicates that the observed plasmid loss following induction (**Figure 5.7b**) can be attributed to the assay's reliance on growth on agar plates. This drawback is especially crucial when conducting plasmid retention assay with cultures that include a VBNC population.

PI staining analysis revealed that the proportion of dead cells was very low throughout fermentation until approximately 23 h post-induction, then increased up to 27% upon ceased of cell growth, higher than observed in shake flasks.

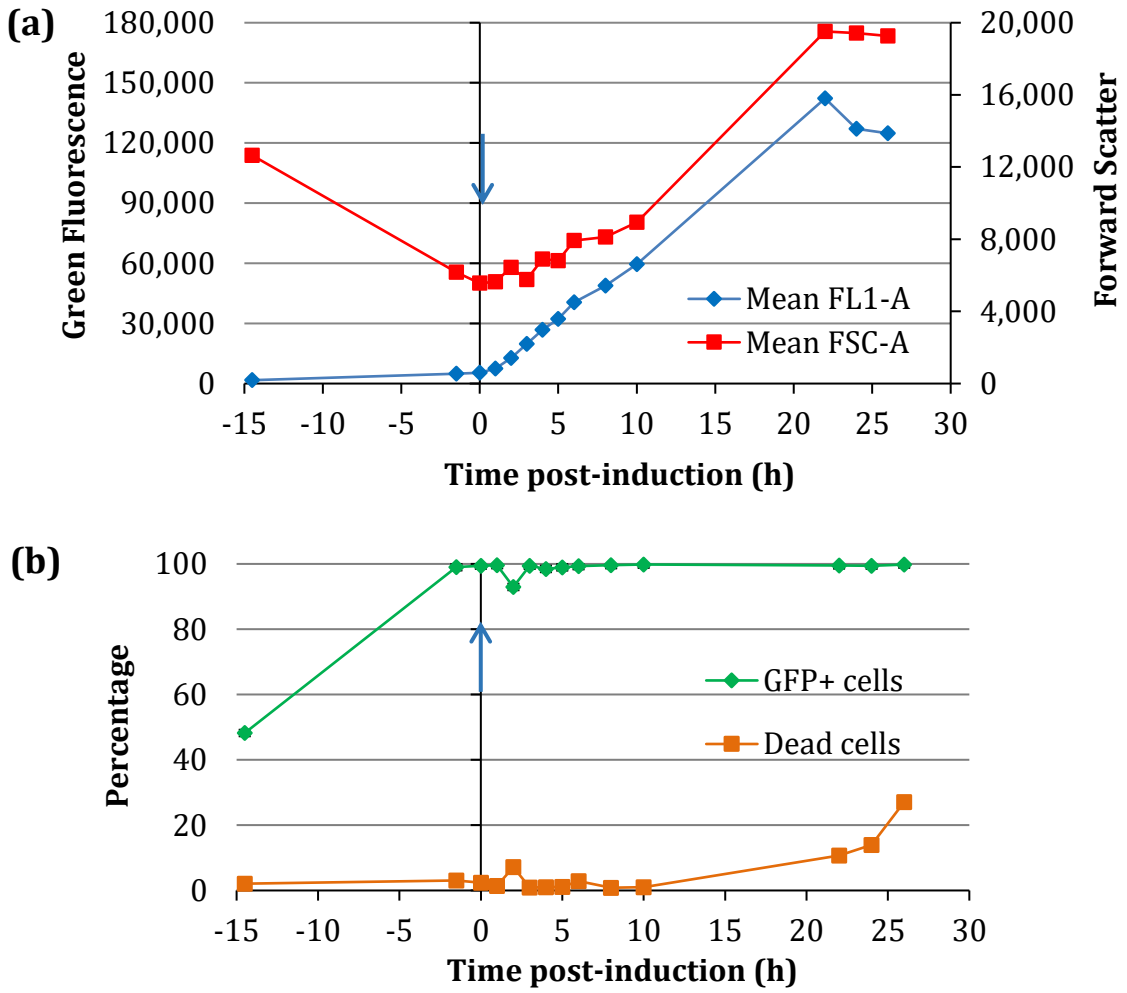


Figure 5.8: FCM data measured for *E. coli* BL-21(T7)-pORT(T7)-TNFα-GFP growth in fed-batch culture for TNFα-GFP production

Recombinant human TNFα fused to GFP (rhTNFα-GFP) was expressed from an arabinose-induced T7 promoter system; induction at mid exponential phase at OD_{650} ≈ 40-50 (approximately after 14 h post-inoculation; indicated by blue arrow), concurrently with feeding at a rate of 67.5 mL/h. (a) Mean green fluorescence (Mean FL1-A) versus mean forward scatter (Mean FSC-A) plots of cells taken from culture samples at interval time points of incubation. (b) Percentage of cells that producing TNFα-GFP (GFP+ cells) and percentage of dead cells (PI+ cells) versus time.

In general, the observed data exhibited an apparent similarity to the TNFα-GFP shake flask cultures (**Figure 3.2a**), in which the mean green fluorescence levels was comparatively lower than that of the GFP cultures, indicating that the production and folding of GFP is easier and more efficient than that of TNFα-GFP.

5.3.2 SDS-PAGE for fed-batch culture of pORT(T7)-TNFα-GFP

Densitometry data for the percentage of TNFα-GFP that was in the soluble or insoluble fraction was shown in **Figure 5.9** and expressed as estimated protein concentrations (mg/mL) in cell extracts (**Figure 5.10**).

Overall, the TNFα-GFP production was seemed to gradually increased over the time but mostly insoluble. Although pORT(T7)-TNFα-GFP cultures showed the high percentage of GFP+ cells of more than 98% throughout the fermentation, but still did not increase green fluorescence measured. This is probably due to misfolding of TNFα-GFP and illustrates the fact that TNFα-GFP production and folding is less straightforward than that of GFP.

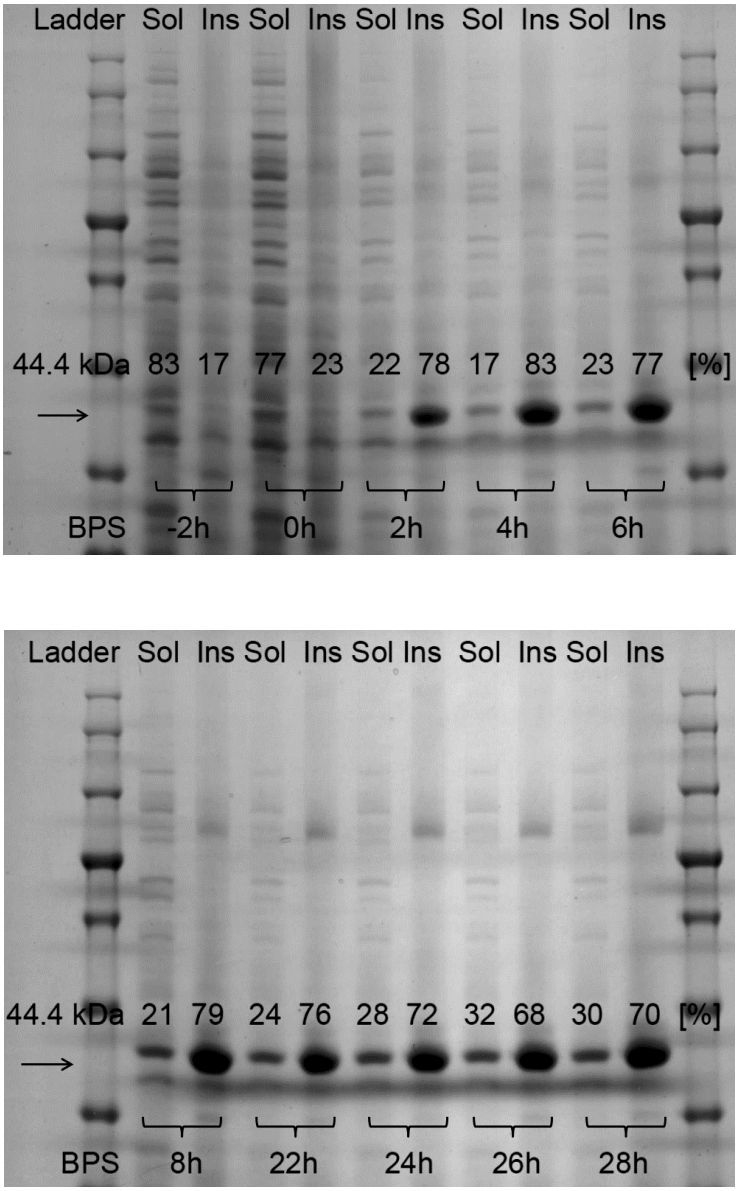


Figure 5.9: The accumulation of TNFα-GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in fed-batch of pORT(T7)-TNFα-GFP cultures

SDS-PAGE gel picture with arrow showing TNFα-GFP (44.4 kDa) and the ratio percentage of soluble versus insoluble TNFα-GFP proteins accumulated inside the cell extracts of pORT(T7)-TNFα-GFP **, as** shown above the band. All sample bands at times post-induction were compared to BPS ladder band. [**BPS**: Marker used (Blue Protein Standard), broad range of 11-190 kDa; **Sol**: Soluble TNFα-GFP; **Ins**: Insoluble TNFα-GFP]

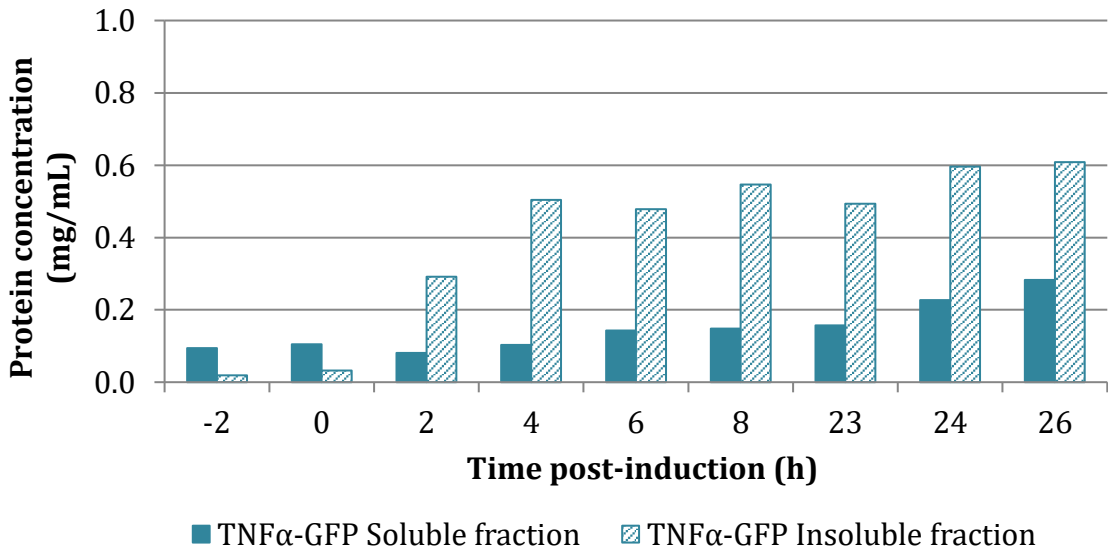


Figure 5.10: The expression of protein solubility in cell extracts measured by SDS-PAGE for TNFα-GFP production in fed-batch cultures of pORT(T7)-TNFα-GFP

Estimated protein solubility of soluble and insoluble fractions in cell extracts of TNFα-GFP at concentration of mg/mL at times post-induction.

5.4 Fed-batch of BL21(A)-p*araBAD*-GFP fermentation

The last fed-batch fermentation was the growth of *E. coli* BL21(A)-p*araBAD*-GFP culture in the 5 L bioreactor and grown at 30°C throughout. Feeding began when DOT increased, reflecting the depletion of initial glycerol in the batch medium, approximately 14 h post-inoculation. Culture was induced with arabinose simultaneously when feeding started (mid logarithmic induction), at an OD_{650} of around 40 (**Figure 5.11c**).

Similar to the observation of prior T7 fermentation cultures, the DOT exhibited an initial high level during the early stages of growth, followed by a gradual decrease as the biomass and subsequent oxygen demand rose. The agitation increased as the DOT reached 24% to increase DOT. The agitation at 800 rpm effectively maintained the DOT at a level above 15% until approximately 20 h following induction, in which the DOT exhibited a rise, while the agitation rate demonstrated a decrease, indicating a decline in oxygen demand. These observations suggest a reduction in metabolic activity (**Figure 5.11a**).

Data obtained from pH probe and pump volume showed that there was an instantaneous blip occurred to the pH, that was observed to drop to 5.35 at approximately 6 h post-inoculation, and that was the starting point at which the bases were added into the vessel medium (**Figure 5.11b**). pH remained at the set point of 6.3±1 following induction throughout the fermentation, and started to increase at 20 h post-induction.

Corroborating to online measurements of FerMac controller data, the growth of p*araBAD*-GFP cultures seemed to be slower at the beginning of incubation, followed by a steady increase upon feeding (**Figure 5.12a**), reaching the final OD_{650} of 85 at 22 h post-induction. The growth of p*araBAD*-GFP culture was the lowest compared to the other fed-batch fermentation cultures of T7 system. The observation of the CFU showed no reduction throughout the times post-induction and plasmid retention maintained at nearly 100% throughout fermentation (**Figure 5.12b**). This indicated that the cultures within the bioreactor did not negatively affected in terms of culturability, reflecting the better-tolerated and well-regulated of *araBAD* promoter.

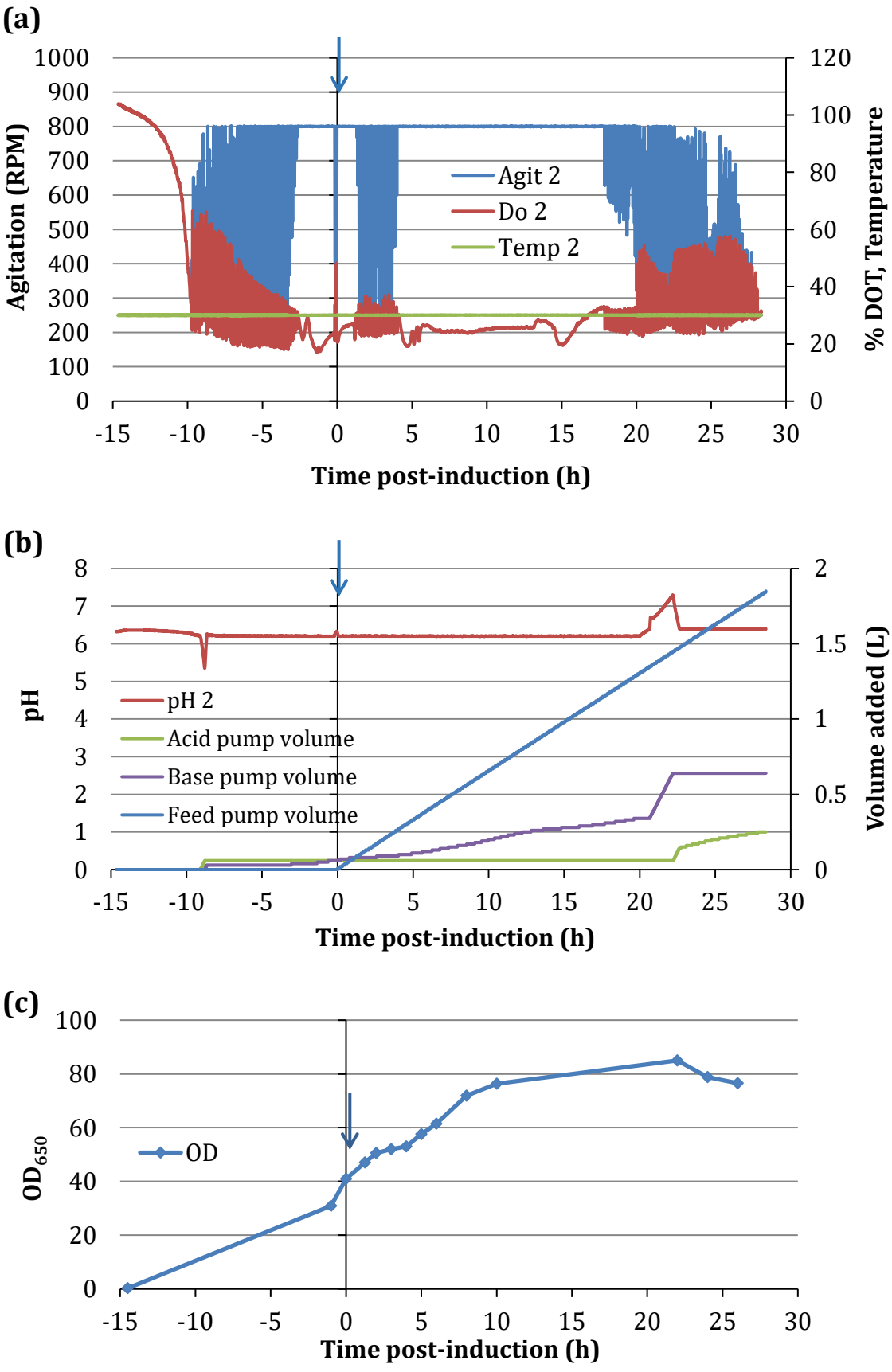

(a)
Agitation (RPM)
% DOT, Temperature
Agit 2
Do 2
Temp 2
Time post-induction (h)
(b)
pH
Volume added (L)
pH 2
Acid pump volume
Base pump volume
Feed pump volume
Time post-induction (h)
(c)
OD650
OD
Time post-induction (h)

Figure 5.11: FerMac controller data of *E. coli* BL-21(A)-p*araBAD*-GFP growth in the study of fed-batch culture for GFP production

E. coli BL21(A) cultures carrying the plasmid p*araBAD*-GFP were grown at 30°C in 5 L vessel and induced with arabinose at mid exponential phase at OD_{650} ≈ 40-50 (approximately after 14 h post-inoculation; indicated by blue arrow), concurrently with feeding at a rate of 67.5 mL/h.

(a) The data of agitation (blue) and percentage of DOT (red), with the temperature of 30°C (green) throughout the fermentation. (b) pH (red), volumes of acid (green), base (purple) and feed (blue) added to the vessel. (c) OD_{650} (blue) measurement at time points intervals.

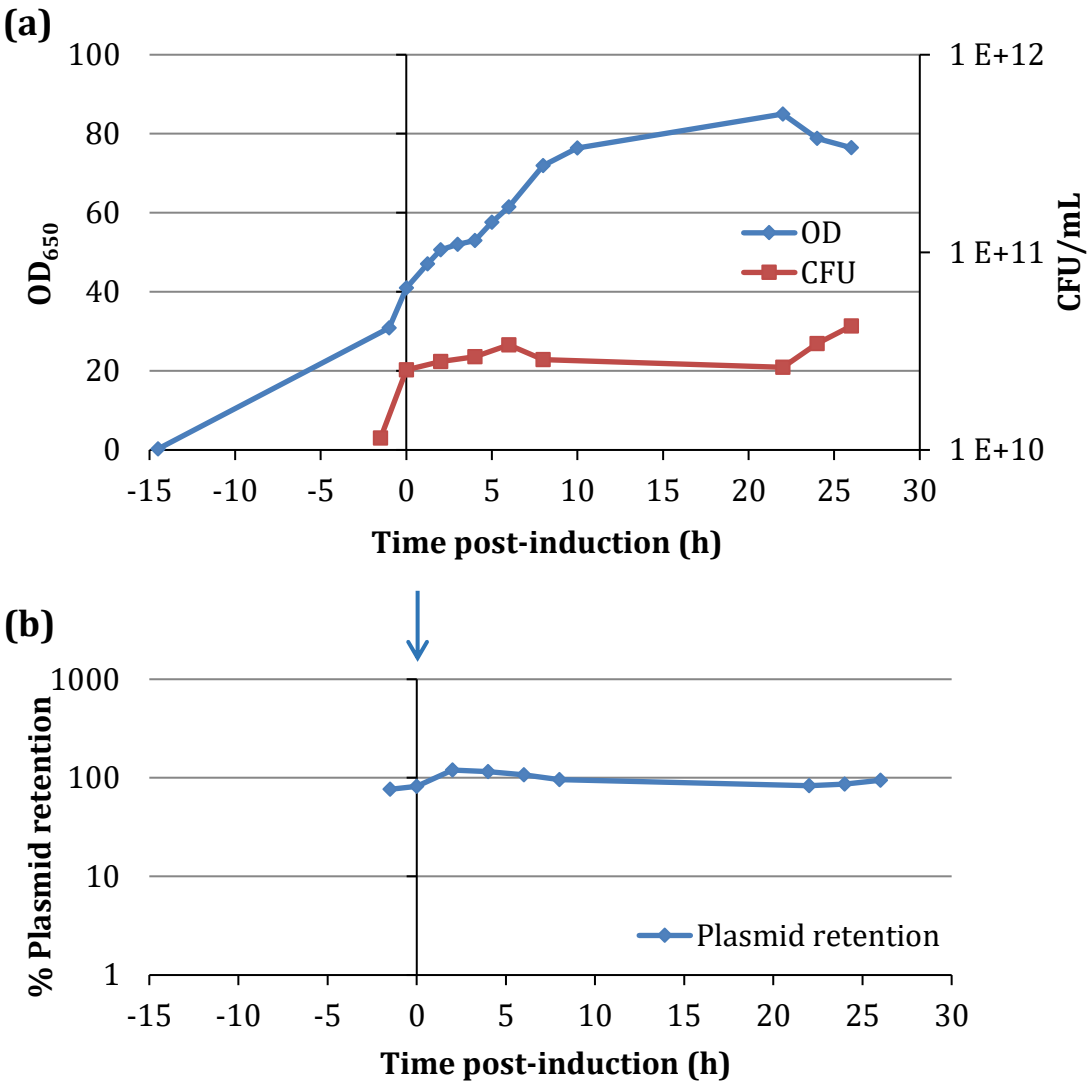


Figure 5.12: Correlation of *E. coli* BL21(A)-p*araBAD*-GFP growth, viability and plasmid retention in fed-batch culture for GFP production

E. coli BL21(A) cultures carrying the plasmid p*araBAD*-GFP were grown at 30°C in 5 L vessel and induced with arabinose at mid exponential phase at OD_{650} ≈ 40-50 (approximately after 14 h post-inoculation; indicated by blue arrow), concurrently with feeding at a rate of 67.5 mL/h. (a) The OD_{650} was measured at intervals and serial dilution of culture samples was plated onto NA agar for indication of cell viability (CFU/mL counting), and (b) were replica-plated onto NA agar supplemented with Kanamycin (50 µg/mL) to reveal plasmid retention.

5.4.1 FCM for fed-batch culture of p*araBAD*-GFP

FCM analysis revealed that the mean green fluorescence per cell in p*araBAD*-GFP culture started to increase after induction (**Figure 5.13a**), up to 30, 000 fluorescence units at 4 h post-induction, but then no further increase in green fluorescence measured. The mean forward scatter FSC-A, which served as a measure of cell size, also did not show any increase following induction, in line with the low green fluorescence produced.

The GFP+ trace of p*araBAD*-GFP culture in fed-batch fermentation reflects the fact that most cells were not productive before induction. The percentage data of GFP+ cells was obviously shown that this culture only had GFP+ cells after induction (**Figure 5.13b**). The proportion of cells that exhibited GFP expression (GFP+) remained consistently over 98% to the end of fermentation. This clearly showed that p*araBAD* promoter strictly regulated the expression of protein, in which it will only occur when there was the presence of arabinose. In contrast to T7 cultures that previously showed that there was also an increase in GFP+ cells before induction. This strongly suggested that T7 promoter system has a leakier expression of protein compared to *araBAD* promoter system. The results obtained from PI staining indicated that the proportion of dead cells remained constantly below 2% following induction until the end of the fermentation.

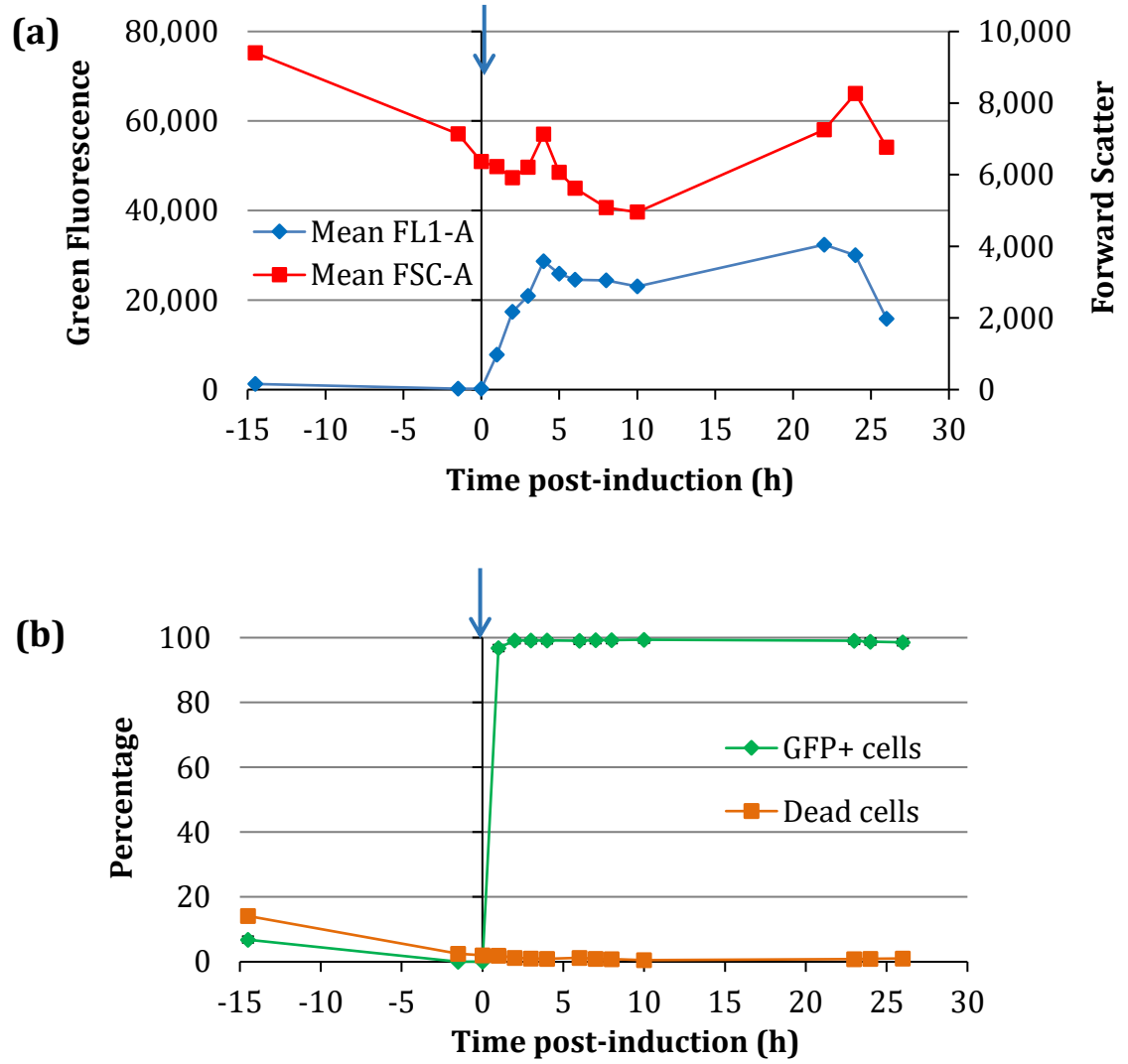


Figure 5.13: FCM data measured for *E. coli* BL21(A)-p*araBAD*-GFP growth in fed-batch culture for GFP production

Green fluorescent protein (GFP) was expressed from an arabinose-induced *araBAD* promoter system; induction at mid exponential phase at OD_{650} ≈ 40-50 (approximately after 14 h post-inoculation; indicated by blue arrow), concurrently with feeding at a rate of 67.5 mL/h. (a) Mean green fluorescence (Mean FL1-A) versus mean forward scatter (Mean FSC-A) plots of cells taken from culture samples at interval time points of incubation. (b) Percentage of cells that producing GFP (GFP+ cells) and percentage of dead cells (PI+ cells) versus time.

5.4.2 SDS-PAGE for fed-batch culture of p*araBAD*-GFP

Densitometry data for the percentage of GFP that was in the soluble or insoluble fraction was shown in **Figure 5.14** and expressed as estimated protein concentrations (in mg/mL) in cell extracts (**Figure 5.15**). Overall, the data showed that the paraBAD-GFP culture was hardly produced GFP over the time. The soluble GFP was the highest measured at 24 h post-induction with 0.284 mg/mL. Yield of GFP accumulation in paraBAD-GFP culture was much lower compared to the cultures of pORT(T7)-GFP and pORT(T7)-TNFα-GFP. This explained the low mean green fluorescence measured in FCM.

Though it was thought that the p*araBAD* promoter was more stable and might end up in better GFP folding to produce more soluble GFP, they didn't generate much green fluorescence in this experiment. This might be because producing GFP drastically following induction puts excessive of a metabolic burden on the cells, yet neither the CFU assay nor the PR assay demonstrated this.

Compare with shake flask growth

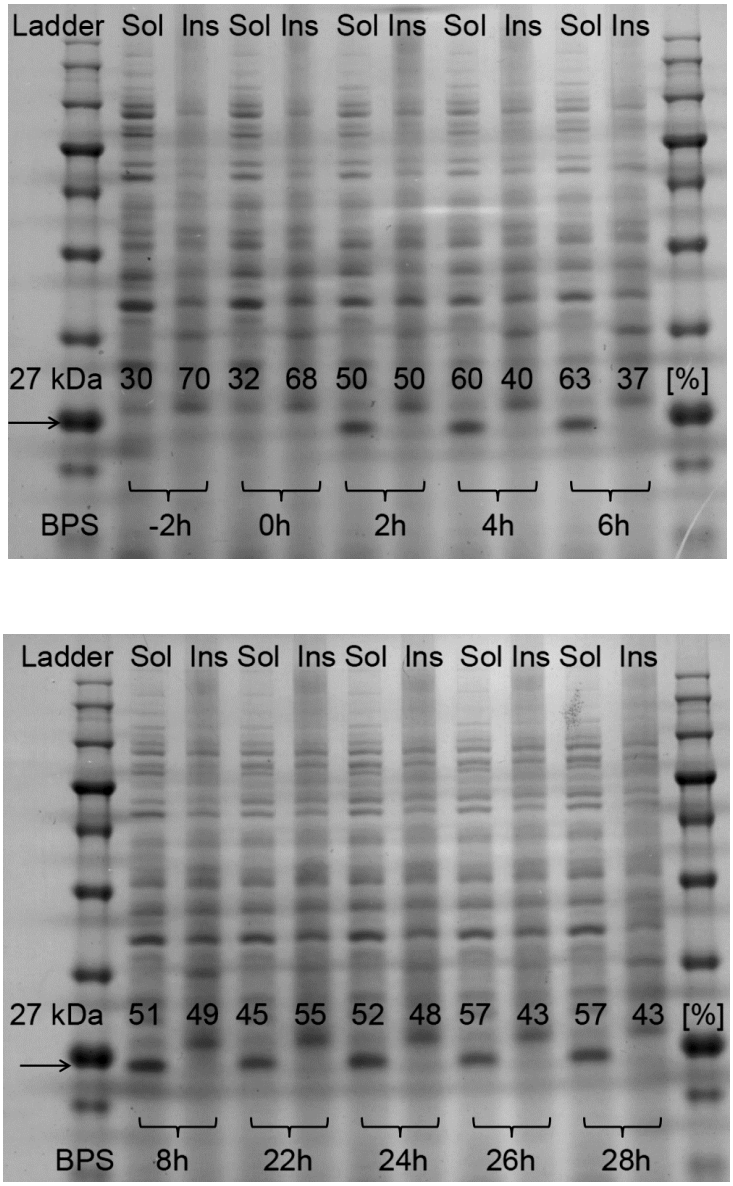


Figure 5.14: The accumulation of GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in fed-batch of p*araBAD*-GFP cultures

SDS-PAGE gel picture with arrow showing GFP (27 kDa) and the ratio percentage of soluble versus insoluble GFP proteins accumulated inside the cell extracts of p*araBAD*-GFP, as shown above the band. All sample bands at times post-induction were compared to BPS ladder band. [**BPS**: Marker used (Blue Protein Standard), broad range of 11-190 kDa; **Sol**: Soluble GFP; **Ins**: Insoluble GFP]

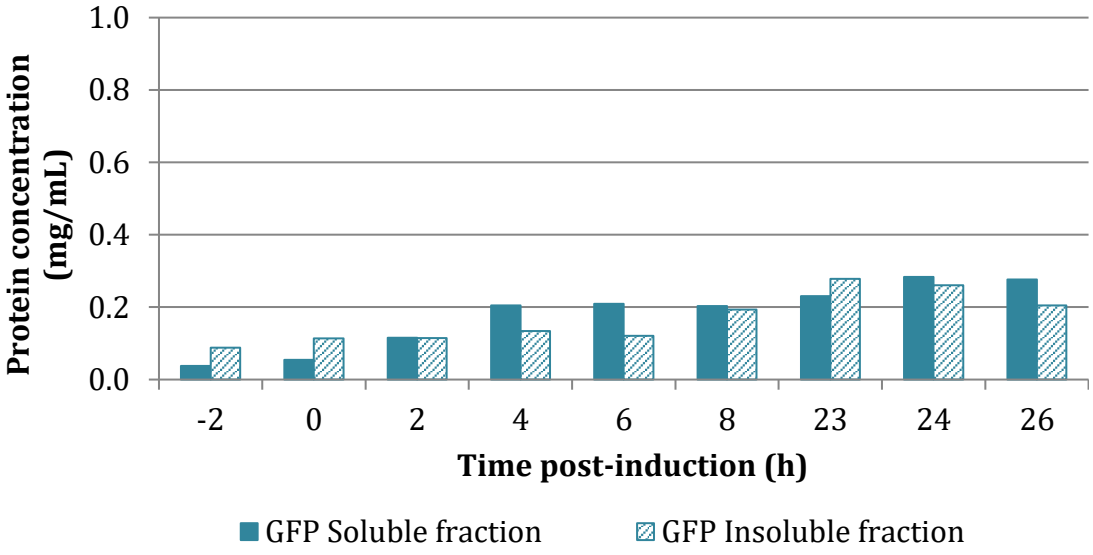


Figure 5.15: The expression of protein solubility in cell extracts measured by SDS-PAGE for GFP production in fed-batch cultures of p*araBAD*-GFP

Estimated protein solubility of soluble and insoluble fractions in cell extracts of GFP at concentration of mg/mL at times post-induction.

5.5 Single cell analysis and fluorescence microscopy of fed-batch fermentations

Single cell FL1-A histogram was analysed for T7 and p*araBAD* expression system cultures grown in fed-batch fermentations. As previously mentioned in **Section 5.1**, *E. coli* BL21(T7) cultures carrying either pORT(T7)-GFP or pORT(T7)-TNFα-GFP, as well as *E. coli* BL21(A) cultures carrying p*araBAD*-GFP, were grown at 30°C in a semi-defined medium in a 5 L bioreactor. Samples were taken at regular intervals and analysed by FCM. The data of FL1-A histogram peaks showed green fluorescence of individual cells.

Green fluorescence histograms of T7-GFP, T7-TNFα-GFP and p*araBAD*-GFP fed-batch fermentations exhibited a single population consistently throughout the growth in all cultures with higher cell densities. This observation indicates that the induction was fully completed, the response was homogeneous, and there was no presence of a plasmid-free GFP- population. These findings suggest that the bioreactor cultures displayed a considerably greater degree of homogeneity compared to the cultures grown in shake flasks. Starting at 7 h post-induction and continuing through the end of the 26 h fermentation, T7-GFP cultures displayed a single population of highly productive GFP-producing (GFP+) cells (on 10^5 fluorescence units), represent higher green fluorescence measured, suggesting cells producing functional recombinant protein GFP (**Figure 5.16a**), whereas T7-TNFα-GFP cultures demonstrated such population only at 22 h post-induction (**Figure 5.16b**). There was no multiple population observed as in earlier batch culture studies.

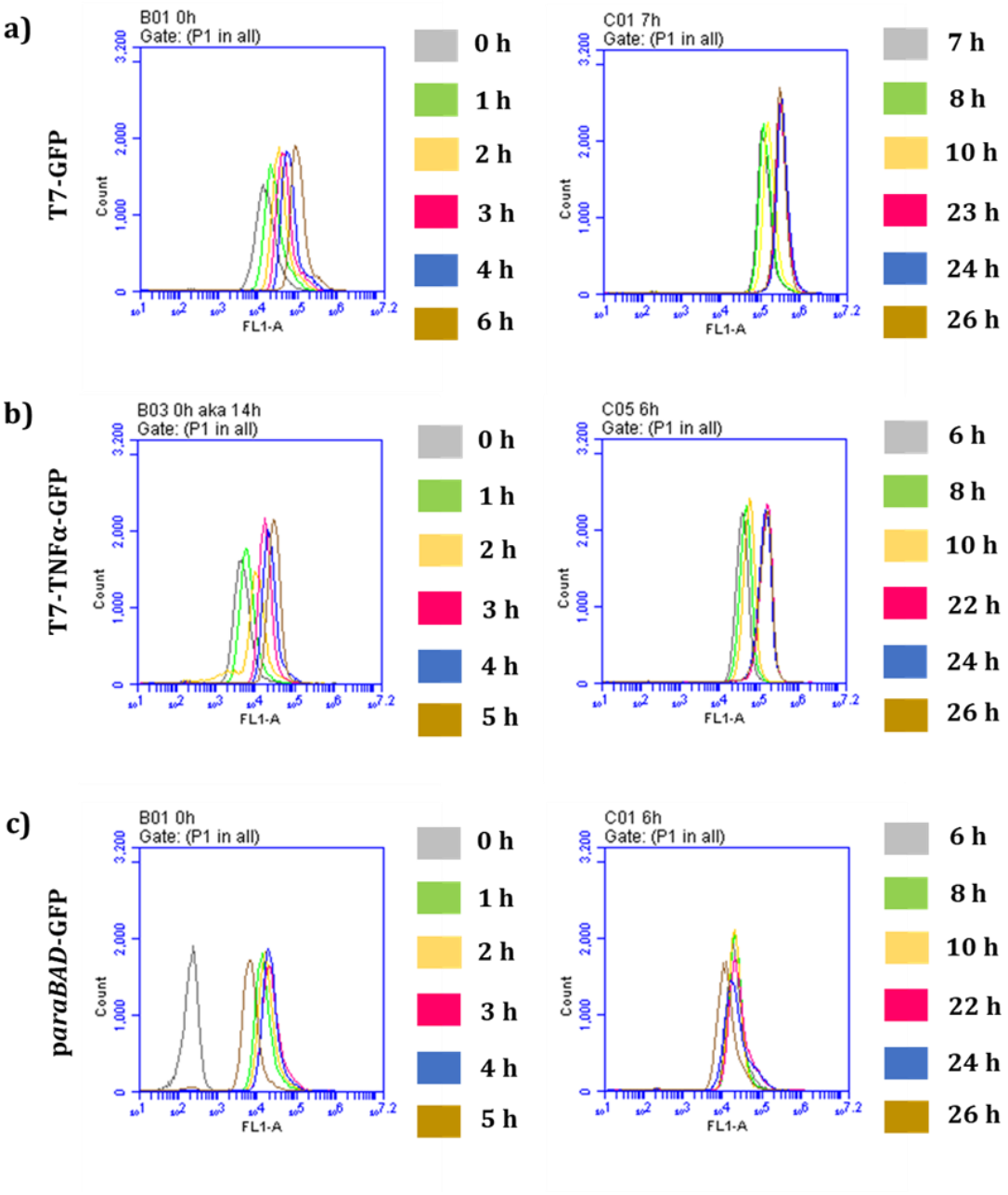


Figure 5.16: Single cell analysis of T7 and p*araBAD* expression system cultures in fed-batch fermentation

Single cell analysis of T7 and p*araBAD* expression system cultures in 5 L bioreactors. *E. coli* BL21(T7) cultures carrying either (a) pORT(T7)-GFP or (b) pORT(T7)-TNFα-GFP; and (c) *E. coli* BL21(A) cultures carrying p*araBAD*-GFP were grown at 30°C in a semi-defined medium and glycerol feed as described in the text. Samples were taken at regular intervals and analysed by FCM. Histograms show green fluorescence of individual cells.

Similar to the T7-TNFα-GFP culture, the green fluorescence histograms exhibited a single population of GFP-producing cells (GFP+) that were highly productive (positioned at 10^5 fluorescence units), however at a later time point of 22 h after induction. Consequently, the T7-TNFα-GFP culture showed a lower mean green fluorescence produced than the GFP culture, despite achieving a higher final OD_{650} compared to GFP cultures (>200 versus 133). T7-TNFα-GFP cultures had predominantly intermediate fluorescence (ranging from 10^3 to 10^4 fluorescence units) from the time of induction until 10 h post-induction, with a decrease in green fluorescence level. However, the population of intermediate fluorescence soon developed, and it seemed that cells transitioned from this population to one of high fluorescence between 22 and 26 h after induction.

The fed-batch p*araBAD*-GFP cultures produced a pattern of green fluorescence similar to shake-flask cultures at the start of the batch fermentation, indicating the absence of productive GFP-producing (GFP+) cells at pre-induction and the present of an intermediate population of GFP+ cells at induction (**Figure 5.16c**). This once again demonstrated the strict regulation of the *araBAD* promoter. Yet, in shake flask, p*araBAD*-GFP cultures demonstrated the presence of a single subpopulation of highly productive GFP-producing (GFP+) cells (on 10^5 fluorescence units) after 24 h of batch culture fermentation, however during fed-batch fermentation, these cultures did not achieve that level until after 26 h of growth. A prior study had addressed the nature of the intermediate expression levels observed at induction of the *araBAD* promoter (Siegele and Hu, 1997). In

fed-batch fermentation, p*araBAD*-GFP culture didn't produce as much GFP as T7 cultures.

GFP is the perfect reporter for in vivo studies. FCM and fluorescence microscopy are two widely used methods for detecting the GFP signal. FCM is an effective and sensitive method for analysing fluorescent intensity quantitatively, while fluorescent microscopy may visualize the subcellular location and expression of GFP. A fluorescent microscope is an optical microscope that excite fluorescence at the specific parts of the cells to observe the structure of fluorescence cells with different physical properties. **Figure 5.17** showed the fluorescence images obtained from observation under fluorescent microscope for BL21(T7)-pORT(T7)-GFP, BL21(T7)-pORT(T7)-TNFα-GFP and BL21(A)-p*araBAD*-GFP fed-batch fermentation cultures. Fluorescence microscopy was used to visualise T7 and p*araBAD* promoter cultures in bioreactors that contained cell populations exhibiting different fluorescence levels. These cell populations had higher green fluorescence, intermediate green fluorescence, or lower green fluorescence, which reflected different RP accumulation and folding states.

FSC-A versus FL1-A plots data obtained from FCM demonstrated that the population with higher FL1-A values had higher FSC-A value, which indicating bacterial size, suggests that cells producing GFP or TNFα-GFP were bigger in size. In a previous work on RPP, Wyre and Overton (2014b) found that bacterial size increased with the mean green fluorescence produced.

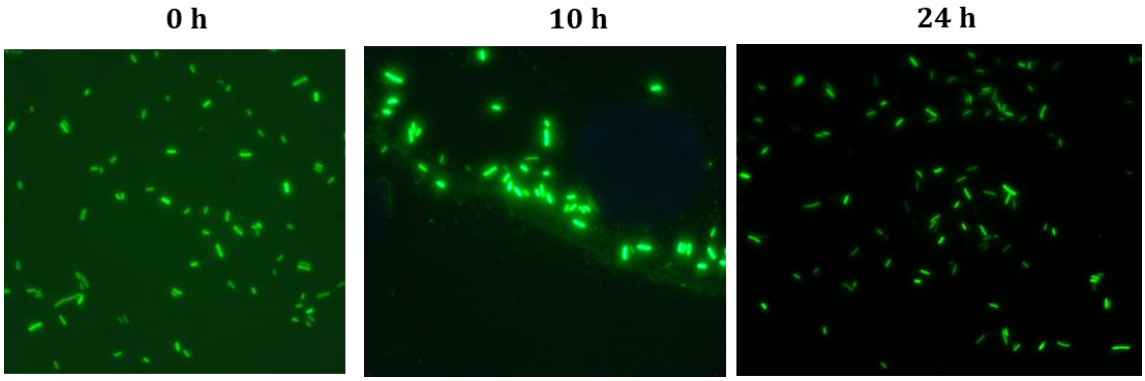
Observation of T7-GFP cultures under fluorescence microscopy revealed that this culture, at 0 h pre-induction, exhibited subpopulations of cells that either

productive or non-productive fluorescence-producing and became more fluorescence over time following induction, in which each cell acquired a greater amount of fluorescence, elongated and increased in size (**Figure 5.17a**). This reflected to the highly production of green fluorescence measured by FCM for this culture (**Figure 5.3**).

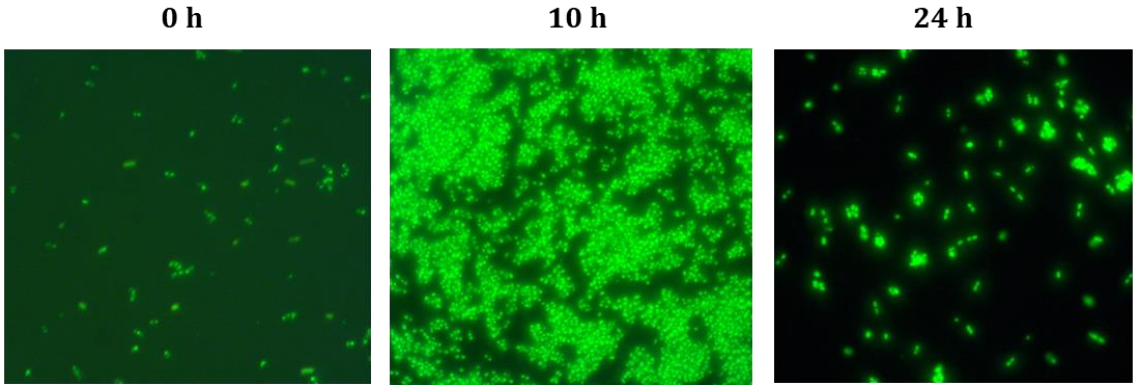
In contrast to T7-TNFα-GFP culture, which displayed smaller and less fluorescent cells compared to T7-GFP at 0 h pre-induction, cells that were denser at 10 h post-induction, exhibited size smaller than T7-GFP culture and produced less fluorescence (**Figure 5.17b**). This was indicative of the intermediate level of green fluorescence generated. The cells grew bigger and were more elongated at 24 post-induction. This reflects the mean green fluorescence value that was recorded for this culture at an intermediate level; at the end of fermentation, the value rose to a greater level (**Figure 5.8**).

While for p*araBAD*-GFP culture, observation under fluorescence microscopy showed that cells did not produce any fluorescence at 0 h pre-induction, and just reflected the autofluorescence of *E. coli* BL21(A). This culture exhibited intermediate fluorescence following induction at 10 h to 24 h post-induction, and the cells size was observed to be smaller and uniform over the time, although few cells showed elongation (**Figure 5.17c**). This was comparable to the FSC-A reading recorded (**Figure 5.13a**) for this culture, which barely showed any increase in FSC-A value, suggesting the size was similar over the time, along with the low mean green fluorescence recorded.

(a) BL21(T7)-pORT(T7)-GFP



(b) BL21(T7)-pORT(T7)-TNFα-GFP



(c) BL21(A)-p*araBAD*-GFP

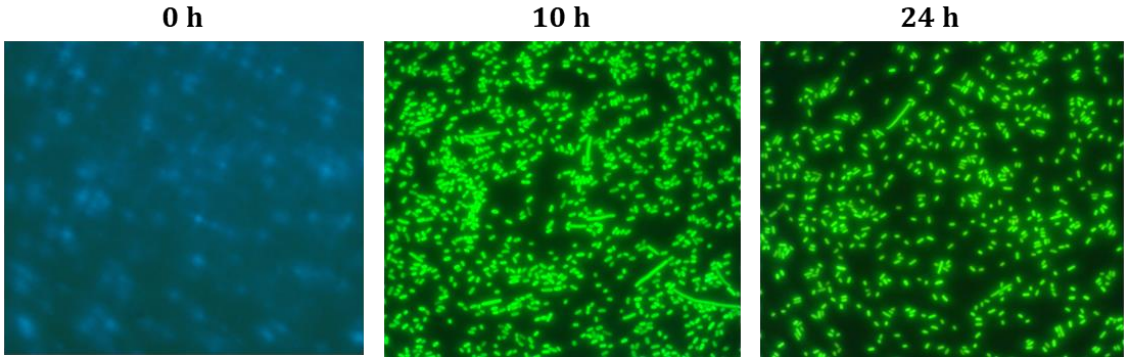


Figure 5.17: Fluorescence microscopy images of T7 and p*araBAD* expression system cultures in fed-batch fermentation of bioreactor

Observation of fluorescence microscopy in fed-batch fermentations. The fluorescence micrographs correspond to (a) BL21(T7)-pORT(T7)-GFP cultures, (b) BL21(T7)-pORT(T7)-TNFα-GFP cultures, and (c) BL21(A)-p*araBAD*-GFP cultures. Culture samples at 0 h, 10 h and 24 h post-induction, respectively, were visualised under fluorescent microscope. These fluorescence microscopy images correspond to Figure 5.16.

5.6 Discussion and conclusion

Comparing bioreactor and shake-flask growth, maximum mean green fluorescence values were broadly comparable for T7 promoter cultures although the bioreactor cultures produced fluorescence much more steadily than the shake-flasks, more resembling the uninduced T7 cultures or the p*araBAD* system. However, when compared to shake-flask, the p*araBAD*-GFP culture in fed-batch fermentation showed a decrease in terms of growth profile and GFP accumulated.

Far greater OD_{650} values were seen in T7 cultures in comparison to p*araBAD* culture. Notably, T7-GFP culture exhibited the maximum OD_{650} value of 133, whereas T7-TNFα-GFP culture displayed the highest OD_{650} value exceeding 200. Conversely, the p*araBAD*-GFP culture demonstrated an OD_{650} of 85. A dramatic drop in CFU was observed for both T7 cultures as they adapting to the medium in bioreactor and productively produce GFP/ TNFα-GFP. While p*araBAD*-GFP culture maintained the CFU throughout the growth, reflecting the better-tolerated of *araBAD* promoter system, leading to a lesser negative impact on bacterial physiology, as showed earlier in shake flask studies.

The T7-TNFα-GFP culture also showed more plasmid loss than T7-GFP culture following induction (**Figure 5.7b**). The phenomenon of plasmid loss was observed as a consequence of stress, mostly attributed to the metabolic burden imposed on the cells by the production of TNFα-GFP and the incidence of protein misfolding. These factors collectively contribute to the loss of plasmids.

Although low plasmid retention was observed in TNFα-GFP culture, the histogram peaks showed a single population of productive GFP-producing (GFP+) cells throughout the growth after induction. This observation implies that the observed decrease in plasmid presence after induction can be attributed to the necessity of agar plate growth in the experiment. Such issues were commonly observed in CFU and plasmid retention assays, particularly when the VBNC population was present. Fortunately, FCM lacks these problems (Fernández-Castané et al., 2016). T7-TNFα-GFP culture loss a huge amount of plasmid compared to T7-GFP culture, indicating more stress on *E. coli* BL21(T7) to produce TNFα-GFP.

Following induction, the majority of the cells in p*araBAD*-GFP culture were in an intermediate fluorescence population (from 10^3 to 10^4 fluorescence units) with lower green fluorescence. Although plasmid retention remained over 85% throughout the fermentation and there was no reduction in CFU after induction, strongly suggesting that the majority of cells remained productive throughout the fermentation (Wyre and Overton, 2014a), yet a low OD growth reading (OD_{650} 85; **Figure 5.12a**) was recorded, the lowest among the other cultures in fed-batch fermentations.

The GFP+ trace of this culture revealed that the majority of the cells were not productive before induction, and the proportion of cells that exhibited GFP expression (GFP+) grew drastically from 2% at 0 h prior to induction to 97% at 1 h following induction. This proportion remained constantly above 98% until the end of fermentation. The maximum green fluorescence value seen was only 32, 000 fluorescence units, reported at 22 h post-induction, indicating that they were

unable to produce much green fluorescence in fed-batch fermentation. Previous study had reported on the nature of the intermediate expression levels of the *araBAD* promoter (Siegele and Hu, 1997).

It was interesting to note that the plasmid retention of p*araBAD*-GFP remained greatly stable, and the proportion of productive GFP-producing (GFP+) cells after induction was also very high, however, the measured mean green fluorescence was considerably lower, indicating a reduced production of GFP per unit biomass. This was probably a consequence of the very low quantity and quality of GFP being produced by *araBAD* promoter in fed-batch fermentation (Wyre and Overton, 2014b).

Given that flow cytometry is a technique used for analysing individual cells, the measurement of green fluorescence is performed on a per-cell basis. Consequently, the bioreactor's overall yield would be substantially higher compared to shake flasks, mostly due to the considerably higher concentration of biomass. The proportions of dead cells exhibited a considerable degree of similarity between the shake flasks and bioreactor, suggesting that the transition to bioreactor growth did not negatively affect cell physiology.

CHAPTER 6

RESTRICTION-FREE CLONING OF Fumarase B PROTEIN

This chapter focuses on the generation of soluble and functional *E. coli* Fumarase B (FumB) enzyme. To generate FumB enzyme, the fumarase B (*fumB)* gene were cloned into the pET41c plasmid. The recombinant plasmid was then transformed into E. coli BL21(DE3)* for further growth optimisation in batch culture and analysis using FCM and SDS-PAGE. The challenge was not only because FumB is an iron-sulphur (FeS) protein, but also that aggregation of inclusion bodies (IBs) might be a major problem as a consequence of protein misfolding.

6.1 Introduction

This study attempts to clone the fumarase B (*fumB*) gene into the pET41c plasmid. The targeted *fumB* gene was obtained from the chromosomal DNA of *E. coli* K-12 MG1655 by PCR amplification (1st round of PCR), and the technique used for cloning is known as restriction-free (RF) cloning. This technique offers a simple and effective process for inserting DNA fragments exactly into a specific location on a circular plasmid without altering any segment of the plasmid or targeted genes. It is also independent of restriction sites or ligation (Van Den Ent and Lowe, 2006). This cloning process is based on the amplification of a specific gene in which a mega-primer is synthesised for linear amplifications of both the plasmid and gene insert (Unger et al., 2010). In this technique, PCR amplification products encoding *fumB* were used as a pair of primers in the linear amplification (2nd round of PCR) around the region of the pET41c plasmid. Conventional molecular cloning techniques rely on enzymatic digestions of the gene insert and a vector for cloning, thus often facing issues in choosing appropriate restriction sites.

6.2 Methods preparation

6.2.1 PCR Amplification

An attempt has been made to understand the fundamental principles of polymerase chain reaction (PCR) and gel electrophoresis. The gene encoding fumarase B (*fumB* gene) was amplified from the chromosomal DNA of *Escherichia coli* (*E. coli*) K-12 MG1655 using PCR. Further optimisation of *fumB* gene amplification were carried out using PCR and QIAquick column PCR purification kit.

6.2.2 Cloning of *fumB* genes into pET41c plasmid

The *fumB* gene were cloned into pET41c plasmid due to its T7 RNA polymerase-dependent promoter and kanamycin resistance marker. These potential ligated pET41c-*fumB* constructs were treated with DpnI to digest methylated parental cells and transformed into DH5α competent cells by using heat-shock transformation methods.

6.2.3 Transformation of BL21(DE3)* with pET41c-*fumB*

The recombinant pET41c plasmid containing the *fumB* gene was transformed into BL21(DE3)* competent cells and grown on NA supplemented with 50 μg/mL Kanamycin (Details in **Section 2.8.5**).

6.3 Polymerase Chain Reaction (PCR)

6.3.1 First (1st) round PCR Amplification

The targeted *fumB* gene was first isolated from the genomic DNA of *E. coli* by PCR amplification (1st round PCR). The band size of the extracted genes was expected to be 1647 bp (base pairs), which was the size specifically identified as the fumarase B gene. The first few attempts resulted in PCR product band with the right size, but also form a primer dimer (**Figure 6.1**).

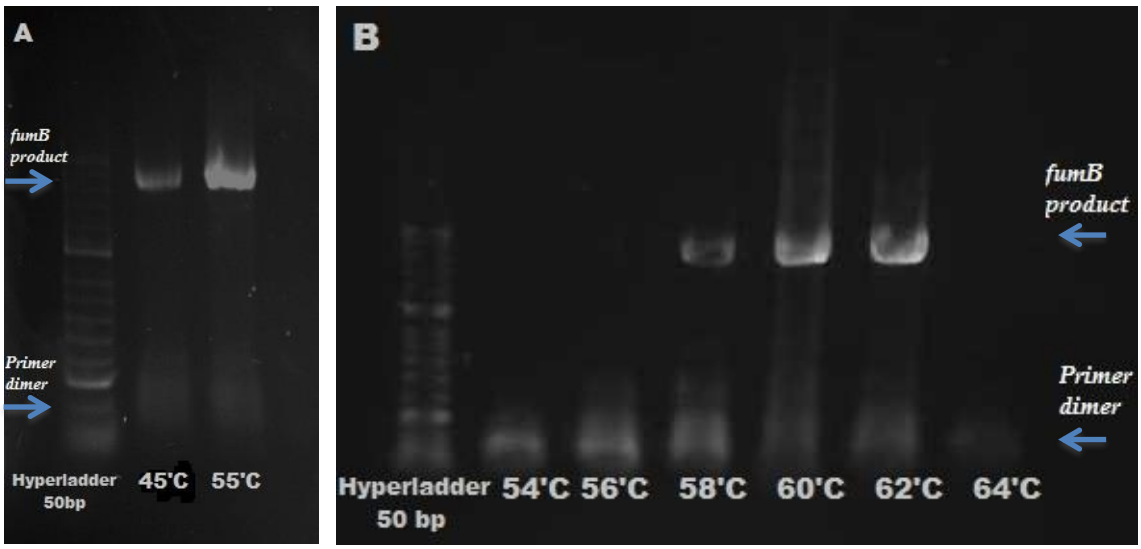


Figure 6.1: Gel electrophoresis of targeted *fumB* gene on 1% agarose gel

(A) *fumB* gene initially extracted from chromosomal *E. coli* K-12 MG1655. (B) PCR optimisation using six different annealing temperatures in gradient. The expected size of the PCR products of *fumB* gene were 1647 bp. The marker used in the first lane (A, B) was Hyperladder II (New England Biolabs, Hitchin, UK) with the maximum band size was 2, 000 bp.

The next step was an attempt to produce the extraction of abundant and pure *fumB* gene evidenced by the production of single and clear bands on gel electrophoresis. Determining the optimal annealing temperature in the 1st round PCR amplification has been made as the first optimisation step. There were six different annealing temperatures used, which were 54 °C, 56 °C, 58 °C, 60 °C, 62 °C and 64 °C. Side products such as primer dimer can be avoided by optimising the annealing temperature of PCR amplification by gradient PCR (Benoit et al., 2006).

The marker used in the first lane (**Figure 6.1a** and **6.1b**) was Hyperladder 50bp (previously known as Hyperladder II; New England Biolabs, Hitchin, UK) with the maximum band size was 2, 000 bp. In **Figure 6.1a**, both bands were obtained from the extracted *fumB* gene from *E. coli* genomic DNA, with different annealing temperature 45°C and 55°C, respectively, with the formation of primer dimer. Higher annealing temperature in gradient, which was 60°C, has been proved by gradient PCR as the best temperature to use for annealing step in 1st round PCR at the beginning of the study (**Figure 6.1b**).

Next, the extraction of *fumB* gene in PCR amplification was run using different volumes of 10 times diluted 1st round PCR products, as the component for DNA template. However, none of them worked well. Next, the volume of genomic DNA and the primer concentration used in 1st round PCR were also optimised. A pair of freshly prepared primers (both forward and reverse primer) were used in all experiments, and the only annealing temperature used in these experiments was 55°C. Primers used during this study are listed in **Table 2.7** (**Section 2.8.1**).

A set of three different volume of genomic DNA was used, which were 1 µL, 0.5 µL and 0.1 µL (**Figure 6.2a**). From the gel electrophoresis, it showed that 0.5 µL of genomic DNA worked well and produced a clearer band compared to the others. While for the primer concentration, there were also three different sets of primer concentration studied, which were 100 µM, 50 µM and 10µM (**Figure 6.2b**). The final concentration of the primer used was the best at 50µM.

After various optimisation were done, the best condition to run 1st round PCR amplification for the extraction of targeted *fumB* gene in this research project, were as followed; 20 µL 5x Hifi buffer, 1 µL 100 mM dNTPs, up to a final volume of 100 µL with sterile distilled water, 50 µM primer mix (forward and reverse), 0.5 µL genomic DNA and 2 µL of velocity DNA polymerase.

We have obtained large quantities of *fumB* with less primer dimer formation when amplified under the best condition of 1st round PCR in the lab, and this was approved by gel electrophoresis, which showed the clear bands at 1647 bp that was exactly the right size for *fumB* gene (**Figure 6.3a**). Instead of only use QIAquick PCR purification kit to purify the PCR products directly from the PCR tubes, an approach of extracting pure DNA from the agarose gel was also used (**Figure 6.3b**). Gel electrophoresis was run by loading 30 µL of PCR products inside a large well in 2% agarose gel prior to extract them from the gel. In total, a volume of 60 µL of *fumB* PCR product was obtained.

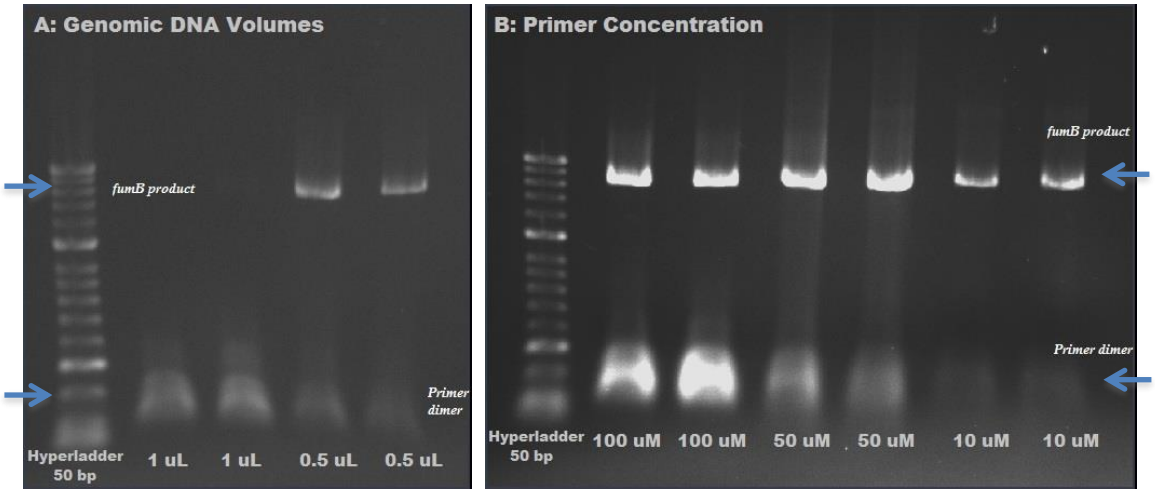


Figure 6.2: Optimisation of (A) genomic DNA and (B) primer concentration in 1st round PCR amplification

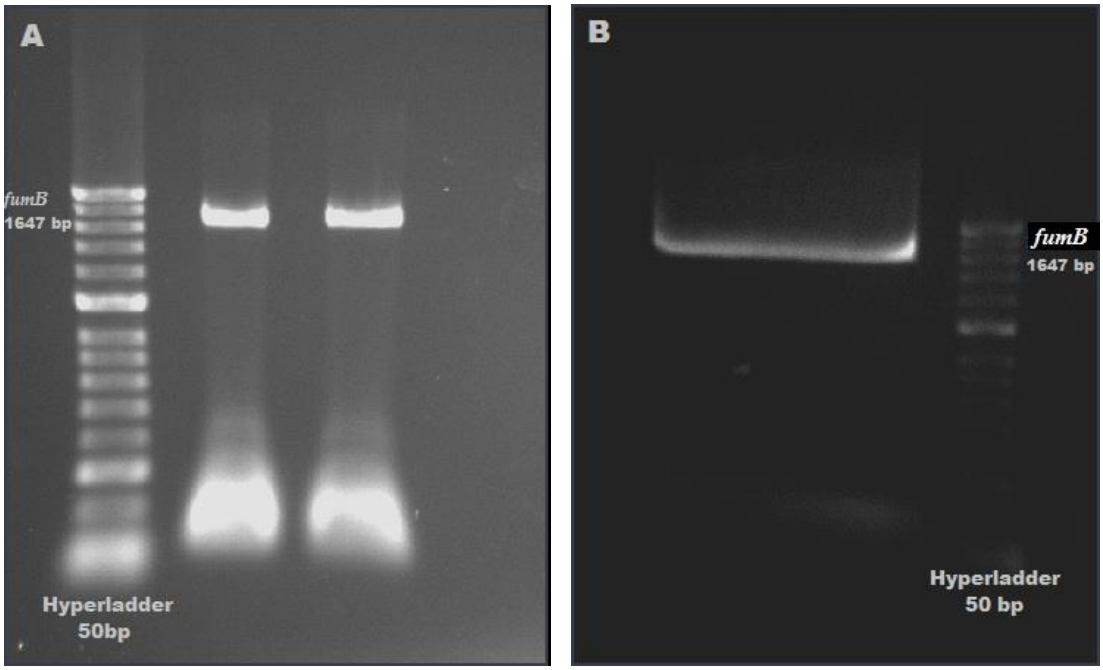


Figure 6.3: Gel electrophoresis showed clear bands at 1647 bp at the optimum condition of PCR amplification

6.3.2 Linear Amplification Reaction (Restriction-free cloning)

Restriction-free (RF) cloning was used in this study for linear amplification reaction in 2nd round PCR. The technique used the product of 1st round PCR (purified *fumB* gene) as a pair of primers in a linear RF amplification, which were introduced into the destination vector pET41c. These targeted genes, which produces a primer pair that, once annealed to the vector of interest, is extended in a linear amplification reaction (Van Den Ent and Lowe, 2006).

All the clones that were amplified were expected to reach the approximate size of 6.6 kb, as the combination size of 1647 bp *fumB* gene with approximately 5 kb size of pET41c vector. In RF amplification method, the total size of insert and vector is limited to 15 kbp, however longer constructs can still be made but with a lower efficiency (Zhou and Gomez-Sanchez, 2000). The introduction of larger insertions of up to 1.1 kb has been demonstrated before as being a possibility (Geiser et al., 2001).

The reaction mixture then was treated with DpnI prior to transformation. DpnI, cleaving the methylated DNA, digests the parental plasmid. Average yields obtained using this method are about 90% (Van Den Ent and Lowe, 2006). Numerous possibilities were being explored to optimize RF cloning in linear amplification reaction by varying the concentration of mega-primers and pET41c vector, as well as the type of DNA polymerase used. In many cases, one PCR product was used at 100 ng per reaction. The efficiency of integration can be

improved by altering the ratio between the mega-primers used (Unger et al., 2010). **Figure 6.4** shows schematic representation of restriction-free (RF) cloning.

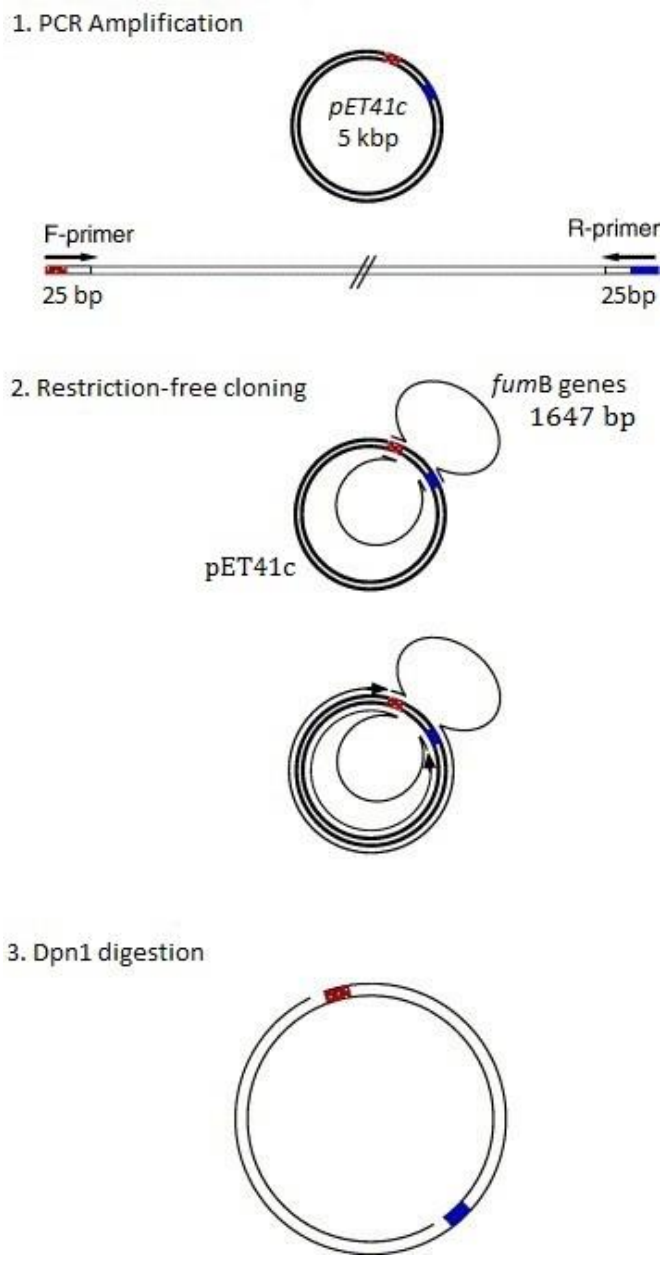


Figure 6.4: Schematic representation of restriction-free (RF) cloning

(1) A circular pET41c vector has unique priming sites (indicated in red and blue) and is combined with a PCR product encoding the gene of interest (*fumB* gene) attached by priming sites homologous to those in the vector. (2) The PCR product acts as a primer pair in a linear amplification reaction. Once annealed to the vector, DNA polymerase (Velocity Enzyme) extends and incorporates the gene into a circular DNA molecule. (3) The methylated parental vector is digested with Dpn1 and the circular, double-stranded DNA is transformed into a suitable host cell such as DH5α competent cells. Figure adapted from Van Den Ent and Lowe (2006).

6.4 Transformation

6.4.1 Transformation of DH5α with pET41c

The pET41c vectors used in this study were produced in large quantities through a transformation of pET41c into ultra-competent DH5α cells. Heat-shock technique used for this purpose has resulted in the growth of many transformant cells. Nutrient agar used to grow the culture samples from heat-shock transformation were supplemented with 50 μg/mL Kanamycin antibiotics as pET41c vector is the Kanamycin-resistant construct (Finkelstein et al., 2003).

For positive control Kanamycin-resistant nrfa (*nrfa::kan*), plenty of colonies grew well on the plate, which was very good because it indicated that transformation protocol used was correct, competent cells used were really competent and nutrient agar plates used were correctly prepared with appropriate kanamycin concentration. The negative control containing only DH5α alone (without any vector inserted) did not result in any colonies. This was as expected because these competent cells did not confer resistance to kanamycin.

Few single colonies from DH5α-pET41c on Na-Kan agar plates were purified, by re-streaking onto a fresh Na-Kan agar plate. Their plasmids were isolated, and EcoRI enzyme was used to linearize them. When run through gel electrophoresis, all the plasmid that were isolated had showed similar size which was approximately 5 kb (**Figure 6.5**). The marker used was Hyperladder 1 kb (previously known as Hyperladder I; New England Biolabs, UK) with the maximum

band size was 10,000 bp (10 kb). The band size indicated that they were specifically pET41c vector that can be combined with *fumB* gene, attached by unique priming sites homologous to those in the vector. The ability to produce these vectors in large amounts in the laboratory made the study becomes easier and cost saving.

6.4.2 Transformation of *E. coli* BL21(DE3)* with pET41c-*fumB*

The recombinant plasmid was then transformed into *Escherichia coli* (*E. coli*) BL21 (DE3) for further growth optimisation in batch culture and analysis using flow cytometry and SDS-PAGE.

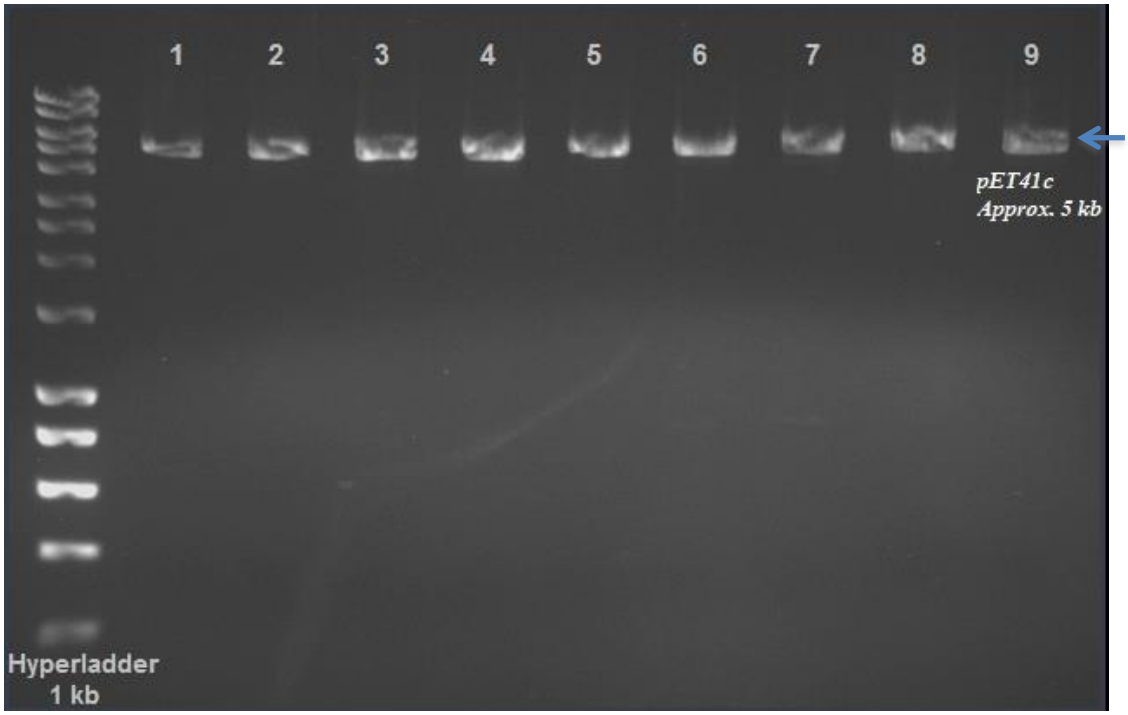


Figure 6.5: Gel electrophoresis showed pET41c linearized plasmid at size approximately 5 kbp

6.5 Standard and Improved Conditions for FumB production

6.5.1 Batch Culture of BL21(DE3)*-pET41c-FumB

Both standard and improved protocols have been designed to grow the cells in shake flasks in batch culture to improve FumB solubility and thereby overall yield of functional FumB. The details of both protocols for FumB production in batch culture was summarised in **Table 6.1**, with several different measurable parameters for comparison purposes. The parameters investigated in this study were the growth temperature and the concentration of inducer (IPTG). Both protocols were using Terrific broth (TB) as the growth medium supplemented with Kanamycin (50 μg/mL) to maintain plasmid retention and induction point set at mid-logarithmic phase of $OD_{650} \approx 0.5$. Glycerol was used as the carbon source.

Based on **Table 6.1**, the temperature used was either 37°C, then reduced to 25°C after IPTG induction for standard protocol (ST); or 25°C for improved protocol (IM). The temperature of 25°C was chosen for both protocols as an approach to increase the FumB folding by minimising the stress, achieved by reducing the production rate of recombinant proteins. Sevastsyanovich et al., (2009) believed that the higher temperature could lead to rapid production of recombinant protein which resulted in the accumulation of high yield insoluble protein that often aggregates as inclusion bodies. IPTG was used as an inducer in this study, at concentrations of 8 μM and 500 μM (0.5 mM).

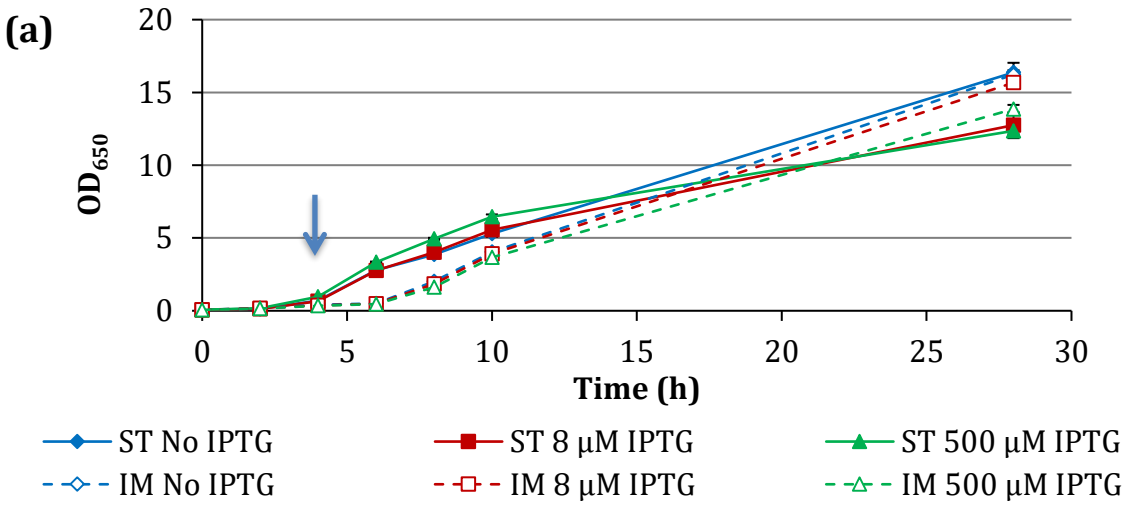
Table 6.1: Standard versus Improved conditions for FumB production

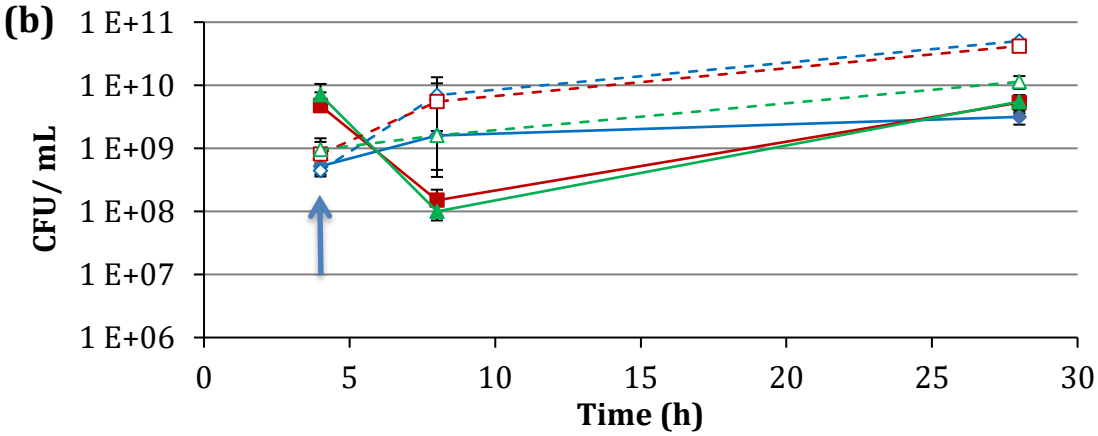
Condition	*Standard protocol*	*Improved protocol*
Media used	Terrific broth	Terrific broth
Temperature	37°C, then reduced to 25°C after IPTG induction	25°C
IPTG concentration	0.5 mM (500 µM) and 8 µM	0.5 mM (500 µM) and 8 µM
Induction point	Mid-logarithmic phase (OD_{650} ≈ 0.5)	Mid-logarithmic phase (OD_{650} ≈ 0.5)
Expectation	High temperature and IPTG concentration will lead to the production of little soluble protein, most FumB aggregating into inclusion bodies	Low temperature and IPTG concentration will lead to the production of soluble recombinant FumB protein

Cultures were sampled at regular intervals. The optimum growth condition obtained from these experiments were measured by OD_{650} and CFU. Plasmid retention was measured using replica plating. Cell samples were analysed with flow cytometry (FCM) using bis-oxonol (BOX) and propidium iodide (PI) to determine the viability levels of cells. SDS-PAGE and subcellular fractionation was also run to determine the solubility of FumB protein generated.

Recombinant protein production was induced when the OD_{650} reached around 0.5 (mid-logarithmic phase), with either 8 µM or 500 µM IPTG. An uninduced control with no IPTG was also included for both protocols. Then, the temperature of standard protocol was reduced to 25°C. All experiments in shake flask were executed in duplicate.

As can be observed in **Figure 6.6**, the use of different temperature for ST and IM had slightly affected the bacterial growth, in which the growth was slower overall in IM cultures upon induction than the ST cultures (**Figure 6.6a**).


(a)
20
15
10
5
0
OD650
0
5
10
15
20
25
30
Time (h)
ST No IPTG
ST 8 µM IPTG
ST 500 µM IPTG
IM No IPTG
IM 8 µM IPTG
IM 500 µM IPTG


(b)
1 E+11
1 E+10
1 E+09
1 E+08
1 E+07
1 E+06
CFU/ mL
0
5
10
15
20
25
30
Time (h)

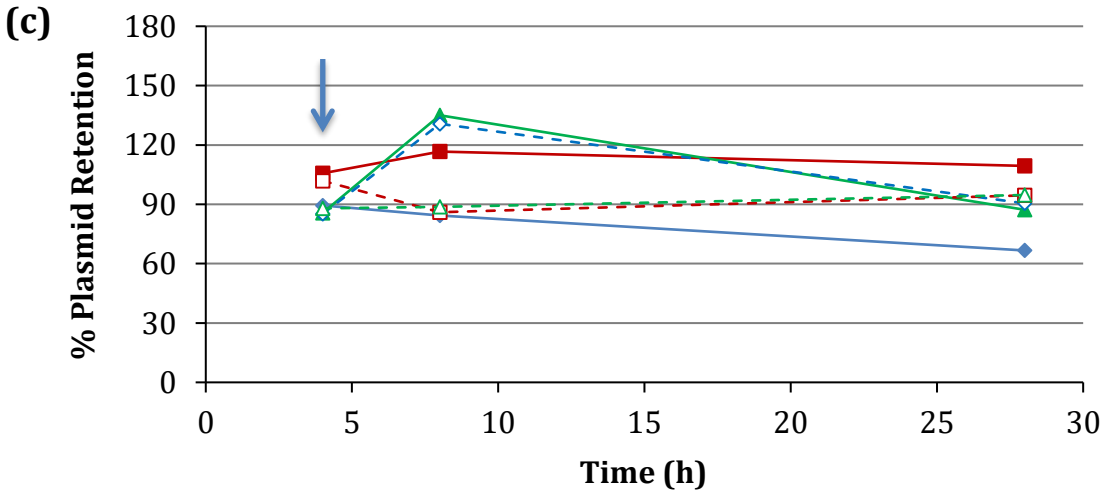

(c)
180
150
120
90
60
30
0
% Plasmid Retention
0
5
10
15
20
25
30
Time (h)

Figure 6.6: Correlation of *E. coli* BL21(DE3)* growth for the study of optimisation for FumB production, evaluated by (a) OD_{650} measurement, (b) CFU/mL counting, and (c) plasmid retention

E. coli BL21(DE3) cultures carrying the plasmid pET41c were grown at standard (ST) and improved (IM) conditions with three different concentrations of IPTG at mid logarithmic phase at $OD_{650} \approx 0.5$ (blue arrow). The OD_{650} was measured at intervals (a). Serial dilution of culture samples at 0 h, 4 h and 24 h post-induction was plated onto NA agar for indication of cell viability (b) and were replica-plated onto kanamycin-supplemented NA agar to reveal plasmid retention (c). All data shown are mean values from replica flasks, error bars are ±1 standard deviation.

Non-induced cultures of both conditions reached a higher cell density than the induced cultures, probably as the result of overgrowth plasmid-free cells. All induced cultures of ST had a lower culturability with less than $1x10^8$ CFU at 4 h post induction (**Figure 6.6b**) compared to IM cultures (more than $1x10^9$ CFU). This might be happened due to the reduction of temperature occured, from 37 °C reduced to 25°C in ST, forcing the cell to adapt to the temperature change that occured in the surrounding. All IM cultures maintain the cell viability throughout the incubation time which signified suitable with the temperature environment of 25°C. In term of plasmid retention, the use of IM protocol with 25°C and IPTG induction led to a drastic plasmid loss around 10% (**Figure 6.6c**). In contrast, both induced cultures of ST maintain their plasmid retention at more than 100% until at least for 4 h post induction. ST culture induced with 8 μM IPTG maintain the plasmid retention until the end of the experiment, however the 500 μM IPTG induced ST culture showed the loss of plasmid retention of more than 10% at 24 h post induction. Obviously for non-induced ST culture had a great loss of one third plasmid retention (more than 30%) after 28 h of incubation.

6.5.2 Flow cytometry analysis for FumB production

As for flow cytometry data, in addition to measure the dead cells with propidium iodide (PI), Bis-oxonol (BOX) was also used to determine cells that were injured. The protocol used to measure injured cells is as stated by Wyre and Overton (2014). This dual staining system of PI and BOX often used as fluorescent dyes in FCM to determine cellular viability. Therefore, all different subpopulations; live or dead, and even injured cells can also be identified using flow cytometry for BL21(DE3)*-pET41c-*fumB* cultures that were producing FumB proteins.

The results in **Figure 6.7** showed the data percentage of injured and dead cells during the incubation time, in which cultures were sampled at regular intervals. **Figure 6.7a** showed the percentage of injured cells were insignificantly measured (less than 0.1 %) in all ST and IM cultures which indicated that the cells were either live or totally dead, no injured cells. Among the two, ST showed increasing percentage of dead cells as early as 4 h post induction especially for cultures with 500 μM IPTG induction and this data is equivalent to the measurement of CFU on these cultures where the cellular viability declined severely at 4 h post induction (**Figure 6.8b**). The 500 μM IPTG induced culture of ST also showed the highest dead cells percentage at the end of 28 h of incubation with around 20 % of dead cells. All IM cultures showed less than 5 % of dead cells during incubation time. This showed clearly the survival of cells were better in IM condition, in which the level of culturability of these cells also showed equivalent data with more than $1x10^9$ CFU throughout the incubation (**Figure 6.7b).**

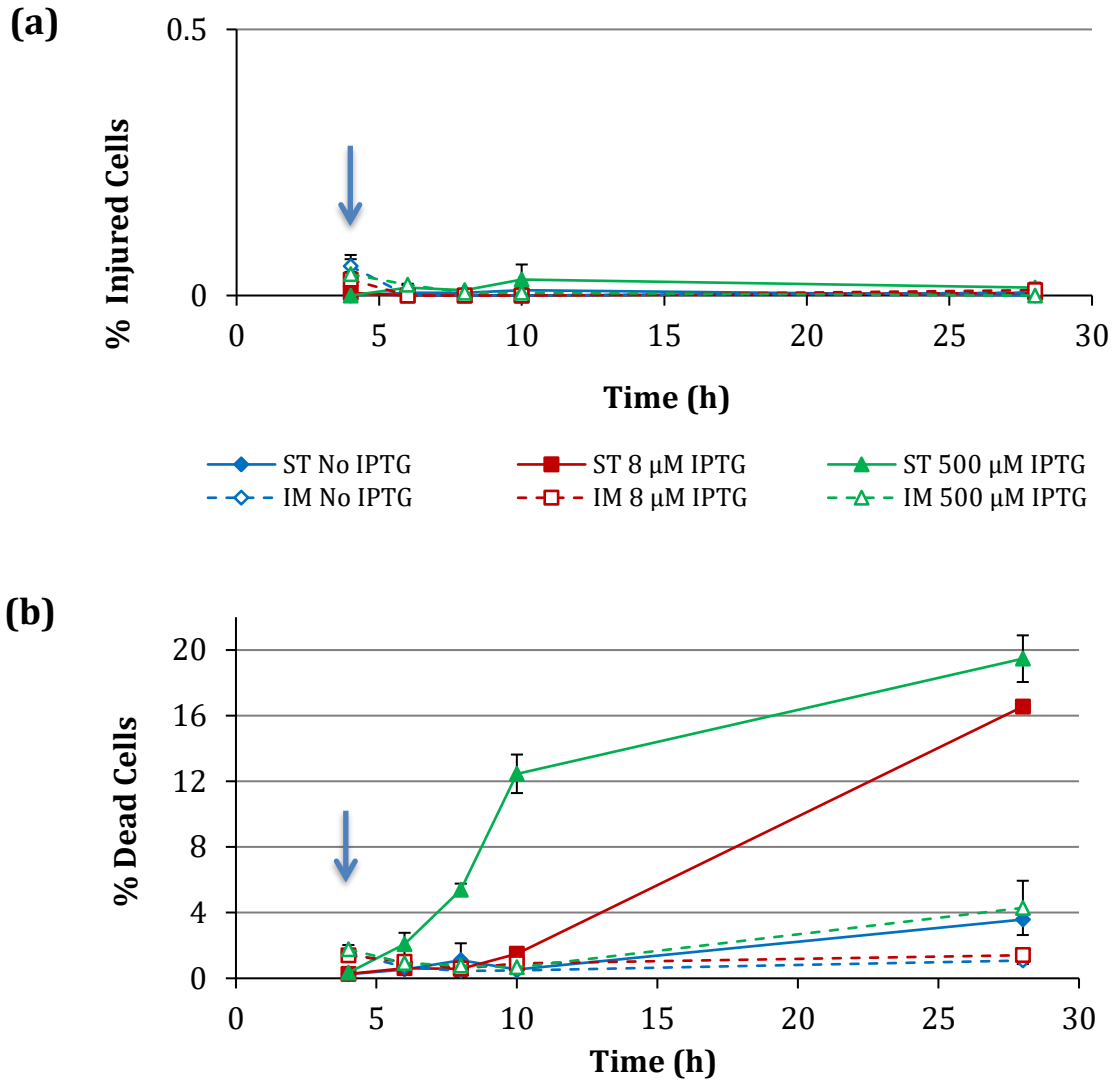


Figure 6.7: Cellular viability measured by flow cytometry for BL21(DE3)-pET41c-*fumB* cultures in batch culture with ST and IM conditions

FumB protein was expressed from an IPTG-induced pET41c plasmid with different concentrations of IPTG (8 µM and 500 µM) and compared to two different protocols. (a) Percentage of injured cells after staining with BOX. (b) Percentage of dead cells after staining with BOX+PI. (c) FCM dot plot data of FL1-A versus FL3-A showed subpopulations in 500 µM IPTG induced cultures of (c.i) ST and (c.ii) IM after 24 hours post-induction. Data are mean values of duplicate samples, error bars are ±1 standard deviation.

By using the gating strategy (Details in **Figure 2.4**; Section **2.9.4**), these data can also be seen more clearly as in **Figure 6.8a** and **6.8b**. With propidium iodide staining, these data not only showed the percentage of dead cells, but also the right percentage of live and injured cells of all populations, in both ST and IM conditions after 24 h post-induction.

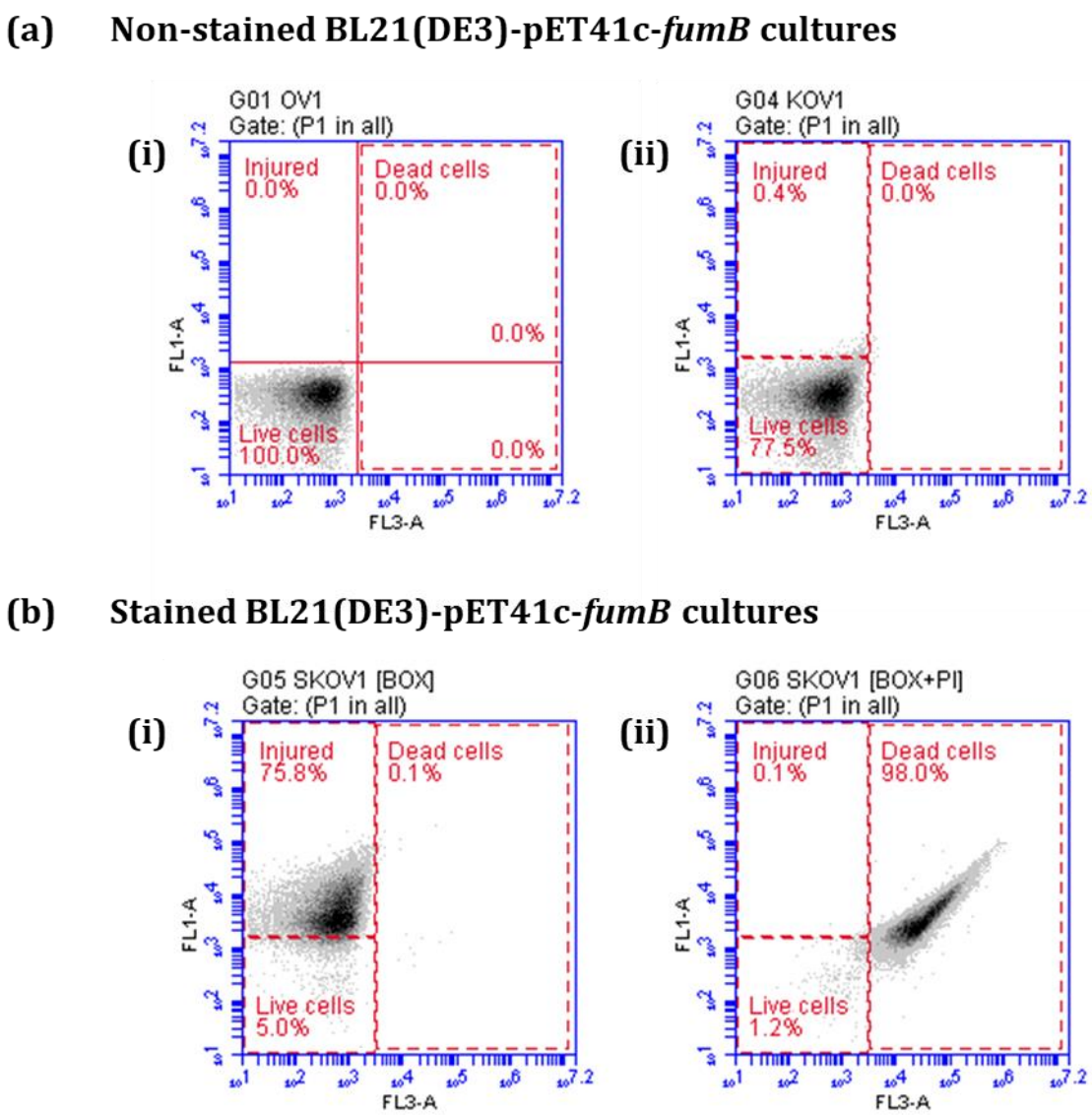


Figure 6.8: Gating strategy to identify live/injured/ dead cells population

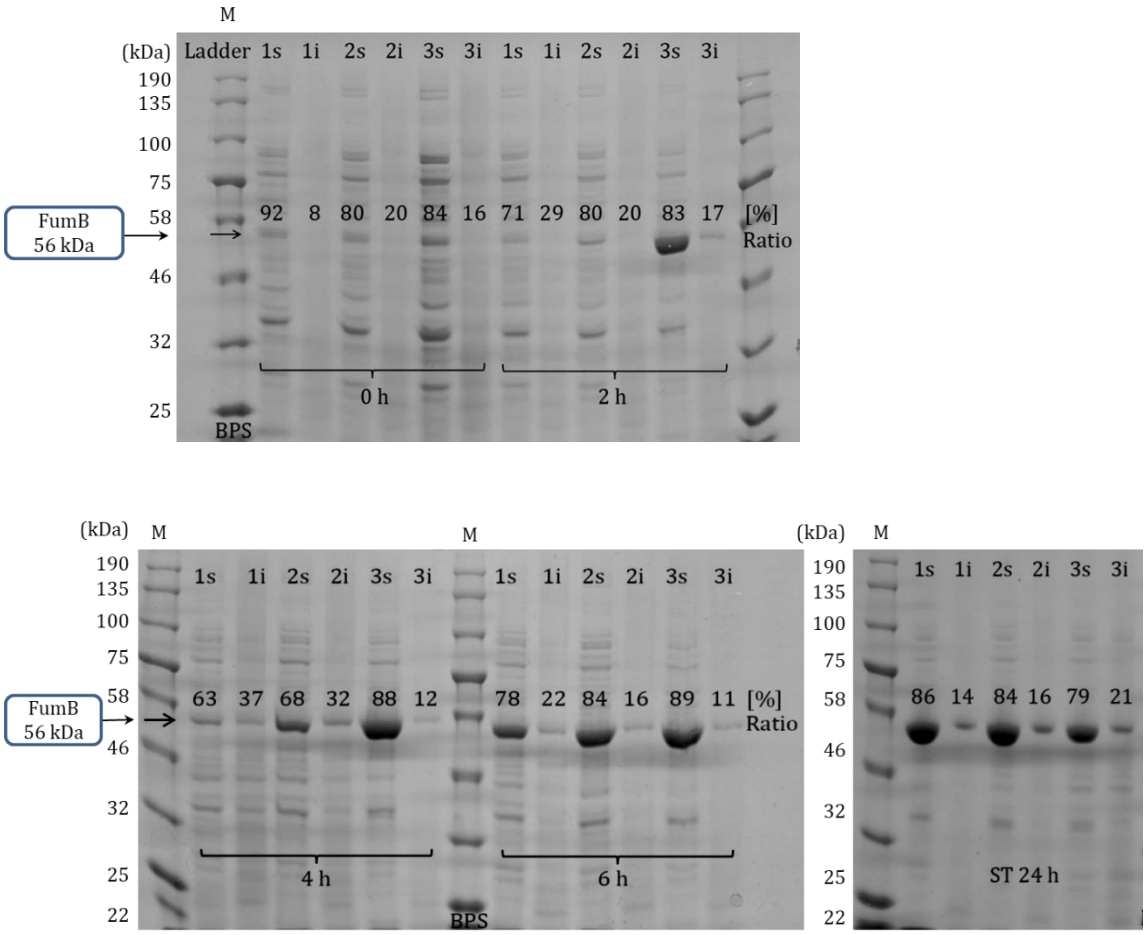
FCM dot plot data of FL1-A versus FL3-A showed subpopulations in the cultures of positive control that represents heterogeneity after 24 hours post-induction of BL21(DE3)-pET41c-*fumB*. (a) is live cells data of non-stained overnight cultures (a.i) before and (a.ii) after boiling in 99 °C water bath for 10 min. (b) is data (a.ii) when (b.i) stained with BOX to determine subpopulation of injured cells and (b.ii) stained with BOX+PI to identify population percentage of dead cells.

6.5.3 SDS-PAGE for FumB production

The production of Fumarase B (FumB) protein was evaluated by SDS-PAGE from the samples that were collected at time points 0 h, 2 h, 4 h, 6 h and 24 h post induction. Bacterial pellets obtained from the samples were fractionated to obtain soluble and insoluble protein fractions using freeze-thaw method. Lysozyme was added into the mixture for cell lysis.

Protein bands were detected in samples obtained from both standard and improved protocol cultures, carrying the pET41c plasmid containing the sequence coding for FumB (pET41c-*fumB*) at around 56 kDa, which was smaller than expected (60 kDa). The protein band intensity increased over the time after induction, mostly for the ones with soluble fractions.

Densitometry data for the percentage ratio of soluble versus insoluble FumB accumulated inside the cell extracts were shown on the SDS-PAGE gel pictures in **Figure 6.9** and **Figure 6.10**. The estimated protein solubility of soluble and insoluble FumB fractions in cell extracts of standard (ST) condition (**Figure 6.11a**) and improved (IM) condition (**Figure 6.11b**) at concentration of mg/mL at times post-induction.



M: Marker used; Blue Protein Standard (BPS), broad range of 11-190 kDa;

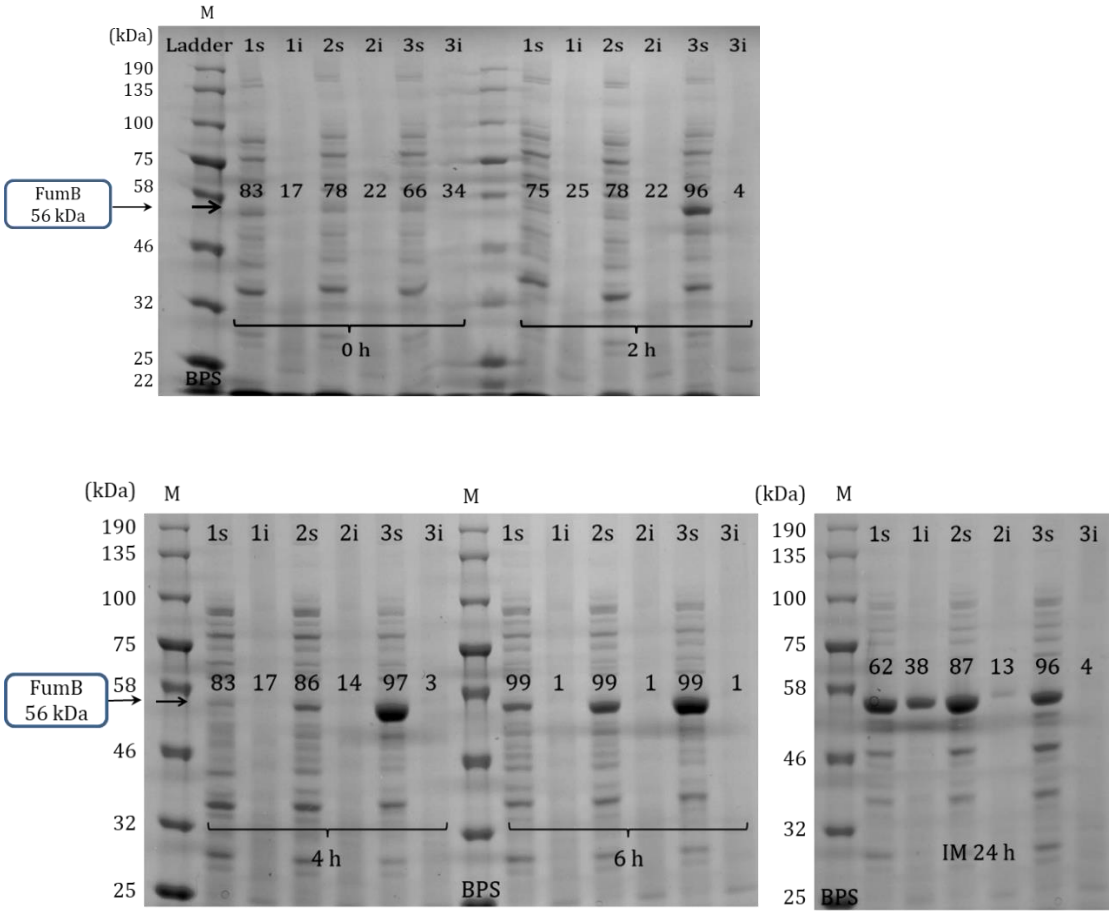
1s: ST No IPTG (Soluble); **1i**: ST No IPTG (Insoluble);

2s: ST 8µM IPTG (Soluble); **2i**: ST 8µM IPTG (Insoluble);

3s: ST 500µM IPTG (Soluble); **3i**: ST 500µM IPTG (Insoluble)

Figure 6.9: The accumulation of FumB subcellular fractionation in cell extracts expressed by SDS-PAGE for standard (ST) condition

SDS-PAGE gel picture with arrow showing FumB (56 kDa) and the ratio percentage of soluble versus insoluble FumB proteins accumulated inside the cell extracts of BL21(DE3)*-pET41c-*fumB* with standard (ST) condition, as shown above the band. All sample bands at times post-induction were compared to BPS ladder band.



M: Marker used; Blue Protein Standard (BPS), broad range of 11-190 kDa;

1s: IM No IPTG (Soluble); **1i**: IM No IPTG (Insoluble);

2s: IM 8µM IPTG (Soluble); **2i**: IM 8µM IPTG (Insoluble);

3s: IM 500µM IPTG (Soluble); **3i**: IM 500µM IPTG (Insoluble)

Figure 6.10: The accumulation of FumB subcellular fractionation in cell extracts expressed by SDS-PAGE of improved (IM) condition

SDS-PAGE gel picture with arrow showing FumB (56 kDa) and the ratio percentage of soluble versus insoluble FumB proteins accumulated inside the cell extracts of BL21(DE3)*-pET41c-*fumB* with improved (IM) condition, as shown above the band. All sample bands at times post-induction were compared to BPS ladder band.

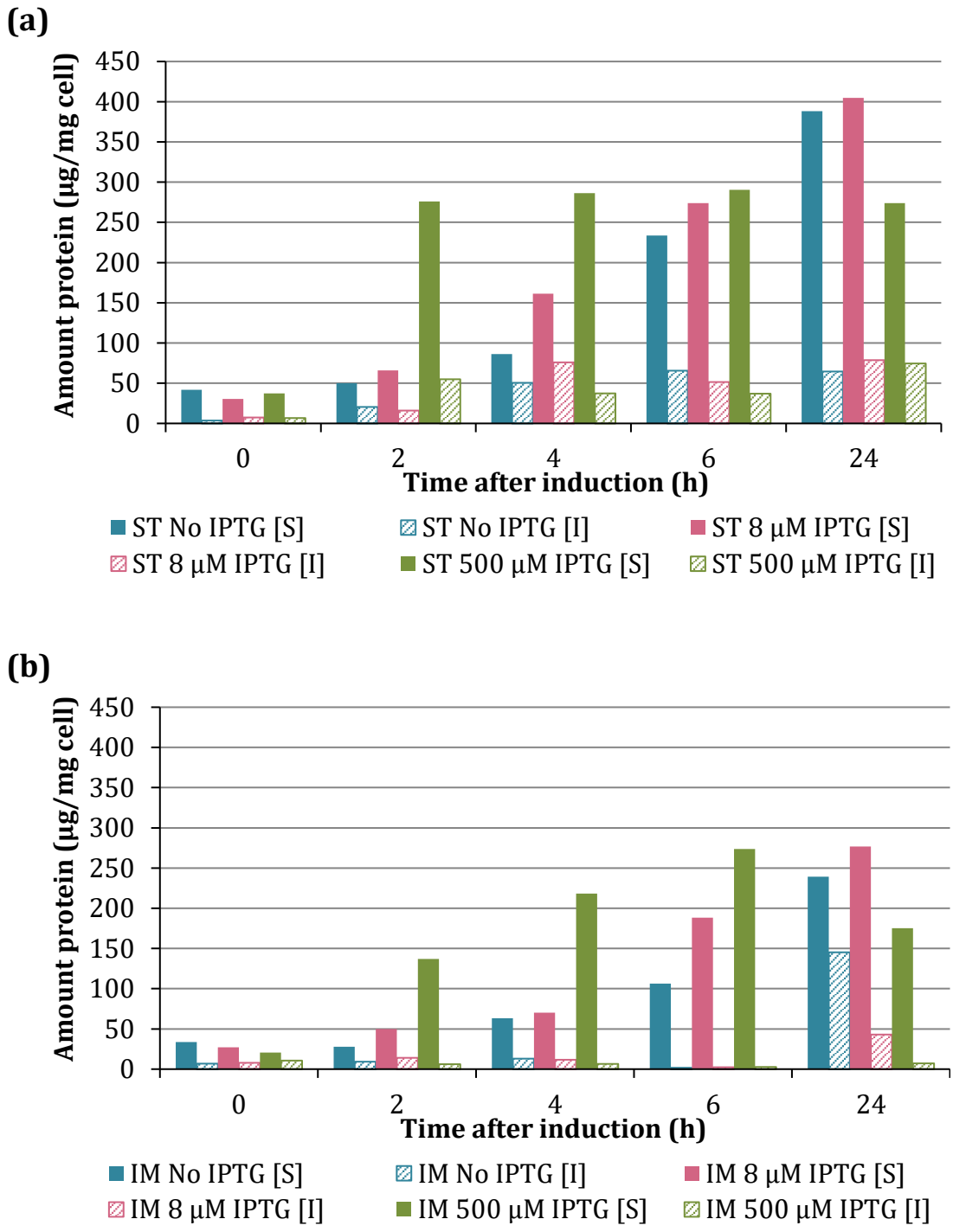


Figure 6.11: The expression of protein solubility in cell extracts measured by SDS-PAGE for FumB production in batch culture

Estimated protein solubility of soluble and insoluble FumB fractions in cell extracts of (a) standard (ST) condition and (b) improved (IM) condition at concentration of mg/mL at times post-induction.

In general, for ST condition, the decrease in the temperature from 37 °C to 25 °C post induction showed a great production of FumB, especially for 500 μM IPTG induced cultures, which started to increase at 2 h post induction (**Figure 6.12**). All induced cultures of ST showed a high plasmid retention (**Figure 6.8c**); thus, the accumulation of FumB was also higher compared to IM cultures (**Figure 6.13**). Castiñeiras (2017) claimed that lower cultivation temperatures in combination with a lower concentration of antibiotics minimised the physiological stress imposed on the cells.

Therefore, lowering the temperature to 25°C after the induction on ST protocol was a wise decision that could increase the synthesis of soluble FumB. However, the use of 25°C temperature during incubation (from starting until the end of incubation) as designed for IM protocol was not seen to be able to produce a better amount of FumB protein, yet the amount of protein produced was much lower.

While for IPTG concentration used for induction to the cultures, the usage of 8 μM IPTG in the experiment for both protocols might be too low for the cultures from 2 h to 6 h post induction, thus accumulation of FumB was also lower than when higher concentrations of IPTG were used, not being enough to fully induce the T7 promoter. However, after 24 h of induction, induced cultures of 8 μM IPTG of both ST and IM conditions showed a greater improvement of FumB production when compared with results obtained for 500 μM IPTG induced cultures.

It can be argued that even though the graph in **Figure 6.9b** showed a higher percentage of dead cells (nearly 20 %) in both induced cultures of ST, and they also showed a lower cellular culturability in CFU at 4 h post induction, but they still produced a great amount of FumB (274 – 404 μM/mg cell) compared to IM cultures (175 – 277 μM/mg cell). This might be because they tend to maintain their plasmid retention throughout the experiment (**Figure 6.8c**), thus able to produce FumB protein before they died. Though, the plasmid retention of 500 μM IPTG induced cultures of the same condition decreased more than 10 % after 24 h post induction, causing FumB protein production to decrease. As cell stress can often be found associated with plasmid loss, the increment of the inducer concentration also had a negative effect on plasmid maintenance.

Contradictory to the results for IM condition, even though the FCM showed a lower percentage of dead cells in all IM cultures (less than 5 %), which indicated low stress profile, but the cultures were still unable to produce higher amount of FumB compared to ST cultures. This was obviously because they had lost more than 10% of their plasmid throughout the incubation especially those were induced with IPTG (**Figure 6.8c**). Still, the production of FumB in these cultures were mostly soluble (**Figure 6.13b**).

Overall, the standard protocol (ST) with the beginning temperature of 37°C, then reduced to 25°C post induction at mid-logarithmic phase (OD_{650} ≈ 0.5) and a concentration of IPTG of 8 μM were found to be optimal for the better production of soluble FumB.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

This thesis stands out for its novel contribution because it discusses how a TNFα-GFP fusion was synthesised within *E. coli* bacteria. Although there has been extensive study on the generation of recombinant TNFα in *E. coli*, it is interesting to note that none of these studies have explored the utilisation of a GFP fusion technique, positioning our work at the forefront of novel approaches in the field. Furthermore, it is important to highlight that the majority of studies have used GFP fused to recombinant proteins as a way to meticulously fine-tune both growth parameters and the selection of expression vectors. While a handful of studies have indeed delved into the realm of RP-GFP fusions, it is interesting to note that none investigated how RP-GFP fusion expression compared to standalone GFP expression. This particular aspect highlights the novelty in the present work, further adding a clear layer of originality.

The most ideal host for RPP often *E. coli*, which is frequently recognised for its extraordinary ability in the field of bioprocessing. *E. coli* has an outstanding tendency for faster growth, which results in attaining of significantly higher cellular densities. This bacterium exhibits additional characteristics that set it apart, particularly its ability to thrive in cultures that utilise cost-effective growth medium. The biopharmaceutical industry finds this particular aspect extremely beneficial for industrial applications. Maintaining cellular physiology and reducing metabolic demands are of utmost importance while attempting to produce RPP. As a result, the production of energy becomes of utmost importance for the host cells, particularly in the process of making RP (Hoffmann et al., 2004). The rates of

expression and the proper folding of recombinant proteins are influenced by both the degree of induction and the stability of mRNA (Castiñeiras, 2017).

This study employed two different promoter systems, T7 and p*araBAD* promoter. The T7 promoter system was employed for the BL21(T7) strain, whilst the *araBAD* promoter system was utilised for the BL21(A) strain, for the production of GFP and TNFα-GFP by batch culture and fed-batch fermentation. The T7 promoter has been identified as a promoter with leaky characteristics, which resulted the occurrence of leaky expression. While the *araBAD* promoter, was identified for its stringent regulatory mechanisms, do not initiate the production of RP in the absence of the inducer arabinose. Based on the findings of this work, the utilisation of the T7 promoter system resulted in the generation of multiple populations, reflecting the heterogeneity of the cultures, particularly towards the end stages of incubation. In contrast to *araBAD* promoter system, a homogeneous population of cells was seen throughout the fermentation process.

The T7 expression system stands out as an excellent choice for expediting RPP, exhibiting a high peak in fluorescence values corresponding with increased levels of GFP/TNFα-GFP synthesis within an interval of 4 to 6 h after induction, particularly in the presence of glycerol. Growth in the presence of glucose seems to extend productivity. Conversely, the p*araBAD* expression system exhibits a contrasting profile, providing a more enduring platform for sustained RPP. In particular, it exhibits increased repression in the presence of glucose. Remarkably, the T7 promoter system exhibits a higher level of leakiness, as evidenced by uninduced cultures showing fluorescence levels akin to induced cultures after 28 h

of growth. *araBAD* promoter system, on the other hand, demonstrates a precise regulation of RPP, ensuring a stricter regulation of protein expression levels. This characteristic highlights its ability to effectively control and attain desired levels of production in a regulated and ideal manner.

As determined by SDS-PAGE and subcellular fractionation analysis, the evaluation of GFP/ TNFα-GFP solubility within both T7 and p*araBAD* cultures, whether in a soluble or non-soluble state, perfectly matches the levels of green fluorescence measured by FCM. This integration of findings emphasises the importance of FCM in this field of study and highlights its significant relevance in defining the solubility profiles of proteins. FCM stands as an indicator of high reliability in experimental analysis. FCM provides an unexpected degree of clarity and speed in revealing the complexities of the RPP. Cultures producing TNFα-GFP generate less fluorescence than those that are solely focused on GFP production. This observation highlights the complex process of TNFα-GFP synthesis and folding, illustrating a pathway less straightforward than that traversed by GFP.

Our extensive research efforts have definitively demonstrated that the incorporation of FCM alongside with GFP fusions is a highly valuable approach in the field of RPP optimisation investigation. The utilisation of green fluorescence, serving as a fast and precise indicator, is crucial for rapidly determining the production and solubility of recombinant proteins. Additionally, the utilisation of GFP as a measurable indicator has facilitated our exploration of the complex domain of promoter activity, providing an understanding of the regulatory

mechanisms that control gene expression within the particular expression system. Moreover, a thorough investigation of cultures using GFP and TNFα-GFP has yielded a comprehensive understanding of the mechanisms by which TNFα impacts both RP productivity and host cell physiology, further enhancing our ability to fine-tune the bioprocessing strategies employed.

Using this methodology, we conducted a comprehensive comparative analysis between the T7 and p*araBAD*-based expression systems, concluding that the latter exhibits superior regulatory accuracy, exerting a more minimal impact on the physiological balance of *E. coli.* This important finding highlights the possibility for improving RPP strategies in addition to demonstrating the many dynamics of these expression platforms. Furthermore, the scope of this research extends beyond its current purview, offering a fertile ground for the exploration of diverse recombinant protein-GFP fusions. Moreover, the versatility of fluorescence measurement renders it amenable to high-throughput methodologies, facilitating the swift optimisation of expression profiles across an expansive experimental landscape, whether through the utilisation of multiwell plates or the implementation of multi-vessel bioreactor setups.

The generation of the functional *E. coli* Fumarase B (FumB) enzyme by batch culture was another aspect that was investigated in this study. Following restriction-free (RF) cloning of the *fumB* gene into the pET41c plasmid, the recombinant plasmid was then transformed into *E. coli* BL21(DE3)* for further growth optimisation in batch culture and analysis using FCM and SDS-PAGE to produce soluble and functional FumB protein. By utilising this strain as the

microbial host, together with recombinant pET41c-*fumB* plasmid in this study, we were able to increase the production of the highly soluble and functional FumB protein. This cloning technique did provide a quick and efficient way to insert DNA fragments precisely into a specific location on the plasmid without changing any of the genes or plasmid segments that were being targeted. The recombinant proteins FumB are frequently utilised in many laboratory research projects as reagents for analysing protein synthesis, structure, and function in cells and organisms, therefore this finding indeed is very beneficial.

REFERENCES

Ahmad, I., Nawaz, N., Darwesh, N.M., et al. (2018) Overcoming challenges for amplified expression of recombinant proteins using *Escherichia coli*. ***Protein Expression and Purification***, 144 (October 2017): 12–18.

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

Alfasi, S., Sevastsyanovich, Y., Zaffaroni, L., et al. (2011) Use of GFP fusions for the isolation of *Escherichia coli* strains for improved production of different target recombinant proteins. ***Journal of Biotechnology*****,** 156 (1): 11–21.

Ali, S.A., Chew, Y.W., Omar, T.C., et al. (2015) Use of FabV-Triclosan Plasmid Selection System for Efficient Expression and Production of Recombinant Proteins in *Escherichia coli*, pp. 1–15.

Anane, E., López C, D.C., Neubauer, P., et al. (2017) Modelling overflow metabolism in *Escherichia coli* by acetate cycling. ***Biochemical Engineering Journal***, 125: 23–30.

Anuar, M., Muhammad, N., Munir, A., et al. (2012) Optimization of a Heterologous Signal Peptide by Site-Directed Mutagenesis for Improved Secretion of Recombinant Proteins in *Escherichia coli*, pp. 48–58.

Babu, K.R., Swaminathan, S., Marten, S., et al. (2000) Production of interferon-α in high cell density cultures of recombinant *Escherichia coli* and its single step purification from refolded inclusion body proteins. ***Applied Microbiology and Biotechnology***, 53 (6): 655–660.

Baeshen, N.A., Baeshen, M.N., Sheikh, A., et al. (2014) Cell factories for insulin production. ***Microbial Cell Factories***, 13 (1): 141.

Baeshen, N.A., Ramadan, H.A.I., Saini, K.S., et al. (2015) Production of Biopharmaceuticals in *E. coli*. ***Current Scenario and Future Perspectives***, 25 (7):

953–962.

Baneyx, F. (1999) Recombinant protein expression in *Escherichia coli.* ***Current Opinion in Biotechnology*****,** 10 (5): 411–421.

Bao, R.M., Yang, H.M., Yu, C.M., et al. (2016) An efficient protocol to enhance the extracellular production of recombinant protein from *Escherichia coli* by the synergistic effects of sucrose, glycine, and Triton X-100. ***Protein Expression and Purification*****,** 126: 9–15.

Barbara, M.., van der Weel, L., Hagen, W.R., et al. (2013) Biochemical similarities and differences between the catalytic [4Fe-4S] cluster containing fumarases FumA and FumB from *Escherichia coli*. ***PloS ONE***, 8 (2): e55549.

Bashir, H., Ahmed, N., Khan, M.A., et al. (2016) Simple procedure applying lactose induction and one-step purification for high-yield production of rhCIFN. ***Biotechnology and Applied Biochemistry*****,** 63 (5): 708–714.

Benoit, R.M., Wilhelm, R.N., Scherer-becker, D., et al. (2006) An improved method for fast , robust , and seamless integration of DNA fragments into multiple plasmids. ***Protein expression and purification*****,** 45: 66–71.

Betts, J.P.J., Warr, S.R.C., Finka, G.B., et al. (2014) Impact of aeration strategies on fed-batch cell culture kinetics in a single-use 24-well miniature bioreactor. ***Biochemical Engineering Journal*****,** 82: 105–116.

Bhatwa, A., Wang, W., Hassan, Y.I., et al. (2021) Challenges Associated With the Formation of Recombinant Protein Inclusion Bodies in *Escherichia coli* and Strategies to Address Them for Industrial Applications. ***Frontiers in Bioengineering and Biotechnology***, 9: 1–18.

Briand, L., Marcion, G., Kriznik, A., et al. (2016) A self-inducible heterologous protein expression system in *Escherichia coli*. ***Scientific Reports*****,** 6: 1–11.

Broger, T., Odermatt, R.P., Huber, P., et al. (2011) Real-time on-line flow cytometry for bioprocess monitoring. ***Journal of Biotechnology*****,** 154 (4): 240–247.

Buchholz, K. and Collins, J. (2013) The roots - A short history of industrial microbiology and biotechnology. ***Applied Microbiology and Biotechnology*****,** 97 (9): 3747–3762.

Burdette, L.A., Leach, S.A., Wong, H.T., et al. (2018) Developing Gram-negative bacteria for the secretion of heterologous proteins. ***Microbial Cell Factories*****,** 17 (1): 1–16.

Carswell, E.A., Old, L.J., Kassel, R.L., et al. (1975) An endotoxin-induced serum factor that causes necrosis of tumors. ***Proceedings of the National Academy of Sciences*****,** 72 (9): 3666–3670.

Castiñeiras, T.S. (2017) Development of innovative screening procedures and fermentation processes for the production of recombinant proteins in *E . coli*. PhD Thesis. University of Birmingham.

Castiñeiras, T.S., Williams, S.G., Hitchcock, A., et al. (2018a) Development of a generic β-lactamase screening system for improved signal peptides for periplasmic targeting of recombinant proteins in *Escherichia coli.* ***Scientific Reports*****,** 8 (1): 1–18.

Castiñeiras, T.S., Williams, S.G., Hitchcock, A., et al. (2018b) Optimizing host cell physiology and stress avoidance for the production of recombinant human tumour necrosis factor α in *Escherichia coli*. ***Microbiology (United Kingdom)*****,** 164 (4): 440–452.

Choi, J.H., Keum, K.C. and Lee, S.Y. (2006) Production of recombinant proteins by high cell density culture of *Escherichia coli.* ***Chemical Engineering Science*****,** 61 (3): 876–885.

Clark, D.P. and Pazdernik, N.J. (2016) Avoiding Toxic Effects of Protein Overproduction. ***Biotechnology*** *(Second Edition)*.

Clark, E.D. (2001) Protein refolding for industrial processes. ***Current opinion in biotechnology***, 12 (2): 202–7.

Constantinidou, C., Hobman, J.L., Griffiths, L., et al. (2006) A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL, and NarQP as *Escherichia coli* K12 adapts from aerobic to anaerobic growth. ***The Journal of biological chemistry*****,** 281: 4802–4815.

Cormack, B.P., Valdivia, R.H. and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). ***Gene***, 173 (1): 33–38.

Costa, A.R., Rodrigues, M.E., Henriques, M., et al. (2014) Feed optimization in fed-batch culture. ***Methods in Molecular Biology*****,** 1104: 105–116.

Cronan, J.E. (2006) A family of arabinose-inducible *Escherichia coli* expression vectors having pBR322 copy control. ***Plasmid***, 55 (2): 152–157.

Croset, A., Delafosse, L., Gaudry, J.-P., et al. (2012) Differences in the glycosylation of recombinant proteins expressed in HEK and CHO cells. ***Journal of biotechnology*****,** 161 (3): 336–48.

Davey, H.M. and Kell, D.B. (1996) Flow Cytometry and Cell Sorting of Heterogeneous Microbial Populations: The Importance of Single-Cell Analyses. ***Microbiological Reviews***, 60 (4): 641–696.

Demain, A.L. and Vaishnav, P. (2009) Production of recombinant proteins by microbes and higher organisms. ***Biotechnology Advances***, 27 (3): 297–306.

Deutscher, J., Francke, C. and Postma, P.W. (2006) How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria. ***Microbiology and Molecular Biology Reviews*****,** 70 (4): 939–1031.

Díaz, M., Herrero, M., García, L.A., et al. (2010) Application of flow cytometry to industrial microbial bioprocesses. ***Biochemical Engineering Journal*****,** 48 (3): 385–407.

Doekel, S., Eppelmann, K. and 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 Microbiology Letters*****,** 216(2): 185–

191.

Doglia, S.M. and Lotti, M. (2014) Protein Aggregation in Bacteria: Functional an Structural Properties of Inclusion Bodies in Bacterial Cells.

Dowlati, Y., Herrmann, N., Swardfager, W., et al. (2010) A meta-analysis of cytokines in major depression. ***Biological psychiatry*****,** 67 (5): 446–57.

Dvorak, P., Chrast, L., Nikel, P.I., et al. (2015) Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway. ***Microbial Cell Factories*****,** 14 (1): 1–15.

Eiberle, M.K. and Jungbauer, A. (2010) Technical refolding of proteins: Do we have freedom to operate? ***Biotechnology Journal***, 5 (6): 547–559.

Eiteman, M.A. and Altman, E. (2006) Overcoming acetate in *Escherichia coli* recombinant protein fermentations. ***Trends in Biotechnology***, 24 (11): 530–536.

Van Den Ent, F. and Lowe, J. (2006) RF cloning: A restriction-free method for inserting target genes into plasmids. ***Journal of biochemical and biophysical methods***, 67 (1): 67–74.

Faust, G., Janzen, N.H., Bendig, C., et al. (2014) Feeding strategies enhance high cell density cultivation and protein expression in milliliter scale bioreactors. ***Biotechnology Journal***, 9 (10): 1293–1303.

Faust, G., Stand, A. and Weuster-Botz, D. (2015) IPTG can replace lactose in auto-induction media to enhance protein expression in batch-cultured *Escherichia coli*. ***Engineering in Life Sciences*****,** 15 (8): 824–829.

Fernández-Castané, A., Li, H., Franzreb, M., et al. (2016) Using flow cytometry to monitor cell physiology of magnetotactic bacteria in fermentation processes. ***New Biotechnology***, 33: S19.

Fernández-Castané, A., Li, H., Thomas, O.R.T., et al. (2017) Flow cytometry as a rapid analytical tool to determine physiological responses to changing O_2 and iron

concentration by *Magnetospirillum gryphiswaldense* strain MSR-1. ***Scientific Reports***, 7 (1): 1–11.

Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J.L., et al. (2009) Microbial factories for recombinant pharmaceuticals. ***Microbial cell factories*****,** 8: 17.

Finkelstein, J., Antony, E., Hingorani, M.M., et al. (2003) Overproduction and analysis of eukaryotic multiprotein complexes in *Escherichia coli* using a dual-vector strategy. ***Analytical Biochemistry***, 319: 78–87.

García-Fruitós, E. (2010) Inclusion bodies: A new concept. ***Microbial Cell Factories*****.** 9 (1) p. 80.

García-Fruitós, E., González-Montalbán, N., Morell, M., et al. (2005) Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. ***Microbial Cell Factories***, 4: 1–6.

Gasser, B., Saloheimo, M., Rinas, U., et al. (2008) Protein folding and conformational stress in microbial cells producing recombinant proteins: A host comparative overview. ***Microbial Cell Factories*****,** 7: 1–18.

Gatti-Lafranconi, P., Natalello, A., Ami, D., et al. (2011) Concepts and tools to exploit the potential of bacterial inclusion bodies in protein science and biotechnology. ***FEBS Journal*****,** 278 (14): 2408–2418.

Geiser, M., Cebe, R., Drewello, D., et al. (2001) Integration of PCR Fragments at Any Specific Site within Cloning Vectors without the Use of Restriction Enzymes and DNA Ligase. ***BioTechniques*****,** (31): 88–92.

Geng, J., Beloin, C., Ghigo, J., et al. (2014) Bacteria Hold Their Breath upon Surface Contact as Shown in a Strain of *Escherichia coli ,* Using Dispersed Surfaces and Flow Cytometry Analysis., ***PloS ONE***, 9 (7): 1–10.

Gill, R.T., Valdes, J.J. and Bentley, W.E. (2000) A comparative study of global stress gene regulation in response to overexpression of recombinant proteins in *Escherichia coli.* ***Metabolic Engineering*****,** 2 (3): 178–189.

Goeddel, D. V., Kleid, D.G. and Bolivar, F. (1979) Expression in *Escherichia coli* of chemically synthesized genes for human insulin. ***Proceedings of the National Academy of Sciences of the United States of America*****,** 76 (1): 106–110.

Gonçalves, A.M., Pedro, A.Q., Maia, C., et al. (2013) *Pichia pastoris*: A recombinant microfactory for antibodies and human membrane proteins. ***Journal of Microbiology and Biotechnology*****,** 23 (5): 587–601.

Gopal, G.J. and Kumar, A. (2013) Strategies for the production of recombinant protein in *Escherichia coli*. ***The protein journal*****,** 32 (6): 419–425.

Gräslund, S., Nordlund, P., Weigelt, J., et al. (2008) Protein production and purification. ***Nature Methods*****,** 5 (2): 135–146.

Gupta, S.K. and Shukla, P. (2016) Advanced technologies for improved expression of recombinant proteins in bacteria: perspectives and applications. ***Critical Reviews in Biotechnology*****,** 36 (6): 1089–1098.

Guzman, L.M., Belin, D., Carson, M.J., et al. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose p*(BAD)* promoter. ***Journal of Bacteriology***, 177 (14): 4121–4130.

Haddadin, F.T. and Harcum, S.W. (2005) Transcriptome profiles for high-cell-density recombinant and wild-type *Escherichia coli*. ***Biotechnology and Bioengineering*****,** 90 (2): 127–153.

Hadj Sassi, A., Trigui-Lahiani, H., Abdeljalil, S., et al. (2017) Enhancement of solubility, purification and inclusion-bodies-refolding of an active pectin lyase from *Penicillium occitanis* expressed in *Escherichia coli*. ***International Journal of Biological Macromolecules***, 95: 256–262.

Hannig, G. and Makrides, S.C. (1998) Strategies for optimizing heterologous protein expression in *Escherichia coli*. ***Trends in biotechnology*****,** 16 (2): 54–60.

Heim, R. and Tsien, R.Y. (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy

transfer. ***Current Biology*****,** 6 (2): 178–182.

Herzberg, M., Kaye, I.K., Peti, W., et al. (2006) YdgG (TqsA) controls biofilm formation in *Escherichia coli* K-12 through autoinducer 2 transport. ***Journal of Bacteriology*****,** 188 (2): 587–598.

Hoffmann, F., Van Den Heuvel, J., Zidek, N., et al. (2004) Minimizing inclusion body formation during recombinant protein production in *Escherichia coli* at bench and pilot plant scale. ***Enzyme and Microbial Technology*****,** 34 (3–4): 235–241.

Horga, L.G., Halliwell, S., Castiñeiras, T.S., et al. (2018) Tuning recombinant protein expression to match secretion capacity. ***Microbial Cell Factories*****,** 17:199.

Horiuchi, T., Mitoma, H., Harashima, S., et al. (2010) Transmembrane TNF-alpha: structure, function and interaction with anti-TNF agents. ***Rheumatology (Oxford, England)*****,** 49: 1215–1228.

Hothersall, J., Godfrey, R.E., Fanitsios, C., et al. (2021) The PAR promoter expression system: Modified *lac* promoters for controlled recombinant protein production in *Escherichia coli*. ***New Biotechnology*****,** 64: 1–8.

Hothersall, J., Lai, S., Zhang, N., et al. (2022a) Inexpensive protein overexpression driven by the NarL transcription activator protein. ***Biotechnology and Bioengineering*****,** 119 (6): 1614–1623.

Hothersall, J., Osgerby, A., Godfrey, R.E., et al. (2022b) New vectors for urea-inducible recombinant protein production. ***New Biotechnology*****,** 72: 89–96.

Hsu, C.C., Thomas, O.R.T. and Overton, T.W. (2016) Periplasmic expression in and release of Fab fragments from *Escherichia coli* using stress minimization. ***Journal of Chemical Technology and Biotechnology***, 91 (3): 815–822.

Huang, C.-J., Lin, H. and Yang, X. (2012) Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. ***Journal of industrial microbiology & biotechnology*****,** 39 (3): 383–99.

Ishii, J., Izawa, K., Matsumura, S., et al. (2009) A simple and immediate method for simultaneously evaluating expression level and plasmid maintenance in yeast. ***Journal of Biochemistry*****,** 145 (6): 701–708.

Jana, S. and Deb, J.K. (2005) Strategies for efficient production of heterologous proteins in *Escherichia coli*. ***Applied microbiology and biotechnology*****,** 67 (3): 289–98.

Jensen, E.B. and Carlsen, S. (1990) Production of recombinant human growth hormone in *Escherichia coli*: Expression of different precursors and physiological effects of glucose, acetate, and salts. ***Biotechnology and Bioengineering***, 36 (1): 1–11.

Jia, B. and Jeon, C.O. (2016) High-throughput recombinant protein expression in Escherichia coli : current status and future perspectives. ***Open Biol***. 6: 160196.

Jozala, A.F., Geraldes, D.C., Tundisi, L.L., et al. (2016) Biopharmaceuticals from microorganisms: from production to purification. ***Brazilian Journal of Microbiology*****,** 47: 51–63.

Kamionka, M. (2011) Engineering of therapeutic proteins production in *Escherichia coli*. ***Current pharmaceutical biotechnology*****,** 12 (2): 268–74.

Kangwa, M., Yelemane, V., Polat, A.N., et al. (2015) High-level fed-batch fermentative expression of an engineered *Staphylococcal* protein A based ligand in *E. coli*: purification and characterization. ***AMB Express***, 5 (1): 1–10.

Kasem Balasiny, B. (2016) Regulation and Sources of Nitric Oxide in *Escherichia coli*. PhD Thesis. University of Birmingham.

Kasli, I.M., Thomas, O.R.T. and Overton, T.W. (2019) Use of a design of experiments approach to optimise production of a recombinant antibody fragment in the periplasm of *Escherichia coli*: selection of signal peptide and optimal growth conditions. ***AMB Express*****,** 9 (1): 5.

Katzke, N., Arvani, S., Bergmann, R., et al. (2010) A novel T7 RNA polymerase

dependent expression system for high-level protein production in the phototrophic bacterium *Rhodobacter capsulatus*. ***Protein Expression and Purification***, 69 (2): 137–146.

Kaur, J., Kumar, A. and Kaur, J. (2018) Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements. ***International Journal of Biological Macromolecules***, 106: 803–822.

Kesik-Brodacka, M. (2018) Progress in biopharmaceutical development. ***Biotechnology and Applied Biochemistry***, 65 (3): 306–322.

Khan, M.M.T., Pyle, B.H. and Camper, A.K. (2010) Specific and Rapid Enumeration of Viable but Nonculturable and Viable-Culturable Gram-Negative Bacteria by Using Flow Cytometry. ***Applied and Environmental Microbiology,*** 76(15): 5088-5096.

Khan, S., Ullah, M.W., Siddique, R., et al. (2016) Role of recombinant DNA technology to improve life. ***International Journal of Genomics*****,** 2016.

Korz, D.J., Rinas, U., Hellmuth, K., et al. (1995) Simple fed-batch technique for high cell density cultivation of *Escherichia coli*. ***Journal of Biotechnology***, 39 (1): 59–65.

Lagendijk, E., Koppert, B. V, Validov, S., et al. (2010) Genetic tools for tagging Gram-negative bacteria with mCherry for visualization in vitro and in natural habitats, biofilm and pathogenicity studies. ***FEMS Microbiol Lett***, 305(2010) 81-90.

Lebendiker, M. and Danieli, T. (2014) Production of prone-to-aggregate proteins. ***FEBS Letters*****,** 588 (2): 236–246.

Lecina, M., Sarró, E., Casablancas, A., et al. (2013) IPTG limitation avoids metabolic burden and acetic acid accumulation in induced fed-batch cultures of *Escherichia coli* M15 under glucose limiting conditions. ***Biochemical Engineering Journal*****,** 70: 78–83.

Lee, S.Y. (1996) High cell-density culture of *Escherichia coli*. ***Trends in Biotechnology***, 14 (3): 98–105.

Lehtinen, J., Nuutila, J. and Lilius, E.M. (2004) Green fluorescent protein-propidium iodide (GFP-PI) based assay for flow cytometric measurement of bacterial viability. ***Cytometry Part A***, 60 (2): 165–172.

Lichenstein, H.S., Hamilton, E.P. and Lee, N. (1987) Repression and catabolite gene activation in the *araBAD* operon. ***Journal of Bacteriology***, 169 (2): 811–822.

Locksley, R.M., Killeen, N. and Lenardo, M.J. (2001) The TNF and TNF receptor superfamilies: Integrating mammalian biology. ***Cell***, 104 (4): 487–501.

Low, K.O., Mahadi, N.M. and Md. Illias, R. (2013) Optimisation of signal peptide for recombinant protein secretion in bacterial hosts. ***Appl Microbiol Biotechnol,*** 97:3811–3826

Luo, Q., Shen, Y.-L., Wei, D.-Z., et al. (2006) Optimization of culture on the overproduction of TRAIL in high-cell-density culture by recombinant *Escherichia coli*. ***Appl Microbiol Biotechnol***, 71: 184–191

Marschall, L., Sagmeister, P. and Herwig, C. (2016) Tunable recombinant protein expression in *E. coli*: enabler for continuous processing? ***Applied Microbiology and Biotechnology***, 100 (13): 5719–5728.

Marti, G.E., Stetler-Stevenson, M., Bleesing, J.J., et al. (2000) Introduction to flow cytometry: A Learning Guide. ***BD Biosciences*.** p. 1: 5.

Mayer, S., Junne, S., Ukkonen, K., et al. (2014) Lactose autoinduction with enzymatic glucose release: Characterization of the cultivation system in bioreactor. ***Protein Expression and Purification***, 94: 67–72.

Mühlmann, M., Forsten, E., Noack, S., et al. (2017) Optimizing recombinant protein expression via automated induction profiling in microtiter plates at different temperatures. ***Microbial Cell Factories***, 16 (1): 1–12.

Müller, S. and Nebe-von-Caron, G. (2010) Functional single-cell analyses: flow cytometry and cell sorting of microbial populations and communities. ***FEMS microbiology reviews*,** 34 (4): 554–87.

Narayanan, N. (2009) Molecular and Genetic Strategies to Enhance Functional Expression of Recombinant Protein in *Escherichia coli*. PhD Thesis. University of Waterloo.

Nebe-Von-Caron, G., Stephens, P.J., Hewitt, C.J., et al. (2000) Analysis of bacterial function by multi-colour fluorescence flow cytometry and single cell sorting. ***Journal of Microbiological Methods*****,** 42 (1): 97–114.

Overton, T.W. (2014) Recombinant protein production in bacterial hosts. ***Drug Discovery Today*****,** 19 (5): 590–601.

Peng, L. and Shimizu, K. (2003) Global metabolic regulation analysis for *Escherichia coli* K12 based on protein expression by 2-dimensional electrophoresis and enzyme activity measurement. ***Applied Microbiology and Biotechnology*****,** 61 (2): 163–178.

Peubez, I., Chaudet, N., Mignon, C., et al. (2010) Antibiotic-free selection in *E. coli:* new considerations for optimal design and improved production. ***Microbial cell factories*****,** 9: 65.

Phillips, G.J. (2001) Green fluorescent protein - a bright idea for the study of bacterial protein localization. ***FEMS microbiology letters***, 204 (1): 9–18.

Rahmen, N., Fulton, A., Ihling, N., et al. (2015) Exchange of single amino acids at different positions of a recombinant protein affects metabolic burden in *Escherichia coli*. ***Microbial Cell Factories*****,** 14 (1): 1–18.

Rameez, S., Mostafa, S.S., Miller, C., et al. (2014) High-throughput miniaturized bioreactors for cell culture process development: Reproducibility, scalability, and control. ***Biotechnology Progress***, 30 (3): 718–727.

Ramon, A., Senorale-Pose, M. and Marin, M. (2014) Inclusion bodies: Not that bad... ***Frontiers in Microbiology***, 5 (FEB): 2010–2015.

Remington, S.J. (2006) Fluorescent proteins: maturation, photochemistry and photophysics. ***Current Opinion in Structural Biology***, 16 (6): 714–721.

Retamal, C., Dewasme, L., Hantson, A.L., et al. (2018) Parameter estimation of a dynamic model of *Escherichia coli* fed-batch cultures. ***Biochemical Engineering Journal*****,** 135: 22–35.

Rinas, U., Garcia-Fruitós, E., Corchero, J.L., et al. (2017) Bacterial Inclusion Bodies: Discovering Their Better Half. ***Trends in Biochemical Sciences***, 42 (9): 726–737.

Rodríguez-Carmona, E., Cano-Garrido, O., Seras-Franzoso, J., et al. (2010) Isolation of cell-free bacterial inclusion bodies. ***Microbial Cell Factories*****,** 9 (1): 71.

Rosano, G.L. and Ceccarelli, E.A. (2014) Recombinant protein expression in *Escherichia coli*: Advances and challenges. ***Frontiers in Microbiology***, 5: 1–17.

Rosano, G.L., Morales, E.S. and Ceccarelli, E.A. (2019) New tools for recombinant protein production in *Escherichia coli*: A 5-year update. ***Protein Science***, 28 (8): 1412–1422.

Sandomenico, A., Sivaccumar, J.P. and Ruvo, M. (2020) Evolution of *Escherichia coli* expression system in producing antibody recombinant fragments. ***International Journal of Molecular Sciences***, 21 (17): 1–39.

Schaepe, S., Kuprijanov, A., Simutis, R., et al. (2014) Avoiding overfeeding in high cell density fed-batch cultures of *E. coli* during the production of heterologous proteins. ***Journal of Biotechnology*****,** 192 (Part A): 146–153.

Schmidt, F.R. (2004) Recombinant expression systems in the pharmaceutical industry. ***Applied Microbiology and Biotechnology*****,** 65 (4): 363–372.

Sevastsyanovich, Y., Alfasi, S. and Cole, J. (2009a) Recombinant protein production: a comparative view on host physiology. ***New Biotechnology*****,** 25 (4): 175–180.

Sevastsyanovich, Y., Alfasi, S., Overton, T., et al. (2009b) Exploitation of GFP fusion proteins and stress avoidance as a generic strategy for the production of high-quality recombinant proteins. ***FEMS microbiology letters*****,** 299 (1): 86–94.

Sevastsyanovich, Y.R., Alfasi, S.N. and Cole, J. a (2010) Sense and nonsense from a

systems biology approach to microbial recombinant protein production. ***Biotechnology and applied biochemistry***, 55 (1): 9–28.

Sevastsyanovich, Y.R., Leyton, D.L., Wells, T.J., et al. (2012) A generalised module for the selective extracellular accumulation of recombinant proteins. ***Microbial cell factories*****,** 11: 69.

Shaner, N.C., Patterson, G.H. and Davidson, M.W. (2007) Advances in fluorescent protein technology. ***Journal of Cell Science*****,** 120 (24): 4247–4260.

Shih, Y.-P., Kung, W.-M., Chen, J.-C., et al. (2002) High-throughput screening of soluble recombinant proteins. ***Protein Science*****,** 11 (7): 1714–1719.

Siegele, D.A. and Hu, J.C. (1997) Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. ***Proceedings of the National Academy of Sciences of the United States of America*****,** 94 (15): 8168–8172.

Singh, A., Upadhyay, V., Upadhyay, A.K., et al. (2015) Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. ***Microbial Cell Factories***. 14 (1) p. 41.

Sivashanmugam, A., Murray, V., Cui, C., et al. (2009) Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. ***Protein Science*****.** 18(5):936-48.

Slouka, C., Kopp, J., Spadiut, O., et al. (2019) Perspectives of inclusion bodies for bio-based products: curse or blessing? ***Applied Microbiology and Biotechnology*****,** 103 (3): 1143–1153.

Sohoni, S.V., Nelapati, D., Sathe, S., et al. (2015) Optimization of high cell density fermentation process for recombinant nitrilase production in *E. coli.* ***Bioresource Technology***, 188: 202–208.

Song, C.W., Kim, D.I., Choi, S., et al. (2013) Metabolic engineering of *Escherichia coli* for the production of fumaric acid. ***Biotechnology and bioengineering***, 110 (7):

2025–34.

Sørensen, H.P. and Mortensen, K.K. (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. ***Journal of biotechnology*****,** 115 (2): 113–28.

Stanbury, P.F., Whitaker, A., Hall, S.J., et al. (2017) "Chapter 1 – An introduction to fermentation processes." ***Principles of Fermentation Technology*****.** pp. 1–20.

Studier, F.W. (1991) Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. ***Journal of Molecular Biology*****,** 219 (1): 37–44.

Su, L., Wu, S., Feng, J., et al. (2019) High-efficiency expression of *Sulfolobus acidocaldarius* maltooligosyl trehalose trehalohydrolase in *Escherichia coli* through host strain and induction strategy optimization. ***Bioprocess and Biosystems Engineering***, 42 (3): 345–354.

Sundström, H., Wållberg, F., Ledung, E., et al. (2004) Segregation to non-dividing cells in recombinant *Escherichia coli* fed-batch fermentation processes. ***Biotechnology Letters***, 26 (19): 1533–1539.

Swardfager, W., Lanctôt, K., Rothenburg, L., et al. (2010) A meta-analysis of cytokines in Alzheimer's disease. ***Biological psychiatry*****,** 68 (10): 930–41.

Terpe, K. (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. ***Applied microbiology and biotechnology*****,** 72 (2): 211–22.

Tseng, C., Yu, C., Lin, H., et al. (2001) Oxygen- and Growth Rate-Dependent Regulation of *Escherichia coli* Fumarase (FumA , FumB , and FumC) Activity. ***Journal of bacteriology***, 183 (2): 461–467.

Tseng, C.P. (1997) Regulation of fumarase (*fumB*) gene expression in *Escherichia coli* in response to oxygen, iron and heme availability: role of the arcA, fur, and hemA gene products. ***FEMS microbiology letters*****,** 157 (1): 67–72.

Ukkonen, K., Mayer, S., Vasala, A., et al. (2013) Use of slow glucose feeding as supporting carbon source in lactose autoinduction medium improves the robustness of protein expression at different aeration conditions. ***Protein Expression and Purification***, 91 (2): 147–154.

Unger, T., Jacobovitch, Y., Dantes, A., et al. (2010) Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. ***Journal of structural biology*,** 172 (1): 34–44.

Upadhyay, A.K., Murmu, A., Singh, A., et al. (2012) Kinetics of inclusion body formation and its correlation with the characteristics of protein aggregates in *Escherichia coli*. ***PLoS ONE*,** 7 (3).

Upadhyay, A.K., Singh, A., Mukherjee, K.J., et al. (2014) Refolding and purification of recombinant L-asparaginase from inclusion bodies of *E. coli* into active tetrameric protein. ***Frontiers in Microbiology*,** 5: 1–10.

Valešová, R., Hollerová-Sobotková, L., Štěpánek, V., et al. (2004) Optimization of the host-plasmid interaction in the recombinant *Escherichia coli* strains overproducing penicillin G acylase. ***Enzyme and Microbial Technology***, 35 (1): 74–80.

Vallejo, L.F. and Rinas, U. (2004) Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. ***Microbial cell factories***, 3 (1): 11.

Vasina, J.A. and Baneyx, F. (1997) Expression of aggregation-prone recombinant proteins at low temperatures: A comparative study of the *Escherichia coli* cspA and tac promoter systems. ***Protein Expression and Purification*,** 9 (2): 211–218.

Vera, A., Gonza´lez-Montalba´n, N., Arı´s, A., et al. (2007) The Conformational Quality of Insoluble Recombinant Proteins Is Enhanced at Low Growth Temperatures. ***Journal of Biotechnology and Bioengineering*,** 96 (6): 1101–1106.

Villaverde, A. and Carrió, M.M. (2003) Protein aggregation in recombinant bacteria: biological role of inclusion bodies. ***Biotechnology letters*,** 25 (17): 1385–95.

Vizcaino-Caston, I., Wyre, C. and Overton, T.W. (2012) Fluorescent proteins in microbial biotechnology--new proteins and new applications. ***Biotechnology letters*****,** 34 (2): 175–86.

Walsh, G. (2010) Biopharmaceutical benchmarks 2010. ***Nature biotechnology***, 28 (9): 917–24.

Walsh, G. (2014) Biopharmaceutical benchmarks 2014. ***Nature Biotechnology*****,** 32 (10): 992–1000.

Walsh, G. (2018) Biopharmaceutical benchmarks 2018. ***Nature Biotechnology***, 36 (12): 1136–1145.

Wandrey, G., Bier, C., Binder, D., et al. (2016) Light-induced gene expression with photocaged IPTG for induction profiling in a high-throughput screening system. ***Microbial Cell Factories*****,** 15 (1).

Wang, J., Wen, B., Xu, Q., et al. (2015) Optimization of carbon source and glucose feeding strategy for improvement of L-isoleucine production by *Escherichia coli*. ***Biotechnology and Biotechnological Equipment*****,** 29 (2): 374–380.

Want, A., Thomas, O.R.T., Kara, B., et al. (2009) Studies related to antibody fragment (Fab) production in *Escherichia coli* W3110 fed-batch fermentation processes using multiparameter flow cytometry. ***Cytometry Part A***, 75 (2): 148–154.

Wilms, B., Hauck, A., Reuss, M., et al. (2001) High-cell-density fermentation for production of L-N-carbamoylase using an expression system based on the *Escherichia coli* rhaBAD promoter. ***Biotechnology and Bioengineering,*** 73 (2): 95–103.

Wurm, D.J., Quehenberger, J., Mildner, J., et al. (2018) Teaching an old pET new tricks: tuning of inclusion body formation and properties by a mixed feed system in *E. coli.* ***Applied Microbiology and Biotechnology*****,** 102 (2): 667–676.

Wurm, D.J., Veiter, L., Ulonska, S., et al. (2016) The *E. coli* pET expression system

revisited - mechanistic correlation between glucose and lactose uptake. ***Applied Microbiology and Biotechnology*****,** 100 (20): 8721–8729.

Wycuff, D.R. and Matthews, K.S. (2000) Generation of an AraC-*araBAD* promoter-regulated T7 expression system. ***Analytical Biochemistry*****,** 277 (1): 67–73.

Wyre, C. (2014) Recombinant Protein Production in *Escherichia coli:* Optimisation of Improved Protocols. PhD Thesis. University of Birmingham.

Wyre, C. and Overton, T.W. (2014a) Flow cytometric analysis of *E. coli* on agar plates: implications for recombinant protein production. ***Biotechnology letters*****,** 36 (7): 1485–94.

Wyre, C. and Overton, T.W. (2014b) Use of a stress-minimisation paradigm in high cell density fed-batch *Escherichia coli* fermentations to optimise recombinant protein production. ***Journal of Industrial Microbiol & Biotechnol.***

Yamaguchi, H. and Miyazaki, M. (2014) Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. ***Biomolecules***, 4 (1): 235–251.

Ytterberg, A.J., Zubarev, R.A. and Baumgarten, T. (2019) Posttranslational targeting of a recombinant protein promotes its efficient secretion into the *Escherichia coli* periplasm. ***Applied and Environmental Microbiology*****,** 85 (13): 1–9.

Yu, C.H., Dang, Y., Zhou, Z., et al. (2015) Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. ***Molecular Cell*****,** 59 (5): 744–754.

Zhang, C., Xing, X.H. and Lou, K. (2005) Rapid detection of a gfp-marked *Enterobacter aerogenes* under anaerobic conditions by aerobic fluorescence recovery. ***FEMS Microbiology Letters*****,** 249 (2): 211–218.

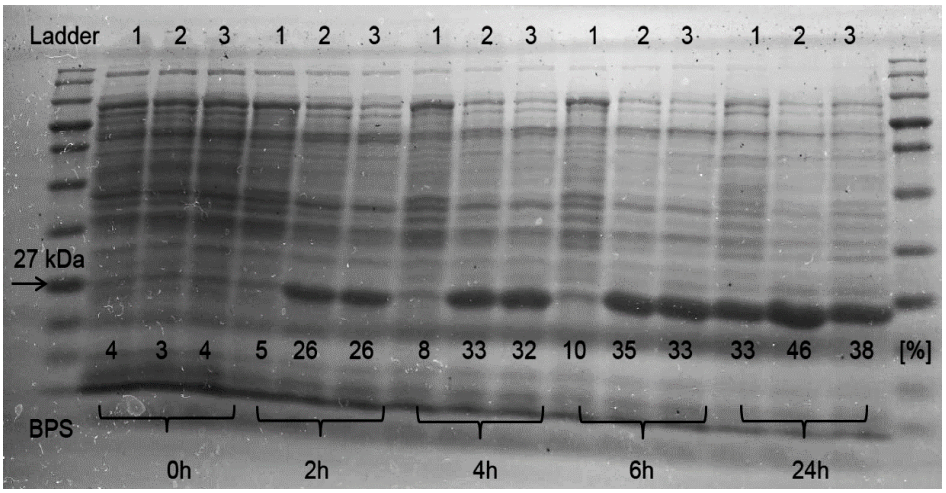
Zhang, J.D., Li, A.T. and Xu, J.H. (2010) Improved expression of recombinant cytochrome P450 monooxygenase in *Escherichia coli* for asymmetric oxidation of sulfides. ***Bioprocess and Biosystems Engineering*****,** 33 (9): 1043–1049.

Zhang, Z., Kuipers, G., Niemiec, Ł., et al. (2015) High-level production of membrane proteins in *E. coli* BL21(DE3) by omitting the inducer IPTG. ***Microbial Cell Factories*****,** 14 (1): 142.

Zhou, M.Y. and Gomez-Sanchez, C.E. (2000) Universal TA cloning. ***Current issues in molecular biology***. 2 (1) pp. 1–7.

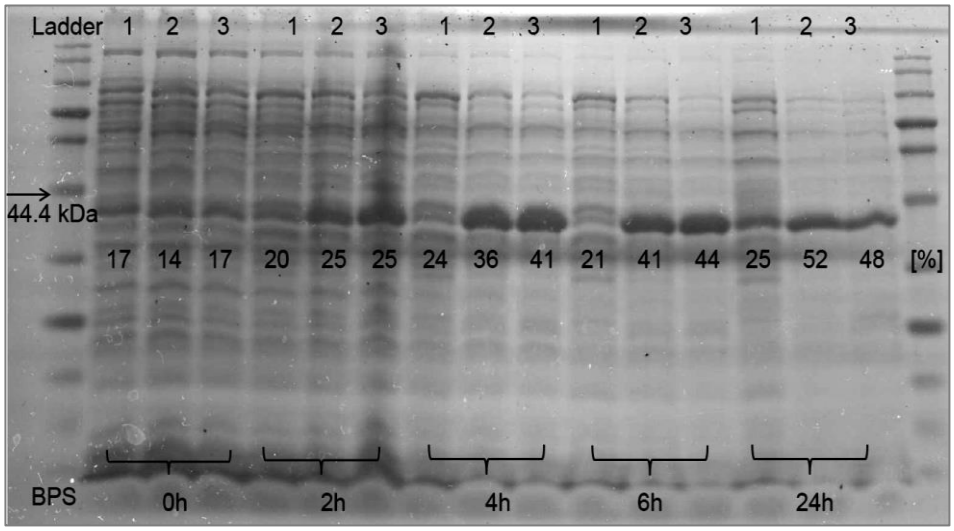
APPENDIX

(a)



1: GFP No Ara; **2**: GFP 0.015 % Ara; **3**: GFP 0.2 % Ara

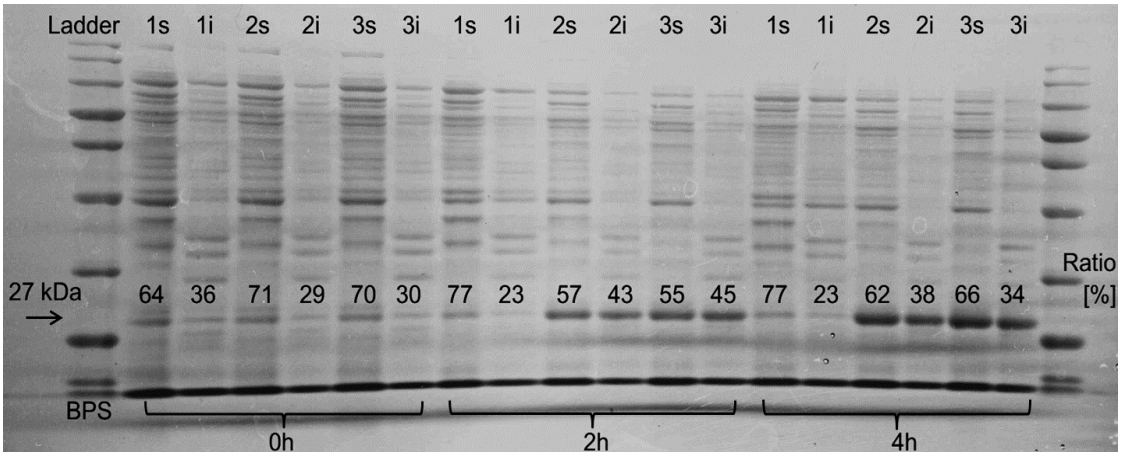
(b)

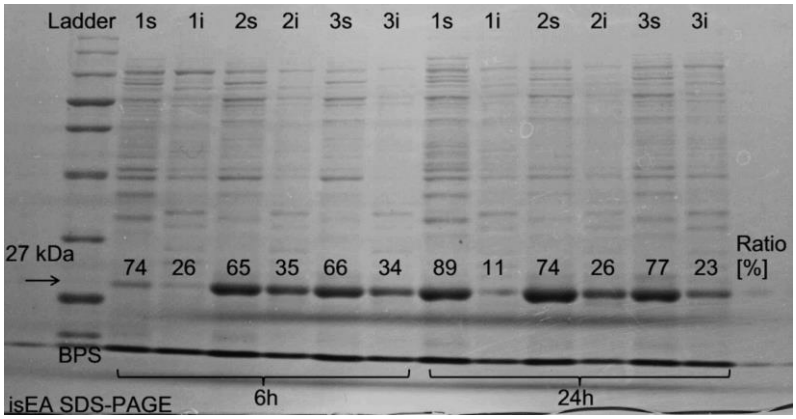


1: TNFα-GFP No Ara; **2**: TNFα-GFP 0.015 % Ara; **3**: TNFα-GFP 0.2 % Ara

Appendix 3.1: The accumulation of GFP and TNFα-GFP in cell extracts expressed by SDS-PAGE in batch culture with 0.8% glucose at 30°C

SDS-PAGE gel picture with arrow showing (a) GFP (27 kDa) and (b) TNFα-GFP (44.4 kDa) with the percentage of total GFP/TNFα-GFP proteins expressed inside the cell extracts. Sample bands at times post induction were compared to BPS ladder band.

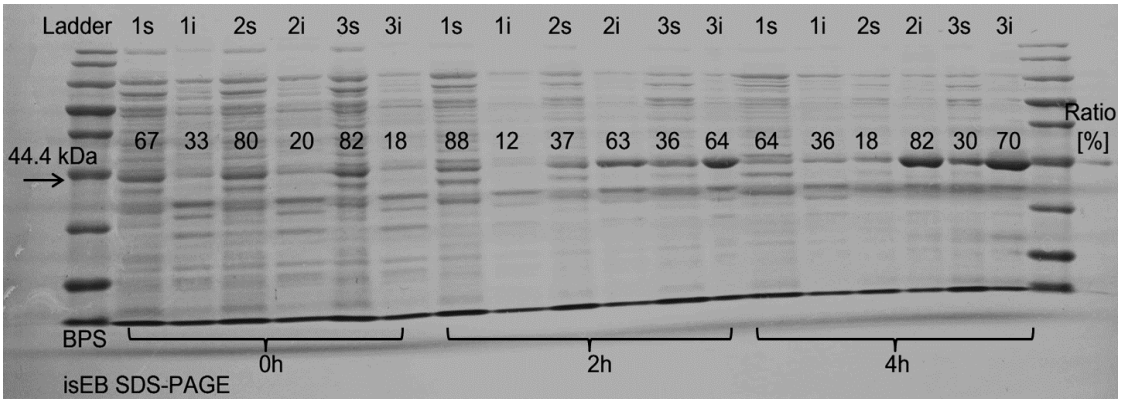


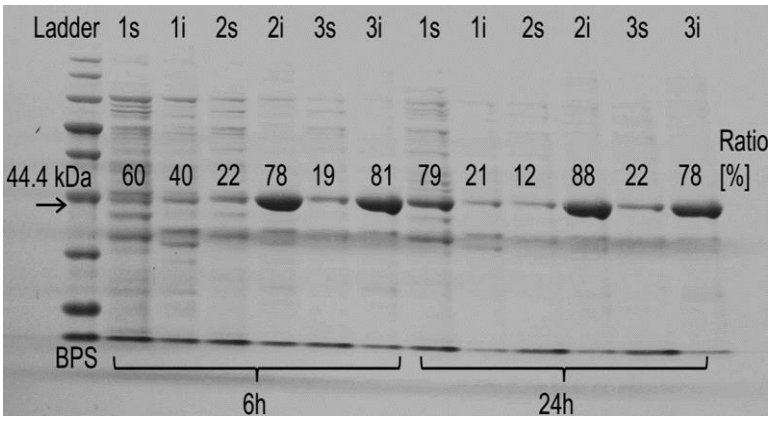


1s: GFP No Ara (Soluble); **1i**: GFP No Ara (Insoluble);
2s: GFP 0.015 % Ara (Soluble); **2i**: GFP 0.015 % Ara (Insoluble);
3s: GFP 0.2 % Ara (Soluble); **3i**: GFP 0.2 % Ara (Insoluble)

Appendix 3.2: The accumulation of GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in batch culture with 0.8% glucose at 30°C

SDS-PAGE gel picture with arrow showing GFP (27 kDa) and the ratio percentage of soluble versus insoluble GFP proteins accumulated inside the cell extracts of pORT(T7)-GFP, as shown above the band. All sample bands at times post induction were compared to BPS ladder band. [1s: Sol GFP No Ara; 1i: Insol GFP No Ara; 2s: Sol GFP 0.015% Ara; 2i: Insol GFP 0.015% Ara; 3s: Sol GFP 0.2 % Ara; 3i: Insol GFP 0.2 % Ara]



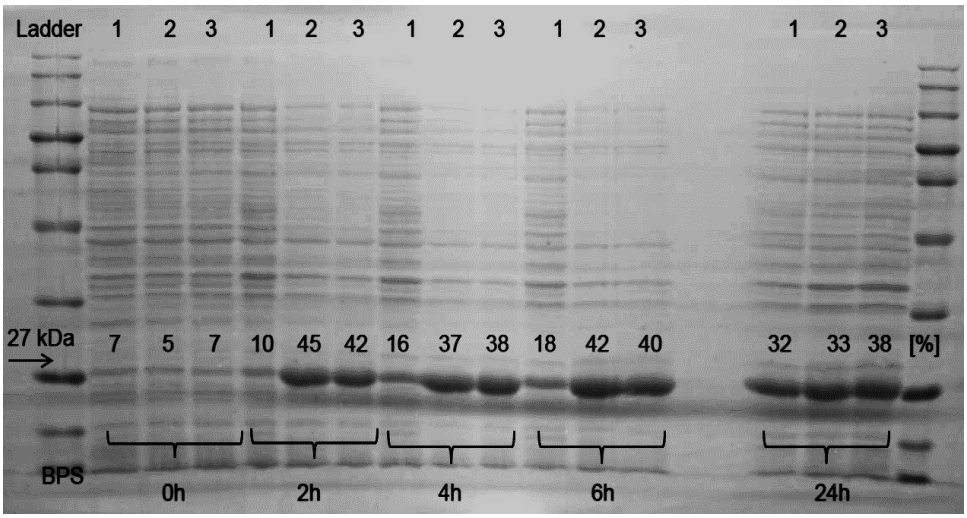


1s: TNFα-GFP No Ara (Soluble); **1i**: TNFα-GFP No Ara (Insoluble);
2s: TNFα-GFP 0.015 % Ara (Soluble); **2i**: TNFα-GFP 0.015 % Ara (Insoluble);
3s: TNFα-GFP 0.2 % Ara (Soluble); **3i**: TNFα-GFP 0.2 % Ara (Insoluble)

Appendix 3.3: The accumulation of TNFα-GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in batch culture with 0.8% glucose at 30°C

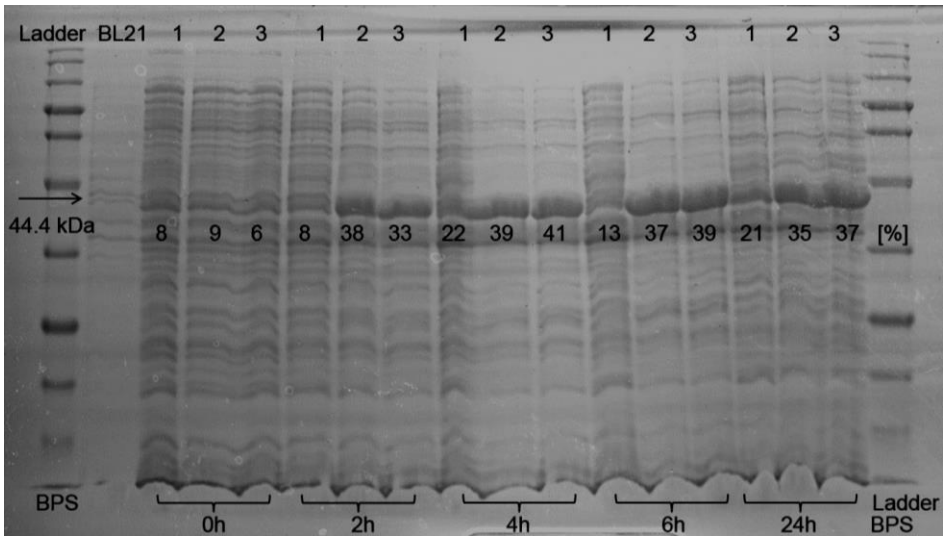
SDS-PAGE gel picture with arrow showing TNFα-GFP (44.4 kDa) and the ratio percentage of soluble versus insoluble TNFα-GFP proteins accumulated inside the cell extracts of pORT(T7)-TNFα-GFP, as shown above the band. All sample bands at times post induction were compared to BPS ladder band. [1s: Sol TNFα-GFP No Ara; 1i: Insol TNFα-GFP No Ara; 2s: Sol TNFα-GFP 0.015% Ara; 2i: Insol TNFα-GFP 0.015% Ara; 3s: Sol TNFα-GFP 0.2 % Ara; 3i: Insol TNFα-GFP 0.2 % Ara]

(a)



1: GFP No Ara; **2**: GFP 0.015 % Ara; **3**: GFP 0.2 % Ara

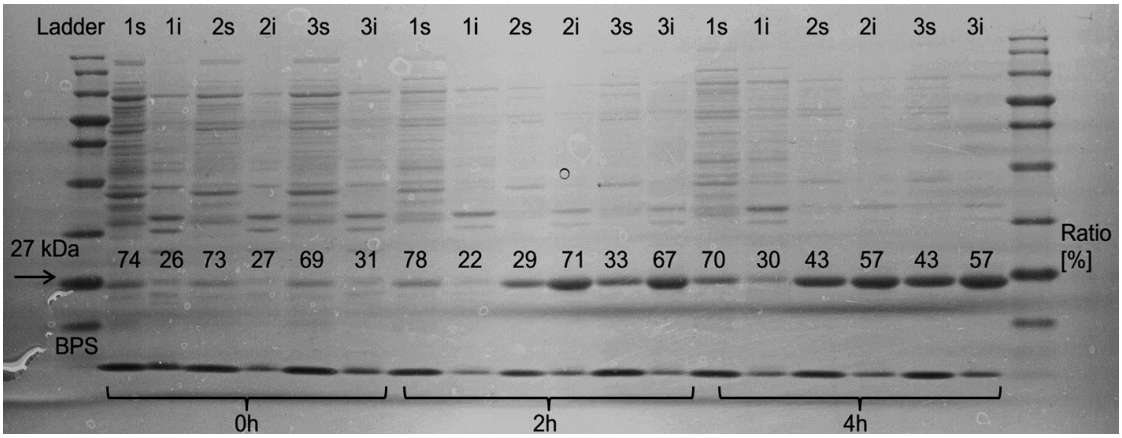
(b)

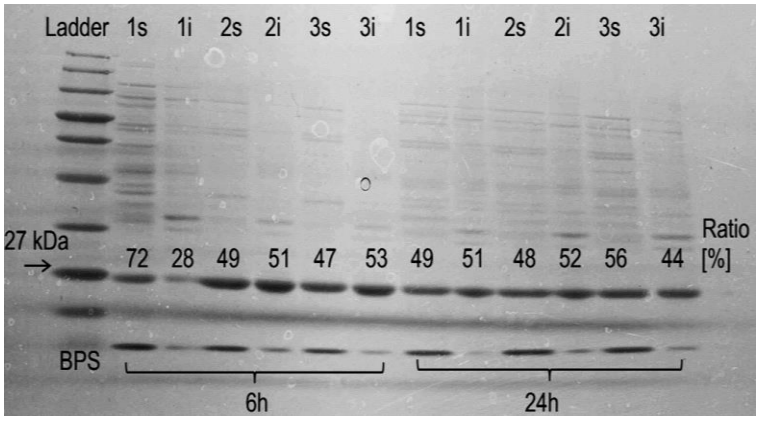


1: TNFα-GFP No Ara; **2**: TNFα-GFP 0.015 % Ara; **3**: TNFα-GFP 0.2 % Ara

Appendix 3.4: The accumulation of GFP and TNFα-GFP in cell extracts expressed by SDS-PAGE in batch culture with 0.4% glycerol at 37°C

SDS-PAGE gel picture with arrow showing (a) GFP (27 kDa) and (b) TNFα-GFP (44.4 kDa) with the percentage of total GFP/TNFα-GFP proteins expressed inside the cell extracts. Sample bands at times post induction were compared to BPS ladder band.

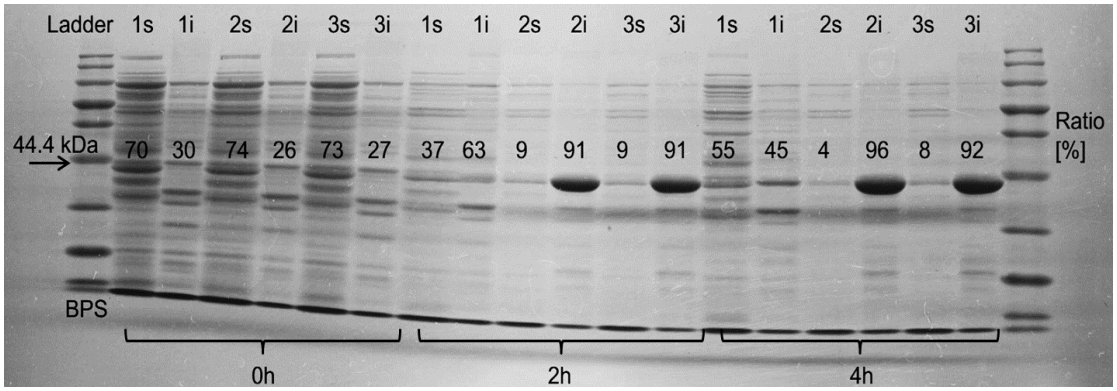


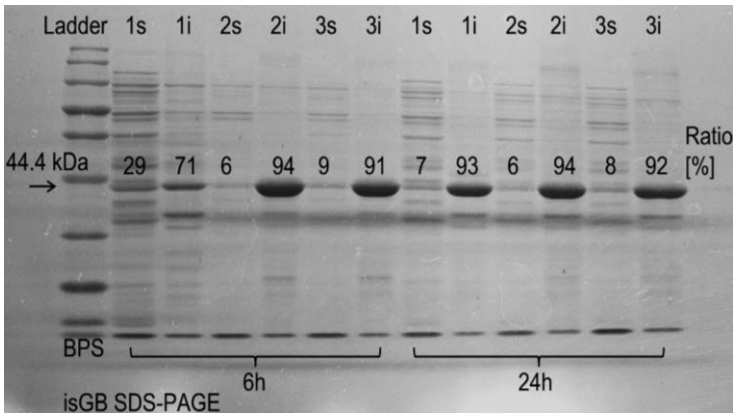


1s: GFP No Ara (Soluble); **1i**: GFP No Ara (Insoluble);
2s: GFP 0.015 % Ara (Soluble); **2i**: GFP 0.015 % Ara (Insoluble);
3s: GFP 0.2 % Ara (Soluble); **3i**: GFP 0.2 % Ara (Insoluble)

Appendix 3.5: The accumulation of GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in batch culture with 0.4% glycerol at 37°C

SDS-PAGE gel picture with arrow showing GFP (27 kDa) and the ratio percentage of soluble versus insoluble GFP proteins accumulated inside the cell extracts of pORT(T7)-GFP, as shown above the band. All sample bands at times post induction were compared to BPS ladder band. [1s: Sol GFP No Ara; 1i: Insol GFP No Ara; 2s: Sol GFP 0.015% Ara; 2i: Insol GFP 0.015% Ara; 3s: Sol GFP 0.2 % Ara; 3i: Insol GFP 0.2 % Ara]

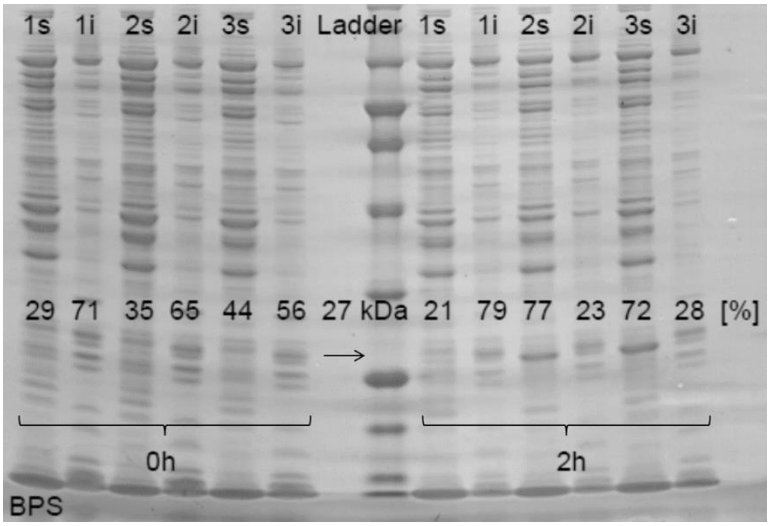


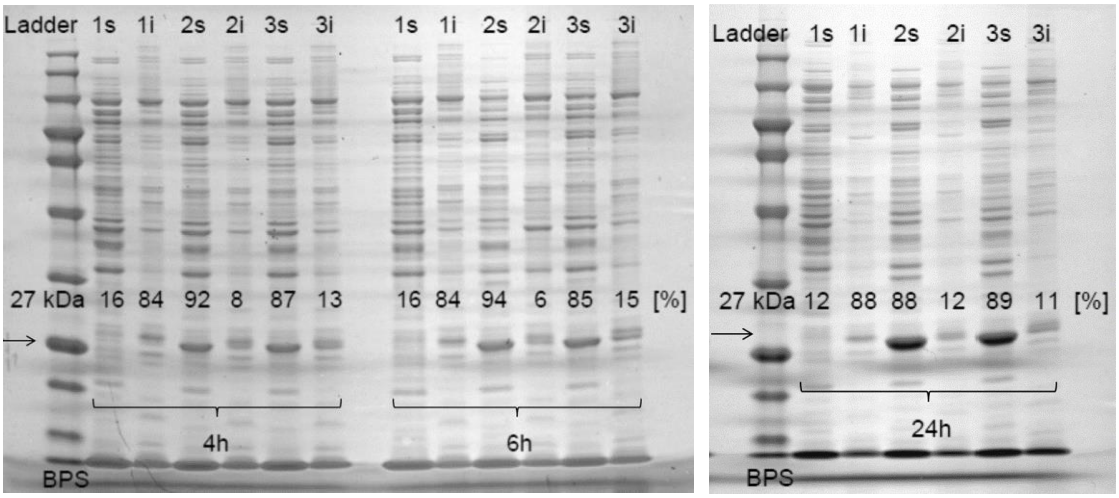


1s: TNFα-GFP No Ara (Soluble); **1i**: TNFα-GFP No Ara (Insoluble);
2s: TNFα-GFP 0.015 % Ara (Soluble); **2i**: TNFα-GFP 0.015 % Ara (Insoluble);
3s: TNFα-GFP 0.2 % Ara (Soluble); **3i**: TNFα-GFP 0.2 % Ara (Insoluble)

Appendix 3.6: The accumulation of TNFα-GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in batch culture with 0.4% glycerol at 37°C

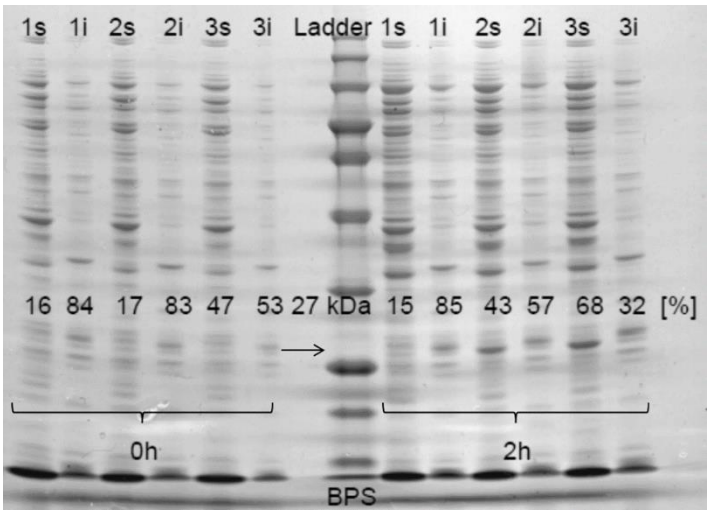
SDS-PAGE gel picture with arrow showing TNFα-GFP (44.4 kDa) and the ratio percentage of soluble versus insoluble TNFα-GFP proteins accumulated inside the cell extracts of pORT(T7)-TNFα-GFP, as shown above the band. All sample bands at times post induction were compared to BPS ladder band. [1s: Sol TNFα-GFP No Ara; 1i: Insol TNFα-GFP No Ara; 2s: Sol TNFα-GFP 0.015% Ara; 2i: Insol TNFα-GFP 0.015% Ara; 3s: Sol TNFα-GFP 0.2 % Ara; 3i: Insol TNFα-GFP 0.2 % Ara]

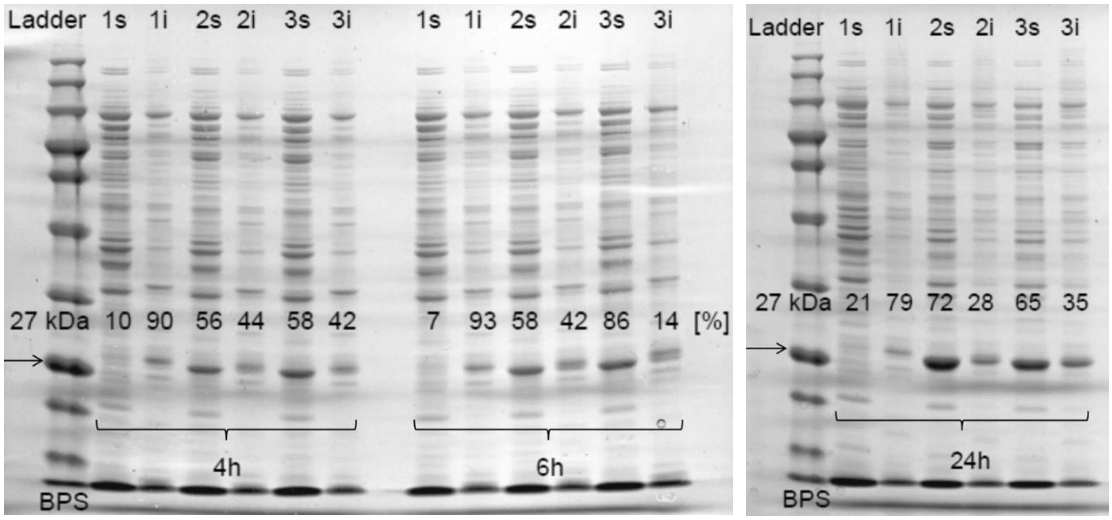




Appendix 4.1: The accumulation of GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in p*araBAD*-GFP cultures with 0.4% glucose at 30°C

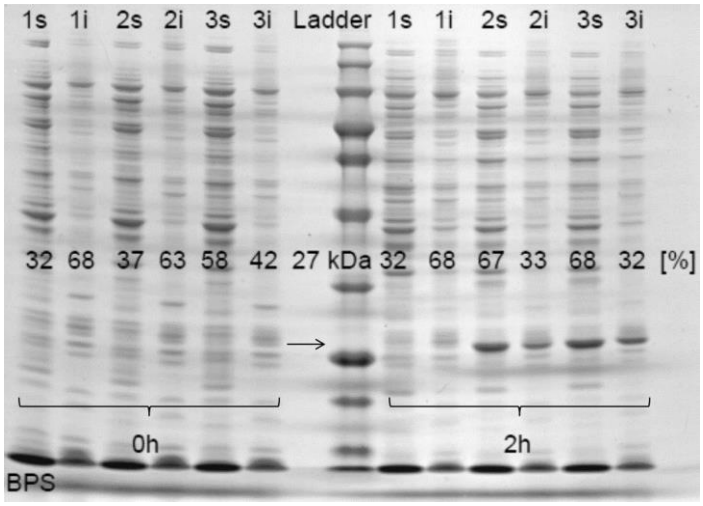
SDS-PAGE gel picture with arrow showing GFP (27 kDa) and the ratio percentage of soluble versus insoluble GFP proteins accumulated inside the cell extracts of p*araBAD*-GFP, as shown above the band. All sample bands at times post-induction were compared to BPS ladder band. [**1s**: GFP No Ara (Soluble); **1i**: GFP No Ara (Insoluble); **2s**: GFP 0.015% Ara (Soluble); **2i**: GFP 0.015% Ara (Insoluble); **3s**: GFP 0.2% Ara (Soluble); **3i**: GFP 0.2% Ara (Insoluble)]

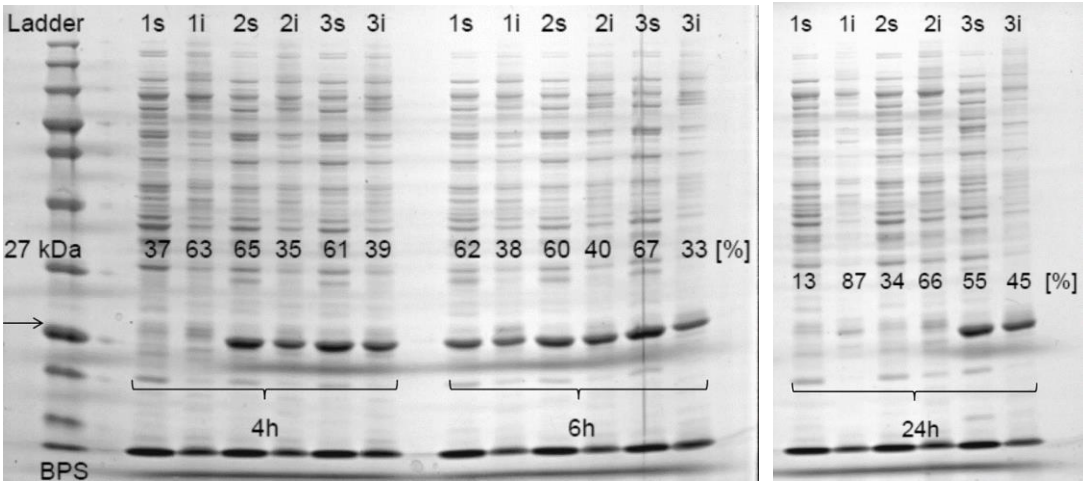




Appendix 4.2: The accumulation of GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in p*araBAD*-GFP culture with 0.8% glucose at 30°C

SDS-PAGE gel picture with arrow showing GFP (27 kDa) and the ratio percentage of soluble versus insoluble GFP proteins accumulated inside the cell extracts of p*araBAD*-GFP, as shown above the band. All sample bands at times post-induction were compared to BPS ladder band. [**1s**: GFP No Ara (Soluble); **1i**: GFP No Ara (Insoluble); **2s**: GFP 0.015% Ara (Soluble); **2i**: GFP 0.015% Ara (Insoluble); **3s**: GFP 0.2% Ara (Soluble); **3i**: GFP 0.2% Ara (Insoluble)]





Appendix 4.3: The accumulation of GFP subcellular fractionation in cell extracts expressed by SDS-PAGE in p*araBAD*-GFP culture with 0.4% glycerol at 37°C

SDS-PAGE gel picture with arrow showing GFP (27 kDa) and the ratio percentage of soluble versus insoluble GFP proteins accumulated inside the cell extracts of p*araBAD*-GFP, as shown above the band. All sample bands at times post-induction were compared to BPS ladder band. [**1s**: GFP No Ara (Soluble); **1i**: GFP No Ara (Insoluble); **2s**: GFP 0.015% Ara (Soluble); **2i**: GFP 0.015% Ara (Insoluble); **3s**: GFP 0.2% Ara (Soluble); **3i**: GFP 0.2% Ara (Insoluble)]

Appendix 5

Publication derived from this work

Zulkifly NAH, Selas Castiñeiras T and Overton TW (2023) Optimisation of recombinant TNFα production in *Escherichia coli* using GFP fusions and flow cytometry. ***Front. Bioeng. Biotechnol***, 11:1171823. doi:10.3389/fbioe.2023.1171823