

MONITORING BACTERIAL PHYSIOLOGY
DURING RECOMBINANT PROTEIN PRODUCTION
USING REPORTER GENE TECHNOLOGY

by

Isaac Vizcaino Caston

A thesis submitted to the
University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

This work presents an evaluation of the applicability of gene reporter technology to monitor *Escherichia coli* stress in industrial conditions with special interest in recombinant protein production. Different reporter plasmids containing promoter sequences of genes of the heat-shock response were utilized to monitor chaperone expression upon different sources of stress such as exposure to chemicals, temperature and anaerobic growth. Activation of the heat shock response was monitored by β -galactosidase activity from the reporter plasmid pQF50groE. Cultures responded to heat-shock, anaerobiosis and β -mercaptoethanol by increasing the expression of σ^{32} -related genes. The performance of fluorescence reporters containing varieties of GFP was measured by fluorimetry and flow cytometry. Low copy number plasmids were demonstrated to be better suited than medium-high copy plasmids to report stress in industrial conditions. Reporter plasmids containing the promoters of the chaperones DnaK and GroES were utilized to measure *E. coli* stress in reducing environments and during recombinant protein production. It was demonstrated that the production strategy caused an impact in the host physiology which determined the outcome of the process. Flow cytometry showed excellent potential to obtain reliable measurements providing data of reporter activity cell death and cell morphology.

Dedicated to the memory of my father
José Vizcaíno López (1953-2001)

Acknowledgements

First of all I would like to thank my supervisor Dr Tim W. Overton for great guidance, the dedication and the support he has given me during this project. Thanks also go to Professor Colin Thomas for his attention. I would also like to thank my group colleagues Amir Anvarian and Christopher Wyre; they have always been supportive and a source of interesting discussions, this gratitude extends to all the people I worked with in Biochemical Engineering. Special thanks to the technicians Dave, Elaine and Hazel for their help during this project. I would like to further acknowledge the help provided by the staff in the School on Biosciences, for their kind contribution with biological materials, specially Yanina Sevastsyanovich and her colleagues.

Many thanks to my friends Angel, Kris, Miao, Ourania, Rafael, Ruben and Yao for their support through these years, and in particular to Da for have taken good care of me “all a time”. I would like to thank my neighbours [REDACTED], for their support, especially during the writing up period. Last but not least I want to thank my mother and my brother or all their support and encouragement.

Contents

Chapter One: Introduction	1
1. Preface and project aims	2
1.1 Industrial recombinant protein products: A biopharmaceutical focus	4
1.1.1 The rise of the biogenics	4
1.1.2 Expressions systems	7
1.1.3 <i>Escherichia coli</i> as host	7
1.1.4 <i>E. coli</i> cells as means of production	15
1.2 Microbial physiology under industrial conditions	18
1.2.1 Response to stress: Gene expression regulation occurs at transcription initiation	19
1.2.2 Regulation at transcription initiation by σ factors	22
1.2.2.1 The master stress response: σ^S	26
1.2.2.2 The heat shock response: σ^{32}	28
1.2.2.3 The envelope stress response: σ^E	31
1.2.3 Regulation of gene expression at gene transcription initiation by transcription factors	34
1.2.3.1 Sensing the environment: Two component signal transduction	35
1.2.4 Protein misfolding	41
1.2.4.1 Fundamentals of protein folding <i>in vitro</i>	41
1.2.4.2 Protein folding and misfolding in <i>E. coli</i>	44
1.2.4.2.1 Proteins folding proteins: the chaperones of the heat shock response	46
1.3 Gene Reporter Technology	55
1.3.1 Fundamentals	55
1.3.2 Green Fluorescent Protein (GFP) as reporter protein	67
1.3.2.1 GFP folding and mutant variants	70
1.3.2.2 Advantages and limitations of using GFP as a reporter	75
1.3.3 GFP in gene reporter technology. Applications in microbiology	77
1.4 Microbial Flow Cytometry	82
1.4.1 The flow cytometry principle	82
1.4.2 Parameters in flow cytometry: Light scatter and fluorescence	85
1.4.3 From the photon to the file	88
1.4.4 Cell physiology is not homogeneous throughout the sample	89
1.4.5 Strategies for observation of physiological conditions	90

Chapter Two: Materials and Methods	92
2.1 Materials	93
2.2 Buffers and Solutions	94
2.3 Bacterial methods	96
2.4 Recombinant DNA techniques	107
2.4 Protocols of transformation	113
2.5 Analysis of proteins	116
2.6 Computational methods and software	118
Chapter Three: Measuring <i>groE::lacZ</i> activity as means of monitoring σ^{32} activity	120
3.1 Introduction	121
3.2 Results	122
3.2.1 Expression of <i>groE::lacZ</i> in relation to temperature	122
3.2.1.1 Responsiveness of <i>groE::lacZ</i> to consecutive temperature variations during growth	125
3.2.1.2 Reduction of activity in the absence of temperature changes	131
3.2.2 Response of <i>groE::lacZ</i> reporter to β -mercaptoethanol addition during growth	133
3.2.3 Involvement of σ^E on <i>groE::lacZ</i> expression upon β -mercaptoethanol addition	143
3.3 Discussion	151
3.3.1 Reservoirs of σ^{32} : the effect of temperature	151
3.3.2 β -mercaptoethanol effect triggered a heat-shock like response	153
3.3.3 The σ^{32} response triggered by β -mercaptoethanol has been shown to be independent of σ^E	156
3.3.4 Cell conditioning	157
3.5 Conclusions and future work	158
Chapter Four: Comparison of bulk measurements from GFP reporter systems with single cell analysis data for monitoring host response to reducing environments	161
4.1 Introduction	162
4.2 Results	163
4.2.1 Investigation of the heat-shock effect on pDNAK and pGROE gene expression	163
4.2.2 Investigation of DTT effect on the <i>rpoH</i> gene expression	166
4.2.3 Comparison of bulk measurements with single cell analysis of <i>rpoH::gfpmut3</i> activation in pA213 in <i>E. coli</i> BL21*	171
4.2.4 Comparative fluorescence study of GFP reporter based on plasmid copy number	182
4.2.5 Comparison of bulk measurements with single cell analysis of <i>groE::gfpmut2</i> activation in pGROE	187
4.2.6 Comparison of bulk measurements with single cell analysis of <i>dnaK::gfpmut2</i> activation in pDNAK reporter	191
4.2.7 Partial characterization of <i>yidQ</i> regulation using pYIDQ	200
4.4 Discussion	206
4.5 Conclusions and future work	215
	216

Chapter Five: Monitoring <i>groE</i> and <i>dnaK</i> expression in <i>E. coli</i> BL21* during recombinant protein production with gene reporter technology	
5.1 Introduction	218
5.2 Results	220
5.2.1 Synthesis of GFP by pGROE as stress reporter upon pET-CCP activation	220
5.2.2 Synthesis of GFP by pDNAK as stress reporter upon pET-CCP activation	230
5.2.3 Controls and further growth conditions.	241
5.2.4 Batch fermentation	253
5.3 Discussion	266
5.3.1 The chaperone expression was greater under standard than under improved protocol	268
5.3.2 <i>groE</i> and <i>dnaK</i> promoters reported stress differently	267
5.3.2 The reporter response can be masked	270
5.3.3 Fluorimetry versus flow cytometry for monitoring of reporter activity during bioprocess	271
5.4 Conclusions and future work	275
Chapter Six: General conclusions	280
References	281

List of figures

Chapter One: Introduction

Fig. 1.1 Interaction of RNAP holoenzyme with DNA	21
Fig. 1.2 Structural groups and phylogenetic characteristics of σ^{70} family	23
Fig. 1.3 Probable secondary structure of σ^{32} mRNA	30
Fig. 1.4 Scheme of σ^E regulation by proteolysis	32
Fig. 1.5 Paradigmatic structure of two-component system	36
Fig. 1.6 CpxP regulation of CpxA/R system	39
Fig. 1.7 Energetic landscape of spontaneous protein folding	43
Fig. 1.8 Structure and cycle of DnaK	49
Fig. 1.9 Molecular structures of GroEL and GroES	51
Fig. 1.10 Integration of the heat shock proteins on a chaperone network	53
Fig. 1.11 Basic diagram of a reporter plasmid	58
Fig. 1.12 Cycle of the aequorin reaction	67
Fig. 1.13 Crystal structure of GFP and chromophore maturation	69
Fig. 1.14 Diagram of the basic elements of a flow cytometer	84

Chapter Two: Materials and Methods

Fig 2.1 pQF50KgroE promoter and restriction map of pUA66 promoterless vector	100
Fig. 2.2 Nucleotide sequence of the promoter of pGROE	101
Fig. 2.3 Nucleotide sequence of the promoter of pDNAK	102
Fig. 2.4 Nucleotide sequence of the promoter of pYIDQ	103

Chapter Three: Measuring *groE::lacZ* activity as means of monitoring σ^{32} activity

Fig. 3.1 Heat-shock during aerobic growth	124
Fig. 3.2 Heat-shock during anaerobic growth	126
Fig. 3.3 β -galactosidase activity during aerobic growth at variable temperature	129
Fig. 3.4 β -galactosidase activity during anaerobic growth at variable temperature	130
Fig. 3.5 β -galactosidase activity at constant temperature	132
Fig. 3.6 <i>E. coli</i> JM109 pQF50KgroE response to β -mercaptoethanol as stress agent	135
Fig. 3.7 Combination of β -mercaptoethanol with heat-shock stress in aerobic cultures	137
Fig. 3.8 Combination of β -mercaptoethanol with heat-shock stress in anaerobic cultures	140
Fig. 3.9 Reporter activity in presence of β -mercaptoethanol in late stationary phase	142
Fig 3.10 Influence of σ^E on β -galactosidase activity	145
Fig. 3.11 Reporter expression at 30°C and at 37°C of <i>E.coli</i> ECA101 ($\Delta\sigma^E$) strain and BW25113 (WT) harbouring pQF50KgroE upon β -mercaptoethanol addition (50mM)	147
Fig. 3.12 Reporter expression at 42°C of <i>E.coli</i> ECA101 ($\Delta\sigma^E$) strain and BW25113 (WT) harbouring pQF50KgroE upon β -mercaptoethanol addition (50mM)	150

Chapter Four: Comparison of bulk measurements from GFP reporter systems with single cell analysis data for monitoring host response to reducing environments

Fig. 4.1 pGROE activity upon heat shock	165
Fig. 4.2 pDNAK activity upon heat shock	167
Fig. 4.3 The effect of DTT on pAA213 (<i>rpoH::gfpmut3</i>) during aerobic growth	169
Fig. 4.4 pAA213 (<i>rpoH::gfpmut3</i>) expression in <i>E. coli</i> BW25113 during anaerobic growth	179
Fig. 4.5 Bulk measurements of expression in <i>E. coli</i> BL21* transformed (<i>rpoH::gfpmut3</i>) cultures in the presence of reductants.	173
Fig. 4.6 Control experiment to calibrate PI staining of live and dead cells	175
Fig. 4.7 Single cell analysis of <i>E. coli</i> BL21* harbouring pAA213 pre-addition of thiols during aerobic growth	177
Fig. 4.8 Single cell analysis of <i>E. coli</i> BL21* harbouring pAA213 two hours after thiol addition during aerobic growth	178
Fig. 4.9 Single cell analysis of <i>E. coli</i> BL21* harbouring pAA213 four hours after thiol addition during aerobic growth	180
Fig. 4.10 Single cell analysis of <i>E. coli</i> BL21* harbouring pAA213 two hours after thiol addition in anaerobiosis	181
Fig. 4.11 Single cell analysis of <i>E. coli</i> BL21* harbouring pAA213 four hours after thiol addition in anaerobiosis	183
Fig. 4.12 RF and growth of <i>E. coli</i> BW25113 transformed with pFVP25 based reporters	185
Fig. 4.13 RF and growth of <i>E. coli</i> BW25113 transformed with pUA66 based reporters	186
Fig. 4.14 pGROE expression in <i>E. coli</i> BL21* in the presence and absence of DTT, β -mercaptoethanol and oxygen	189
Fig 4.15 Single cell analysis of pGROE (<i>groE::gfpmut2</i>) expression two hour after addition	190
Fig 4.16 Single cell analysis of pGROE (<i>groE::gfpmut2</i>) expression four hours after addition	193
Fig. 4.17 pDNAK expression in <i>E. coli</i> BL21* in the presence and absence of DTT, β -mercaptoethanol and oxygen	195
Fig. 4.18 Single cell analysis of <i>E. coli</i> BL21*harbouring pDNAK two hours after addition	196
Fig 4.19 Single cell analysis of <i>E. coli</i> BL21*harbouring pDNAK four hours after addition	199
Fig. 4.20 Catabolite repression of pYIDQ	203
Fig. 4.21 CpxR and CpxA regulation of pYIDQ	205
Fig. 4.22 Proposed diagram for pYIDQ regulation by the CpxAR system.	213

Chapter Five: Monitoring *groE* and *dnaK* expression in *E. coli* BL21* during recombinant protein production with gene reporter technology

Fig. 5.1 RF, growth curves and SDS-PAGE obtained using pGROE obtained during growth of <i>E. coli</i> BL21* pGROE/pET-CCP in standard and improved protocols.	221
Fig 5.2 Single cell analysis of samples from standard and improved protocols of <i>E. coli</i> BL21* pET-CCP/pGROE	223
Fig 5.3 Single cell analysis of samples from standard and improved protocols of <i>E. coli</i> BL21* pET-CCP/pGROE (continuation)	228
Fig. 5.4 RF, growth curves and SDS-PAGE obtained using pDNAK obtained during growth of <i>E. coli</i> BL21* pDNAK/pET-CCP in standard and improved protocols.	231
Fig 5.5 Single cell analysis of samples from standard and improved protocols of <i>E. coli</i> BL21* pET-CCP/pDNAK	233
Fig 5.6 Single cell analysis of samples from standard and improved protocols of <i>E. coli</i> BL21* pET-CCP/pDNAK (continuation)	236
Fig. 5.7 RF and growth curves obtained using the reporter pUA66 in standard and improved protocols	240
Fig 5.8 Results from pET11c vector co-transformed with the reporter plasmids	243
Fig. 5.9 Reporting of stress during recombinant protein production in anaerobic incubation	246
Fig 5.10 <i>E. coli</i> BL21* pDNAK/pET-CCP double transformants at different concentrations of IPTG	248
Fig. 5.11 Single cell analysis of <i>E. coli</i> BL21* pDNAK/pET-CCP two hours post-induction with 0, 0.05, 0.5 and 5mM IPTG	250
Fig 5.12 Single cell analysis of <i>E. coli</i> BL21* pDNAK/pET-CCP two hours post-induction with 0, 0.05, 0.5 and 5mM IPTG	251
Fig 5.13 Single cell analysis of <i>E. coli</i> BL21* pDNAK/pET-CCP two hours post-induction with 0, 0.05, 0.5 and 5mM IPTG	252
Fig 5.14 Off-line sample analysis during batch fermentation	255
Fig. 5.15 Single cell analysis during batch fermentation	258
Fig. 5.16 Single cell analysis during batch fermentation (continuation)	259
Fig. 5.17 Single cell analysis during batch fermentation, last 12 hours, showed a decrease in filamentation	260
Fig. 5.18 Filamentous morphology of <i>E. coli</i> BL21* pDNAK/CCP and RF-Flow cytometry comparison	262
Fig 5.19 On-line monitoring during batch fermentation	265

List of tables

Table 1.1 Biopharmaceutical products approved in the last three decades	5
Table 1.2 Examples of glycosylated recombinant protein products	12
Table 1.3 Summary of main biological considerations when utilizing <i>E. coli</i> as host for recombinant protein production	17
Table 1.4 Consensus sequences recognized for some sigma factors	25
Table 1.5 Gene expression σ factors related to common stress responses	27
Table 1.6 Heat shock proteins families conserved across kingdoms	48
Table 1.7 Origin of replication and copy number of various plasmids	59
Table 1.8 Summary of some of most common reporter proteins	61
Table 1.9 Bioluminescence systems discussed based on luciferin/luciferase reactions	64
Table 1.10 Classification of GFP derivate mutants	73
Table 1.11 Variety of dyes utilized in flow cytometry	87
Table 2.1 Strains used in this study	97
Table 2.2 List of plasmids utilized in this study	99
Table 2.3 Fluorescence settings for luminescence spectrophotometer	106
Table 2.4 Flow cytometer settings	106
Table 2.5. Oligonucleotide primers used in this study	109
Table 2.6 Thermocycler configuration for PCR	111

Abbreviations

bp	Base pair
CCP	Cytochrome c peroxidase
CDC	Carbon dioxide consumed
CV	Coefficient of variation
DO	Dissolved oxygen
DTT	Dithiothreitol
FR	Relative fluorescence
FS	Forward scatter
Gm	Geometric mean
IAA	Indole-acrylic acid
IPTG	Isopropyl β -D-1-thiogalactopyranoside
LB	Lennox broth
λ_{ex}	Excitation wavelength
λ_{em}	Emission wavelength
Md	Median
OD _x	Optical density at x nm
ONPG	Orto-Nitrophenyl- β -galactoside
OXC	Oxygen consumed
PI	Propidium iodide
RQ	Respiratory quotient
SS	Side scatter
TBE	Tris-borate-EDTA buffer
WT	Wild type

Chapter One

Introduction

1. Preface and project aims

Preface

The introduction comprehends a literature review aimed to provide detailed explanations of the different aspects of the science, which were brought together during this study. The first section consists of an overview of past, present and future trends of industrial recombinant protein production in bacteria highlighting the increased interest on bringing into account microbial physiology in process design leading to better product yields. The second section explores the relevant aspects of *E. coli* physiology in industrial environments focusing on gene regulation by sigma factors and the molecular mechanisms underlying the alteration of host homeostasis by protein misfolding.

The third section, gene reporter technology, is introduced as the main strategy to observe *E. coli* physiology *in vivo*. Green fluorescent protein use as reporter protein is reviewed highlighting the existence of mutant variants optimized for *E. coli* expression and folding. The past, present and future trends of this technology are reviewed in this section. The fourth section aims to explain why flow cytometry is an ideal tool to measure reporter gene expression on bacteria. Multiparameter flow cytometry principles are explained in this section including its utilization with chemical dyes.

In summary, the following literature review formulates the industrial issues and their underlying molecular causes in the host physiology. Gene reporter technology is put forward as a strategy to detect the activation of specific molecular mechanisms in the host. Finally, flow cytometry is presented as the preferred equipment to monitor such activation.

Aims

- To review current strengths and weaknesses of gene reporter technology focusing on industrial fermentation applications.
- To monitor the heat-shock response during *E. coli* growth as a consequence of stresses common in bioprocesses including temperature variations, chemical stressors, and recombinant protein production.
- To compare the activation of gene reporter systems based on their different promoter structures in order to assess their suitability for industrial uses.
- To assess and compare biochemical assays, relative fluorescence, and multi-parameter flow cytometry, as methods to measure reporter protein in industry.
- To utilize gene reporter data to demonstrate that misfolding protein stress could be a contributing cause of the differences in recombinant protein productivity previously observed between different industrial protocols of fermentation.

1.1 Industrial recombinant protein products: A biopharmaceutical focus

In 1978 the successful synthesis of recombinant human insulin in *Escherichia coli* was announced. It took a process of four years of scientific development, involving process scale up and clinical trials, within an uncharted regulatory framework. This first biopharmaceutical product was licensed by the FDA (Food and Drug Administration) as a commercial drug (Humulin) in 1982 (Johnson, 1983). Biopharmaceuticals are traditionally defined as biological pharmaceuticals (generally large molecules) produced by biotechnological methods. This broad definition is being adapted to new technologies and narrowed to products which are obtained from genetically engineered organisms; a modern definition excludes antibiotics and vaccines manufactured in unmodified organisms ('natural products'). The evolution of these definitions (biopharma and biotechnology) affects how such concepts are utilized by the legal instances which regulate the sector (Rader, 2008). Based on the modern definition, the major share of biopharmaceuticals is composed of recombinant proteins manufactured in genetically modified hosts. These products have higher molecular weights than chemical drugs and depend on their complex structure to perform their function. As a consequence of this complexity, they cannot be manufactured using the same techniques as chemical drugs (Crommelin et al., 2003). In the last 30 years, new companies and new signature biological products were introduced, such as [Table 1.1]: enzymes; hormones; cytokines; clotting factors; vaccines and monoclonal antibodies, generating sales worth \$54.5 billion in 2007 in total, with expectations of rising ~25% by 2012 (Kresse, 2009).

1.1.1 The rise of the biogenerics

The purification of recombinant human growth hormone (rhGH) from *E. coli* as an active

Recombinant protein product and description	Manufacturer	Indication/Use	Approval
Humulin. (Human insulin produced in <i>E. coli</i>)*	Eli Lilly	Diabetes mellitus	1982 (US)
Protropin. (Growth hormone (GH) differing from human GH only in containing an additional N-terminal methionine residue; produced in <i>E. coli</i>)*	Genentech	hGH deficiency in children	1985 (US)
Roferon A (human Interferon-alpha-2a produced in <i>E. coli</i>)#	Hoffmann–La Roche	Hairy-cell leukemia	1986 (US)
Humatrope (Human GH produced in <i>E. coli</i>)*	Eli Lilly	hGH deficiency in children	1987 (US)
Actimmune (Human Interferon-gamma-1b; produced in <i>E. coli</i>)#	Genentech	Chronic granulomatous disease	1990 (US)
Leukine (Granulocyte Macrophage-Colony Stimulating Factor, differing from the native human protein by one amino acid, Leu23; produced in <i>E. coli</i>)#	Immunex (now Amgen)	Autologous bone marrow transplantation	1991 (US)
Proleukin (Interleukin-2, differing from human molecule in absence of an N-terminal alanine and a C125S substitution produced in <i>E. coli</i>)#	Chiron	Renal-cell carcinoma	1992 (US)
Betaseron (Interferon-beta1b, differing from human protein by Cys17Ser substitution; produced in <i>E. coli</i>)#	Berlex Labs (Richmond, CA, USA)/Chiron (Emeryville, CA, USA)	Relapsing/remitting multiple sclerosis	1993 (US)
ReoPro (abciximab, Fab fragments derived from a chimeric monoclonal antibody (mAb),directed against the platelet surface receptor GPIIb/IIIa)†	Centocor	Prevention of blood clots	1994 (US)
Betaferon (Interferon-beta1b differing from human protein by C17S substitution; produced in <i>E. coli</i>)#	Schering AG	Multiple sclerosis	1995 (EU)
Ecokinase (reteplase, tissue plasminogen activator (tPA) produced in <i>E. coli</i> ; differs from human tPA in that three of its five domains have been deleted)‡	Galenus Mannheim (Germany)	Acute myocardial infarction	1996 (EU)
Infergen (Interferon-alpha, synthetic type I interferon produced in <i>E. coli</i>)#	Amgen (US), Yamanouchi Europe (Leiderdorp, The Netherlands) (EU)	Chronic hepatitis C	1997 (US), 1999 (EU)
Rituxan (rituximab chimeric mAb directed against CD20 antigen found on the surface of B lymphocytes) †	Genentech/Biogen/IDEC	Non-Hodgkin lymphoma	1997 (US)
Beromun (Tumour necrosis factor-alpha produced in <i>E. coli</i>) #	Boehringer Ingelheim	Adjunct to surgery for subsequent tumor removal, to prevent or delay amputation	1999 (EU)
Lymrix (OspA, a lipoprotein found on the surface of <i>Borrelia burgdorferi</i> ; produced in <i>E. coli</i>)•	GSK	Lyme disease vaccine	1998 (US)
Remicade (infliximab, chimeric mAb directed against TNF-alpha)†	Centocor (Horsham, PA, USA)	Treatment of Crohn disease	1998 (US), 1999 (EU)
PegIntron A (PEGylated Interferon-alpha-2b produced in <i>E. coli</i>)#	Schering-Plough (Kenilworth, NJ, USA)	Chronic hepatitis C	2000 (EU), 2001 (US)
Viraferon (human interferon-alpha-2b produced in <i>E. coli</i>)#	Schering-Plough	Hepatitis B, C	2000 (EU)
Kineret (anakinra; Interleukine-1 receptor antagonist produced in <i>E. coli</i>)#	Amgen	Rheumatoid arthritis	2001 (US)
Pegasys (pegylated interferon alpha-2a produced in <i>E. coli</i>)#	Hoffman–La Roche (Nutley, NJ)	Hepatitis C	2002 (EU and US)
Forteo teriparatide (shortened form of human parathyroid hormone produced in <i>E. coli</i> ; also sold as Forsteo in the EU)*	Eli Lilly	Established osteoporosis in postmenopausal women	2003 (EU)
Fortical (salmon calcitonin produced in <i>E. coli</i>)*	Upsher-Smith Laboratories (Minneapolis, MN, USA)/ Unigene (Fairfield, NJ, USA)	Postmenopausal osteoporosis	2005 (US), 2003 (EU)

Table 1.1 Biopharmaceutical products approved in the last three decades

The table shows some representative products approved either by the FDA (US) or the EMEA (EU) manufactured in *E. coli*. Adapted from Walsh (2006) .Legend: (*) Hormone; (#) Cytokine; (†) Antibody; or antibody fragment; (§) Clotting factor; (•) Vaccine

product was first reported in 1981 (Olson et al., 1981), leading to its approval as commercial drug in 1985 manufactured by Genetech under the name of Protropin®. It was given orphan drug status which granted privileges as an incentive for the development of treatments for rare diseases; in this case GH deficiency in children. However, other uses of the hormone resulted in revenue of ~\$47million one year after its launch. With that market perspective, in 1987 Eli Lilly & Co. obtained FDA approval for Humatrope® which triggered a legal battle with Genetech (Ezzell, 1987b). It consisted of an rhGH with 1 less amino acid in its structure, therefore was different. The Court ruled it a new drug (Ezzell, 1987a). This was an example of the absence of legal tools to define what the drug effectively was. Due to inherent complexity a biopharmaceutical may present small variations in its structure while retaining functionality which caused an increase of claims in the patents aimed to protect the biobdrug and its activity-blocking competitors. This brings us to the key issue of patent protection. Many of the biopharmaceuticals shown in [Table 1.1] have had their patent protection expire or is about to expire. In 2009, biopharmaceuticals with a market worth of approximately \$10bn lost patent protection (Walsh, 2006). EMEA (European Medicines Agency) arranged regulations on November 2010 for the licensing of rDNA (recombinant DNA) derived drugs for which the patent had expired. Evidence of similarity is to be determined by pharmacokinetics and pharmacodynamics studies, assessed by *in vitro* and non-clinical *in vivo* tests. These requirements were set lower than expected, opening the door to a \$36.4bn biogenerics market (Sinha, 2011). One year after the approval of the 'Patient Protection and Affordable Care Act' in the US, the FDA still did not provide a regulatory framework for biosimilars like its European counterpart. In 2015, with a market of \$6.6bn on its own, the monoclonal antibody (mAb) Rituxan® will lose patent privileges in US (Carey, 2011). Competition amongst producers may require reduced manufacturing costs where the choice of expression system has a major impact.

1.1.2 Expression systems

In 2009 there were 151 recombinant rDNA based drugs licensed by FDA and EMEA. The production system seemed to be divided in two major blocks: microorganisms (73); and mammalian cells (59). *Escherichia coli* has been the preferred microbe for protein production accumulating 28 products (Ferrer-Miralles et al., 2009). On the other hand, Chinese hamster ovary cells (CHO) are the most utilized mammalian host with three major advantages: they are easily transfectable; they grow in suspension to high density cultures; and they can growth in serum-free medium (Trill, Shatzman, and Ganguly, 1995). Still, the continuous development of industrial processes with CHO cells since 1995 did not affect a parallel development of *E. coli* processes. This was thought to occur due to the elevated cost of mammalian systems and the achievable product quality of microbial systems. In addition there was an increase in the number of drug products that do not require glycosylation or postranlsational modifications thus making bacterial synthesis possible. However, the current rDNA technologies seem have exhausted conventional manufacturing (Ferrer-Miralles et al., 2009). Understanding *E. coli* physiology may deliver tools for improving yields of high quality product making the difference during the present and future struggles for a share of the biopharmaceutical market.

1.1.3 *Escherichia coli* as host

From model organism to industrial standard

Escherichia coli is a rod-shaped Gram-negative bacterium, a facultative anaerobe of the family Enterobacteriaceae within the Gamma subdivision of Proteobacteria (Madigan, 2009).

E. coli was first isolated in 1885 and its genus was renamed from *Bacterium* to *Escherichia* in 1919. It is present in the colon of healthy individuals and it has been utilized in qualitative tests for drinking water to detect fecal contamination. The strain K-12 was first isolated in 1922 from human faeces of a diphtheria patient and, due to its continuous growth in laboratory conditions, has lost its ability to colonize the human gut. During the second half of the 20th century it has been utilized as model organism for diverse studies accumulating vast amounts of data (Lederberg, 2004). The 4,639,221 base pairs of the *E. coli* K-12 chromosome were sequenced (Blattner et al., 1997a) and made available through databases such as EcoGene (Rudd, 2000) and EcoCyc (Keseler et al., 2005).

Advantages of producing recombinant protein in E. coli

The wealth of biochemical knowledge on *E. coli* has been one of the key advantages of its utilization. Since then, many of the concerns regarding its safety have been cleared through the success of numerous products. *E. coli* K-12 strain was implemented in the first commercial process of rDNA products where comprehensive documentation regarding its safety was provided (Johnson, 1983). In addition, further strains and optimized expression vectors have been developed to cope with the industrial demands. *E. coli* shows fast growth on inexpensive media; reaches elevated cell density; and it is able to accumulate high protein density (Baneyx, 1999). While many of previous obstacles have been addressed some major considerations persist.

Considerations for producing recombinant protein in E. coli

Endotoxins are heat stable macromolecules of Gram-negative bacteria found in the outer membrane. These lipopolysacharides are released to the media during growth and cell death reaching high concentration in lysates. Endotoxins can cause acute immune response in the patient when the biological drug is administered intravenously (Heumann and Roger, 2002). In the instances where *E. coli* is utilized as a host, part of the downstream processing efforts is directed to validate a guaranteed amount of endotoxin removal (Petsch and Anspach, 2000).

Synthesis of recombinant protein of mammalian origin is likely to require codons which otherwise are rarely required in bacteria. These types of cognate aminoacyl-tRNAs are present in a very low concentration and a major issue is that lack of availability causes the translation process to stall (Kane, 1995). This problem has been addressed with site-directed mutagenesis of the recombinant gene (codon optimization) or by co-transforming with a plasmid able to express the tRNA to the levels required. Both solutions have been successful. In the latter case there are commercial plasmids available such as pR.A.R.E. from Novagen (Sorensen and Mortensen, 2005). After solving the problem of tRNA depletion, mRNA stability and unwanted secondary structures were observed to gain significance as major bottlenecks (Wu et al., 2004).

The mRNA of *E. coli* is stable for between 30 seconds to 20 minutes. During overexpression of recombinant protein mRNA stability is cause for concern (Baneyx, 1999). The main enzyme for RNA degradation is RNase E. In order to cleave RNA, RNase E recognizes two points of the mRNA: a monophosphate ribonucleoside in the 5' extreme; and a non-terminal nucleoside. Experiments showed reduced activity on 5' extremes with triphosphate ribonucleosides. It was speculated that a terminal triphosphate group acts limiting the

degradation (Mackie, 1998). In some cases mRNA folds on itself (hairpin structure) causing a reduction of the degradation rates in vitro. These structures, derived from the leader sequence of the *ompA* gene, were utilized to increase the stability of mRNAs. When utilized as UTRs (Untranslated regions) the hairpin structure disabled RNase attack on the preferred 5' terminal. Translation was achieved because of downstream Shine-Dalgarno sequences enabling ribosome binding (Carrier and Keasling, 1999). In spite of these advances, a universal tool to control mRNA stability in bacteria has not been developed (Sorensen and Mortensen, 2005).

Posttranslational modifications: Glycosylation

Production of antibodies and their fragments constitutes a large share of biopharmaceutical revenues and can be used as example to illustrate some of the key issues of utilizing *E. coli* as a host. The first of these issues is glycosylation. During protein synthesis in mammalian cells, oligosaccharides are attached to glycosylation points on some proteins during their processing in the Golgi apparatus. When an antibody is digested with papain it yields two fragments. While the Fc (Fragment crystallisable) fragment interacts with the cell surface, the Fab (Fragment antigen binding) fragment consists of the antigen binding portion the antibody. These fragments can be synthesized as stand-alone products. Glycosylation is necessary in the case of the Fc of IgG (Immunoglobulin G) for performing its biological function where the oligosaccharide is required for the correct conformation of the protein (Jefferis, 2009). In addition, glycosylation is important for many therapeutic proteins [Table 1.2] by improving stability during formulation and increasing half-life *in vivo* (Sola and Griebenow, 2009).

There are studies which show some degree of bacterial glycosylation (Benz and Schmidt, 2002; Durr et al., 2010; Kowarik et al., 2006) but an *E. coli* strain capable to perform glycosylation consistently and in a human-like pattern has not yet been engineered. On the other hand, *E. coli* is an advantageous host in the cases where an excessive or modified glycosylation patterns triggers immune response on the user (De Groot and Scott, 2007). *E. coli* is a feasible host if glycosylation it is not required for the protein function. Antibody fragments with a short half-life can be useful for example in tumor imaging (Andersen and Reilly, 2004). The pharmacokinetics of Fab expressed in *E. coli* can be improved by attachment of a polyethyleneglycol (PEG) chain to the polypeptide sequence without loss of efficacy which could enable it to compete with the circulation time of recombinant IgG expressed in mammalian cell (Leong et al., 2001).

Cytoplasmic environment

The reducing environment of the cytoplasm of *E. coli* is not favorable to disulfide bond formation. In the cytoplasm, the formation of stable disulfide bonds is rare and seems to be confined to a small group of proteins such as ribonucleotide reductase or the transcription regulator OxyR (Stewart, Aslund, and Beckwith, 1998). It was shown that the cytoplasmic disulfide bond formation in alkaline phosphatase can be catalyzed by thioredoxins which were also shown to participate in the regulation of cytoplasmic homeostasis (Bessette et al., 1999). This characteristic of *E. coli* physiology constitutes a bottleneck during manufacture of insulin, growth hormone, antibodies and antibody fragments (Fc and Fab) due to the requirement for structural disulfide bonds. If the disulfide bonds are not formed correctly, the protein has a high tendency to misfold and aggregate with others in an amorphous insoluble

Brand Name (International Non-proprietary name)	Indication	Effects of Glycosylation	# of glycans
Replagal. (α -galactosidase)	Treatment of Fabry disease	Protects against aggregation and precipitation	3
Myozyme. (α -glucosidase)	Treatment of Pompe disease	Protects against thermal denaturation	6
Prolastin. (Alpha 1-antitrypsin)	Treatment of congenital Alpha 1-antitrypsin deficiency with emphysema.	Protects against chemical and thermal denaturation	3
Bucelipase alfa (cholesterol esterase) Merispase. (Meristem therapeutics)	Treatment of lipid mal-absorption related to exocrine pancreatic insufficiency	Protects against proteolytic degradation	11
Gonal-F. (Corifollitropin alfa)	Treatment of infertility	Protects against thermal denaturation	10
Xigris. (Drotrecogin alfa)	Treatment of severe sepsis	Protects against proteolytic degradation	4
Epogen. (Epoetin alfa)	Treatment of anemia associated with chronic renal failure (CRF)	Protects against oxidation, thermal, chemical, and pH denaturation, kinetic	3
IgG-like antibodies	Multiple indications	Protects against proteolysis and thermal denaturation	2

Table 1.2 Examples of glycosylated recombinant protein products

This table show some commercial recombinant protein products and the extent they benefit from diverse levels of glycosylation. Adapted from Sola and Griebenow (2009).

mass called inclusion bodies (Williams et al., 1982). This requires purification of inclusion bodies, dissolution and refolding (Patra et al., 2000). On the other hand, the periplasm is an oxidizing environment that contains the enzymes necessary for disulfide bond facilitation (DsbA, DsbB, DsbC and DsbD) (Kadokura, Katzen, and Beckwith, 2003). To exploit this characteristic, periplasmic secretion systems were devised for several protein products such as growth hormone (Becker and Hsiung, 1986). It was shown that Fab can be produced in the *E. coli* periplasm (Humphreys et al., 2002). Complete antibodies could be assembled by independent secretion of polypeptide chains to the periplasm in an appropriate ratio which was shown to be crucial (Simmons et al., 2002). For example, the production of each polypeptide chain of the antibody was regulated at transcriptional level by the promoter in the production plasmid and the rate of secretion regulated at the post-transcriptional level by TIRs (Translation initiation regions) (Mazor et al., 2008). The key development in this strategy was the incorporation of a technique which enable the detection and isolation of individual bacteria which produce the antibody chains in the correct ratio. APEX (Ancored periplasmic expression) consists of the display of scFv (single chain variable fragment) attached to the periplasmic membrane by a directly fused anchor domain (Harvey et al., 2004). Full antibodies are displayed in *E. coli* periplasm (*E*-clonal antibodies) by a variation of the APEX system. The Fc portion of the antibody is connected to the anchor domain by an Fc-binding protein. The candidates with expression systems which resulted in the correct ratio, display functional antibodies and could be screened by Fluorescent Activated Cell Sorting (FACS). The expression system could then be transferred to a strain lacking the gene encoding the Fc-binding protein resulting in full antibody secretion (Sidhu, 2007). This technique was utilized to produce a form of IgG and constitutes a promising approach for producing full glycosylated forms of therapeutic antibodies (Mazor et al., 2008).

Inclusion body formation

The aggregation of protein in an insoluble amorphous mass (inclusion bodies) occurs naturally as a biophysical consequence of misfolding or denaturation of protein within living organisms. Under certain conditions, yeast (Cousens et al., 1987) and animal cells (Thomas, Booth, and Roy, 1990) may also produce inclusion bodies. In natural conditions 15-25% of *E. coli* proteins are thermolabile, a leading cause of protein denaturation (Mogk et al., 1999). Chaperones are several families of proteins which evolved to interact with other proteins or peptides in order to prevent or resolve aggregation. For example DnaK interacts with nascent protein to prevent aggregation, IbpA is a small holding chaperone which shields hydrophobic aminoacid R groups and GroE refolds denatured proteins. These proteins form a coordinated network which performs quality control on protein folding status (Baneyx and Mujacic, 2004). Heat shock proteins are reviewed in section **1.2.4.2**. Some of the issues in relation to recombinant protein overexpression in *E. coli* are summarized in [Table 1.3]. In combination with biophysical properties these factors result in saturation of the chaperone network and as consequence much of the recombinant protein aggregating in inclusion bodies (Rinas et al., 2007). From a practical point of view, the differential centrifugation of inclusion bodies is a useful purification stage when the products accumulate in the cytoplasm in such form. Inclusion bodies expressed in the periplasm represent a more difficult problem due to entwinement with peptidoglycan (Swartz, 2001). Since the inclusion bodies can be resolved into native forms (Carrio and Villaverde, 2001), if refolding rates are economically feasible, inclusion body production may be beneficial for the process. This is most likely in the case of small peptides or proteins that are relatively easy to refold. Larger proteins are more complex to refold correctly and tend to be softer meaning that can be easy to unravel by solvent attack. These factors reduce the refolding yields and the production strategy has to be reassessed.

These strategies consist of parallel approaches developed during the last 20 years to deliver robust methods for producing other soluble recombinant proteins in *E. coli*. Cultivation at lower temperatures has been shown to prevent aggregation (Cabilly, 1989). Other alterations of the culture environment such as high osmolarity (Blackwell and Horgan, 1991) or presence of sucrose in the medium (Kipriyanov, Moldenhauer, and Little, 1997) also showed a positive influence on solubility. Secretion of the recombinant protein is also discussed as another method to avoid aggregation by accumulation of product in the cytoplasm (Choi and Lee, 2004). A different focus consists on engineering the strain such as the double mutant *trxB/gor* commercialized by Novagen. This strain, *E. coli* Origami DE3, is selected when an oxidizing cytoplasm favoring disulfide bond formation is needed (Lehmann et al., 2003; Sonoda et al., 2011). Interestingly, (Sonoda et al., 2011) work also showed that co-expression of chaperone GroE enhanced the yield of soluble scFV pointing to natural chaperones as folding enhancers during production. Co-expression of chaperones to increase solubility of product was tested previously by (de Marco and De Marco, 2004; Luo and Hua, 1998; Schlieker, Bukau, and Mogk, 2002). However chaperone overexpression does not always enhance folding and has unintended consequences reviewed recently by (Martinez-Alonso et al., 2010).

1.1.4 *E. coli* cells as means of production

Microbial cells are indeed analogous to factories. They contain large enzymatic macrostructures such as the translation initiation complex functioning as a self assembled process operation unit. They are self powered and capable of exploiting resources in their own immediate environment. The production output is affected by intrinsic critical control points such as disulphide bond establishment and fail-safe mechanisms are commonly in place such as the chaperone network. Much of the efforts to elucidate the functioning of these

mechanisms in *E. coli* have yielded solutions for many of its shortcomings as host, and understanding of some intricacies of its physiology. *E. coli* is a living organism which evolved to utilize resources for improving survival, growth and reproduction. Most of the strategies applied for high level gene expression (Hartley, 2006) and high cell density culture (Choi, Keum, and Lee, 2006) seek to exploit *E. coli* as a cell factory which originates different types and levels of stress responses on the host (Boor, 2006). Process monitoring is key in order to determine if the cell factory function is causing unwanted consequences on its biology characterized by stress, cell damage, growth arrest and ultimately cell death. On-line monitoring is achieved with a sensor in the production line which can provide instant measurements. During off-line monitoring samples are taken from the production line for analysis yielding a delayed result. Most of the monitoring systems utilized have their origin in chemical engineering under the expectation that by maintaining certain variables, it is possible to maintain production in terms of yield and/or quality (Clementschi and Bayer, 2006). The biochemical complexity of the *E. coli* as cell factory is such that chemometric variables need to be combined with mathematical models to estimate biological variables (Clementschi and Bayer, 2006). The *E. coli* physiology is an integration of these biological variables. It has been discussed that improved physiology of the host leads to greater production yields (Chou, 2007; Sevastyanovich, Alfasi, and Cole, 2009). Process monitoring could be aided by biological tools to measure complex biological variables that are difficult to predict or quantify in order to maintain *E. coli* homeostasis during production. Protein folding status is such a type of complex biological variable triggered by a combination of chemical, physical and biological factors. It is determined by the physiological condition of the cell during production and ultimately, influences the process performance. For monitoring

	Causes	Consequences	Possible solutions	References
Endotoxin Levels	These lipopolisaccharides are ubiquitous in Gram-negative bacteria	If passed on the final product may harm the patient	Validation of removal downstream processing	(Petsch and Anspach, 2000)
Depletion of specific tRNAs	Codon usage is different between eukaryotes and prokaryotes	Rare tRNAs may not be sufficient causing traslation to stall	Codon optimization/additional plasmid encoding scarce tRNAs	(Sorensen and Mortensen, 2005)
mRNA stability	mRNA in <i>E. coli</i> is rapidly degraded	Production is rate limited by transcription	Incorporation of RNA leader sequence to reduce RNase action	(Carrier and Keasling, 1999)
Protein glycosylation	<i>E. coli</i> cannot reproduce glysosylation patters on proteins	Aglycosylated products may be functional. However the stability <i>in vivo</i> decreases	In some cases site specific attachment of PEG to the protein enhanced stablility without loss of efficacy	(Leong et al., 2001)
disulfide bond formation	the cytoplasmic environment strongly disfavors disulfide bond formation	Peptide folds incorrectly and aggregates in inclusion bodies	<i>In vitro</i> refolding/periplasmic secretion	(Patra et al., 2000) (Humphreys et al., 2002)

Table 1.3 Summary of main biological considerations when utilizing *E. coli* as host for recombinant protein production

First column shows the typical considerations when using *E. coli* as a host.

protein folding it is necessary to have knowledge of *E. coli* physiology in order to understand how the cell senses this stress situation, and then an adequate target can be selected for measurement. The incorporation of a technology, such as reporter gene technology, able to recognize the cell interpretation and to produce a measurable signal is crucial to this development. Finally, a suitable detection system is needed for quantifying the signal and collecting data.

1.2 Microbial physiology under industrial conditions

Industrial conditions are very different to the natural environments where microbes have evolved and adapted to. *E. coli* is usually grown industrially at cell densities that exceed natural scenarios. It is altered with foreign DNA, forced to resist selective pressure and to produce foreign proteins upon induction. At certain stages oxygen and nutrient shortage may occur. It was shown that these factors combine to cause of growth arrest in bioreactors (Konz, King, and Cooney, 1998; Matin, 1990). The effect on *E. coli* growth could be correlated with amount of recombinant protein production, its synthesis being the cause of metabolic burden to the host (Bentley et al., 1990) In addition, on-line monitoring of O₂ consumption, CO₂ profiles and biomass showed that increased induction of protein production originated a sequence of events which caused energy inefficiency towards protein production (Hoffmann and Rinas, 2001). Metabolic burden can be summarized as the quantitative effect of diverse factors on the host physiology caused by the presence and activation of recombinant production systems.

Microarray technology has been utilized for monitoring gene expression (Schena et al., 1995) which enabled the observation and comparison of transcription levels of all genes in the host.

This became very useful for observing transcriptional response to stress. For example, (Zheng et al., 2001) revealed that many genes became activated in response to hydrogen peroxide. It was shown that mutants lacking the major regulator of peroxide response, OxyR, expressed some stress genes, indicating an OxyR-independent response. Interestingly, the results also revealed a certain degree of overlap of stress responses by up-regulation of heat shock protein genes *ibpA* and *ibpB* in wild type and $\Delta oxyR$ mutants. Such experiments have shown that stressful situations trigger a series of events involving the coordinated regulation of expression of many genes, frequently spanning across different regulatory networks. Upon a stimulus, a chain of events leads to activation of specific regulatory genes which orchestrate coordinate responses of groups of genes scattered around the genome. In industry, pH probes, thermometers, and galvanic sensors for dissolved oxygen serve to monitor media conditions during fermentations. Such sensors convert a physical signal from the detector into an electrical signal (Clarke et al., 1985). Subsequently the electrical signal is interpreted into digital data for the operator. Biological elements of stress regulation operate in an analogous manner to sense these parameters outside and inside the cell. Understanding these mechanisms of the diverse stress responses is essential for designing the tools which serve to monitor them and monitor from the perspective of the host stress caused by industrial fermentation conditions.

1.2.1 Response to stress: Gene expression regulation occurs at transcription initiation

RNA polymerase

Restoration and maintenance of the cell homeostasis during physiological challenges requires an alteration of gene expression patterns. The major control point of gene expression occurs at

transcription initiation (Murakami et al., 2002). DNA is transcribed to RNA by RNA polymerase (RNAP). This enzymatic complex is highly conserved in all living organisms and, in bacteria, it is the target of most transcriptional regulators (Ebright, 2000). The RNAP complex is formed by the subunits $\alpha_2\beta\beta'\omega$ and the dissociable σ subunit **[Fig. 1.1]** The α subunits dimerize, bind DNA and collaborate in the assembly of the β and β' subunits, which have the ability to interact with two polynucleotide chains; the DNA template and the nascent RNA. The ω subunit does not participate in the transcription process but it was suggested to function as a stabilization factor of the RNAP complex (Hampsey, 2001). While the RNAP core ($\alpha_2\beta\beta'$) possesses catalytic activity for RNA polymerization, it is the σ subunit, also called a σ factor that recognizes and bind promoter DNA. When the RNAP core binds a σ subunit it forms a holoenzyme ($E\sigma$)

The promoter structure

The transcription of all genes is under the control of promoters. A promoter consists in a DNA region upstream the gene which presents specific sequences that recruit and interact with the structure of the RNAP complex **[Fig. 1.1(c)]**. For example, the σ^{70} family recognizes the promoter elements at -35 and -10 which are 35 and 10 base pairs (bp) before the transcription initiation point. The -10 element, also called Pribnow box (Pribnow, 1975) and the -35 element have the consensus sequences for σ^{70} TATAAT and TTGACA respectively; which allowing variation, exist as homologous structures (Hawley and McClure, 1983). In fact, the ability to bind -10 and -35 elements has been shown to be highly conserved across species. The σ^{70} subunit binds the two main elements of promoter, the -35 and -10 hexamers

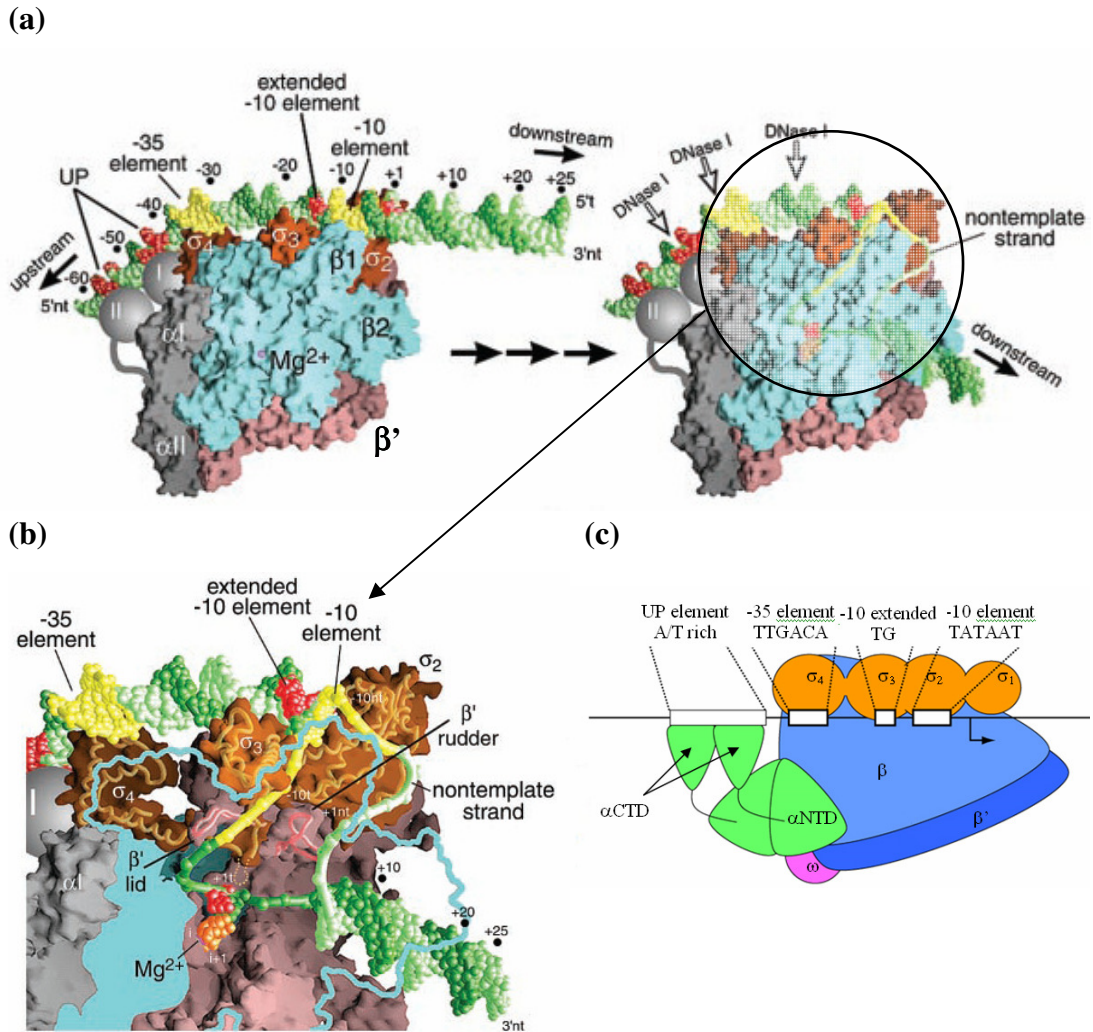


Fig. 1.1 Interaction of RNAP holoenzyme with DNA

In (a) it is shown the interaction of the RNAP holoenzyme with the DNA and the subsequent melting of the double strand structure at the promoter site. The site of DNA melting is detailed in (b) where the subunit β removal from the model is depicted with a cyan line enables the observation of the transcription bubble and the intervention of Mg^{2+} in the process. Models shown in (a) and (b) have been produced from x-ray crystallography data. Adapted from Murakami et al. (2002) (c) Scheme of the different subunits in the RNAP holoenzyme showing the promoter features that interact with the different domains of the σ subunit expressed by subscript. From Busby and Savery (2001).

through σ_4 and σ_2 domain, Other important elements are: the TG motif (or extended -10 motif) which interacts with the σ_3 domain; and the UP element, which interacts with the α subunits and is located approximately 20bp upstream of the [-35] hexamer (Busby and Ebright, 1994). These four elements contribute to RNAP recruitment but the specific contribution of each varies from promoter to promoter. Assembly of the RNAP complex on the promoter provokes an unwinding of the DNA from positions -10 to +2 bp enabling the RNA synthesis to proceed [Fig. 1.1 (b)].

1.2.2 Regulation at transcription initiation by σ factors

The σ^{70} family

The general structure of the σ factor consists of four rigid domains connected by flexible linkers which perform the biochemical activities of the σ factor [Fig. 1.2 (a)]: σ_1 , when present, is suggested to perform inhibition of DNA binding modulating transcription (Dombroski, Walter, and Gross, 1993); σ_2 binds to the [-10] element and contributes to promoter melting; σ_3 may bind to the extended -10 element, which may act as substitute for -35 element binding (Campbell et al., 2002a); and σ_4 binds to the [-35] element of the promoter (Murakami et al., 2002). Multiple sequence analysis of different σ factors was performed and visualized on a phylogenic tree (Lonetto et al., 1992). Modern algorithms and further DNA sequencing data revealed two major findings: each species screened contained distinct families of σ factors; and such families were shared amongst species [Fig. 1.2 (b)].

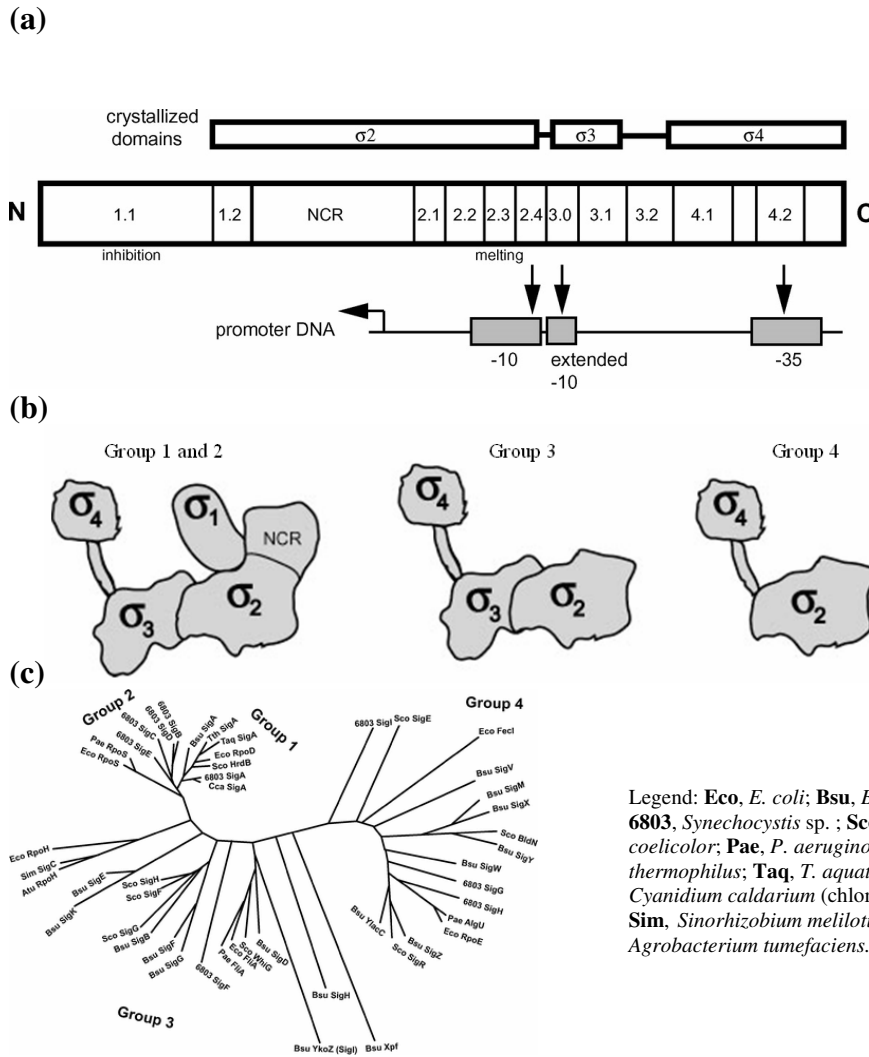


Fig. 1.2 Structural groups and phylogenetic characteristics of σ^{70} family

The diagram (a) is a schematic representation of a typical σ subunit structure based on σ^{70} family. Vertical arrows show the regions interacting with the promoter. NCR stands for non-conserved region (b) shows the domain arrangement of the diverse structural groups based on complete holozyyme structural information. The phylogenetic tree (Fitch-Margoliash) (c) represents the different groups of within the σ^{70} family originated from multiple sequence alignment spanning from 1.2 to 2.4 regions of the σ^{70} subunit of different species. Images from Gruber and Gross (2003).

These groups present structural differences in the protein domains which have implications in the affinity for the RNAP core ($\alpha_2\beta\beta'$ ω) or binding sites for transcription factors (Gruber and Gross, 2003).

There are approximately 4000 genes in the *E. coli* genome (Blattner et al., 1997b). This emphasizes the necessity of a selection mechanism to decide which genes are to be transcribed. The interchange and reuse of specific σ subunits by the RNAP complex (Travers and Burgess, 1969) and the competition among σ subunits for the RNAP core (Maeda, Fujita, and Ishihama, 2000) are thought to be the most common control points. There are six members in the σ^{70} family in *E. coli*: σ^{70} , known as housekeeping factor; and the alternative sigma factors: σ^S ; σ^H ; σ^E ; σ^F ; and σ^{FecI} . Each σ factor competes to bind the RNAP core on a weighted manner based on affinity differentials. It was shown that a 16 fold affinity difference exists between the highest affinity (σ^{70}) and the lowest affinity (σ^S). Taking into account that σ^{54} belongs to a different σ family (Gruber and Gross, 2003) and does not bind -35 element, binding experiments *in vitro* yielded the following hierarchy of binding preference to RNAP core: $\sigma^{70} > \sigma^{54} > \sigma^F > \sigma^{32} / \sigma^{FecI} > \sigma^E > \sigma^S$ obtained in conditions of 2:1 saturation of σ subunits and RNAP core (Maeda, Fujita, and Ishihama, 2000). Modulation of the relative abundances of each σ factor type is decisive on which genes are to be transcribed. Each σ^{70} family member possesses different affinities for specific variations in the sequences of the -35 and -10 elements [Table 1.4], thus displaying preference to bind specific promoters. Most of the growth related and housekeeping genes are transcribed by $E\sigma^{70}$ (RNAP with σ^{70}). The alternate σ factors are involved in expressing many genes related to diverse stress responses. In the right conditions, an increase of a determined σ factor or a modulation of its affinity for

σ^{70} -family	Name, synonyms	Consensus sequence (*)		
		-35 Element	Spacer	-10 Element
Group 1 Primary σ -factors	σ^{70} , RpoD (<i>E. coli</i>)	TTGACA	16–18	TATAAT
Nonessential primary-like σ-factors				
Group 2 Stationary-phase σ -factor	σ^S , σ^{38} RpoS (<i>E. coli</i>)	TTGACA		CTATACT
Alternative σ-factors				
Group 3 Flagella σ -factors	σ^{28} , σ^F , FliA (<i>E. coli</i>)	TAAA	15	GCCGATAA
Group 3 ECF σ -factors	σ^E , σ^{24} (<i>E. coli</i>)	GAACCTT	16–17	TCTRA
Group 3 Heat shock σ -factors	σ^{32} , σ^H , RpoH (<i>E. coli</i>)	CTTGAAA	11–16	CCCATnT
	σ^B , SigB (<i>Bacillus subtilis</i>)	GTTTAA	12–14	GGGTAT
Group 3 Sporulation σ -factors	σ^H , SpoOH (<i>Bacillus subtilis</i>)	AGGAWWT	12–14	RGAAT
	σ^F , SpoIIAC (<i>Bacillus subtilis</i>)	WGCATA	14–15	GGnRAYAMTW
	σ^E , SpoIIGB (<i>Bacillus subtilis</i>)	GKCATATT	13–15	CATACAMT
	σ^G , SpoIIIG (<i>Bacillus subtilis</i>)	TGAATA	17–18	CATACTA
	σ^K , SpoIIIC (<i>Bacillus subtilis</i>)	AC	16–17	CATAnAnTA
Group 4 Ferric citrate transport (addition from (Enz, et al., 1995))	σ^{FecI}	AAGGAAAAT	17	TCCTTT
σ^{54}-family				
	σ^N , σ^{54} , RpoN (<i>E. coli</i>)	TGGCAC	5	TTGCW

*Ambiguous codes: N, any base; R, A or G; W, A or T; Y, C or T; M, A or C; K, G or T.

Table 1.4 Consensus sequences recognized for some sigma factors

The table indicates the consensus sequences for some of the sigma factors of eubacteria. The group number makes reference to the structural types in the phylogenetic tree from [Fig 1.2 (c)]. This table was adapted from Wösten, (1998). “Spacer” is the distance in base pairs between -35 and -10 elements, “ECF”= Extracytoplasmic factor

the RNAP core causes displacement of σ^{70} (Ishihama, 2000). While σ^{FecI} transcribes specific genes to deal with iron starvation (Visca et al., 2002), and σ^{F} activates genes related to the flagellum production (Chilcott and Hughes, 2000), the most relevant σ factors regarding microbial physiology in bioprocesses are σ^{S} , σ^{H} and σ^{E} .

1.2.2.1 The master stress response: σ^{S}

The σ^{S} factor is the master regulator of the stress response, recognizing about 70 promoters related with resistance genes for diverse stress conditions such as high temperature, ethanol, oxidative stress, low pH, UV radiation, and elevated osmolarity (Hengge-Aronis, 2002a).

Initially called σ^{38} (it is 38kDa in size) it has a major role in the expression of many stationary phase genes (Lange and Hengge-Aronis, 1991), thus was redefined as the σ^{S} factor. The gene *rpoS* (**R**NAP **p**olymerase **S**) encodes σ^{S} and is upregulated during the change from exponential growth to stationary phase (Lange and Hengge-Aronis, 1994). Due to its lower affinity for the RNAP core it is important that σ^{S} is expressed in proportion to compete with other σ subunits. During stationary phase σ^{S} was shown to compete with σ^{70} for the RNAP core (Farewell, Kvint, and Nystrom, 1998). The presence of σ^{S} ranges from undetectable during exponential growth to near one third of the amount of σ^{70} in stationary phase (Jishage et al., 1996). The anti- σ factor Rsd (**R**egulator of **s**igma **D**; σ^{D} is a synonym of σ^{70}) contributes to the increase of σ^{S} titre binding σ^{70} during stationary phase (Jishage and Ishihama, 1998). The involvement of σ^{S} on the expression of such a variety of stress situations requires a very complex regulatory network at different levels. The regulation of *rpoS* at transcriptional level is very complex involving diverse transcription factors [Table 1.5].

Gene	Promoters	σ Factor	Transcriptional regulation	
			Activators	Repressors
<i>rpoS</i>	(P1, P2)*	σ^{70}	GadX (Glutamate decarboxylase sytem)	ArcA-P (Anaerobic respiration control)
	P3		σ^S is the σ factor of stationary phase	Fur-[Fe2+] (Ferric uptake regulation) cAMP-CRP (cAMP receptor protein)
<i>rpoE</i>	P1	σ^E	σ^E is the σ factor of bacterial envelope stress	CpxR-P (Conjugative pilus expression)
	P2	σ^{70}		
<i>rpoH</i>	P1	σ^{70}, σ^S	CpxR-P	ATP-Dna (Dna is a protein which plays a key function in bacterial DNA replication (Ogawa, et al., 2002))
	“P2”	#	IhfB-IhfA (Integration host factor)	CytR (Cyt repressor is modulated by Cytidine binding and operates as anti-activator binding CRP (Kallipolitis and Valentin-Hansen, 1998))
	P3	σ^E		
	P4	σ^{70}		
	P5	σ^{70}	σ^{32} is the σ factor of the heat-shock response	
	P6	σ^{54}		
			cAMP-CRP acts as modulator	

Table 1.5 Gene expression σ factors related with common stress responses

The table shows the σ subunits that recognize the different promoters expressing the three most relevant σ factors in microbial physiology during bioprocess. (*) regarding *rpoS* expression, P1 and P2 are promoters of the operon *NlpD-rpoS* (NlpD: new lipoprotein). These promoters contribute to basal *rpoS* expression independently of growth phase (Lange and Henggeaaronis, 1994). (#) in *rpoH* P2 was originally believed to be a minor promoter site for σ^{32} (Fujita and Ishihama, 1987), in fact such site has been shown to be present in *Caulobacter crescentus* (Wu and Newton, 1997). In *E. coli*, it has been demonstrated that σ^{32} does not recognize *rpoH* thus “P2” is not accounted as promoter. This table was produced with information made available from EcoCyc database (Keseler, et al., 2011).

Promoter recognition and expression of σ^S genes

The main promoter structural elements -10 and -35 which binds σ^S are highly similar to those of σ^{70} . However, σ^S seems to be able to bind degenerate hexamers in the -35 element which can result in emergence of selectivity in a non exclusive manner (Gaal et al., 2001). The [-10] extended element displaying a cytosine at [-13] is specific for σ^S binding (Hengge-Aronis, 2002b). Based upon the ability of σ^{70} to sense the spacer region (Dombroski et al., 1996), it has been suggested that sequence variability between σ factors could be a cause for the selectivity by alteration of the spacer region in certain promoters (Hengge-Aronis, 2002b). Recent reports acknowledged the importance of the spacer region in sequence recognition based on mutated σ^{70} studies (Singh et al., 2011).

1.2.2.2 The heat shock response: σ^{32}

The sigma factor σ^{32} (32kD) is the master coordinator of the heat shock response (Arsene et al, 2000). In microbial physiology the heat shock response constitutes a strategy for coping with damage in proteins, involving the coordinated expression of heat shock proteins (such as IbpA); chaperones (such as GroEL; DnaK; and their cohorts); and ATP-dependant proteases (such as ClpAB) (Yura and Nakahigashi, 1999). The factor σ^{32} is encoded by the *rpoH* gene which is regulated by 5 promoters (P1; P3; P4; P5; and P6). The promoters P1; P4; and P5 are recognized by the σ^{70} subunit while P3 is recognized only by σ^E . At 30°C, promoters P1 and P4 account for the 90% of the transcription (Kallipolitis and Valentin-Hansen, 1998). P1 can alternatively recruit σ^S depending on the growth phase (Janaszak et al., 2009). P6 contains a binding site for σ^{54} (Janaszak et al., 2007). The promoter P5 is regulated by the transcription factor cAMP-CRP while P3 and P4 have shown to be regulated negatively by DnaA

(Kallipolitis and Valentin-Hansen, 1998). This complex regulation ensures expression at diverse physiological conditions including temperatures up to 50°C mediated by σ^E .

Regulation of σ^{32}

A major regulation point of σ^{32} activity occurs at the post-transcriptional level. While the transcription rate of *rpoH* did not increase upon temperature up-shift, the amount of mRNA hybridized increased up to five fold at 42°C (Kamath-Loeb and Gross, 1991). The σ^{32} mRNA is accumulated in the cytoplasm folded in a secondary structure. This structure has been predicted utilizing the folding algorithm Mfold (Jaeger, Turner, and Zuker, 1990) and tested in detail. The ribosome binding site becomes blocked in a hairpin structure at low temperature [Fig. 1.3]. Temperature increase melts this secondary structure enabling transcription (Morita et al., 1999). This property of σ^{32} mRNA was described as a microbial thermosensor (Klinkert and Narberhaus, 2009).

Once σ^{32} has been translated it is subjected to further regulation where stability of σ^{32} is a chief property of its own regulation with a half life of ~1 minute. The membrane-bound protease FtsH has been shown to be critical for its degradation by observing an increase of in the half-life of σ^{32} up to 2h in a *ΔftsH* mutant (Tatsuta et al., 1998). On the other hand, the chaperones DnaK and its co-chaperone DnaJ play a major role by reversibly binding to σ^{32} . This ATP-dependant inhibition is partially reverted by GrpE in a regulatory cycle important for the maintenance of σ^{32} in relation to protein misfolding (Gamer et al., 1996).

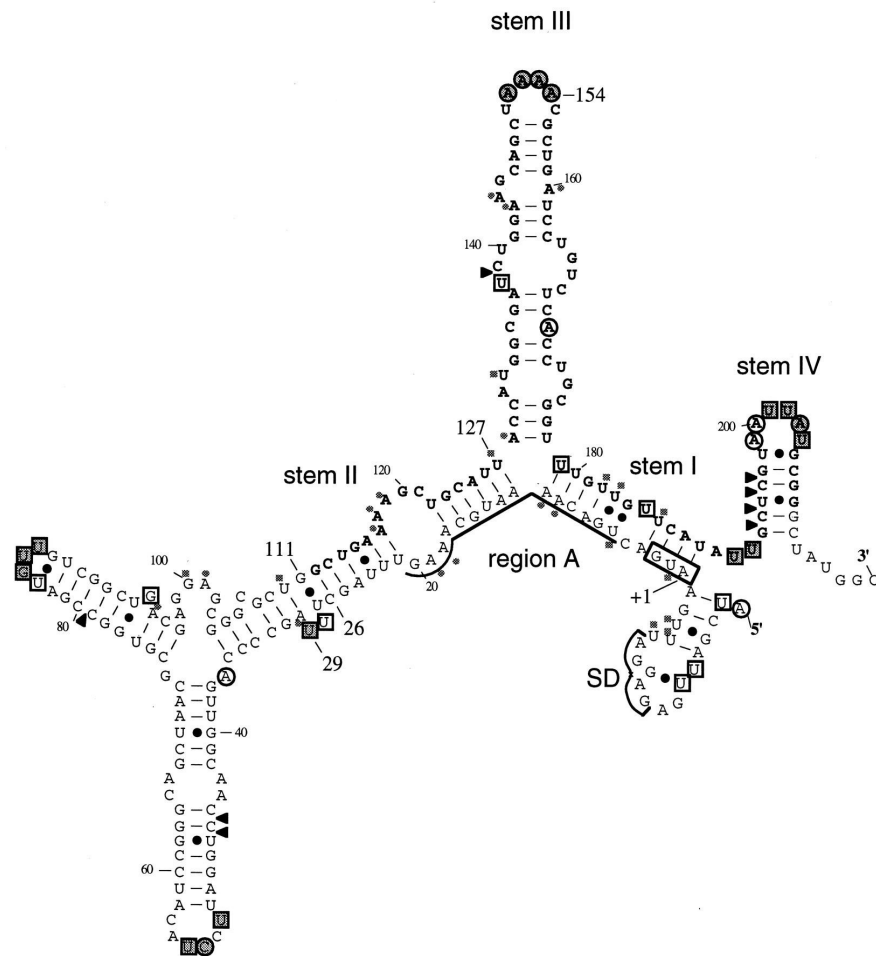


Fig. 1.3 Probable secondary structure of σ^{32} mRNA

This structure was obtained from structure probing with chemical probes that modify single stranded ribonucleotides. Boxed bases, outlined open, and nonoutlined shaded symbols grade the reactivity to the probes from strong, modest, and weak respectively. This conformation is probable at low temperatures which disables it from being translated. SD is the ribosome binding site. From Morita et al. (1999)

1.2.2.3 The envelope stress response: σ^E

Sigma E was originally discovered as a novel factor transcribing the *rpoH* promoter P3. While all the other promoters are function at diverse temperatures, σ^E (24 kD; also called σ^{24}) was the only factor which remained active on extreme heat shock (~50°C). Initially, σ^E was thought to be part of a second heat shock response to extreme temperature (Erickson and Gross, 1989). σ^E was shown to transcribe 43 other genes. However, its main function is to recognize promoters of the battery of genes related to the correct functioning of the bacterial envelope (Dartigalongue, Missiakas, and Raina, 2001). The most relevant examples are DegP and FkpA. DegP is an essential protein of the periplasmic space for survival at elevated temperatures (Lipinska et al., 1989), and has dual functionality, proteolysis of misfolded proteins and chaperone activity. DegP was shown to contribute to α -amylase (an enzyme which degrades maltodextrins) refolding *in vitro*. While protease activity was undetectable below 22°C, it rose to 8 fold from 28°C to 42°C showing the dual function of this widely conserved protein (Spiess, Beil, and Ehrmann, 1999). Protein-protein interaction is mediated by PDZ domains (Harris and Lim, 2001), which are structural domains of ~80-90 aminoacids in the protein. DegP associates in multimers of 6, 12 or 24 subunits. Trimer or hexameric forms did not show any these activity being mostly resting forms. Association in 12 subunit multimer exposed proteolytic sites. When DegP formed 24 subunit multimer, changes in the conformation caused by the multimerization resulted in deactivation of proteolytic sites and activation of the chaperone function (Krojer et al., 2008). Besides its quality control function through proteolytic activity, DegP is crucial for safe transport of partially folded outer membrane proteins through the periplasm (Sklar et al., 2007). FkpA is an isomerase with chaperone activity related to folding proteins of the envelope (Narayanan and Chou, 2008). These findings redefined it as the envelope stress σ factor, a member of the ECF (extra

cytoplasmic factor) subfamily (Raivio and Silhavy, 2001). The bacterial envelope is a complex and dynamic structure subject to constant changes during bacterial growth and environmental variations. These include mechanism to manage stress responses caused by chemicals such as ethanol, extreme temperature and the folding status of envelope related proteins.

Regulation of σ^E

The expression of *rpoE* involves a positive feedback loop where σ^E transcribes its own gene through the promoter P2 (Rouviere et al., 1995) as the first gene in the *rpoE* operon, followed by *rseA*, *rseB*, and *rseC* (**regulator of sigma e**) (Missiakas et al., 1997). The regulation of σ^E occurs mostly at the posttranslational level where several protein interactions are involved [**Fig. 1.4**]. RseA is an inner membrane protein which acts as an anti- σ^E , binding it by its N-terminal region. RseB is found in the periplasm acting as a misfolding detector. Under normal conditions RseB is bound to RseA in a transmembrane complex which sequesters σ^E . In the event of a rise of misfolded proteins in the periplasm, RseB is titrated out triggering the release of σ^E by the cytoplasmic domain of RseA (DeLasPenas et al, 1997). Crystallized RseB has shown a structure which enable interaction with RseA and a feature which could enable interaction with lipoproteins (Wollmann and Zeth, 2007). The previous observations are compatible with the most recent model of σ^E regulation. It was suggested that if RseB is titrated away from RseA by misfolded protein, RseA would be exposed to proteolysis with consequent increase of free σ^E (Ades et al., 1999) [**Fig. 1.4 (b)**] In this proteolytic cascade, the proteins RseP, DegS, Clp, and Lon participate with different degrees of involvement (Chaba et al., 2007). RseP (also referred as YaeL or EcfE) degrades RseA upon activation

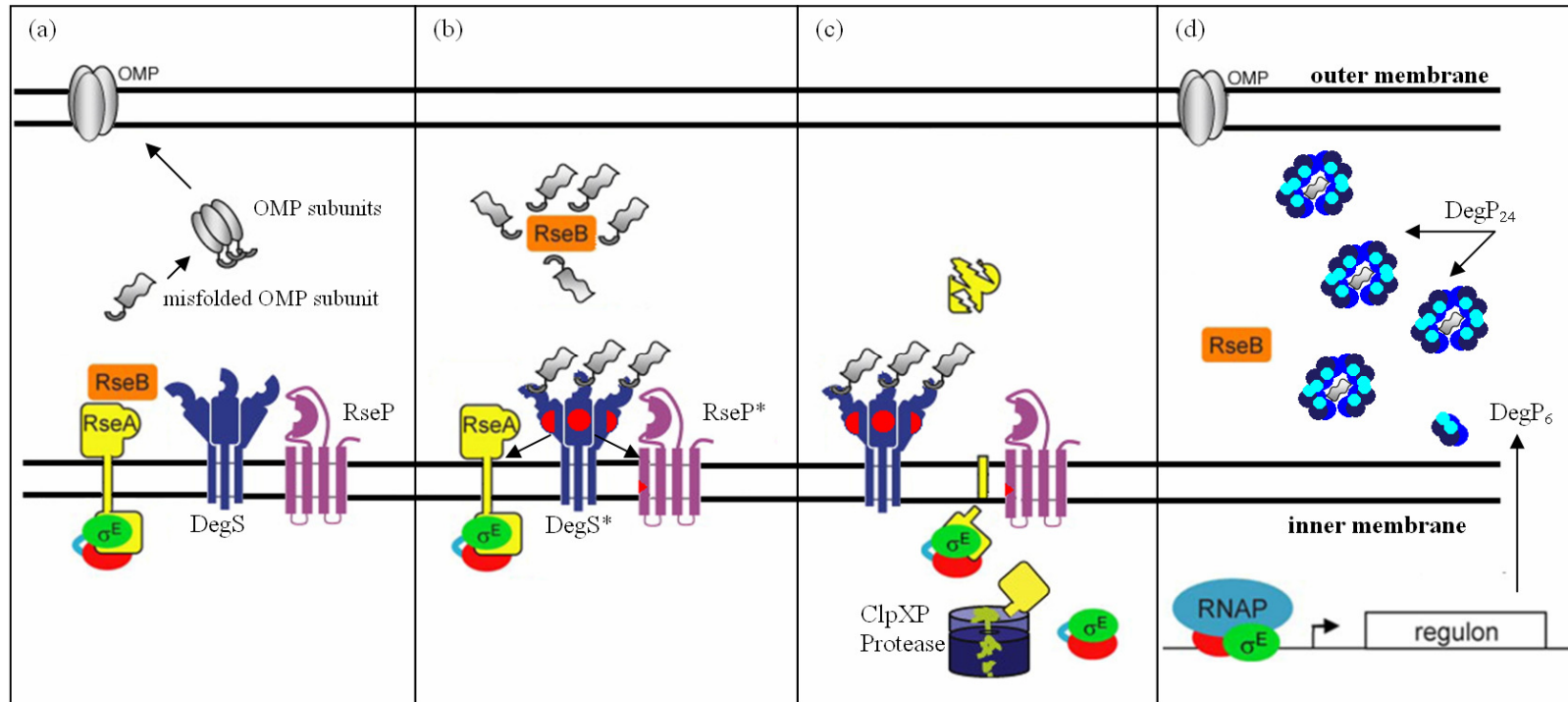


Fig. 1.4 Scheme of σ^E regulation by proteolysis

The box (a) illustrates the uninduced status of the system during presence of correct folding of OMP subunits. (b) If misfolded OMP subunits accumulate, they activate DegS (DegS*) proteolytic activity and sequester RseB leaving RseA vulnerable to proteolysis. Proteolytic activity of DegS* degrades the periplasmic domain of RseA and activate RseP. (c) RseP* cleaves the transmembrane domain of RseA exposing the cytoplasmic domain to destruction by ClpXP releasing σ^E . (d) σ^E activates transcribe genes of the σ^E regulon, amongst them DegP. DegP multimerizes into DegP₂₄ capturing misfolded OMP and contribute to its transport through the periplasmic space allowing an increase of RseB titre which restarts the regulatory cycle. Image adapted from (Ades, 2008).

(Kanehara et al., 2002). DegS is a regulatory protease which has shown two properties: (1) it is activated by misfolded outer membrane proteins (OMPs) (Walsh et al., 2003); and (2) possesses proteolytic activity able of specific cleavage of RseP exposing the active centre and enabling RseA degradation (Chaba et al., 2007); **[Fig. 1.4 (b)]**. DegS presents a PDZ domain which consists of a sequence of ~80 residues. Crystallized structures of DegS associated to an OMP have shown that interaction through the PDZ domain triggers a conformational change which exposes the proteolytic sites of DegS (Wilken et al., 2004).

1.2.3 Regulation of gene expression at gene transcription initiation by transcription factors

TFs (transcription factors) are proteins which possess DNA binding domains, enabling them to interact with DNA in specific sites, activating or repressing the recruitment of the RNA polymerase complex. It has been found that seven of these proteins (CRP, FNR, IHF, FIS, ArcA, NarL and Lrp) modulate more than half of all the *E. coli* genes (Martinez-Antonio and Collado-Vides, 2003).

Catabolite repression

CRP is one of the major regulators of metabolic flux. CRP (**c**AMP **r**eceptor **p**rotein) is activated by cAMP which causes a conformational change that enables dimerization of two identical subunits and DNA binding and. With a total mass of 45kD, cAMP-CRP (or CAP, **c**atabolite **a**ctivator **p**rotein), may recruit the RNAP through interaction with the α subunits through the C-terminal domain (α CTD) achieving modulation of gene expression (Busby and Ebright, 1999). Glucose inhibits the enzyme adenylate cyclase, lowering the levels of cAMP

(Peterkofsky and Gazdar, 1974). It was shown that glucose also reduces the expression of CRP by auto-regulation (Ishizuka et al., 1994).

Regulator of fumarate and nitrate reduction (FNR)

FNR (30kD) is a member of the same TF structural family as CRP (Green, Scott, and Guest, 2001) and it coordinates the necessary metabolic changes in order to utilize alternate electron acceptors in anaerobic conditions (Spiro and Guest, 1990). In anaerobic conditions, the N-terminal sequence of FNR contains 4 cysteine residues that bind a [4Fe-4S] cluster, which promotes the dimerization of FNR, causing a conformation change that exposes DNA binding sites similar to CRP (Moore, Mettert, and Kiley, 2006). Presence of oxygen in the cytoplasm converts the cluster to $[2\text{Fe-2S}]^{2+}$ which inactivates FNR (Khoroshilova et al., 1997). The metabolic changes directed by FNR regulation are coordinated with ArcA and NarL (Unden and Bongaerts, 1997).

1.2.3.1 Sensing the environment: Two component signal transduction

Two-component signal transduction enables the host to sense the environment and correctly respond by transmitting information from a molecular sensor to the DNA. As reviewed by Stock et al. (2000), the general structure and function features two components: a sensor kinase protein in the cytoplasmic membrane and a cytoplasmic response regulator (RR) [Fig. 1.5]. Particular environmental parameters produce an effect on the sensor kinase which mediates the activation of the cytoplasmic response regulator. The sensor has a histidine kinase (HK) domain while the RR an asparagine phosphotransfer domain. Environmental sensing triggers autophosphorylation of the sensor. The HK domain transfers the phosphoryl

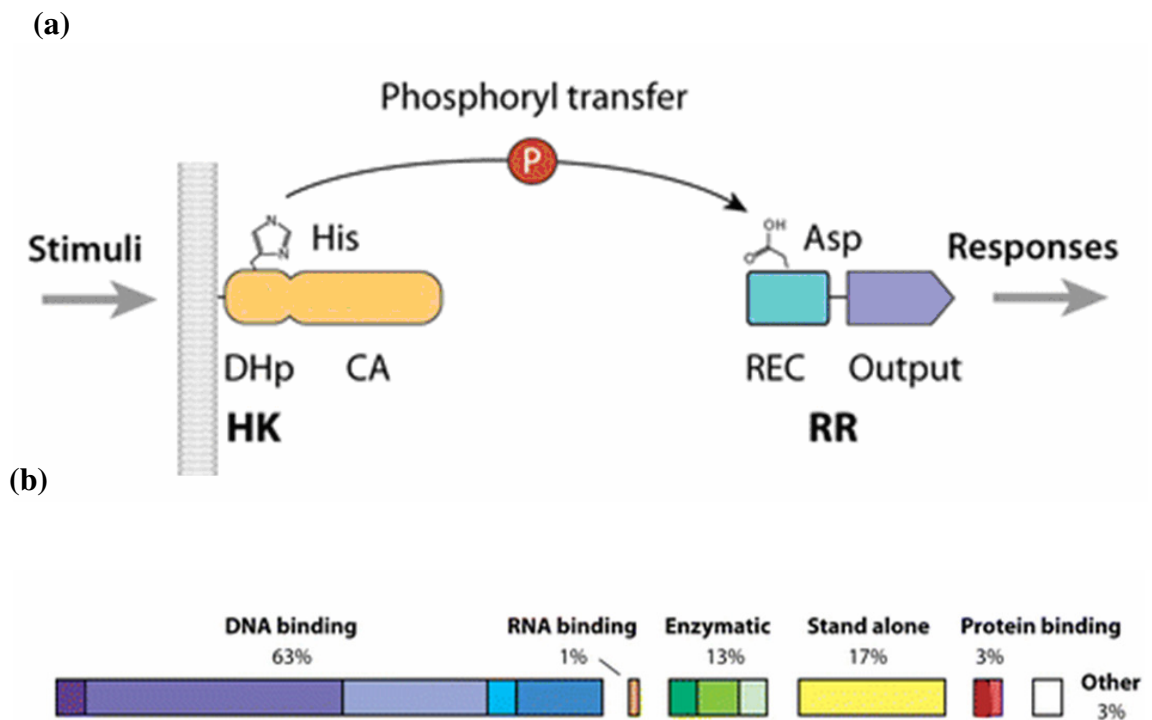


Fig. 1.5 Paradigmatic structure of two-component system

Diagram (a) shows the element of a typical histidine kinase (HK). The histidine phosphotransferase domain (DHp) and the catalytic ATPase (CA) form the kinase core. The in the response regulator (RR) asparagines is shown on the receiver domain (REC). (b) Shows the diversity of the effector domains in the response regulator. The percentage indicates the distribution of each functional class. The ‘Stand alone’ portion represent probable response regulator of unknown function. Adapted from Gao and Stock (2009)

group to the response regulator protein through its phosphotransfer activity, producing a conformational change in the RR which enables it to perform its activity in the cytoplasm. Most of the RRs in *E. coli* function as transcription factors (TFs) (Mizuno, 1997). TFs present DNA binding domains which enable them to interact with DNA at specific sites, activating or repressing the formation of the RNA polymerase complex. This transduction system is a common mechanism to transduce external stimuli into cytoplasmic effects.

The aerobic respiration control ArcA/B

ArcA/B regulates expression of genes related with aerobic metabolism (Sawers, 1999). This two-component system consists of a membrane bound sensor protein, ArcB and a DNA binding protein ArcA which in its phosphorylated form regulates up to 175 different operons related to growth in diverse redox conditions (Salmon et al., 2005). The sensor protein ArcB possesses an intramolecular phosphorelay mechanism. During anaerobic conditions, autokinase and phosphotransferase activities of ArcB result in active ArcA-P. In aerobic conditions, ArcB dimerization disables the kinase activity which results in dephosphorylation of ArcA-P by ArcB, yielding inactive ArcA (Krell et al., 2010). Quinones are electron carriers associated to the membrane which participate in electron transfer. They have been postulated to reduce ArcB forming the disulfide bonds necessary for dimerization. (Georgellis, Kwon, and Lin, 2001) Evidence showed that ubiquinone-0 (soluble analog of ubiquinone-8) inhibited auto-phosphorylation of ArcB while molecular O₂ and H₂O₂ did not affect ArcB; implying that a determined redox state of the ubiquinone pool of the respiratory chain may lead to gene regulation through ArcA~P (Malpica et al., 2004). ArcA/B regulation is critical during growth under low oxygen conditions, regulating carbon flux through TCA-cycle (tri-carboxylic acid) (Alexeeva et al., 2003). It has been shown that ArcB can function

as sensor kinase for RssB, branching into the σ^S stress route adding a regulatory effect on σ^S based on aerobic/anaerobic sensing (Mika and Hengge, 2005).

Envelope stress sensing by CpxA/CpxR

The CpxA/R (conjugative pilus expression) was first reported as a regulatory system of bacterial conjugation through pilus formation upon study of *cpx* mutants (McEwen and Silverman, 1980). Signals that affect the Cpx A/R regulation include: high pH (Nakayama and Watanabe, 1995); alterations in membrane composition (Mileykovskaya and Dowhan, 1997); increase of the OMP NlpE, key in bacterial adhesion (Hirano et al., 2007); conjugative pilus subunits (Jones et al., 1997); increase of osmolarity (Jubelin et al., 2005); indole (Hirakawa et al., 2005); and copper in a NlpE dependent manner (Gupta et al., 1995; Yamamoto and Ishihama, 2006). It was suggested that most of the physico-chemical challenges may lead to the same consequence, accumulation of misfolded protein in the periplasm (DiGiuseppe and Silhavy, 2003). Three main proteins are involved this pathway [Fig. 1.6]: CpxA is integrated in the cytoplasmic membrane functioning as sensor kinase; CpxR is the cytoplasmic response regulator activated by CpxA; and the periplasmic protein CpxP, a small stress combative protein (Danese and Silhavy, 1998; DiGiuseppe and Silhavy, 2003). Experiments have shown that CpxP is degraded along with misfolded pilus subunits by DegP, a member of the σ^E regulon, suggesting that it functions as an adapter to recognize misfolded protein in the periplasm (Isaac et al., 2005). The recently obtained crystal structure of CpxP supported this possible function and provided a structural basis to the inhibitory mechanism of CpxP on CpxA during absence of stress (Zhou et al., 2011). The response regulator CpxR has shown to be activated upon entry in stationary phase in $\Delta cpxA$ mutants by

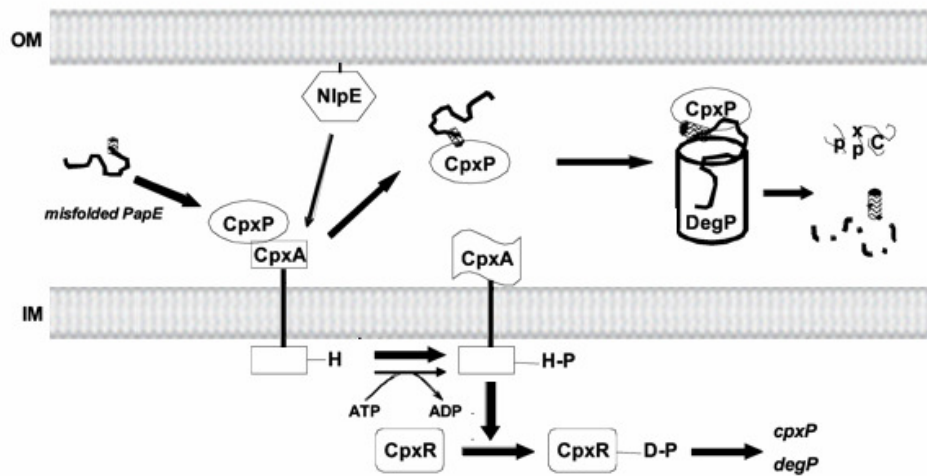


Fig. 1.6 CpxP regulation of CpxA/R system

CpxP is bound by misfolded protein (in this case PapE, a pilus subunit) which decouples it from CpxA making it susceptible to activation by the adhesion sensor NlpE. CpxA activation triggers phosphorylation of CpxR resulting in active CpxR-P which upregulates expression of *cpxP* and *degP* genes. DegP proteolytic function degrades the CpxP which is bound complexed with misfolded protein. Adapted from Isaac et al. (2005)

mechanisms which are not clear (Wolfe et al., 2008). However, crosstalk of EnvZ and CpxR in absence of CpxA has been observed (Siriyaporn and Goulian, 2008). CpxR~P has the ability to bind DNA and its main targets are the genes *degP*, *dsbA* (a periplasmic protein required for disulfide bond formation (Pogliano et al., 1997)) and *cpxP* (Wolfe et al., 2008). Genomic analyses of the conserved binding sites and experiments on mutants have shown Cpx involvement in positive regulation of the *rpoE-rseABC* operon and other genes of the σ^{32} regulon (De Wulf et al., 2002). The CpxA/R two-component regulatory network overlaps the σ^E regulon constituting a second stress response to protein misfolding in the periplasm (Raivio and Silhavy, 1999). The CpxA/R regulation is also involved in the regulation of the porins OmpC and OmpF. While σ^E decreases the expression of both porins, CpxA/R activation results in an increase in *ompC* and a strong decrease of *ompF* gene expression (Batchelor et al., 2005).

EnvZ-OmpR sense and regulate in response to osmolarity

EnvZ is a transmembrane protein which senses the environment and is capable of phosphorylation and dephosphorylation of the transcriptional regulator OmpR (outer membrane protein regulator) sensing osmolarity of the environment. (Cai and Inouye, 2002). Phosphate activated OmpR~P regulates the expression of the major OMPs or porins OmpC and OmpF involved in the passive diffusion of small molecules modulating sensitive to osmolarity. The levels of these porins are also regulated by EnvZ-OmpR upon pH variations (Heyde and Portalier, 1987) and nutrient limitation (Liu and Ferenci, 1998). While OmpC is expressed in high osmolarity, OmpF is mostly expressed in low osmolarity conditions (Alphen and Lugtenberg, 1977). The sensing mechanism of EnvZ remains unexplained (Kenney, 2010).

1.2.4 Protein misfolding

1.2.4.1 Fundamentals of protein folding *in vitro*

50 years ago, protein denaturation experiments were performed with β -mercaptoethanol and urea on bovine ribonuclease (Anfinsen et al., 1961). After restoration of solvent conditions, the polypeptide performed spontaneous folding which featured self-arrangement of intermediary states. These results contributed to formulate the hypothesis of thermodynamic protein folding where polypeptide sequence contains the information necessary for correct folding, which received a Nobel Prize Award in 1972 (Anfinsen, 1972; Anfinsen, 1973). *In vitro*, polypeptide chains acquire their conformation by a combination of two factors: the polypeptide sequence; and the physicochemical properties of the solvent. The three-dimensional structure is determined by the topology of the native protein in solution centering the discussions on the folding mechanisms in an interface between physics, chemistry, and biology (Baker, 2000; Shakhnovich, 2006). The amino acids of the polypeptide are thermodynamically driven by the solvent to turn and bend in a most energetically favorable state. These movements are rotations governed by two angles in an amino acid: phi (ϕ), which represents the torsion of the atomic bound between C_α and the amino terminus; and psi (ψ), which represents the torsion at the C_α carboxylic terminus. The angle limitations are based on physics of atomic interactions caused by each amino acid (Ramachandran et al., 1963). For example, the simplest amino acid Glycine can generate more possible angles than the cyclic imino acid Proline. The “allowed” angles for each amino acid are represented bi-dimensionally in degrees of ψ (ordinates) against ϕ (abscises), in a Ramachandran plot and it representing energy regions of every possible conformation for each amino acid. The low energy regions contain the most likely bond angles of the amino acid studied, contributing to

shape the protein structure. These plots are commonly utilized to appraise the inferred three-dimensional structure of a protein due to its simplicity and sensitivity to detect physical impossibility in suggested conformations (Kleywegt and Jones, 1996). Furthermore, this visualization technique has been improved by introducing new calculations of atomic distances and reviewing the assumptions regarding charge, hydrogen bonds and Van der Waals forces (Ho et al., 2003).

Even under the assumption that every amino acid acquires bond angles on the lower energy regions, the possible combinations are enormous (2^{100} for a small 100 amino acid peptide). Given the number of protein molecules in a sample, each could be in a different conformation. If the fraction of time to produce a right angle is as low as 10^{-11} seconds (there are faster rates known), the folding process would take 10^{52} years as a process of “trial and error” (Karplus, 1997). This is known as the Levinthal’s Paradox (Levintha, 1968) which suggested that each protein has an established folding pathway marking the starting point of modern protein folding theory. In the different models postulated, a common feature is the kinetic control or energy pathway (Karplus, 1997). With the exception of complicated folding chemistry (such as proline isomerization) different unfolded intermediates equilibrate between different conformations and free energy upon folding which is in agreement with experimental data (Creighton, 1988). During the different stages of the protein folding the free energy levels of the molecule are represented in a funnel-like graph from high energy, as a lineal polypeptide, to low energy, as a compact folded native protein (Dill and Chan, 1997). This theoretical graph is commonly known as the energy landscape of a protein [**Fig. 1.7**], where each peptide follows its own folding pathway quickly, satisfying Anfinsen’s thermodynamic hypothesis and Levinthal’s concerns. For example, a rough energy landscape has narrow pathways to correctly fold native protein involving kinetic traps and energy barriers where protein folding

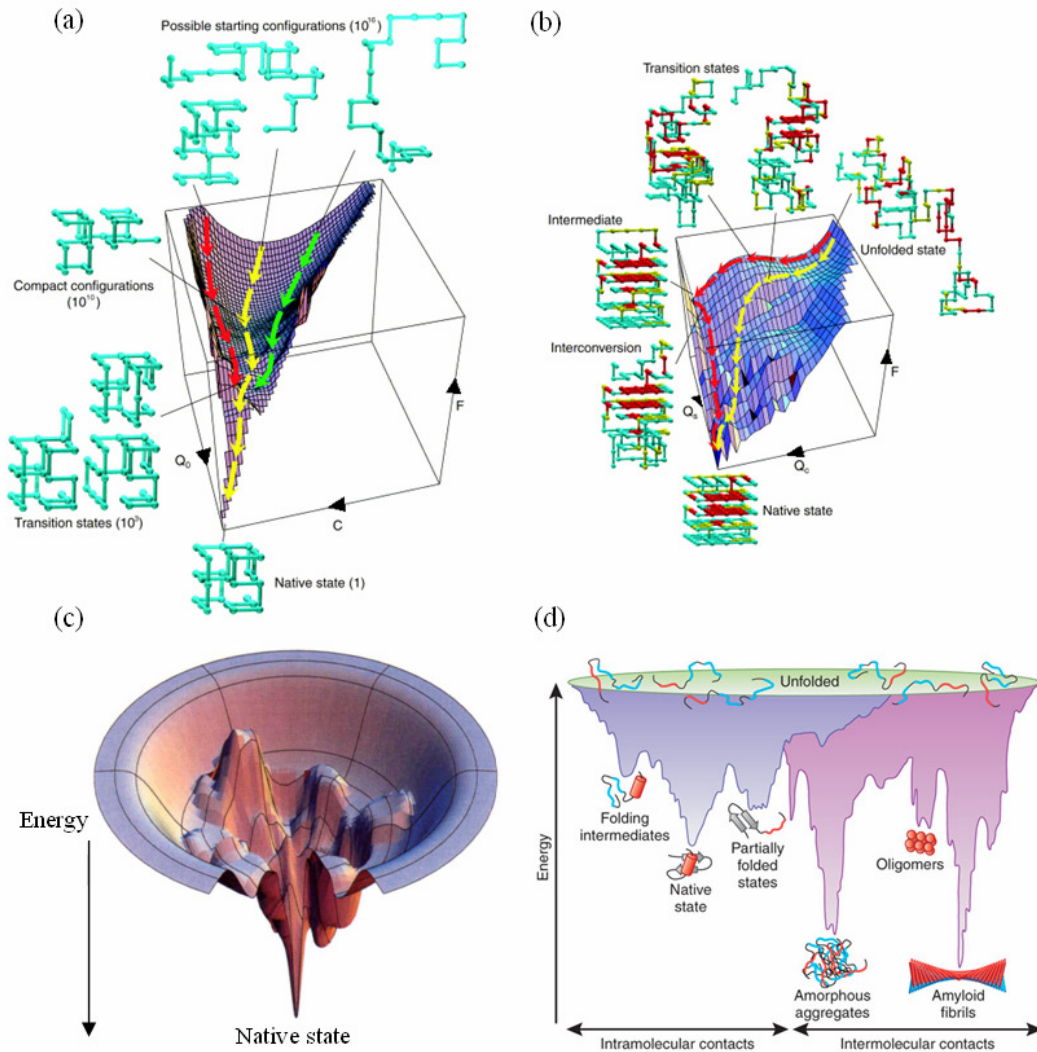


Fig. 1.7 Energetic landscape of spontaneous protein folding

The diagrams are diverse representations of the physics of protein folding. **(a)** Shows the simulation of spontaneous folding of a 20 element ideal protein. Q_0 represents shows the number of native state contacts while C represents any contacts established. **(b)** Is shows the simulation for a protein which needs specific contacts to form an active (Q_c) centre shown in red, and a set of surface contacts (Q_s). **(c)** shows the theory of energy funnels in which the numbers kinetic traps increase with the complexity of the protein, such as multimeric proteins **(d)** involving intermolecular contacts. Amorphous aggregates are trapped in stable energetic pits. **(a)** and **(b)** are adapted from (Dinner, et al., 2000), **(c)** is adapted from Dill and Chan (1997) and **(d)** from Hartl and Hayer-Hartl (2009)

generates multiple transitional intermediates (Dill and Chan, 1997). The two-state kinetics model defines each “state” (unfolded-folded) as the space containing the conformational possibilities in agreement with the observed appearance of intermediate forms (Zwanzig, 1997). The free energy of each intermediate form can be inferred by models and simulations based on the physics of the polypeptide in solution (Dinner et al., 2000).

1.2.4.2 Protein folding and misfolding in *E. coli*

Proteins with complex topology and long range interactions within the peptide increase the chances of falling in kinetic traps or metastable states (Brockwell and Radford, 2007). In contrast with the controlled environments *in vitro*, the *E. coli* cytoplasm is crowded with macromolecules, mainly proteins, nucleic acids, and complex sugars, up to 400g l⁻¹ occupying from 5% to 40% of the total volume (Ellis and Minton, 2003). While the physics principles governing protein folding *in vitro* are the same (Brockwell and Radford, 2007), the behavior of the unfolded peptide in a buffer is not a good estimate of its behavior *in vivo* (Ellis, 2007). The combination of these factors originates a tendency for aggregation of metastable unfolded proteins, exposing of hydrophobic regions, into stable amorphous structures called inclusion bodies (Hartl and Hayer-Hartl, 2009). It has been suggested that protein folding rates are not extensively subjected to natural selection. While overall structure and function depends on the amino acid sequence, which is subjected to evolutionary forces, the protein folding rates have shown being more dependent to the native three-dimensional topology which are subjected mostly to physicochemical forces (Alm and Baker, 1999; Baker, 2000; Kim, Gu, and Baker, 1998). Living organisms are equipped with diverse methods and strategies to enhance protein folding during and after its synthesis.

Protein synthesis

Proteins are synthesized by ribosomes utilizing mRNA as template in a well documented process called translation. The prokaryotic ribosome has a total mass of approximately 2.75×10^6 Da, generally expressed by its sedimentation coefficient 70S (Svedberg), between two subunits: 50S (34 proteins and ~3000 nucleotides rRNA); and 30S (21 proteins and 1500 nucleotides rRNA) (Steitz, 2008). The translation is initiated upon the recruitment of the 30s initiation complex formed by: the 30S ribosome subunit; the three initiation factors (IF1, IF2 and IF3); and GTP (**g**uanine **t**riphosphate). If the SD (Shine-Dalgarno) sequence is present in the untranslated region of the mRNA, it hybridizes with the rRNA of the 30S subunit positioning the start codon close to the P site. This step is blocked if the mRNA is folded, as observed previously in the cases of σ^S and σ^{32} mRNAs. The IF3 acts as proof-read avoiding incorrect Met-tRNA^{met} (transferent RNA charged with methionine) binding. When Met-tRNA^{met} positions over the start codon, IF3 is liberated and the 50S subunit is added to the complex with release of IF1, IF2 and GDP+P_i. (Myasnikov et al., 2009). The 70S ribosome polymerizes the peptide in series of cyclic reactions involving diverse elongation factors. The release factor (RF) terminates the process upon recognition of the termination codon. Folding may start co-translationally before the release of the peptide given that the 20-30 first amino acids of the sequence have left the ribosome. This type of folding was found to occur in prokaryotes spontaneously or mediated by chaperones acting on the nascent chain (Fedorov and Baldwin, 1997).

Inclusion Bodies

The appearance of inclusion bodies has been observed since the first attempts to express eukaryotic protein in *E. coli* (Marston, 1986). However, it was shown overproduction of natural *E. coli* proteins also led to protein aggregation (Botterman and Zabeau, 1985). Inclusion bodies are an amorphous mass of insoluble protein which generally acquires a spheroid shape and may appear in other cell types. In industry, inclusion bodies are easy to purify but the key step remains to achieve efficient refolding *in vitro* (Singh and Panda, 2005). It has been observed with phage subunit overexpression that the inclusion bodies can be very homogeneous displaying affinity for similar hydrophobic patches in similar kinetically trapped polypeptides (Speed, Wang, and King, 1996). This is thought to be analogous to the behavior of some protein misfolding related diseases such as amyloidoses and prion diseases (Fink, 1998). Overexpression of eukaryotic protein may lead to inclusion bodies due to tRNA depletion for specific aminoacids resulting in stalling of mRNA translation, complexity of the protein (multiple domains) or absent of disulfide bond formation (Fischer, Sumner, and Goodenough, 1993).

1.2.4.2.1 Proteins folding proteins: the chaperones of the heat shock response

The macrostructure of some proteins can achieve great complexity at tertiary and quaternary structures. Incorrect folding or assembly occurs in nature due to incorrect interactions giving rise to non-functional structures (Ellis, 1993). In a modern definition, molecular chaperones are a wide group of proteins with the common ability of folding and unfolding of noncovalent interactions and the assembly and disassembly of macromolecular structures without becoming permanently associated during the biological functions of such structures (Ellis,

2006). In *E. coli*, temperature increase and misfolded protein accumulation at low temperatures activates the master regulator of the heat-shock response, σ^{32} , which transcribes a network of chaperones and proteases (Yura and Nakahigashi, 1999). This network of heat-shock proteins monitors the quality of the proteins and controls the heat-shock response (Guisbert et al., 2004). Five classes of heat-shock proteins (Hsp) are conserved in all living organisms [**Table 1.6**] (Hoffmann and Rinas, 2004).

Hsp70 family: DnaK-DnaJ-GrpE

Under physiological heat-stress conditions DnaK is the most versatile chaperone and has shown to prevent aggregation of thermolabile proteins, consisting of up to 10-20% of detected proteins in cell extracts and *in vivo* (Mogk et al., 1999). Inclusion bodies were revealed to contain significant amounts of DnaK, a chaperone (Rinas et al., 2007). The structure of DnaK features two functional domains a polypeptide binding domain; and an ATPase domain [**Fig.1.8 (a)**] The peptide binding domain recognizes specific hydrophobic regions as demonstrated by (Rudiger et al., 1997) where DnaK was found to bind most efficiently to patches of 5 hydrophobic amino acids with a central Leucine, a motif which should be buried in aqueous solution. The ATP nucleotide modulates peptide binding release during the DnaK folding-holding cycle [**Fig 1.8 (b)**]. When DnaK binds ATP it switches to low affinity mode releasing the peptide in seconds, ADP bound DnaK switches to high affinity mode releasing the peptide in minutes, even hours. DnaJ catalyzes the fast γ -phosphate cleavage of the DnaK bound ATP (McCarty et al., 1995). The ATPase activity induced by DnaJ is increased by the presence of misfolded peptide bound to both DnaK and DnaJ. Thus, while DnaK-ATP selects and competes with DnaJ for its own substrate, dual binding of DnaJ and misfolded peptide to

DnaK triggers long term affinity (Han and Christen, 2003). This generates the dual functionality of DnaK as ATP dependent unfolding activity (foldase) and prevention of

Family	Name	kDa	Function	Co-chaperone	ATPase activity
Hsp100	ClpA	84/46	Folding, degradaion	ClpP	yes
	ClpX				
	ClpB	95/80	disaggregation		
Hsp90	HtpG	65.5*	Fold newly synthesised peptide		
Hsp70	DnaK	70	Folding, holding, dissagregation and regulation	DnaJ, GrpE	
Hsp60	GroEL	60	Folding	GroES	
sHsps	IbpA	~16 [#]	Inclusion body binding, Holding		
	IbpB				

Table 1.6 Heat shock proteins families conserved across kingdoms

The roles and molecular weights are characterized from *E. coli*. Adapted from Hoffmann and Rinas (2004). (*) denotes the experimental molecular weight and ([#]) denotes molecular weight estimated from nucleotide sequence (Keseler, et al., 2011).

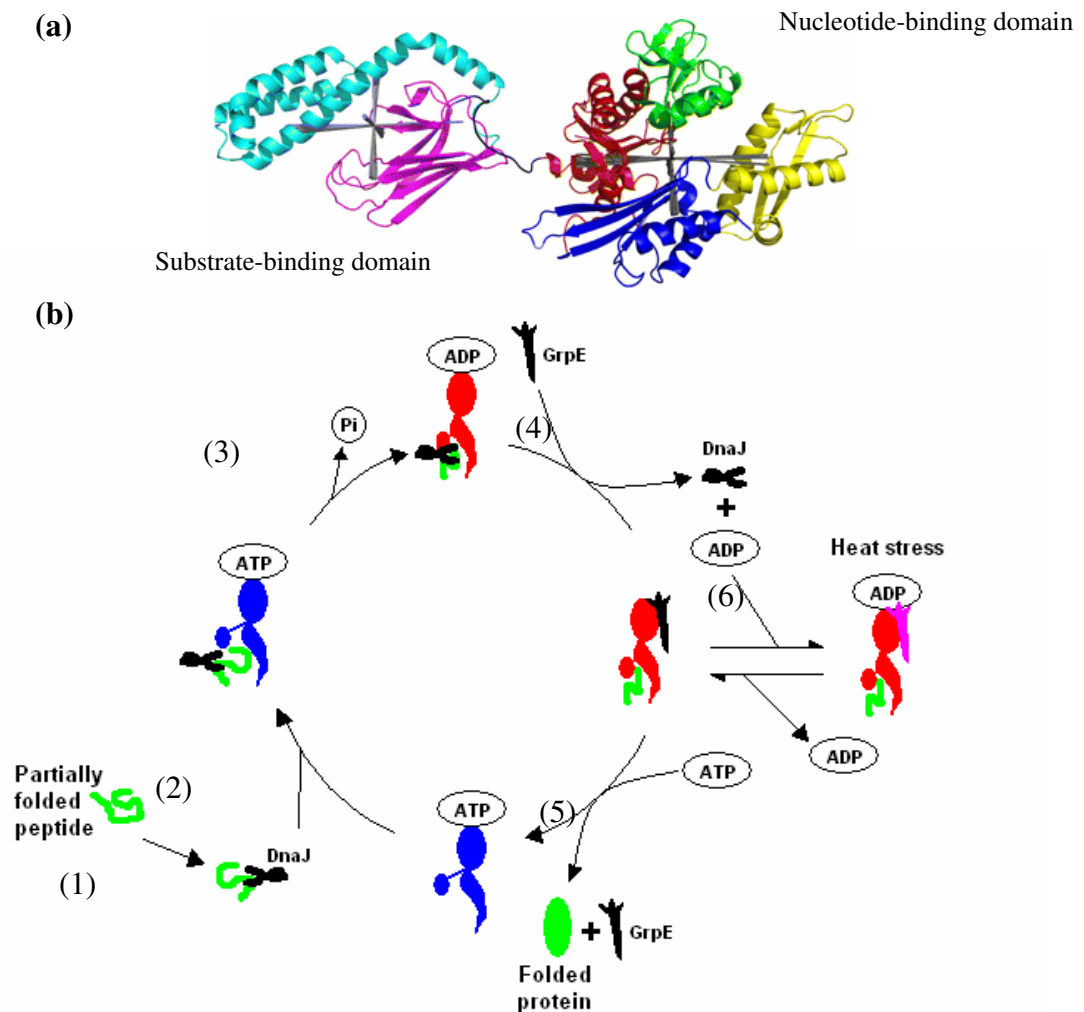


Fig.1.8 Structure and cycle of DnaK

The structure of DnaK is depicted in (a). The substrate-binding domains are shown in magenta (acting as lid) and in cyan. The nucleotide binding domains are shown in red, green, yellow and blue. The structure has been inferred with NMR. The molecule represented is in ADP bound form. Adapted from Bertelsen, et al. (2009). (b) Illustrates the DnaK cycle. (1) DnaJ co-chaperone binds partially folded peptide and transfers it to DnaK-ATP (2). DnaJ stimulates ATP hydrolysis (3) enabling high peptide affinity mode. DnaJ is released (4) and the co-chaperone GrpE binds DnaK-ADP mediating the nucleotide exchange (5) resulting on DnaK-ATP the low peptide affinity form thus the peptide, now folded, is released. Heat stress (6) may cause conformational changes on GrpE (shown in magenta) which slows the nucleotide transfer consequently slowing the cycle providing long term binding of misfolded protein. Figure based on Winter and Jakob (2004).

aggregation by long term binding (holdase) (Slepenkov and Witt, 2002). GrpE binds DnaK and catalyzes the exchange of ADP for ATP, which activates low affinity function. Heat stress partially unfolds GrpE which slows down the ADP/ATP exchange rate from DnaK resulting in long term activation of high affinity mode (Siegenthaler et al., 2004). Long term binding to DnaK targets unfolded proteins to degradation proteases (Sherman and Goldberg, 1992). It has also been shown that DnaK solubilizes aggregates *in vitro* (Diamant et al., 2000). One of the major functions of DnaK is the regulation of the heat-shock response through σ^{32} mediated by DnaJ. In the absence of stress, σ^{32} is targeted by DnaJ and bound by DnaK in a reversible manner by its polypeptide binding domain, thus protecting it from degradation although preventing it from binding RNAP. In case of a rise of misfolded protein occurrence the σ^{32} factor is titrated out from DnaK which triggers a heat shock response. The relative shortage of free DnaK is cause for high sensitivity of the heat-shock response to protein misfolding (Tomoyasu et al., 1998). Recent data support this model showing that DnaJ and DnaK recognize and bind σ^{32} in different sites at the N-terminal region. This supports the mechanism that co-bound σ^{32} to DnaJ and DnaK activates ATPase activity and holdase activity, capturing σ^{32} and down-regulating heat-shock response genes (Rodriguez et al., 2008).

Hsp60: GroEL-GroES

The *E. coli* GroE (aggregation of λ E) was first identified by mutations which interfere with the assembly of viral protein subunits (Georgopoulos et al., 1972; Takano and Kakefuda, 1972). This operon is regulated by σ^{32} as part of the heat-shock response and is composed by two genes *groES* (small) and *groEL* (large), both essential for cell viability, with translation products that assemble in a highly conserved structure (Hemmingsen et al., 1988). X-ray

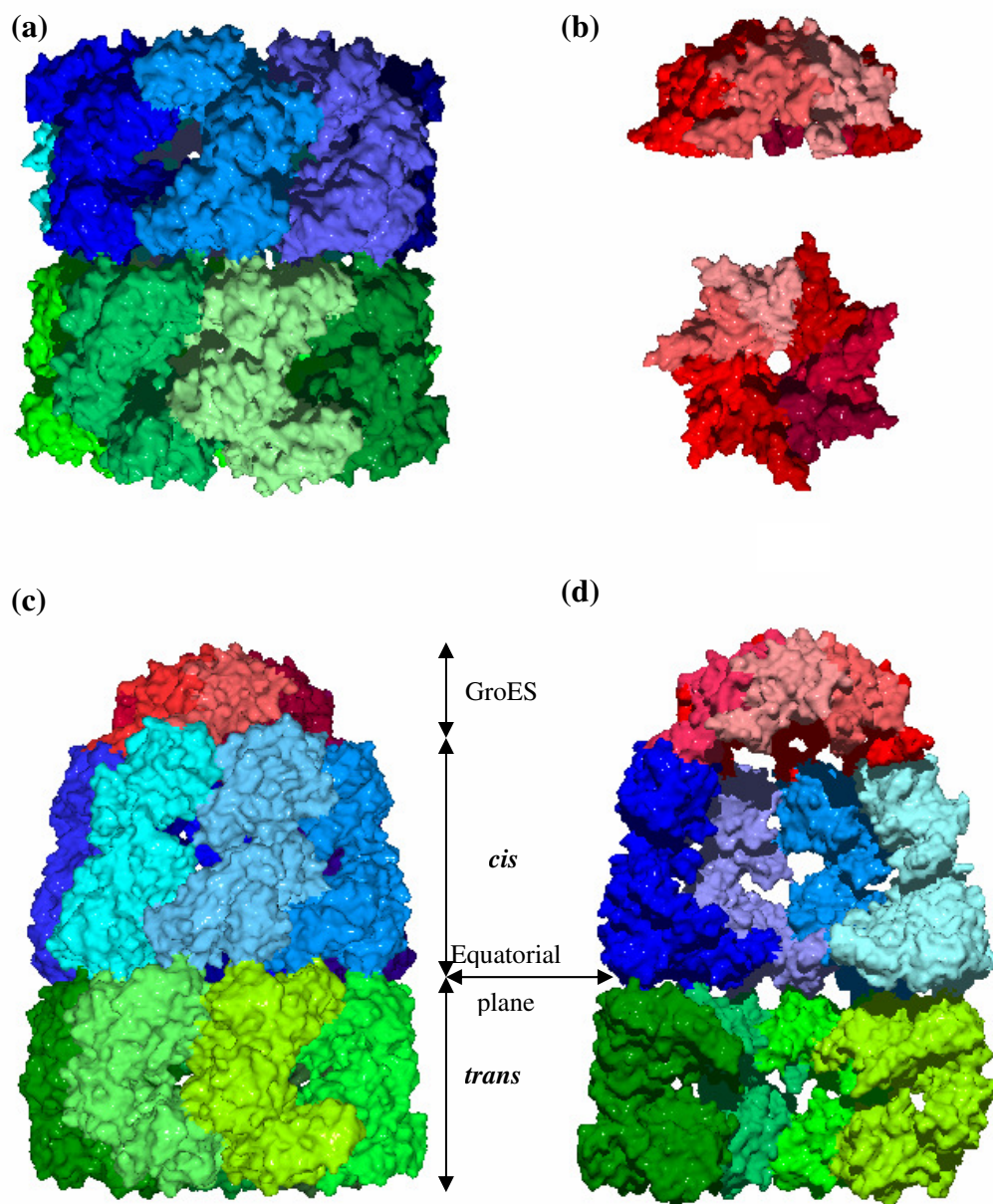


Fig. 1.9 Molecular structures of GroEL and GroES

The subunit two stacked heptameric rings of GroEL are shown in blue and green (a). This conformation binds ATP which enabled the attachment of GroES (b) depicted in red. (c) Attachment of GroES to an ATP bound ring provokes a conformational change *trans* to *cis* which result in a expanded cavity where refolding takes place. (d) three of the GroEL subunits of each ring have been removed to show the cavities of *trans* and *cis* conformations. Surfaces were calculated with DeepView (Guex and Peitsch, 1997) and rendered with POV-ray from the PDB structures: (a) 1KP8 deposited by Braig et al. (1994) and; 1AON, (b) (c) and (d) deposited by Xu et al.,(1997).

crystallography of both proteins revealed a macrostructure [Fig 1.9 (a)] featuring two substructures: two stacked heptameral rings consisting of 14 identical subunits of GroEL (58KDa) which forms two cylindrical chambers of 45Å, presenting weak hydrophobic surfaces, separated by an equatorial plane; and one heptameral ring of 7 GroES (10KDa) subunits (Xu, Horwich, and Sigler, 1997), which are known to complex in the presence of ATP and Mg²⁺ as a lid (Chandrasekhar et al., 1986) [Fig 1.9 (b)]. This association produces structural changes [Fig 1.9 (c)] in the surface of GroEL-ES cavity from hydrophilic to hydrophobic and generates a space of 80Å of diameter and 85Å in height doubling the initial volume before association (Xu et al., 1997) [Fig 1.9 (d)]. The preferred substrates of GroEL are proteins with two or more domains consisting of α -helix and buried β -sheets which display numerous hydrophobic surfaces (Houry et al., 1999). Isolation of GroEL-ES-substrate complexes have shown that of approximately 250 proteins interact with GroEL, ~85 of them are chaperone dependant utilizing up to 75% of the GroEL available, which includes 13 essential proteins (Kerner et al., 2005). The current representation of GroEL/ES mechanism is represented in [Fig 1.10 (a)]. GroEL binds ATP molecules on one of the heptameral rings which enables capture and polypeptide unfolding and GroES binding in this order (Tyagi, Fenton, and Horwich, 2009). This generates a conformational change (*cis*) which creates a hydrophilic environment and confines the peptide on the inner cavity where it refolds in a confined environment (Apetri and Horwich, 2008). Then, the ATP is hydrolyzed allowing a new round of ATP to bind the second ring. Recent studies on ADP release showed that ATP binding to the second ring of GroEL (in *trans* conformation) alters the conformation of the *cis* ring causing ejection of ADP, GroES and refolded peptide (Tyagi, Fenton, and Horwich, 2010), restarting the cycle with the second ring of GroEL.

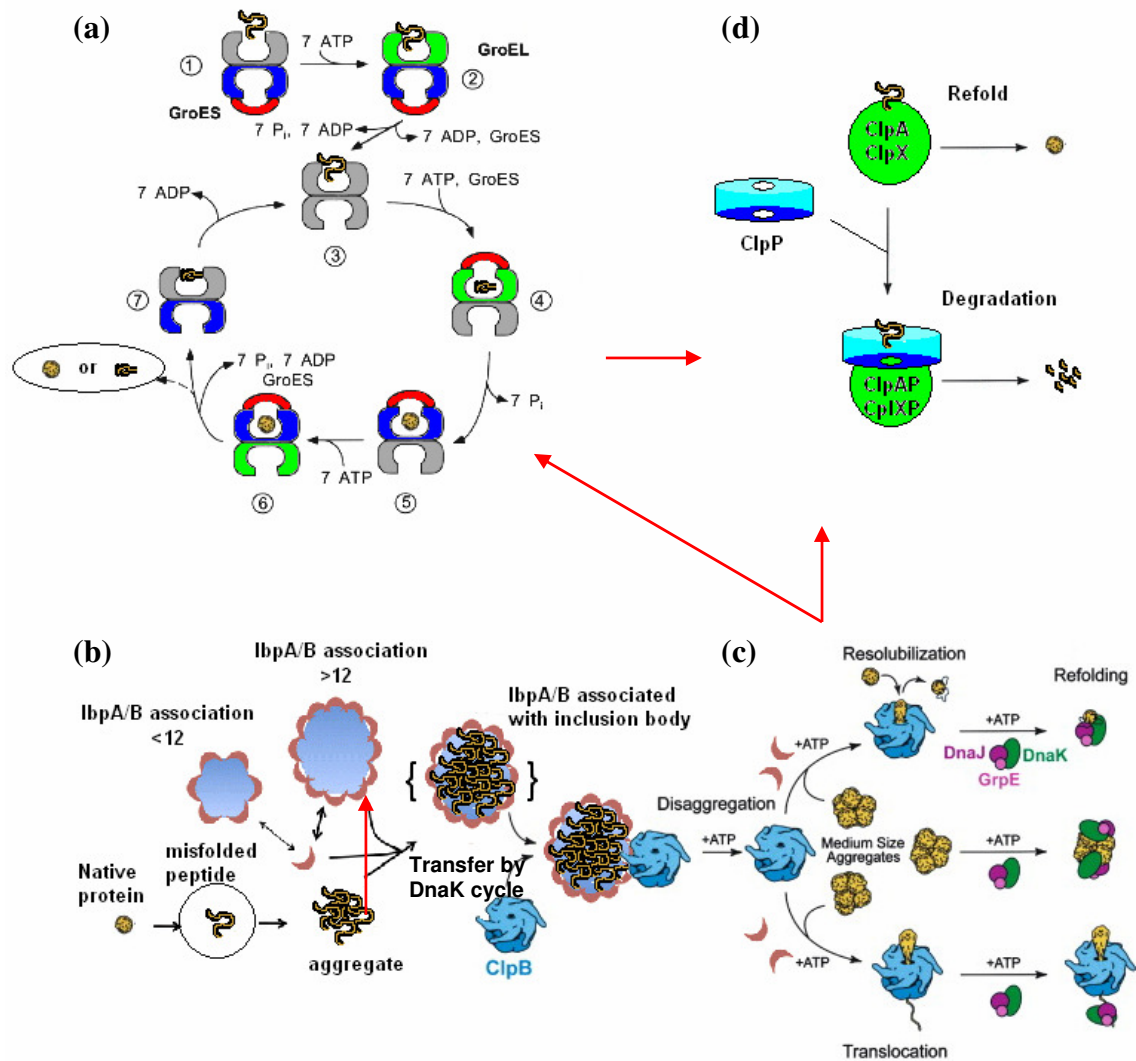


Fig. 1.10 Integration of the heat shock proteins on a chaperone network

Images are arranged anticlockwise. Diagram (a) is a representation of the GroEL/ES system. (1) peptide binds unoccupied ring. (2) ATP binds to the *trans*-ring represented in green and (3) the ATP hydrolysis releases GroES and ADP from the cis-ring. (4) ATP and GroES bind the peptide containing ring. (5) ATP hydrolysis results on ADP bound ring, represented in blue, and a refolding step on the peptide. (6) ATP binds the ring with free nucleotide slots, short lived intermediates which are represented in grey. (7) Hydrolysis in this ring causes a conformational change in the equatorial plane which provokes the ejection of the folded peptide and GroES. If the peptide is not folded it quickly rebinds GroEL restarting the cycle. Adapted from (Martin, 1998) (b) depicts the functioning of the small heat shock proteins associated with the inclusion bodies. Adapted from (Eyles and Gierasch, 2010) (c) shows the action of ATP dependant ClpB on inclusion bodies and its chaperoned activity. It also integrates refolding with the DnaK cycle adapted from (Lee, et al., 2003). (d) is a confluence point on the ClpA and ClpX. Peptides may be refolded by these chaperones but in case of unviable refolding ClpP binds de complex resulting in degradation of the misfolded peptide. Based on (Wickner, et al., 1994) Red arrows show transference of client peptide amongst chaperone systems establishing a network.

sHsps: IbpA, IbpB

These small heat-shock proteins were first found associated with inclusion bodies produced during recombinant protein production and are encoded by the genes *ibpA* and *ipbB* (inclusion body-associate protein). These genes are found contiguously in the chromosome separated by 110bp; they are transcribed together by $E\sigma^{32}$ upon heat shock and translated into proteins by independent SD sequences and terminators (Allen et al., 1992). IbpA and IbpB were found to be associated to the outer membrane based on numerous experiments of cell fractionation; however, upon heat shock they could be found associated to the inclusion body fraction (Laskowska et al., 1996). Overexpression of *ibpAB* resulted in a super-resistant phenotype to oxidative and thermal stress based on survival in extreme conditions (Kitagawa, Matsumura, and Tsuchido, 2000). $\Delta ibpAB$ mutants displayed a weak phenotype based on survival tests when subjected to extreme heat shock temperatures (50°C). However, survival in mutants during mild heat shock was not significantly different from the survival in wild type showing that they are not essential for disaggregation of protein (Kuczynska-Wisnik et al., 2002).

sHsps are conserved in all kingdoms of life and have been shown to interact with each other and with the network of chaperones and proteases upon a rise in protein aggregation [Fig 1.10 (b)] (Matuszewska et al., 2005). Their ability to protect against thermal and oxidative stress has been documented (Matuszewska et al., 2008) and it has been hypothesized that IbpA and IbpB may interact with the lipids of the outer membrane collaborating to reduce its fluidity at high temperature, partially based on experiments on sHsps of lactic acid bacteria (Coucheney et al., 2005; Nakamoto and Vigh, 2007).

Hsp100: Clp proteases

The members of this family of heat shock proteins function as a quality control of the proteins translated in the cytoplasm. These proteases may recognize specific degradation tags of exposure of hydrophobic surfaces (Wickner et al., 1999). A possibility is that the molecular mechanism for substrate selection is interaction with PDZ-like domains (Levchenko et al., 1997). ClpA and ClpX form double hexameric rings and may unfold proteins in an ATP dependant manner. They interact with ClpP forming the protease complexes ClpAP and ClpXP, which rapidly degrade proteins [**Fig 1.10 (c)**] (Wickner et al., 1994). This has been observed on a phage protein, while the non-functional dimmers could be resolved by ClpA/X but would be degraded by ClpAP/XP, thus preventing aggregation (Wickner et al., 1994). It has been shown that ClpP recognizes σ^S and present it to ClpX for degradation being a major regulator of σ^S levels during growth (Schweder et al., 1996). If aggregation should occur, other member of this protease family, ClpB, may contribute to resolving inclusion bodies. Upon observation of its crystal structure it has been proposed that ClpB may mechanically disaggregate high molecular weigh inclusion bodies [**Fig 1.10 (d)**] (Lee et al., 2003). ClpB is known to contribute with DnaK in refolding peptides extracted from aggregates (Doyle and Wickner, 2009).

1.3 Gene Reporter Technology

1.3.1 Fundamentals

A gene reporter system consists of an artificial arrangement of DNA which results in the production of a specific protein in response to determined biochemical factors. This protein

should be measurable by its activity or its quantity, serving as gauge of those physiological factors. Gene reporter technology is the application of reporter systems to measure biological and biochemical parameters within living organisms. There are three main aspects to take into account: the biology of the process to be measured; the choice of reporter protein; and the vector which delivers it into the host. The biochemistry of the stimuli or response to be studied is central in relation to reporter design. The reporter protein to be produced is important in relation to methods and techniques available for monitoring its presence. The vector which carries the system is crucial to proper delivery and performance inside the host. It has been more than 40 years since experimental work involving gene reporter technology started (Casadaban and Cohen, 1980). This original reporter system was carried by a plasmid containing a promoter fused to the *lacZ* reporter gene which encodes β -galactosidase. The effects of the diverse genetic regulatory mechanisms on the promoter were reported by alterations of the reporter gene expression which could be assayed by the enzymatic activity of β -galactosidase (Casadaban and Cohen, 1980). Initially, the promoters were fused to enzymes in order to quantify expression by biochemical assay. The development of this technology has been accelerated by the rise of molecular biology technologies, the development of protein engineering and the implementation of high throughput methods. Experimentation with different reporter proteins in the last years has resulted in an explosion of applications of this technology. Through the years, new reporters based on luminescent or fluorescent proteins have been discovered and perfected for multiple applications.

A reporter system can be designed to fulfill specific goals such as highlight occurrence, location or magnitude of a biochemical process. However, the designed system can be utilized in other experiments to extract different information from its original purpose. Therefore a construct design may generate many applications exploitable through different experimental

approaches. Knowledge of their limitations makes them reusable. Commonly, a reporter system is built on a plasmid featuring an origin of replication, an antibiotic selectable marker and a promoter::reporter gene fusion. The gene reporter system may be integrated in the host genome as done by (Morschhäuser, Michel, and Hacker, 1998) for observing gene expression in *Candida albicans*. However the basis of its mechanism is very similar to those functioning on plasmids. On the other hand its utilization in plasmids independent from the genome provides other benefits due to its facility to introduce in different strains.

Elements in a reporter system

Plasmids are molecules of circular double helix DNA which can replicate autonomously separated from the host chromosome. Plasmids used as vectors to carry reporter systems are called reporter plasmids. The general elements of a reporter plasmid [Fig 1.11] are the origin of replication, the antibiotic resistance, a promoter and a reporter gene. The variation of these elements generates diversity of uses and applications of this technology. The origin of replication and an antibiotic selectable marker are ubiquitous in most plasmid chassis. The origin of replication encodes the instructions to regulate the amount of plasmid copies present in the host (Weisblum et al., 1979). Once the plasmid is inside the host the origin instructs the self-replication to an average number of copies per cell (del Solar et al., 1998). This regulation of the origin of replication may affect the ability of two plasmids of coexisting in the same host (Novick, 1987). Plasmids can be classified by their average copy number [Table 1.7]. Antibiotic resistance is important as selectable marker because it may influence plasmid retention during culture stages.

The promoter is located upstream the reporter gene [Fig 1.11]. Its function is to recruit the

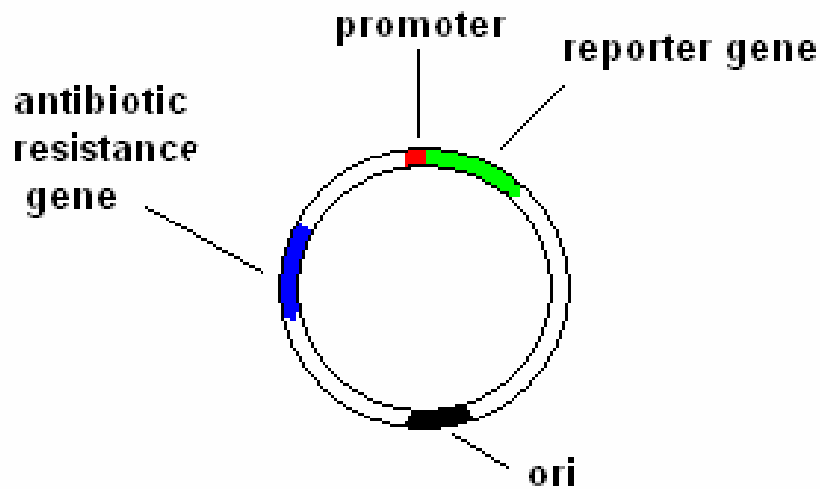


Fig. 1.11 Basic diagram of a reporter plasmid

The diagram shows the basic elements which remain common to most reporter systems. New design may require variation of the promoter, to control expression of the reporter gene. The reporter gene may vary in the design accordingly with the experiment or method of detection. The antibiotic resistance gene can be selected appropriately to the type of host organism or culture conditions. The origin of replication (ori) is the determinant of the number of copies in each cell. Excessive copies may create cytoplasmic burden, diminished number of copies could weaken detection.

Plasmid	Origin of replication	Copy number (copies per cell)	Classification
pUC vectors	pMB1*	500-700	High copy
pBluescript vectors	ColE1	300-500	High copy
pGEM vectors	pMB1*	300-400	High copy
pTZ vectors	pMB1*	>1000	High copy
pBR322	pMB1*	20-30	Low copy
pQE vectors	ColE1	~30	Low copy
pREP4 vectors	p15A	~30	Low copy
pACYC and derivatives	p15A	10-12	Low copy
pSC101 and derivatives	pSC101	~5	Very low copy

Table 1.7 Origin of replication and copy number of various plasmids

(*) The difference in copy number is achieved by mutagenesis of the same origin sequence, closely related to ColE1 Adapted from (Sambrook J., et al., 1989)

factors necessary for transcription, thus gene expression. The specific regulation of the promoter region dictates how much of the reporter protein is produced. In some instances this serves for studying the regulation of unknown promoters in the host, for example (Zaslaver et al., 2006) generated a library of promoters from ORFs of *E. coli* K-12 some of which were previously characterized but some unknown which are now available to be tested. The promoter region of the reporter may have its activity driven by the presence and strength of transcription factor binding sites or contained by the presence of repressors.

The reporter protein is a translation product encoded by the reporter gene. This protein should possess a unique activity or quality easily discernible from any other present in the host cell. It must be able to produce a signal quantifiable with the means available. The strategy used to measure or visualize its presence depends on the biochemistry of the reporter protein.

The reporter proteins: β -galactosidase reporters

There are many types of reporter proteins [Table 1.8]. Metabolic enzymes can be utilized as reporters by addition of a substrate analog to the metabolite they act upon in order to perform a measurable biochemical reaction. Utilization of chloramphenicol acetyltransferase (CAT, *cat* gene) or β -glucuronidase (*uidA*) (Jefferson, Burgess, and Hirsh, 1986) as reporter genes is very common on mammalian cells because β -galactosidase activity is endogenous. The most common metabolic enzymatic reporter in microbiology is the β -galactosidase, encoded by *lacZ* gene.

Reporter protein	Substrate addition	Measurement technique	References
Metabolic enzymes based reporters			
β -galactosidase	ONPG	Spectrophotometry	(Casadaban and Cohen, 1980)
	fluorescein-di-beta-D-galactopyranoside	Flow cytometry	(Plovins et al., 1994)
	Galacton-Plus TM	Chemoluminescence	(Bronstein, et al., 1996)
CAT	acetyl CoA	Radiometric method	(Naylor, 1999)
β -glucuronidase	p-nitrophenyl b-D-glucuronide	Spectrophotometry	(Jefferson et al., 1986)
	Glucuron TM	Chemoluminescence	(Bronstein et al., 1996)
Luminiscence based reporters			
<i>luc</i> gene	ATP + D-Luciferine ATP + benzothiazoyl-thiazole	Luminometer	(Hakkila et al., 2002)
<i>lux</i> gene	Not needed (<i>lux</i> operon includes the substrate)	Luminometer	(Hakkila et al., 2002)
Fluorescence based repoters			
Flavin-mononucleotide binding protein	Not needed (the substrate FMN is present in host cells)	Fluorimetry	(Drepper et al., 2010)
GFP	Not needed	Fluorimetry	(Cormack et al., 1996)

Table 1.8 Summary of some of most common reporter proteins

(TM) indicates registered Trade Mark substrate. These reporter proteins are reviewed individually in the text.

β -galactosidase was the first protein to perform as a reporter (Casadaban and Cohen, 1980). The biological reaction catalyzed by β -galactosidase is the cleavage of the disaccharide lactose to glucose and galactose. It has been expressed successfully as a recombinant protein in prokaryotes and eukaryotes. The attributes that make this enzyme a good reporter could be summarized by its kinetics, stability and assay possibilities. Microbial β -galactosidase is stable at higher pH and temperature than eukaryotic. This enables selective inactivation of host galactosidases when this reporter is being used in a eukaryotic host with high reaction background (Pocsi et al., 1993). On the other hand bacterial and fungal β -galactosidase has high affinity for its substrate reflected by a low Michaelis-Menten constant (K_m) of 13mM (Lederberg, 1950). The substrate used to quantify β -galactosidase activity is O-nitrophenyl-galactopyranoside (ONPG), cleaved by LacZ to produce yellow coloured O-nitrophenol. During the β -galactosidase assay the cells are lysed and their contents exposed to the reaction buffer of the biochemical test. The reaction can be followed by spectrophotometry at the wavelength of 420nm; measuring over time enables activity quantification. ONPG is also an established reactive for diagnostic bacteriology as substrate in qualitative test to determine β -galactosidase presence (Pickett and Goodman, 1966). There are other ways to assay β -galactosidase activity, such as by chemoluminescence and flow cytometry as shown in [Table 1.8].

Addition of further elements to the fusion promoter::*lacZ* enabled the development of useful markers for molecular biology. The α -peptide is a specific region encoded in the *lacZ* gene which is vital for the enzyme activity (Langley et al., 1975). DNA insertion within the sequence of *lacZ* gene displaces the α -peptide in the sequence which results in an inactive product. The lack of β -galactosidase activity reports for successful DNA insertion (Ruther, 1980). This is the basis for X-gal (5-bromo-4-chloro-3-indolyl β -D-galactoside) blue/white

screening. β -galactosidase transforms X-gal into a blue colored compound. Growing on plates with agar containing X-gal, blue colonies have a intact β -galactosidase gene thus no DNA insertion, successful cloning is found in white colonies.

The application and discovery of new reporter proteins coevolved with the development and implementation of high-throughput screening strategies. Bioluminescent or fluorescent properties can be measured on a more direct manner; ideally this should be on a non-invasive manner as defined by van der Meer and Belkin (2010). The discovery and application of bioluminescent and fluorescent reporters are of great significance to the gene reporter technology.

Bioluminescent reporters

Bioluminescence consists of the release of energy in the form of light through enzymatic catalysis. The enzyme responsible for bioluminescence is called luciferase which reacts with a specific substrate called luciferine producing photonic emission. There are several organisms capable of producing such proteins [Table 1.9]. Some unicellular algae, bacteria, coelenterates, beetles and fish have this ability. There are several types of luciferases known by their abbreviation: *lux* from bacteria, *luc* from fireflies and *lfc* from dinoflagellates (Wilson and Hastings, 1998). They are not structurally homologous and present differences in their respective luminescence mechanisms. Interestingly these luciferin-luciferase systems evolved independently from more than 30 different origins based on phylogeny studies (Hastings, 1983). In contrast with *lfc*, the luciferases from bacteria (Lux) and firefly (Luc) have been successfully implemented as reporters. The bioluminescence in bacteria is regulated genetically by the operon *luxCDABE*. LuxAB proteins constitute the subunits α and β of the

Origin	Luciferine	Luciferase	Requirements
Insect	Luc	benzothiazoyl-thiazole (a)	ATP as co-substrate
Bacteria	Lux	long chain aldehyde (b)	FMNH ₂ and molecular O ₂
Coelenterates	Aequorine	Coelenterazine (c)	Ca ²⁺

Table 1.9 Bioluminescence systems discussed based on luciferine/luciferase reactions

(a) Complex molecule that needs to be added the reaction consumes ATP by condensing it into the luciferine. (b) synthesis encoded in the *luxCDE* genes (c) This molecule is ubiquitous in coelenterates. The product, coelenteramide, is regenerated by the organism.

luciferase. The proteins LuxC, D and E form an enzymatic complex which synthesizes a long chain aldehyde to serve as luciferine. The bioluminescence is then produced by the mixed oxidation of reduced flavin mononucleotide and the long chain aldehyde (Meighen, 1988). The LuxAB proteins were utilized as on-line monitoring reporters in bacteria by (Marincs, 2000) where the production of bioluminescence was correlated with living cells. This method was suggested as an on-line tool to monitor living *E. coli* biomass without the limitation of optical methods such as turbidimetry. Bacterial luciferase has been utilized in eukaryotes however its use presented limitations. The reason was that fusions of LuxAB, harnessed to operate under a single promoter, did not fold correctly at temperatures higher than 30°C (Escher et al., 1989). The Luc protein from firefly was used in eukaryotic hosts as a reporter to observe expression patterns. The main drawback is a requirement for benzothiazoyl-thiazole to be added as luciferin and ATP into the luciferase assays which can be expensive.

The coelenterates also produce luciferases. These luciferases are found in *Aequorea spp.* and *Renilla spp.* which utilize coelenterazine as luciferine [Fig 1.12]. The luciferase from *Aequorea spp.*, aequorin, produces light of blue spectrum upon combination with coelenterazine in presence of calcium (Ward and Cormier, 1975). This luciferase is different because it does not require cofactors like FMNH₂ in LuxAB or ATP consumption like Luc. X-ray crystallography experiments (Head et al., 2000) shown that the coelenterazine is found stabilized by hydrogen bonds inside the aequorin as chromophore. As exposed in [Fig. 1.12] when two Ca²⁺ ions bind to the aequorin the protein conformational change releases CO₂, coelenteramide and blue light. This altered conformation of aequorin is called apoaequorin which can reform into aequorin in presence of oxygen and coelenterazine. The cDNA of apoaequorin has been cloned (Inouye et al., 1985) and became a useful indicator of

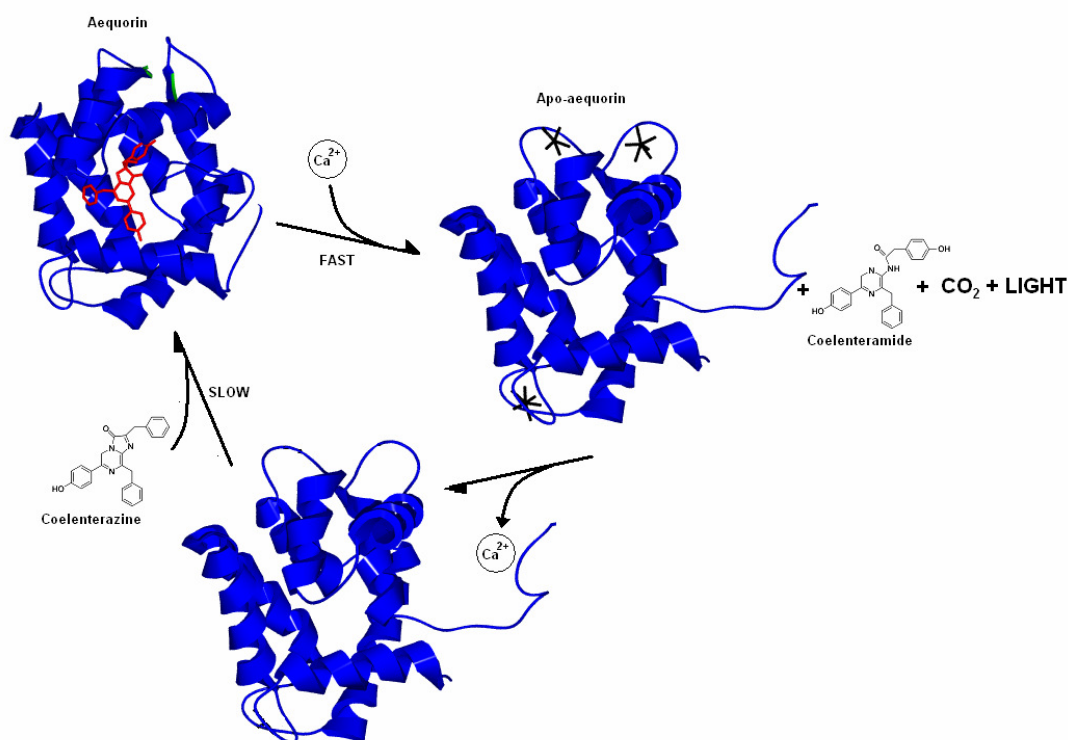


Fig. 1.12 Cycle of the aequorin reaction

Aequorin containing coelenterazine, in red, reacts with calcium (black stars) to produce coelenteramide, CO_2 and light while changing conformation. Calcium release yields apo-aequorine which may rebind coelenterazine restarting the cycle. Based on Shimomura (2005). Images were processed with DeepView (Guex and Peitsch, 1997) and rendered with POV-ray from the PDB structures: aequorine, 1EJ3 (Head, et al., 2000); and apo-aequorine, 1SL8 (Deng, et al., 2005).

intracellular calcium, for example in yeast cells as shown by (Nakajima-Shimada et al., 1991). Other luciferases from coelenterates such as *Renilla spp.* have been used successfully as reporter proteins in mammalian cells (Lorenz et al., 1996) and in plants (Mayerhofer et al., 1995). These luciferases were discovered complexed with a second emitter or protein antenna. The second emitter or protein antenna was named 'green protein' and displayed very interesting characteristics (Shimomura, 2005).

1.3.2 Green Fluorescent Protein (GFP) as reporter

Discovery

'Green proteins' were first discovered by Shimomura et al. (1962) in *Aequorea spp.* forming a complex with Aequorine. Hastings and Morin (1969) utilized the term 'green fluorescent protein' (GFP) to refer to its ability to produce fluorescence. Morise et al. (1974) postulated later that photonic energy released from the aequorine reaction may be transferred by protein-protein interaction to the second emitter which converts it into green fluorescence. When purified aequorine was adsorbed onto a matrix, it generated blue flashes of light upon Ca^{2+} addition. However, when GFP was co-adsorbed the addition resulted in green fluorescence. This mechanism was accepted to occur *in vitro* as *in vivo* and GFP was acknowledged as ultimate emitter in *Aequorea spp.* In addition, the GFP was purified and its absorption and emission properties characterized. Shimomura (1979) proposed the structure of a light emitting element analogous to a luciferine as first suggested by results of studying papain digests of GFP. Improvements in molecular biology allowed cloning the DNA encoding the 238 amino-acids of GFP (Prasher et al., 1992), accelerating the investigations of GFP biochemistry. Chalfie et al. (1994) successfully expressed functional GFP in other organisms

such *E. coli* and *Caenorhabditis elegans* demonstrating that it was possible to visualize the protein by illuminating the organism with an UV source, without addition of cofactors or other proteins. The extraordinary structure of GFP holds the key to its properties. The crystal structure of GFP [Fig. 1.13 (a)] was described by Ormo et al. (1996) as a β -barrel formed by 11 strands resembling a Greek amphora with the chromophore in a central helix shielded from the solvent. The study of GFP folding dynamics and chromophore maturation was greatly assisted by the production and study of new mutant variants of GFP, ultimately unlocking its immense capabilities as a reporter protein.

The chromophore

Expression of functional GFP in many organisms initially suggested that its biosynthesis occurs spontaneously and upon reaction with ubiquitous enzymes and cofactors. Cody et al. (1993) resolved the structure of the chromophore [Fig. 1.13(b)] which consisted in the cyclization and oxidation of the residues Serine⁶⁵-dehydrotyrosine⁶⁶-glycine⁶⁷. This group of residues can be excised from GFP when digested with papain. Heim et al. (1994) postulated that the formation of the chromophore was a spontaneous phenomenon where correct protein folding and correct chromophore maturation were showed to be co-dependent for functional GFP biosynthesis. *E. coli* produced a portion of misfolded GFP as inclusion bodies which did not possess a correctly folded chromophore whereas soluble GFP was functional. While GFP produced in anaerobic conditions did not fluoresce, fluorescence could be recovered after admission of air (Heim et al., 1994). GFP maturation took up to 4 hours long in an oxygen dependent system.

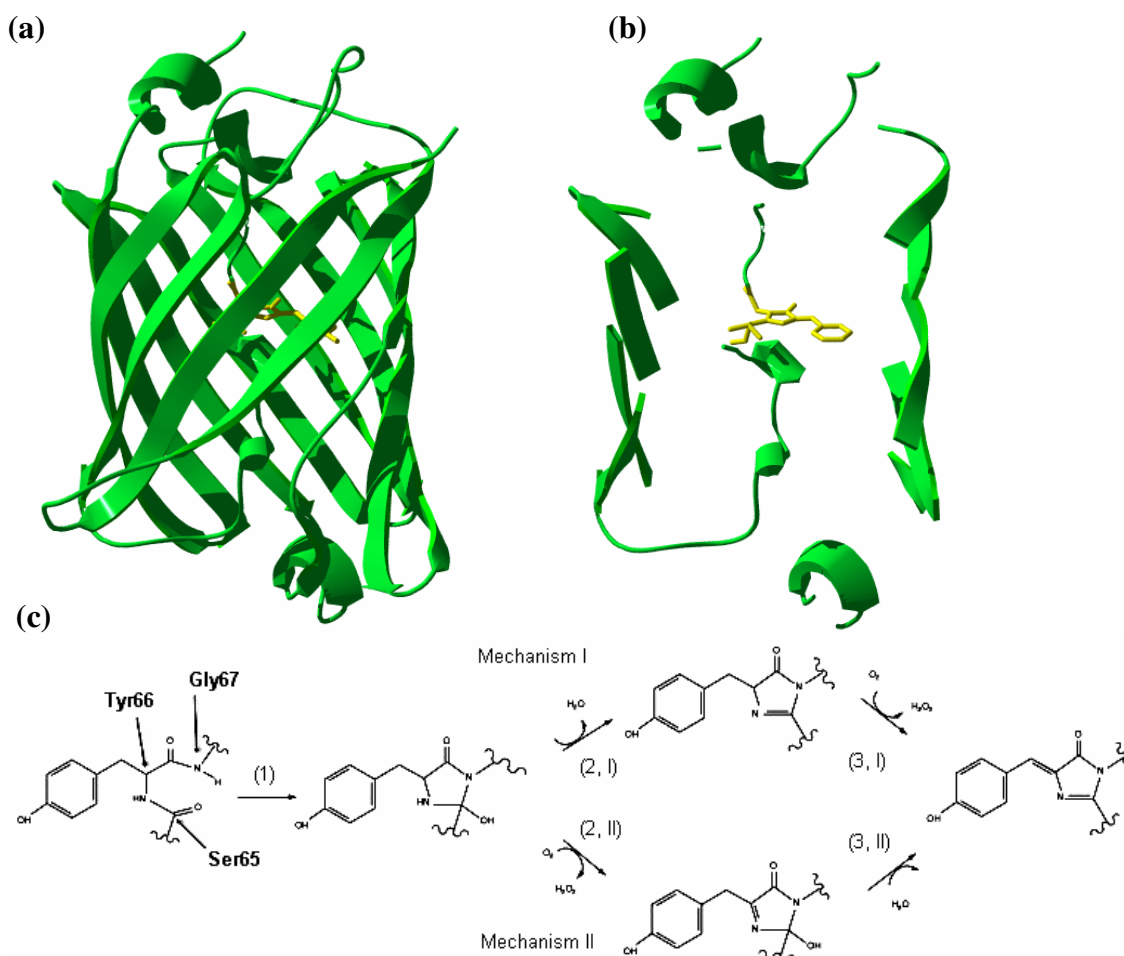


Fig. 1.13 Crystal structure of GFP and chromophore maturation

Image (a) Shows the structure obtained by X-Ray diffraction of GFP from *Aequorea victoria*. the α -helix contains the chromophore, in yellow, and it is encased by the β -strands. Image (b) is a slab of the protein containing the chromophore, highlighted in yellow and the surrounding β -strands. Both images were processed with DeepView (Guex and Peitsch, 1997) and rendered with POV-ray from the PDB structure: 1EMA by Heim, et al., (1995). Diagram (c) shows the steps of chromophore maturation during GFP folding a result of the cyclization (1), dehydration (2, I) and Oxidation (3, I) with consequent release of H_2O_2 . The alternative mechanism II considers the possibility of the oxidation preceding dehydration. Adapted from Terranova and Nifosi (2010).

Reid and Flynn (1997) proposed that the chromophore maturation steps consisted of consecutive cyclization, dehydration and oxidation [Fig.1 13 (c)] of the protein. This proposition was later supported by Barondeau et al. (2006) asserting that dehydration is a key step to lock the cyclic structure. However, it was argued that the oxidation step may precede dehydration (Rosenow et al., 2004). In addition, H₂O₂ generation was detected, before fluorescence emission, thus supporting this theory (Zhang et al., 2006). Jackson et al. (2006) suggested that either mechanism may occur, depending on O₂ concentration or the mutant variety studied. The intricacies of this process are still under debate (Terranova and Nifosi, 2010).

1.3.2.1 GFP folding and mutant variants

GFP folding

The wild-type GFP evolved to be synthesized and folded at oceanic temperatures. It possesses high pH stability but correct folding in *E. coli* at higher temperatures posed some difficulties. Overexpression of wild type GFP in *E. coli* led to manifestation of apoGFP non-fluorescent aggregates in the form of inclusion bodies (Siemering et al., 1996). Experiments consisting of the mechanical unraveling of the β -barrel structure from the N-terminal α -helix concluded with the demonstration of two distinct energetically stable states of unfolded GFP (Dietz, 2004). From a biophysics perspective, these states suggested that transitional forms of unfolded or partially folded GFP can become trapped in energy pockets during overexpression, especially due to chromophore maturation suggesting a rough energy landscape (Andrews et al., 2007). It was shown that the chaperone GroEL/ES can contribute to the correct refolding (Makino et al., 1997). However, GFP folds significantly better during

its translation rather than refolding during *in vitro* experiments (Ugrinov and Clark, 2010). The formation of the β -barrel structure is necessary for the maturation of the chromophore. Conversely, the cyclization of the chromophore amino acids plays a part in the formation of the β -barrel. During translation three β -strands are synthesized before the central α -helix. Then the remaining β -strands are synthesized encasing the α -helix. The GFP from *Aequorea* has a weak tendency to dimerize compared with GFP from *Renilla* which is an obligate dimer (Tsien, 1998). However, GFP from *Aequorea* may form dimers during overexpression due to hydrophobic patches in the β -strands. This leads to deviation of GFP absorption spectra in respect to molar concentration, thus complicating measurement and affecting solubility especially in high ionic strength environments (Yang et al., 1996). This characteristic may also compromise experiments of Fluorescence resonance energy transfer (FRET) (Campbell et al., 2002b) explained subsequently.

Mutant variants can improve folding behavior

Mutations in the β -strands commonly resulted in the modification of GFP properties such as improved folding (Zimmer, 2002). Siemering et al. (1996) produced mutants which showed fluorescence at 37°C. This was achieved by PCR-induced mutations V173A and S175G. Another mutant, GFPCycle 3 was produced by Crameri et al. (1996) through introduction of random mutations. The changes P99S/M153T/V163A were located in three different β -strands. These residues: serine, threonine and alanine substituted the original hydrophobic residues rendering the protein surface more hydrophilic thus less prone to aggregation. It exhibited a 42-fold increase in fluorescence compared with the wild type. It is also called GFPuv due to the presence of an excitation peak at 380nm, in the ultraviolet region. The substitutions A206K, L221K, or F223R (β -strand number 10) preventing dimerization by

exchanging hydrophobic residues for positively charged residues at the dimerization contact. These mutations did not affect the spectral behaviour of the protein (Zacharias et al., 2002). The mutation A206K became ubiquitous in commercial monomeric mutants (Shaner et al., 2005). Nevertheless the expectations of using GFP as a gene expression reporter rest upon the improvement of its folding and maturation times.

Mutations at the chromophore modify spectral behavior

One of the most remarkable aspects of this work was the production of GFP with different spectral properties by mutagenesis. Mutant variant S65T (Heim et al., 1995) consisted of the substitution of serine for threonine in the carboxyl extreme of the chromophore. While the wild type GFP possessed two wide excitation peaks at 395 and emission of 475nm this variant showed one single excitation peak at 490nm with emission at 510nm. This longer wavelength enhanced photostability and produced higher brightness. The maturation time of S65T-GFP was similar to that of wild type GFP and it produced aggregates in *E. coli* at high concentrations. S65T-GFP non-functional aggregates were disaggregated by urea resulting in refolded and fluorescent S65T-GFP *in vitro* demonstrating conclusively that chromophore maturation is an autocatalytic process (Reid and Flynn, 1997). Tsien (1998) classified the GFP mutant derivatives by their chromophore chemistry [Table 1.10] For instance S65T mutation alters the chromophore chemically producing a phenolate anion chromophore (Class II) which results in the increase of excitation wavelength from ultraviolet to blue, approximately 488nm (Heim et al., 1995).

In contrast with previously introduced random mutations the chromophore was also altered with a more rational approach. Threonine-203 interacts with the chromophore; it was

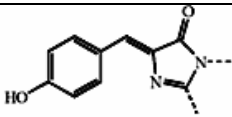
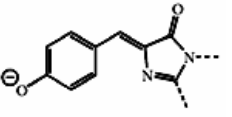
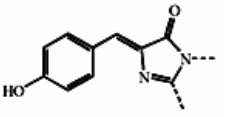
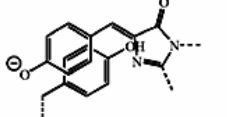
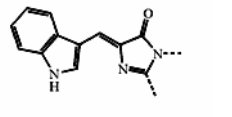
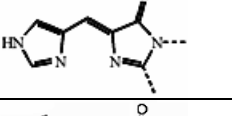
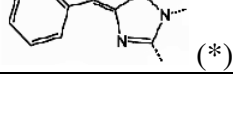
Fluorescence type	Classification	Chromophore description	Typical Chromophore structure	$\lambda_{\text{ex}}/\lambda_{\text{em}}$ in nm	Example	Mutations
Green fluorescence	Class I	Wild Type		395/504	GFPwt	None
	Class II	Phenolate anion		485/507	GFPmut2	S65Q, V68L and S72A
				487-509	Emeral	S65T, S72A, N149K, M153T and I167T
	Class III	Neutral phenol		399/511	H9-40	T203I, S72A, Y145F
Yellow fluorescence	Class IV	Stacked π -electron system		514/527	TOPAZ	S65G, S72A, K79R, T203Y
Cyan fluorescence	Class V	Indole in the chromophore		435/495	W1C	S65A, Y66W, S72A, N146I, M153T, V163A
Blue fluorescence	Class VI	Imidazole in the chromophore		382/446	P4-3	Y66H, Y145F
	Class VII	Phenyl in the cromophore		360/442	Y66F	Y66F

Table 1.10 Classification of GFP derivate mutants

The chromophore structures show attachment to the main peptide by the dashed lines. The maximum excitation and emission wavelength is denoted by λ_{ex} and λ_{em} respectively. Table based on Tsien (1998) (*) complemented from Hasegawa et al. (2007)

suspected that introduction of polar aromatic radicals from tyroxine, histidine or tryptophan would lower the excitation energy of the adjacent chromophore. Providing an extra polarizable π -electron system to the phenolate anion in the chromophore (Class IV) resulted in a split of electrons resonating, effectively sharing the energy. Consequently, the lowered emission energy was reflected as an increase in the excitation-emission wavelengths, distinguishable from GFP, leading to the production a new range of yellow fluorescent proteins. The mutation Y66W adds indole to the chromophore (Class V) which produces emission of higher energy visible in the cyan spectra thus cyan fluorescent proteins. Histidine in the position 66 introduces imidazole (Class VI) and produces a mutant with blue fluorescence (Heim and Tsien, 1996). In recent developments (Mishin et al., 2008) presented a novel variety which featured Serine⁶⁵/Asparagine⁶⁸ mutations, generated by a sequential random mutagenesis; it became the first GFP mutant from *Aequorea spp.* to produce red fluorescence at 529nm.

GFPmut2 is one of the most utilized mutant variant of GFP due to its spectral characteristics and folding properties. It features 3 substitutions: S65Q, V68L and S72A, was first produced and characterized by Cormack et al. (1996). These substitutions resulted in a phenolate ion (class II) active centre with green fluorescence with very similar spectroscopic characteristics to S65T-GFP making GFPmut2 optimal for Fluorescence Activated Cell Sorting (FACS). It has a single excitation peak circa 488nm which is a source commonly available in this type of equipment. The additional changes provided faster folding and better chromophore formation efficiency at 37°C. While wild type GFP matures in 1-4 hours (Heim et al., 1994), GFPmut2 is detectable in *E. coli* in exponential growth after 8 minutes (Cormack et al., 1996) and it shows 100-fold increase in brightness compared to *E. coli* expressing wild type GFP. The peak emission was shown to be in the region of 509-511nm (Abbruzzetti et al., 2005). The

chromophore is ionizable with a pK_a of 6.1 ± 0.1 . Protonated forms absorb at 380nm while unprotonated absorb at 485nm. It has been shown that the protonated and unprotonated forms have the same molar extinction coefficient at 425nm. The presence of this isosbestic point is important for quantifying total GFP. The optimal fluorescence was obtained at pH=7.6 during fluorometric measurements at 485nm_{ex} and 507nm_{em} of protein solutions in different buffers (Chirico et al., 2002).

1.3.2.2 Advantages and limitations of using GFP as a reporter protein

GFP has been shown to be an excellent reporter protein, therefore many of the shortcomings or disadvantages of recombinant expression of the original GFP from coelenterates have been counterbalanced thanks to the development of mutant varieties. These new mutants have shown reduced toxicity caused by accumulation, significantly decreased folding and chromophore maturation time and improved codon usage for each host type. However, some limitations have persisted. One of them is the high background signal resulting from autofluorescence from the host (Niswender et al., 1995). Since excitation and emission peaks are so close in the spectra it is noted that pass-filters are necessary to discern low concentrations from the background of mutants such as GFPmut2 (Cormack et al., 1996). In addition, the chromophore from each protein will emit a fixed quantity of light (Tsien, 1998) therefore GFP emission cannot be amplified as in other reporters such as enzymes which may amplify the signal over noise ratio due to product accumulation over time. The combination of these factors imposes a minimal quantity of protein where a significant increase can be distinguished from background. Nevertheless the close correlation between molar concentration of GFP and quantity of fluorescence relies on this property.

Other inconvenience originates from the necessity of oxygen for chromophore maturation. It has been shown that GFP fluorescence may under perform when exposed to anoxic environments (Tielker et al., 2009). It was proposed to use FMN binding fluorescent protein as a substitution in such conditions (Drepper et al., 2010). However, the singular oxygen chemistry of the chromophore holds useful advantages. It has been previously shown that GFP and some of its mutant variants developed red fluorescence upon being flashed with blue light in absence of oxygen (Elowitz et al., 1997). The photoconversion enabled GFP to absorb at 543nm and emit in the range of 570–630nm. This property was utilized to visualize protein diffusion in bacterial cells (Elowitz et al., 1999) and protein mobility and continuity of yeast organelles (Jakobs et al, 2003).

Another obstacle for the suitability of GFP to certain experiments is its long half-life of up to 24h after maturation or even longer. GFP stability over time is an advantage in some applications but the persistence of GFP may interfere with the correct interpretation of gene expression data. To overcome this limitation new unstable variants have been produced. It was shown that certain sequences have the ability to render the protein susceptible to proteases (Keiler et al., 1996). Andersen et al. (1998) performed several peptide additions to the C-terminal extreme of the GFPmut3* protein resulting in the production of the variant GFP_{AAV} (Alanine-Alanine-Valine tagged) which degrades within a half-life of 60 minutes.

Fluorescent proteins also enabled the study of protein-protein interaction through Fluorescence resonance energy transfer (FRET). Morise et al. (1974) observed this phenomena between the aequorin and the GFP. For efficient FRET some special conditions should be met: the energy of emission of the donor must overlap with the excitation of the acceptor; and the efficiency of FRET decreases with the 6th power of the distance between

donor and acceptor (Stryer and Haugland, 1967), which should be oriented in a manner that maximizes the photon transfer. Miyawaki et al. (1997) utilized this property to design intracellular sensors for the Ca^{2+} ion. It consisted of a fusion of cyan fluorescence protein, the *Xenopus* sp protein calmodulin and green fluorescent protein. Ca^{2+} is bound by the calmodulin which triggers a conformational change which brings CFP and GFP into proximity. This system, called Chameleon-1, reported presence of Ca^{2+} by increase of green fluorescence and decrease of cyan fluorescence. FRET became utilized in many applications designed to monitor protein-protein interactions such as G-proteins (Janetopoulos et al., 2001). In summary, the *Aequorea* spp. based fluorescent proteins possess great advantages and some limitations. Nevertheless experiments or applications can be designed to exploit this type of reporter proteins.

1.3.3 GFP in gene reporter technology. Applications in microbiology

Current applications of GFP as a reporter usually feature a specific mutant of the many spectral varieties (Shaner et al, 2005). This enables coupling experiments with optical and fluorimetric technologies. Fluorescent proteins are an excellent tool for on-line monitoring in comparison with chemoluminescent or metabolic enzymatic reporters. The common applications of GFP based gene reporter systems, recently reviewed by Vizcaino-Caston et al. (2011), include: whole cell biosensors; recombinant protein production quantification; folding efficiency quantification; transcriptional reporters and physiological monitoring of intracellular environment,

Whole cell biosensors

Biosensing consists on the utilization of reporters to monitor processes or environments based on the regulatory elements of an organism to sense its surroundings. This can be achieved by chromosome insertion of the reporter system or by utilization of plasmids. Whole cell biosensors have shown a great potential for industrial and environmental applications. For example, Chalova et al. (2007) produced and characterized a whole cell biosensor strain for dietary lysine. The reporter system consisted of a *GFPmut3* gene inserted in the genome of a lysine auxotroph *E. coli* strain. GFPmut3 production was placed under control of the *lacI* promoter, inducible by IPTG. The auxotrophic strain cannot synthesise lysine therefore in absence of this essential amino acid the organism did not survive. When lysine was available in the environment addition of IPTG induced the expression of GFP in the auxotrophic *E. coli*. Fluorescence reported for quantitative concentrations of lysine by highlighting the presence of modified bacteria, enabling its utilization for monitoring lysine amounts in samples which contained mixed cultures or in non-sterile feedstock.

Good understanding of microbial physiology is necessary for coordinating the reporter system with the environmental parameters to be quantified. The promoter *phnS* from *Burkholderia sartisoli* regulates the response to polycyclic aromatic hydrocarbons (PAHs), a pollutant common in petrochemical spillages, and its expression is blocked by the repressor protein PhnR. The fusion of *phns::egfp* was harnessed by Tecon et al., (2006) and demonstrated that fluorescence could measure exposure to contaminated soil. The system was also utilized by Deepthi et al., (2009) to demonstrate that the PAHs from the Exxon Valdez spillage are bioavailable while the PAHs natural in coals present in the Gulf of Alaska are not. These natural coals used to be a source of interference when pollutants were measured with other methods.

Protein production and protein folding quantification

GFP fusions have been utilized in the past for quantifying protein production. A dual reporter system was first produced and validated for protein production by (Albano et al., 1998). GFP was fused to CAT (chloramphenicol acetyltransferase) under the control of the *araBAD* promoter inducible by arabinose as a transcription unit, being transcribed sequentially. The mRNA produced was translated by different ribosome binding sites yielding the two proteins. *E. coli* JM105 was grown in fed-batch fermentation with periodic assays of CAT activity and GFP detection measurements. The results showed a strong correlation between CAT activity and GFP fluorescence. This same construct was further tested by (DeLisa et al., 1999) during batch and fed-batch fermentation with on-line monitoring with a fluorescent probe. It was concluded that GFP could report for CAT production. In recent developments GFP has been used to report correctly folded protein. Sevastsyanovich et al. (2009) utilized a reporter system based on T7 inducible promoter upstream of a *cheY::gfp* fusion. The production of protein was monitored in *E. coli* growing under different fermentation conditions where the amount of soluble product was indicated by the magnitude of fluorescence produced. The GFP featured the mutations S65T for improved fluorescence and F64L for improved folding in *E. coli* (Waldo et al., 1999).

Transcriptional reporters

Fluorescent reporters have become ubiquitous in new applications inspired by synthetic biology. Synthetic biology is discussed as a form of advanced genetic modification (de Lorenzo and Danchin, 2008). However, this modern approach in molecular biology involves the integration of *de novo* biological systems produced by design. GFP and its derivatives

have become an essential tool in reporter systems for gene transcription in diverse new applications. For instance, a gene oscillator has been made in *E. coli* consisting on a dual feedback circuit (Stricker et al., 2008). On one plasmid, the fusion *araC::yemGFP* was under control of a mixed promoter containing two regions: the operator site of *araBAD* promoter, activated by AraC binding; and the operator region of *lacZ*, repressed by the protein LacI. In the second plasmid, the same mixed promoter was placed upstream the gene *lacI*. Single cell microscopy was utilized to monitor yemGFP production in co-transformed *E. coli* cells upon IPTG and arabinose addition. Fluorescence measures oscillated periodically when cells were grown in 0.7% arabinose and the oscillation period showed to be tuneable by changing IPTG concentration.

Most recently, (Regot et al., 2011) experiments showed how, by cell to cell signalling, GFP production reported the operation of logical gates by yeast cells genetically modified into two types: sensor and reporter. Loaded together on microfluidic platforms, the sensor cells types were able to sense chemical signals and communicate it with a molecular messenger to the reporter cell type which revealed the result. Fluorimetric measurements of GFP production resulted in the correct values for each proposition. For example, GFP was produced by an AND cell type combination when both stimulants were added while fluorescence was absent when only one stimulant was added.

Monitoring biological processes and intracellular environments

Fluorescent transcription reporters have been utilized in industrial approaches amongst them the monitoring of microbial physiology during biological manufacture of diverse products. GFP and its variants enabled the study of physiological parameters in intact cells. Cha et al.

(1999) produced reporter plasmids in order to observe response to different types of stress. The reporters consisted of different pBR322 plasmids containing *dnaK*, *rpoH*, and *clpB* promoters fused to GFPuv. Levels of stress were observed by fluorescence measurements showing that extracellular and intracellular environment may be causes of stress. This became very relevant to improve monitoring of large scale fermentations. While bioreactor instrumentation enabled monitoring of the media, fluorescent measurements enabled monitoring the inner environment of the cell. For example, oxygen limitation is a factor in large scale fermentations that may lead to a metabolic change in the bacteria unfavorable for production. *E. coli* as a facultative anaerobe may alternate to anaerobiosis if there is insufficient dissolved oxygen, a factor that was monitored on-line with GFP. The promoter for nitrate reductase (*narG*) regulates the expression the nitrate reductase operon in anaerobic metabolism (Stewart et al., 2003). Garcia et al. (2009) produced the reporter pNar-GFPuv to exploit the activation of the promoter for measuring oxygen deprivation. The results showed that fermentations performed in stirred tank reactors with a single impeller caused higher relative fluorescence in the host than fermentors equipped with four impellers. It was concluded that the reporter could measure the physiological effect of low oxygen environments originated by inadequate mixing.

On a more general approach the microbial stress response was monitored through studying the expression of the gene *rpoS*. Delvigne et al. (2009) utilized this promoter in the fusion *rpoS::GFPmut2* where the expression levels of *rpoS* were monitored during fed-batch fermentations. Using flow cytometry, it was demonstrated that the microbial population segregated into different populations with different stress status by changing the mixing efficiency. Further experiments on microbial physiology monitoring featured the introduction of dual systems to control reporter activity, consisting two plasmids: one to cause stress upon

induction; and a second plasmid to function as a reporter. For example, Nemecek et al. 2(008) compared the performance of chromosome-integrated reporters and vector-based reporters by monitoring the activity of *dnaK::gfpmut3* fusion upon metabolic overburden. The stress was induced with a pET30a plasmid expressing human superoxide dismutase. Shake flask experiments and fed-batch fermentations were performed in both systems. The results led to three major conclusions: The recombinant protein increased *dnaK* promoter activity; the strains with the chromosomal reporter produced more recombinant protein but the signal from the reporter was too weak; and the double transformant strain produced recombinant protein and reflected stress by an increase in fluorescence, fulfilling the needs of an on-line monitoring tool.

It can be argued that proteomics and transcriptomics are the most complete tools for measuring translational and transcriptional variation in the cells. However, transcriptomics and proteomics processes are too slow and costly to be used as a process analysis techniques in industrially grown cell factories. All these gene reporter systems have to be implemented with the right measurement instrumentation in order to optimize the data throughput.

1.4 Microbial Flow Cytometry

1.4.1 The flow cytometry principle

The first flow cytometer was a particle counter designed to detect bacterial spores in the air during the 2nd World War (Davey and Kell, 1996). In the 1950s, this tool was improved by Wallace Coulter resulting in the Coulter Counter A. Cells were transported in a narrow fluid stream through a detector which measured the cell volume by variations of electrical

conductivity between cells and media. Coulter Counter B was capable of relating the amplitude of this signal upon passing the sampling point with the size of the cell (Davey and Kell, 1996). Modern flow cytometry is founded on the behavior of light and advances in optics during the last 60 years. The basic elements of a modern flow cytometer are depicted in [Fig. 1.14]. The sample is forced through a nozzle into a stream of sheath fluid achieving laminar flow. The pressure of sample and sheath must be adequate in order to generate hydrodynamic focusing, where the sample stream is sufficiently narrowed to direct the particles or cells orderly as single events. The intrinsic light scattering properties and fluorescence of each particle are obtained when illuminated by a light source. Flow cytometers are also equipped with photodiodes coupled to photomultipliers which transform photons in electric signals and amplify such signals respectively. Combined with suitable filters, the amount of specific fluorescence emitted by the sample can be measured.

Flow cytometry has been utilized to screen for abnormalities of blood and bone marrow cells including its application to observe some types of leukemia by studying the morphology of such cells (Diamond, Nathwani, and Rappaport, 1982). Diagnostic applications were increased by the utilization of monoclonal antibodies carrying fluorescent tags. Antibodies were raised against specific membrane proteins displayed by different cell types, including abnormal cells, enabling discrimination of specific cells (Knapp, Strobl, and Majdic, 1994). Bibliometric studies have shown that reports of flow cytometry applied to general microbiology have steadily risen for last 20 years to the present when approximately 8% of flow cytometry related work is applied to microbiology (Muller and Davey, 2009). Bacteria are most difficult to analyze, due to their small size; the alignment process requires more precision, and because their size is similar to noise particles, setting thresholds and discriminators can be more complicated. (Hewitt and Nebe-Von-Caron, 2004). In addition,

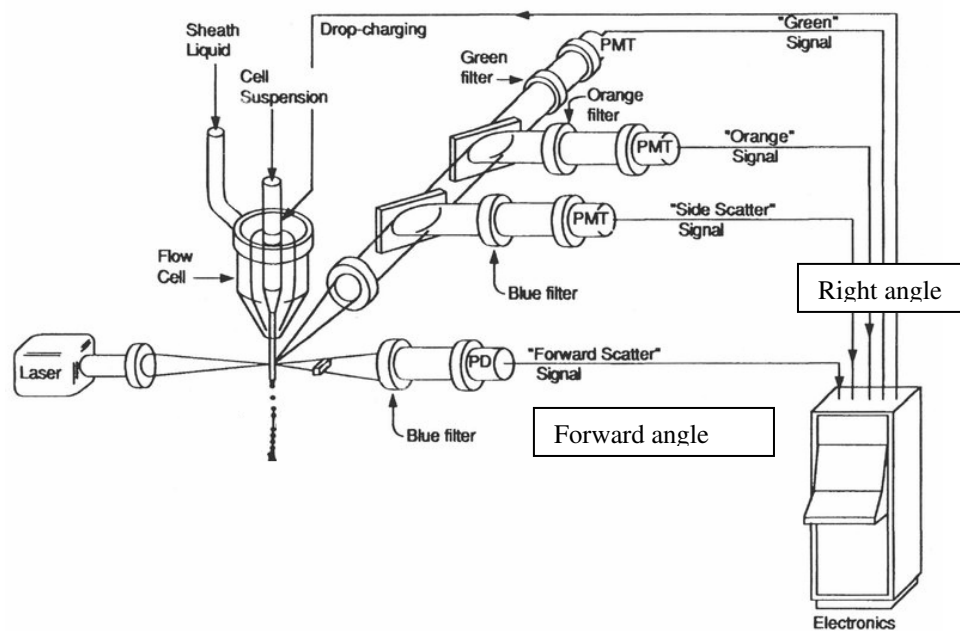


Fig. 1.14 Diagram of the basic elements of a flow cytometer

The cell suspension flows within a stream of sheath liquid and into the flow cell. As a laser light hits the flow cell, different optical signals are produced as result. Light scatter and fluorescence signals are collected in the right angle aided by a set of dichroic mirrors which filter the light of the appropriate wavelength into each PMTs. Signals are then digitalized. Adapted from Givan (1992)

the rate of individuals detected per second is much higher which requires higher laser speed and data management. Some modern flow cytometry allows acquisition speeds of more than 70,000 events per second (Spidlen et al., 2006). Thus the applications of flow-cytometry relies on the optical technologies available to generate and detect light; translate those optical signals in electric pulses; and the computation power and information technologies available manage real-time data acquisition.

1.4.2 Parameters in flow cytometry: Light scatter and fluorescence

The flow cell is illuminated with a light source. While Arc lamps are cheaper, most modern flow cytometers incorporate lasers as light source. When the source light hits the particle to be analyzed, it splits into fragments of light of different wavelengths which are refracted in multiple directions. The pattern of scattering and refraction is inherent to the morphological characteristics of the particle. Forward angle light scatter (FALS) is collected on the same axis as the light source, and collects light fractions generally within the 0° - 10° angle which can be extended up to 20° (Hewitt and Nebe-Von-Caron, 2004). The right angle side scatter (RALS) detector is mounted orthogonally to the light source collecting light fractions scattered sideways (also referred as side scatter or SS) upon interaction with the particle or cell. The amount of scattered light collected in the FALS is generally indicative of the size of the particle while that collected through RALS has its origin in the complexity or granularity of the particle. For most microbial populations the readings in both scatter detectors are correlated and this relationship allows discrimination of the cells from background noise (Hewitt and Nebe-Von-Caron, 2004).

However there are factors which affect the refractive properties of the cell thus the light scatter readings. In the case of rod shaped cells such *E. coli*, it was shown that FALS did not increase linearly with particle volume over the longitudinal dimension of the rod generating a unimodal distribution while RALS data produce a bimodal diffuse cluster. Cells were sorted from different regions of the bimodal cluster which resulted in the same distribution. This was interpreted to be caused by different light scatter detection as the cell rotates as it travels through the laser. As a coccoid organism, *Rhodococcus sp.* produced unimodal clusters both in FALS and RALS (Hewitt and Nebe-Von-Caron, 2004).

It has been shown that accumulation of insoluble protein in inclusion bodies results in an increase of forward and side scatter and it was discussed as a potential method to detect inclusion bodies (Hewitt and Nebe-Von-Caron, 2004; Lewis et al., 2004). However, due to the unpredictability of light scatter behavior combined with other refractive alterations of the cell it was argued that forward scatter is not optimal for monitoring inclusion body formation (Wallberg et al., 2005). Values for these scatter parameters are generally displayed together in bi-dimensional plots; FALS on ordinates and RALS on abscises. This representation is usually sufficient to discriminate populations based on morphology.

The light source emits a spectrum of choice which can be utilized in order to act as excitation energy to generate fluorescence from the sample. The fluorescence is collected by photomultiplier tubes (PMT) on the 90° alignment axis (same as RALS). The appropriate filters are selected in order to collect the photons of the required wavelength range. Several PMTs may be enabled to amplify specific fluorescence which may be intrinsic to the sample or generated by application of fluorescent dyes (Veal et al., 2000). There are many chemical fluorescent dyes to obtain useful information from the sample [Table 1.11].

Cell ligand or substrate		Dye		Ex/Em*
Nucleic acids	DNA, RNA	Ethidium bromide (EB)		510/595
		Propidium iodide (PI)		536/623
		Cyanines	TO-PRO	515/533 (TO-PRO 1) 462/661 (TO-PRO 3)
			TOTO	514 /533 (TOTO 1) 642/660 (TOTO 3)
			SYTO	485-508/498-527
			SYBR	497/520 (Sybr Green)
	DNA	DAPI		350/470
		Hoechst 33342		350/461
		DRAQ5		647/700
Total protein		FITC		495/525
Lipids		Nile Red		551/636
Fluorogenic substrates		ChemChrome dyes (CY, CB, CV6)		488/520
		FDA , CFDA, CFDA/SE, CFDA-AM		492/519
		Calcein-AM		494/517
		CTC		450/630
		FUN-1		480/ 580
Intracellular pH		BCECF-AM		482/520
		SNARF-1		510–580/587–635
Membrane energization		Rh123		507/529
		DiOC ₆ (3)		484/501
		DiOC ₂ (3)		482/497
		DiBAC ₄ (3)		488/525
		JC-1		498,593 / 525,585
Citoplasmic Ca ⁺²		Indo-1		361,330/405, 480
Antibodies or oligonucleotides		BODIPY		503/512
		PE, PE-Cy5 conjugates		490/575, 690
		Alexa Fluor 488		495/519
		Oregon green		496/524
Autofluorescent proteins		GFP and derivatives		489/508

Table 1.11 Variety of dyes utilized in flow cytometry

(*) indicates the wavelengths of excitation and emission. From Díaz et al. (2010)

1.4.3 From the photon to the file

The PMT senses the photons released through FALS, RALS or fluorescence with a photocathode which releases electrons (e^-) into the tube. The amplification is achieved by a cascade of dynodes within the tube where the acceleration of electrons released to the next dynode depends directly on PMT voltage. Increasing the overall voltage of the PMT promotes electron transfer up the dynode ladder, thus achieving signal amplification. Because every dynode releases twice the electrons it receives, linear voltage increases results in exponential increase of sensitivity. There are 10 dynodes in the PMTs of the EPICS Elite (Coulter Corporation). The electrons received in the anode are of a maximum of 1024 (2^{10}) e^- per photon which are transformed into a voltage pulse by the integrator.

There are two types of signal: integral and peak. The voltage pulse of a peak signal is represented by the maximum amount of electrons received when the sample is within the beam. Integral signal is the maximum amount of electrons accumulated as the particle travels throughout the section of the beam. Therefore, while a peak signal is proportional to the highest fluorescence detected from the particle, the integral signal is proportional to the overall fluorescence generated from the particle. The voltage pulse leaves the integrator to be amplified either linearly or logarithmically. The pulses greater than 10 Volts (V) are off scale when digitalized. Logarithmic amplification widens the scale of weak signal and contracts the scale of the strong signals enabling visualization of a greater range of signals in the final histograms.

In order to manage the vast amount of data collected these need to be digitalized enabling their storage in computer files. The analog-to-digital converter (ADC) assesses each pulse

received and assigns it to a channel by its voltage. The ACD utilizes 10 bits to encode the data in channels. There are 1024 (2^{10}) channels each representing a 0.01V increment of peak amplitude where the maximum is 10.24V. Successive approximation is applied to smaller increments.

1.4.4 Cell physiology is not homogeneous throughout the sample

In microbial cultures there are three main sources of heterogeneity (Davey and Kell, 1996; Davey and Winson, 2003). The first cause is genetic and arises by accumulation of random genomic mutations or adaptive mutations through the generations. The second source is dependent on cell cycle based on differences on protein expression during the different stages of the culture and factors which alter the rate individuals move into different growth phase. The third is caused phenotypic variability caused by differences in the immediate microenvironment each bacterium dwells in. Stochasticity, defined as a form of random noise, is inherent in gene expression which may lead to different phenotypes in clonal populations (Elowitz et al., 2002). Stochasticity is biologically advantageous because it enables the population of cells to sample phenotypes and anticipate immediate variations in the environment (Thattai and van Oudenaarden, 2004).

Modern flow cytometers can acquire and manage large amounts of data per second for each single event, enabling a statistical resolution never found in bulk measurements performed for biomass quantification or reporter activity. Multi-parametric analysis enables to infer the physiological status of cells overcoming the disadvantages of CFU (colony forming units) counts by being able to bring into account the non-culturable cells (Hewitt and Nebe-Von-

Caron, 2004). Combined with the right dye is possible to perform population studies and culture heterogeneity assessments.

1.4.5 Strategies for observation of physiological conditions

Fluorescent dyes

As shown in [Table 1.14] there are numerous fluorescent dyes and techniques that enable to screen many possible biological states of the cell. Propidium iodide (PI) and DiBac₃(4) (Bis-(1,3-dibutylbarbituric acid)) are two stains which enable observation of basic physiological status of microbial cultures (Hewitt and Nebe-Von-Caron, 2001; Lopez-Amoros et al., 1995). PI binds DNA and possesses red emission spectra when excited with laser light at 488nm. The PI dye cannot cross an intact cell membrane thus it is utilized to discriminate cells upon membrane integrity. DiBac₃(4) is a lipophilic stain which ingresses into the cell if the membrane is depolarized. This anionic dye only enters the when all active transport have halted, thus these cells may have their homeostasis compromised. DiBac₃(4) emits in the green spectra upon excitation with the blue laser. PI and DiBac₃(4) can be monitored simultaneously and usually are represented together in bi-dimensional plots. These plots can show the heterogeneity of the sample population based on susceptibility of each individual cell to each dye (Lewis et al., 2004). While higher values of DiBac₃(4) fluorescence represent gradual loss of membrane potential, thus unhealthy cells, higher values of PI fluorescence results from DNA staining which implies likelihood of these events being dead cells. This has been widely applied to viability studies during *E. coli* fermentations, for example (Hewitt and Nebe-Von-Caron, 2001; Lewis et al., 2004). These dyes were utilized to study membrane potentials and membrane permeability in other species such as *Staphylococcus aureus* and *Micrococcus luteus* (Novo et al., 2000). The antibacterial effect of silver was tested

comparatively on *E. coli* and *S. aureus* where the discrepancy of flow cytometric data and colony formed units was attributed to active but not culturable cells suggesting that silver compromises homeostasis leading to cell death (Jung et al., 2008). PI was also utilized to assess the effectiveness of antifungals on species of *Candida* and *Aspergillus* (Pinto et al., 2008).

Fluorescent protein production and gene reporter technology

Fluorescent proteins and gene reporter technology allows monitoring gene expression in the culture. Multi-parameter flow cytometry enhances these observations providing single individual resolution of intact cells which allows monitoring the variability within the sample. Therefore flow cytometry is an excellent technique to observe reporter expression levels in situations where heterogeneity in the sample is expected. This technique was utilized to observe horizontal gene transfer of conjugative plasmids. The reporter plasmid carried a *lac::GFP* fusion. The donor strain possessed the gene *lacI^{q1}* encoding the repressor of the *lac* promoter. The transfer could be observed on mixed cultures of donor and acceptor in the presence of IPTG. While the donor repressed reporter activity by *lacI^{q1}* expression, the acceptor (without the repressor gene) would develop fluorescence. Flow cytometry resolved both populations enabling observation and quantification of plasmid transfer (Sorensen et al., 2003). It has also been used combined with reporter gene technology for monitoring plasmid loss (Bahl, Sorensen, and Hestbjerg Hansen, 2004). Industrial fermentations are prone to heterogeneity due to localized microenvironments resulting from gradients of nutrient or dissolved oxygen (Hewitt and Nebe-Von-Caron, 2001). Differential expression and population segregation have been observed on stress generating in relation to mixing conditions (Delvigne et al., 2009); making flow cytometry an ideal technique for monitoring.

Chapter Two

Materials and methods

2.1 Materials

Chemicals utilized were sourced from the following manufacturers unless otherwise stated:

Sigma, Sigma-Aldrich, Difco, Fisher, Bacto, Bhd or Oxoid

Media

LB (Lennox Broth) consisted of 20(gL⁻¹) Tryptone (Bacto), 10(gL⁻¹) yeast extract (Difco) and 10(gL⁻¹) NaCl (Sigma) in distilled water. Glucose and glycerol were utilized alternatively as carbon sources as indicated by the experiment added to the media prior autoclaving. SOC contained 2 % (w/v) tryptone, 0.5 % (w/v) yeast extract, 10 mM NaCl (Sigma), 2.5 mM KCl (Oxoid), 10 mM MgCl₂ (Oxoid), 10 mM MgSO₄ (Oxoid) and 20 mM glucose . Nutrient agar (Oxoid) and LB agar were utilized as solid media. When utilized were dissolved distilled water and sterilized. Solid and liquid media were sterilized in an autoclave for 15 minutes at 121°C and allowed to cool prior to addition of the antibiotic when required. Agar plates were prepared by pouring 25mL of agar into each plate allowing to set at room temperature. Excess moisture was removed by drying the plates 30 minutes at 50°C. Plates were stored at 4°C. Plates containing antibiotics were utilized withing 2 months.

Antibiotics

Antibiotic stocks were prepared in distilled water and filter sterilized at concentrations of: Kanamycin (Sigma), 50(mgmL⁻¹); ampicillin (Sigma) 100(mgmL⁻¹); and carbenicillin (Sigma), 100(mgmL⁻¹). Long term storage was at -20°C in 1mL aliquots. Alternatively, aliquots were kept at 4°C for short term storage.

Chemicals

β -mercaptoethanol (Sigma) addition to the media was calculated using its concentration (14.3M). Indoleacrylic acid (IAA) (Sigma) stock consisted of 2.5(mgmL⁻¹) dissolved in 95% (v/v) of ethanol and 5% (v/v) in distilled water, filter-sterilized (\emptyset 0.22 μ m) and stored at -20°C. IPTG (Isopropyl- β -D-Thiogalactosidase) was utilized to induce the T7 promoter of pET expression vectors. The stock consisted of a 0.5M solution in distilled water filter-sterilized (\emptyset 0.22 μ m) which was kept at 4°C. Ethidium bromide stock at 11mgmL⁻¹ was prepared in distilled water from commercial tablets and kept in an opaque bottle at room temperature. Propidium Iodide (PI, Invitrogen) stock, for cytometric staining of dead cells, was made up at 200(μ gmL⁻¹) in distilled water and stored at 4°C.

2.2 Buffers and Solutions

Phosphate Buffered Saline (PBS)

This buffer was utilized to resuspend or dilute samples of living cells. The composition 'Dulbecco A' was utilized containing: NaCl, 8.0(gL⁻¹); KCl, 0.2(gL⁻¹); Na₂HPO₄, 1.15(gL⁻¹); and KH₂PO₄, 0.2(gL⁻¹); with pH=7.3. This solution was prepared from commercial tablets.

β -galactosidase activity assay

Ortho-nitrophenyl- β -galactopyranoside (Sigma-Aldrich) 13mM (ONPG) was dissolved in A-buffer (0.1 M K₂HPO₄ and 0.1 M KH₂PO₄, pH 7.0). Z-buffer contained 0.75 g of KCl, 0.246 g of MgSO₄.7H₂O, 8.53 g of Na₂HPO₄, 4.866 g of NaH₂PO₄.2H₂O and 2.7 mL of β -mercaptoethanol per litre. Z buffer was adjusted to pH 7.0 before use.

Buffers for DNA electrophoresis

TBE was prepared and stored as a 5x concentrated stock consisting of: 0.445M Tris (pH 8.0); 0.445M boric acid and 10 mM EDTA. The working solution was obtained by diluting it five

times to produce 1xTBE used for separation of DNA by agarose gel electrophoresis. Agarose solutions contained 1% (w/v) type-II agarose (Sigma) in 1 x TBE buffer. The agarose was melted in TBE buffer by briefly boiling it in a microwave oven. Agarose solution was stored at 60°C for up to four weeks in liquid form or allowed to solidify and kept at room temperature for up to three months. DNA blue sample buffer consisted of: 0.025 % (w/v) bromophenol blue; 10% (v/v) glycerol; 10mM Tris-HCl pH 7.5; and 1mM EDTA in distilled water.

Buffers for SDS-PAGE

Protogel (30 % (w/v) acrylamide and 0.8 % (w/v) NN'-methylenebisacrylamide) was used for both the stacking and the resolving gel. Stock stacking gel buffer contained 1.25M Tris-HCl (pH 6.8) and stock resolving buffer contained 0.75M Tris HCl (pH 8.3). Solutions were vacuum filtered with Whatman filter paper (pore size Ø 0.45µm). Ammonium persulphate 80(mg mL⁻¹) was dissolved in distilled water, distributed into 1 mL aliquots and stored at -20°C. SDS-PAGE loading buffer contained 20mL of glycerol, 5mg of bromophenol blue, and 10mL of SDS (sodium dodecyl sulphate) 20% solution (w/v) (Bio-Rad), and was made up with 0.1 x stock stacking gel buffer to a volume of 92mL. This solution was distributed into 1 mL aliquots and stored at room temperature. 87µl of β-mercaptoethanol was added to each 1 mL aliquot of SDS-PAGE sample buffer immediately prior to use. Stock 10 x SDS-PAGE running buffer contained 30g of Tris and 150g of glycine per litre of distilled water. Running buffer was prepared by mixing 100 mL of 10x stock solution with 5 mL of 20 % (w/v) SDS solution (Bio-Rad) and 895 mL of distilled water.

Buffers and solutions for Coomassie blue staining

Coomassie blue stain consisted of 0.2 % (w/v) Coomassie Brilliant Blue R, 50 % (v/v) methanol and 10 % (v/v) acetic acid in distilled water. Before use, the Coomassie stain was vacuum filtered with Whatman filter paper (pore size Ø 0.45 µM) to remove undissolved solids. Fast destain consisted of 40 % (v/v) methanol and 10 % (v/v) acetic acid in distilled water. Shrink solution consisted of 48 % (v/v) methanol and 2 % (v/v) glycerol in distilled water.

2.3 Bacterial methods

E. coli strains utilized in this study are listed in [Table 2.1]. Untransformed strains were stored on agar plates at 4°C and utilized up to four weeks. For longer term storage, exponentially growing cultures were prepared in 1 mL aliquots with addition of glycerol to 25% (v/v) prior storage at -80°C.

Precultures

Single colonies were inoculated into 5mL of LB supplemented with antibiotics as required and cultured overnight (16 hours) at 37°C in universal bottles shaking at 200rpm. 300µl of overnight culture was used to inoculate each culture in each experiment. *E. coli* BL21* overnights were incubated at 30°C for no longer than 14 hours.

<i>E. coli</i>	Genotype	F'	Source/Reference
JM109	<i>endA1, recA1, Δ(lac-proAB)</i>	yes	Promega
BW25113	<i>Δ(araD-araB)567, ΔlacZ4787</i>	no	(Baba et al., 2006)
ECA101	<i>Δ(araD-araB)567, ΔlacZ4787 ΔrpoE</i>	no	(Baba et al., 2006)
JW3883	<i>Δ(araD-araB)567, ΔlacZ4787 ΔcpxR</i>	no	(Baba et al., 2006)
JW3882	<i>Δ(araD-araB)567, ΔlacZ4787 ΔcpxA</i>	no	(Baba et al., 2006)
BL21*	F– <i>ompT</i> (DE3)	no	(Miroux and Walker, 1996)

Table 2.1. Strains used in this study

This table summarizes the characteristics of the *E. coli* strains utilized in this study. *araD-araB* mutations block arabinose metabolism; *endA1*, eliminates the non-specific digestion by Endonuclease I. *recA1* mutation reduces occurrence of unwanted recombination in cloned DNA. *proA/B*, indicates proline requirement. DE3 indicates that the strain contains a prophage carrying the T7 RNA polymerase gene. F' refers to the presence or absence of the F plasmid. *ompT*, indicates the lack of an outer membrane protease

E. coli growth

Bacteria were grown in 30mL of LB supplemented with the appropriate antibiotics as indicated in [Table 2.2] and carbon source as indicated by the experiment description. Aerobic cultures were incubated in 250mL conical flasks shaken at 200rpm whereas anaerobic cultures were incubated in static capped boiling tubes. Incubation temperatures and times were specific of each experiment. Biomass was measured using spectrophotometry of absorbance at 650nm.

Batch fermentation

The seed culture for fermentation consisted of a culture 2% (v/v) in relation to the amount of media in the fermenter vessel. Cells were inoculated from an agar plate into 1L conical flask containing 56mL LB supplemented with the necessary antibiotics for plasmid maintenance. The fermentation system consisted of a bench top fermentor manufactured by Electrolab, model Fermac 310 equipped with an Electrolab vessel of 5L maximum capacity; Dissolved oxygen (DO) probe; pH probe; and temperature probe. The vessel was heated with an electrical heating mat and cooled via cooling loop connected to a thermostatic circulator set at 4°C (2219 Multitemp II, LKB Biomna). The pH probe was calibrated utilizing standard buffers and the responsiveness of the DO probe was set at 100% and tested in water by sparging nitrogen before utilization. Media, 2.8L of LB, were sterilized for 30 minutes at 121°C within the vessel with the probes, the air filters and the tubing attached. The carbon source, 2% glucose, was autoclaved as a 40% (v/v) solution separately, for only 5 at 121°C to avoid caramellization. This solution was added into the vessel prior to inoculation. The pH was controlled at 6.3 by the automated addition of HCl 5% (v/v) and NH₃10% (v/v). 100mgmL⁻¹ Carbenicillin.

Plasmid	Main features	Reference/source
pQF50groE	<i>groE::lacZ</i> , <i>Kan^R</i> (50µgml⁻¹). Ori: <i>pMB1</i> (from pBR322) and <i>ori1600</i> . (Farinha and Kropinski, 1990)[Fig 2.1 (a)]	(Wang and deHaseth, 2003)
pDNAK-FPV	<i>dnaK::gfp</i> , <i>Amp^R</i> (50µgml⁻¹) Ori: <i>ColE1</i> (from pBR322)	(Aertsen et al., 2004)
pRPOH-FPV	<i>rpoH::gfp</i> , <i>Amp^R</i> (50µgml⁻¹) Ori: <i>ColE1</i> (from pBR322)	(Aertsen et al., 2004)
pYIDQ	<i>yidQ::gfpmut2</i> , <i>Kan^R</i> (25µgml⁻¹) Ori: <i>SC101</i> [Fig 2.4]	This study
pUA66	Promoterless <i>gfpmut2</i> , <i>Kan^R</i> (25µgml⁻¹) Ori: <i>SC101</i> [Fig 2.1 (b)]	(Zaslaver et al., 2006)
pDNAK	<i>dnaK::gfpmut2</i> , <i>Kan^R</i> (25µgml⁻¹) Ori: <i>SC101</i> [Fig 2.3]	(Zaslaver et al., 2006)
pGROE	<i>groE::gfpmut2</i> , <i>Kan^R</i> (25µgml⁻¹) Ori: <i>SC101</i> [Fig 2.2]	(Zaslaver et al., 2006)
pET11c	<i>T7</i> promoter, <i>lac</i> , <i>Amp^R</i> (100µgml⁻¹) Ori: <i>ColE1</i>	Novagen
pET-CCP	<i>T7</i> promoter, <i>lac</i> , <i>ccp</i> , <i>Amp^R</i> (100µgml⁻¹) Ori: <i>ColE1</i>	(Sevastyanovich et al., 2009)
pGEM-T	<i>Amp^R</i> (100µgml⁻¹)	Promega

Table 2.2 List of plasmids utilized in this study

Kan^R encoded kanamycin resistance while *Amp^R* encodes resistance to ampicillin and carbenicillin. In bold are shown the concentration of antibiotic added for plasmid maintenance. In the case of *Amp^R* ampicillin was used except for batch fermentation where gentamicin was preferred. Ori, states origin of replication. *lac*, encodes the Lac repressor, *ccp* encodes cytochrome c peroxidase (CCP) from *Neisseria gonorrhoeae*.

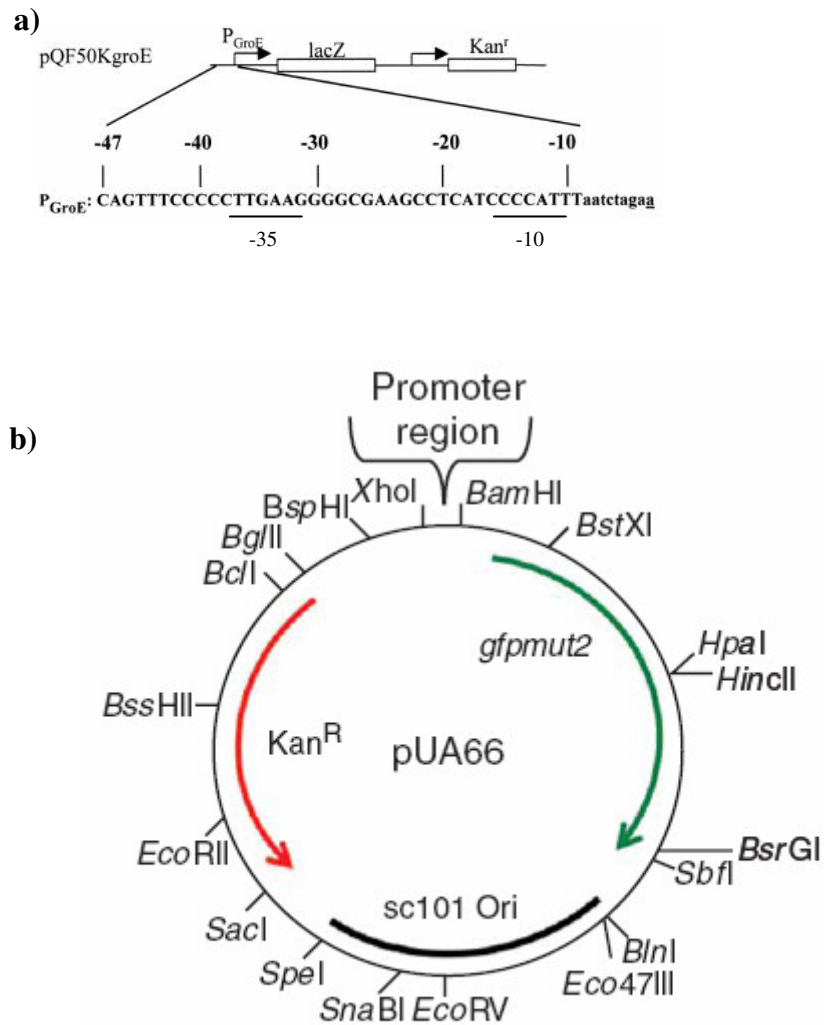


Fig 2.1 pQF50KgroE promoter and restriction map of pUA66 promoterless vector

(a) shows the unique σ^{32} promoter or pQF50KgroE. From (Wang and deHaseth, 2003) (b) shows the restriction map of the promoterless vector pUA66 Kan^R is the gene encoding kanamycin resistance, *gfpmut2* and the origin of replication SC101 are also represented. From (Zaslaver et al., 2006)

5' *yjeH* (YjeH APC transporter)

CTACCAGCGCAGCTAACGCAGGAACGGCAAACACGCCAGTgcctaataatgacgtcgatagcaggccaatgccctgggccagccccagttcttgtttgagtc

actcatggggttgatgtccgattgcgccccaaatTTTGGGCAACTgcgtagattttcgatggtagacaaatcagattcgcttatgacggcgatgaagaaattgcg

atgaaa | CRP-Binding site | | UP ELEMENT | | -35 σ^{32} n=15 -10 |

| distal | proximal |

| A/T motifs |

ttgaca | σ^{70} n=17 |

groS

ctggtcaccagccgggaaccacgtaagctccggcgtcacccataacagatacggactttctcaaggagaggttatcaATGAATATTCGTCCATTGCA

3' *gfpmut2*

TGATCGCGTGATCGTCAAGC

Shows the insert in pGROE containing the whole intergenic region. CRP binding site has been inferred from consensus sequences. The up-element is also shown. In spite of being mostly a σ^{32} -dependant promoter (yellow), σ^{70} may also bind in this site (green). With the arrow, in red is indicated the transcription initiation site, in bold (**ATG**) the translation initiation site. Note that the insert contains several base pairs of the gene *groES*. Underlined is the sequence used by Wade et al.(2006)

pDNAK: Insert size = 457 bp

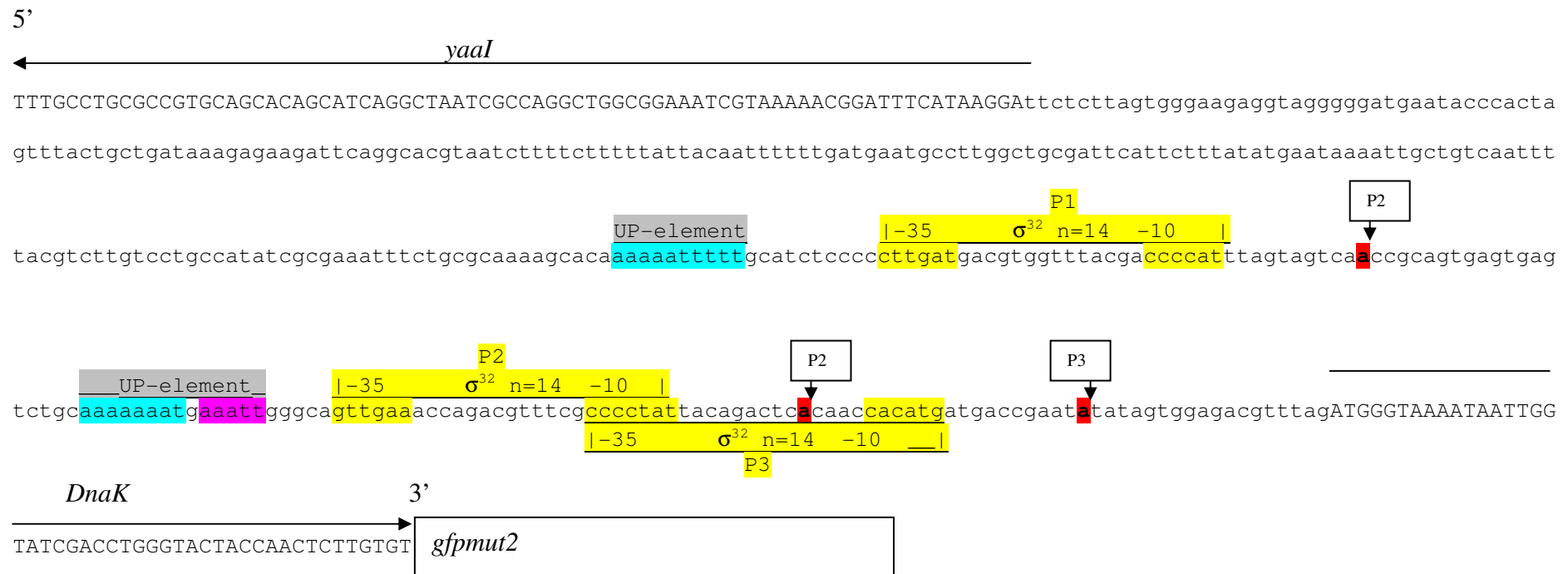


Fig. 2.3 Nucleotide sequence of the promoter of pDNAK

Shows the insert in pDNAK containing the whole intergenic region. The up-element is also shown (distal in cyan, proximal in magenta). The three σ^{32} binding sites (P1, P2 and P3) are indicated in yellow. In red, labelled with an arrow are shown the transcription start sites in bold (**ATG**) is denoted the translation initiation site. Note that the insert contains several base pairs of the gene *dnaK*

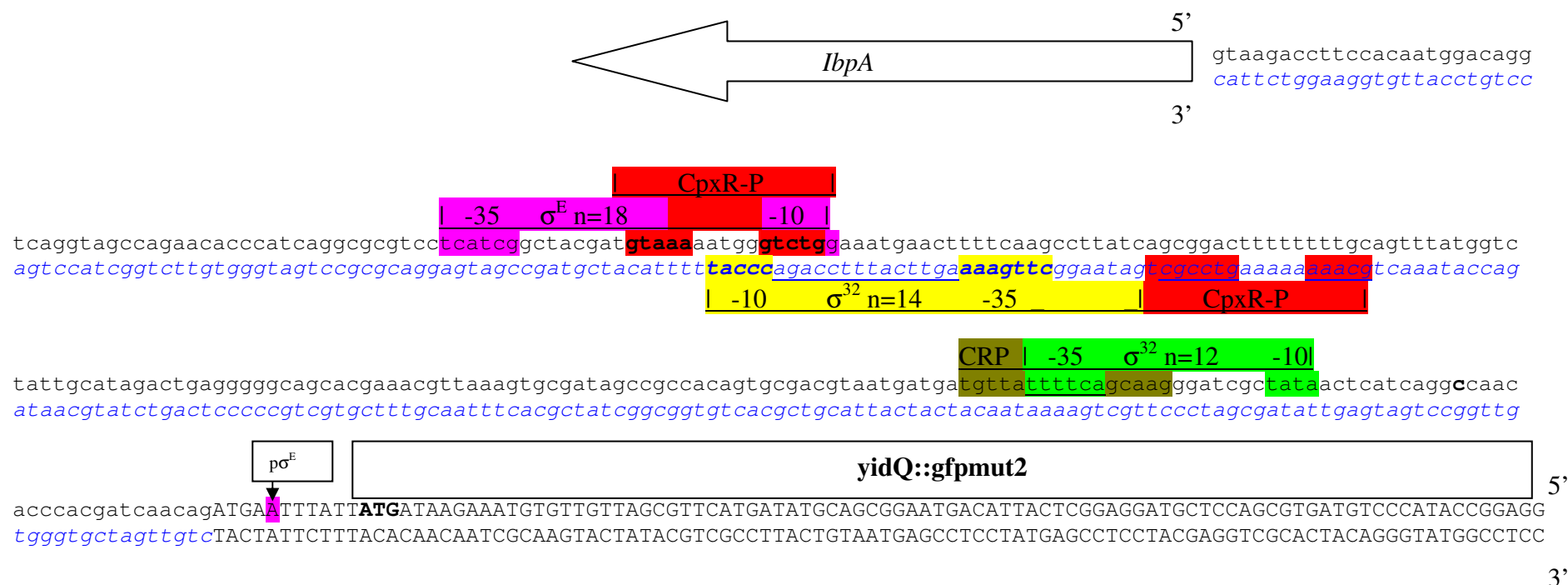


Fig. 2.4 Nucleotide sequence of the promoter of pYIDQ

Shows the insert in pYIDQ. In magenta is shown the σ^E binding and the transcription initiation site. The σ^{32} binding site of *ibpA* promoter in the complementary (chain) indicated in yellow (Kotelnikova et al., 2005). In red, are shown the CpxR-P binding sites inferred from De Wulf et al. (2002). σ^{70} may also bind in this promoter (green). The binding sites for CRP binding are highlighted in brown. Note that the insert contains 96bp into pairs of the *yidQ* gene (gfpmut2 is further downstream). Many of the strong CpxR-P binding sites further upstream (Dartigalongue et al., 2001) in the YidQ promoter were not included in this vector.

and 25mgmL⁻¹ Kanamycin were added for plasmids retention. Temperature was set at 37°C until the culture reached an OD₆₅₀ of ~0.5. Then temperature was set at 25°C and 0.5mM IPTG was added to the fermenter, following the standard protocol described by (Sevastyanovich, Alfasi, and Cole, 2009).

β-galactosidase activity assay.

2mL culture samples were taken into uncapped universal bottles. 30μl of toluene and 30μl of sodium deoxycholate (1%) were added to the samples to trigger cell lysis and shaken for 20 minutes at 37°C. The lysates were kept on ice until needed. To measure the β-galactosidase activity, 200μl of each lysate was transferred into three test tubes containing Z-buffer in a water bath at 30°C. Two test tubes contained 1.8mL of Z-buffer, the third tube contained 2.3 mL of Z-buffer was used as background control. Once all lysates were distributed, the β-galactosidase reaction was started by addition of 0.5mL of 13mM ONPG (prepared in A-buffer) to all tubes except of the background control tubes. The time from ONPG addition was measured with a stopwatch. The reaction was allowed to progress at 30°C for a maximum of two hours or until a noticeable change of colour from transparent to yellow by comparison with the control. Then it was stopped by addition of 1mL of Na₂CO₃ 1M to all tubes raising the pH which inactivates the enzyme. ONPG hydrolyzed by β-galactosidase resulted in change from uncoloured to yellow establishing a colorimetric reaction measurable by absorbance at 420nm. The activity was calculated by the formula:

$$\text{Activity} = (1000 \times (\text{OD}_{420\text{R}} - \text{OD}_{420\text{C}}) \times 3.5) / (\text{time} \times 4.5 \times 0.4 \times V \times \text{OD}_{650})$$

In the formula, OD_{420R} is the arithmetic average of optical density of the sample measured from two independent test tubes at 420nm and OD_{420C} is the background control at 420nm. 3.5 corresponds to the total volume in the test tube in mL, 0.4 is the dry mass of cells, in mg, in 1 mL of a culture of *E. coli* with an optical density at 650 nm of 1.0. OD₆₅₀ is the optical

density of the culture at 650nm. V corresponds to the volume in mL of lysate used. 4.5 is the molar extinction coefficient of ONPG at 420 nm. The activity is expressed in nanomoles of ONPG hydrolyzed per minute per milligram of dry cell weight mass. Method described by Jayaraman et al (1987).

Quantification of fluorescence by fluorimetry

The samples taken from the culture were centrifuged to harvest the cells and the volume of LB from the sample was exchanged by PBS prior to measurements because of the elevated background in the green spectrum produced by the LB media. The suspension of cells was transferred into a reusable precision quartz cuvette (1cm light path) manufactured by Hellma and diluted in PBS to achieve an OD₆₅₀ between 0.3-0.5. The cuvette was kept clean between samples by rinsing it with ethanol 70% (v/v.) Fluorescence from the cells was measured with a luminescence spectrophotometer, manufactured by Pelkin Elmer model LS50B and configured as described in [Table 2.3].

Microbial flow cytometry

Samples taken from the culture were diluted immediately in 2mL of PBS. If cell aggregation was suspected cells were sonicated for 5 seconds prior to analysis. The flow-cytometer utilized was an EPICS Elite manufactured by Coulter Corporation. During analysis it was configured under the settings summarized in [Table 2.4] unless otherwise is stated in the experimental description. This equipment featured a manual alignment unit for positioning of the flow cell in relation to the laser beam and the forward scatter diode. The quality of this alignment was determined with latex beads (Flow-Check Fluorospheres, Beckman-Coulter) with $0.538\mu\text{m} \pm 0.012\mu\text{m}$ of diameter. PI was utilized at a working concentration of $7.5 (\mu\text{g mL}^{-1})$ made from stock $200(\mu\text{g mL}^{-1})$ in PBS. Cells were stained with PI for 1 minute before measurement using flow cytometry.

	Wavelength	Slit with
Excitation	488nm	2.5nm
Emission	509nm	15nm
Integration time	3 seconds	
PMT	775 Volts	

Table 2.3. Fluorescence settings for luminescence spectrophotometer

Photo multiplier tube general settings			
FS-LOG	PMT1 (SS)	PMT2	PMT4
250	700	900	variable
Light source			
Laser type	Wavelength	Power	
Argon gas	488nm	15mW	
Fluidics			
Sheath preassure	5.0 psi (Pounds per square inch)		
Sample preassure	Variable. Adjusted to data rate		
Data rate	1000-2000 events per second		
Maximum events	25000 events		

Table 2.4 Flow cytometer settings

2.4 Recombinant DNA techniques

Isolation of plasmid DNA.

Plasmids used in this study are listed in [Table 2.2]. Plasmid DNA was isolated from bacteria using a Qiagen miniprep kit. The manufacturer recommendations were applied as follows; for medium or high copy number plasmids transformed in *E. coli*, a single colony was inoculated into a wide test tube containing 2mL of LB supplemented with the appropriate antibiotic and incubated overnight at 37°C with vigorous shaking. Alternatively, when purifying very low copy number plasmids, a single colony was inoculated into a 20mL conical flask containing 10mL of LB supplemented with the necessary antibiotic and incubated overnight at 37°C with vigorous shaking. The cells were harvested by centrifugation and processed as specified by the manufacturer. Medium-high copy plasmids were eluted in 50µl of TB buffer (Tris Borate from Qiagen) whereas low copy plasmids were eluted in 30µl TB. Plasmid content was validated by DNA electrophoresis as described subsequently.

Electrophoresis of DNA

The DNA to be separated was mixed in a 4:1 ratio with DNA blue sample buffer and separated by electrophoresis. The equipment utilized, iMyRun from Helixx Technologies, comprised of a gel casting system, an electrophoresis tank and a power

supply. The agarose molten solution was poured into the casting system; plastic combs were suspended used to form wells from 10 to 40 µl capacity. Once set, the gels were submerged in 0.5 x TBE solution in a horizontal electrophoresis tank, the DNA samples to be separated were loaded into the wells and an electric current was passed over the apparatus at 100 V for one hour. DNA was visualised by staining with a 0.5 (µgmL⁻¹) solution of ethidium bromide, which intercalates between bases of double stranded DNA molecules and fluoresces when exposed to ultraviolet light. Gels were visualised using an UVP UV-transilluminator with a wavelength of 304nm.

Oligonucleotide design

Custom oligonucleotide primers were synthesised by Invitrogen. Each primer was resuspended in sterile water and diluted to give a 10 µM solution. Stock solutions of primers were stored at -20 °C. The melting temperature of oligonucleotides was calculated using the formula:

$$81.5 + (G + C)\% \times 0.41 - 675/n$$

Where (G + C)% is the percentage of the nucleotides guanine and cytosine in the primer and n is the total length of the primer. The primers are listed in [Table 2.5]

Amplification of E. coli DNA by the Polymerase chain reaction (PCR)

The polymerase chain reaction is a technique used to amplify specific fragments of DNA. The oligonucleotide primers were specifically designed to anneal to known sequences flanking the desired DNA fragment. During thermal cycling, the DNA is denaturated at 95°C separating the strands in a step called melting. Then, the temperature is reduced and one of the primers hybridizes with the coding strand of the template while the other hybridizes with the complementary strand. This hybridization step is called annealing. The thermostable DNA

Promoter	Oligonucleotide primers
<i>ibpA</i>	Forward 5' ATA TAC TCG AGG TAT GGG ACA TCA CG 3' Reverse 5' ATT AAG <u>GAT CCC</u> CTT CCA CAA TGG AC 3'
<i>groES</i>	Forward 5' ATA TAC TCG AGA GGC TGT TAT TGC CC 3' Reverse 5' TTA CAG <u>GAT CCC</u> TTT GAG AAA GTC CG 3'
<i>dnaK</i>	Forward 5' ATA TAC TCG AGA CGG ATT TCA TAA GG 3' Reverse 5' TAA TAG <u>GAT CCC</u> CAC TAT ATA TTC GG 3'

Table 2.5 Oligonucleotide primers used in this study

This table shows the primers used to amplify the promoter regions. The restriction sites are underlined for *Bam*HI and in bold for *Xho*I.

polymerase then replicates the DNA between the two primers by adding nucleotides to the free 3'terminal (5' to 3' direction). Successive reactions achieved by temperature cycles result in large quantities of DNA fragments being generated. Such cycles were programmed into the Thermal Cycler Px2 (Thermo) to amplify the fragments [Table 2.5].

PCR reaction mix was made up in a total volume of 50 μ L in sterile distilled water. For all polymerase chain reactions, the reaction mixture contained: 1.5 μ L $MgCl_2$ and 10% (v/v) Optibuffer (Supplied with Bio-X-act Kit from Bioline); 5 μ L of deoxyribonucleotides (dATP, dCTP, dGTP and dTTP, from a Quiagen mix); 1 mM of each oligonucleotide primer; and 5 μ L of template DNA, which was obtained from *E. coli* MG1665. Reactions were catalysed by 0.25 μ L Bio-X-Act, the reaction. Success of the PCR reaction was determined by observing the presence of a band at the correct size by DNA electrophoresis. DNA was purified utilizing the Qiagen PCR purification Kit.

pGEM-T cloning

Purified PCR products were cloned into a pGEM vector (Promega) as specified by the manufacturer and transformed into ultracompetent *E. coli* JM109 supplied with the kit. Blue/White screening was utilized to detect successful transformants. Candidates were tested for positive insertion by plasmid purification followed by analysis of restriction fragments in agarose gels. Positive cloning was revealed by the presence of a band of the correct size upon double digestion with *Bam*HI and *Xho*I.

Restriction endonucleases and DNA modifying enzymes

	Step 1		Step 2		Step 3		No of Cycles
	Denaturation		Annealing		Elongation		
Stage 1	95°C	3 minutes	-		-		1
Stage 2	95°C	30 seconds	40°C	30 seconds	72°C	30 seconds	5
Stage 3	95°C	30 seconds	56°C	30 seconds	72°C	30 seconds	25
Stage 4	-	-	-	-	72°C	15 minutes	1

Table 2.6 Thermocycler configuration for PCR

Stage one consists of denaturation of the template DNA which results in strands separation. In stage two the first polymerization of the DNA flanked by the primers occurs; 40°C is the annealing temperature of the bases of the primer excluding the restriction site and the overhang. The extension step occurs at the working temperature of the DNA polymerase as specified by the supplier. The bulk of the DNA amplification occurs in stage 3 in which 56°C is the annealing temperature of the entire primer. In the last stage the elongation step was extended 15 minutes.

Restriction enzymes *Bam*HI and *Xho*I (New England Biolabs) were used to cut DNA. The proportion recommended by the manufacturer of DNA and enzyme units were mixed and incubated at 37 °C for two hours as specified by the manufacturer to release the insert PCR-generated promoter fragment from pGEM. Due to the proximity between restriction targets in pUA66 (6bp), restriction digestion was performed sequentially as specified by Zaslaver et al. (2006). The vector was digested with *Xho*I for 7 hours at 37°C followed by a cleaning step with the PCR purification kit from Qiagen. The product was then digested with *Bam*HI for 5h at 37°C.

Calf intestinal alkaline phosphatase (CIAP) (New England Biolabs)

This enzyme was utilized to prevent the recircularisation of pUA66 during ligation. It was utilized on the vector after digestion with restriction enzymes by incubating for 1 hour at 37°C with 10 units of CIAP in 1xNEB buffer (*New England Biolabs*)

Purification of DNA fragments from agarose gels.

DNA fragments were purified from agarose gel using a Qiagen kit. The band was cut from the gel using a scalpel. The gel slice was then dissolved and the DNA extracted from it and purified according to the manufacturer's protocol.

T4 DNA ligase

This enzyme ligates two pieces of DNA by the formation of a phosphodiester bond between the 5' phosphate group of one strand and the 3' hydroxyl group of the other strand. The DNA fragments to be ligated were incubated overnight at 16 °C with 1 unit of T4 ligase in the manufacturer supplied buffer. The molar ratios of vector to insert were 1:1 and 1:3. The amount of insert necessary was calculated by the formula:

$$\frac{\text{ng of vector} \times \text{size of insert (in kb)}}{\text{size of vector (in kb)}} \times \text{insert:vector ratio} = \text{ng of insert required}$$

The mass in ng of the vector and insert were estimated by comparison with 1kb DNA Ladder standard (New England Biolabs). This was performed with the image densitometry of electrophoretic bands option in the Kodak 1D Gel Electrophoresis Documentation System.

High-throughput cloning screening

The flow cytometer was utilized for screening of transformants capable of producing green fluorescence when compared with that of colonies transformed with promoterless pUA66 as control. This enormously facilitated the selection of candidates. Those colonies showing evidently greater green fluorescence than the control were cultured and harvested. Plasmids were purified, separated in aliquots and double digested with *Bam*HI and *Xho*I for 2 hours at 37°C. These digests were separated by electrophoresis; when a band of approximately the expected size of the insert was observed; the plasmid was taken for sequencing. Sequencing was performed at the sequencing services of the University of Birmingham in the School of Biosciences from aliquots containing the vector and either of the oligonucleotide primers.

2.4 Protocols of transformation

Preparation of competent cells by the CaCl₂ method

0.2mL of an overnight culture was used to inoculate 20 mL of LB containing appropriate antibiotics in a 250mL conical flask. This culture incubated at 37 °C with constant shaking at 200rpm until the OD_{650nm} was in the range of 0.3-0.5. Then, the culture was transferred to two previously cooled test tubes and incubated on ice for ten minutes. The bacteria were harvested

by centrifugation for 5 minutes at 5000g at 4°C. Pellets were gently resuspended in 4 mL of ice-cold 0.1 M CaCl₂ solution and incubated on ice for 20 minutes, before being harvested by centrifugation as previously. Pellets were gently resuspended in 1 mL of ice-cold 0.1 M CaCl₂ and 15 % (v/v) glycerol. The competent cell suspension obtained was stored overnight at 4 °C or at -70 °C for longer periods.

Preparation of competent cells by the Nishimura Protocol

The Nishimura protocol has been shown to be a rapid method to prepare large stocks of competent cells capable of yielding excellent transformation efficiency (Nishimura et al., 1990). 0.5 mL of an overnight culture was used to inoculate 50 mL of LB containing appropriate antibiotics supplemented with 0.2% glucose and 10mM MgSO₄· 7H₂O in a 500mL conical flask and grown at the required temperature (30°C for *E. coli* BL21* and heat sensitive mutants) to mid exponential phase. Then, the culture was split into two ice-cooled 50mL centrifugation tubes and incubated on ice for 10 minutes. Cells were harvested by centrifugation for 10 minutes at 4°C. The cells were resuspended in 0.25mL of fresh pre-cooled culture medium and 1.25mL of a precooled filter sterile (Ø 0.22µm) solution containing 36% glycerol, 12% PEG6000, and 12 mM MgSO₄· 7H₂O in LB broth (pH=7) was added. The suspension of competent cells was divided into 0.1mL aliquots and stored at -80°C.

Heat shock transformation

100 µL of competent cell suspension mixed with 1-3 µL of plasmid DNA was incubated on ice for one hour. Bacteria were heat shocked at 42 °C for one minute. After heat shock, 0.1 mL of pre-warmed at 37°C SOC medium was added. Alternatively, when following Nishimura protocol, 50 µl of competent cells were thaw on ice and incubated up to 30 minutes on ice with

5µL (or approximately 100pg) of plasmid DNA. Cells were induced to take the plasmid with a 1 minute heat pulse at 42°C, chilled on ice 2 minutes and added 0.5 mL of prewarmed LB.

Ultracompetent *E. coli* JM109 cells (Promega pGEM-T Kit) were also transformed with plasmid DNA by the heat shock method as specified by the manufacturer. This procedure was the same as above except that the heat shock was for 45 seconds, and was followed by one minute incubation on ice.

Suspensions of heat-shock transformed cells were incubated at 37 °C for one hour with constant shaking. This period was increased to 2 hours if a low copy plasmid was being utilized to transform. Bacteria were harvested by centrifugation, resuspended in 100 mL of SOC, spread onto agar plates supplemented with appropriate antibiotics and incubated at 37°C overnight (30°C for *E. coli* BL21*). Single transformant colonies were selected and inoculated onto fresh agar plates and supplemented with the necessary antibiotics divided into grid. These plates were incubated at 37°C overnight (30°C for *E. coli* BL21*), stored at 4°C and used as stock for experiments up to two weeks.

Double transformants

This was applied only to *E. coli* BL21*. Following the protocols described, the plasmid pET-CCP was transformed in first place. *E. coli* BL21*-pET-CCP were made competent and transformed with the low copy number pUA66-based plasmid.

2.5 Analysis of proteins

Denaturing polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were separated according to their mass using polyacrylamide gels containing SDS (sodium dodecyl sulphate) following the procedure described by Shapiro et al (1967). The SDS contributes to the denaturation of the ensuring that the proteins are in linear conformation so that mobility depends only on the peptide length. Peptides treated with SDS acquire an overall negative charge, for this, amino acid charges do not affect the mobility. SDS-PAGE gels consist of two parts: a smaller stacking gel containing 6 % (w/v) acrylamide above a larger resolving gel containing 15 % (w/v) acrylamide. The stacking gel concentrates the proteins so they arrive at the resolving gel at the same time. The resolving gel separates the proteins according to size. The casting and running system utilized was OmniPAGE mini from Cleaver Scientific.

The resolving gel was made by mixing 10 mL of stock 30 % acrylamide (Protogel), 10 mL of stock resolving gel buffer and 100 µl of 20 % (w/v) SDS. Ammonium persulphate (APS) is a strong oxidizer of which 200 µl of 8 % (w/v) solution was added to polymerize the acrylamide. Addition of 10 µl of TEMED (N, N, N', N'-tetramethylethylenediamine) catalyzed the polymerization. The solution was mixed and poured immediately between two 100 mm x 100 mm glass plates separated by 0.75 mm spacers until it reached a height of 0.5 cm below the end of the comb. A layer of 0.1 % (w/v) SDS was added to exclude air bubbles. Once the resolving gel had polymerised for 30 minutes the stacking gel was prepared. The stacking gel consisted of 3 mL of stock 30 % acrylamide, 1.5 mL of stacking gel buffer, 11.5 mL of distilled water, 75 µl of 20% (w/v) SDS, 150 µl of APS and 7.5 µl of TEMED. The 0.1 % SDS overlay was removed from the surface of the resolving gel and the glass plates were filled with stacking gel and a comb inserted to form the wells. Acrylamide was left to polymerise for at least one hour. Successful polymerization was confirmed by setting of left over gel. The comb was removed and the wells were washed with distilled water. Excess water was removed from

the wells by shaking the gel. Protein samples prepared from cell pellets were resuspended in SDS-PAGE sample buffer, heated at 100 °C for 3 minutes and then cooled before being loaded into the dry wells. Electrode buffer was applied to the wells with a pasteur pipette and the tank filled with electrode running buffer. An electric current was applied at 110 V for 3 hours.

Coomassie blue staining of SDS-PAGE gels.

Proteins separated by SDS-PAGE were detected by staining using Coomassie Brilliant Blue R (Sigma). The gel was placed in approximately 300 mL of Coomassie stain and incubated at room temperature for 30 minutes with gentle shaking. The gel was washed twice with 200 mL of fast destain for an hour with shaking, placed in shrink solution for 1 hour and dried overnight between two sheets of gel drying film (Promega).

Bug Buster

BugBuster is a reagent from Novagen which in few steps can produce preparations of soluble and insoluble proteins from cell pellets. Initially its use was intended to separate inclusion bodies from soluble protein. However, due to the properties of the BugBuster reagent proteins associated to the cell membrane such as CCP (cytochrome *c* peroxidase) are precipitated with the membranes in the insoluble fraction. While it could not be used to completely separate inclusion bodies from correctly formed protein in the cell it served to concentrate the recombinant protein produced.

Bacterial pellets were re-suspended in 67µL of BugBuster reagent per OD₆₅₀ unit and incubated at room temperature for 10 minutes with gentle shaken. Soluble and insoluble fractions were separated by centrifugation at 8500g at 4°C for 20 min. Separated fractions were

added to the same volume of sample buffer as that used initially with BugBuster and boiled at 100°C for 10 minutes. To ensure the same biomass in all samples, two volumes of the soluble and one volume of the insoluble fractions were loaded in the SDS-PAGE.

2.6 Computational methods and software

Representation of flow cytometry data

The data is written on a file following Flow Cytometry Standard (FCS) for which exist 3 versions. FCS1.0 was created in 1984 in order to facilitate communication and analysis of data (Murphy and Chused, 1984). The newer versions FCS2.0 and FCS3.0 date from 1990 and 1997 respectively (Dean et al., 1990; Seamer et al., 1997). These updates responded to advances in computing which enabled greater information management. ‘List mode’ means that the acquisition program stores the data of the selected parameters for each event in the order received from the ADC thus order of detection. Under FCS standards these data are encoded in 16-bit words. If the researcher needs to see raw data or utilize different analysis software, it needs to be decrypted into ASCII (American Standard Code for Information Exchange) or exported to a spreadsheet which can be considered a weakness of this standard (Redelman, 2004). Each event is shown correlated to each selected parameter with the ADC assigned values in a consecutive list ordered by acquisition time in seconds.

WinMDI

In order to facilitate analysis, the flow cytometry data was represented in graphs; commonly histograms, dot plots or density plots. The histogram represents the count of events in each channel. This originates a data distribution based on the intensity registered. The density plots or dot plots are created by representing one parameter against another. Consequently each data

point is positioned on the plot according to the channel it occupies for each parameter. In order to facilitate visualization and data management density plots were utilized. In this graph, the data point is coloured accordingly to density. These plots can be represented with different resolutions; the 256x256 plot were utilized to resolve every four events sampled.

Web-based analyses.

The gene information and chromosomal context was obtained from the database EcoCyc (Keseler et al., 2011). Specific DNA sequences for primer design were obtained from coliBASE (Chaudhuri et al., 2008). BLAST (basic local alignment search tool) of the National Library of Medicine was utilized to identify sequenced DNA. Needle, an alignment tool from the EBI (European Bioinformatics Institute) was utilized to identify the best alignment of two given sequences. Chromas (<http://www.technelysium.com.au>) was used to read DNA sequencer output.

Chapter Three

Measuring *groE::lacZ* activity as means of monitoring σ^{32} activity

3.1 Introduction

The cytoplasmic heat shock response is coordinated by the alternate sigma factor σ^{32} which in such situations competes with other sigma factors for the RNA polymerase core (Gruber and Gross, 2003). Recombinant protein production has been shown to originate a heat-shock like response (Gamer et al., 1996). For this reason there was an interest in studying the levels of σ^{32} in various situations relevant to growth conditions utilized in recombinant protein production. The plasmid pQF50KgroE, harbouring a *groE::lacZ* fusion, was originally designed for studying the sequence determinants of the binding site of σ^{32} (Wang and deHaseth, 2003). Therefore, its promoter consists only of the σ^{32} -binding portion of the *E. coli groES/EL* promoter [Fig. 2.1]. Presence of functional σ^{32} can be monitored *in vivo* by measuring expression of the *lacZ* gene resulting in β -galactosidase, quantifiable by biochemical assay. This reporter system was utilized to study levels of σ^{32} in different conditions of aeration and temperature fluctuation relevant to bioprocess conditions.

Previous studies on mammalian cells have shown that membrane permeating thiol reductants are a leading cause of protein misfolding (Gelman and Prives, 1996; Lodish and Kong, 1993). β -mercaptoethanol is known to permeate through cell membranes (Malpica et al., 2004) and was utilized in this study to observe if β -mercaptoethanol is able to activate the σ^{32} regulon in *E. coli*. Data collected was utilized to differentiate states of rest, stress and recovery delivering useful information for future studies. Since transcription of the *rpoH* gene encoding σ^{32} is regulated by σ^E , a mutant with a deletion of the *rpoE* gene (referred here as $\Delta\sigma^E$) was utilized to observe possible interference of σ^E .

3.2 Results

3.2.1 Expression of *groE::lacZ* in relation to temperature

These experiments were designed to observe the activity of *groE::lacZ* fusion in conditions known to trigger a σ^{32} response. Diaz-Acosta et al. (2006) measured GroEL and DnaK protein levels using [³⁵S]-methionine radiolabeling in SDS-PAGE demonstrating that chaperone levels increase under aerobic conditions upon heat-shock. In anaerobic growth, this chaperone induction was more rapid and more sustained than aerobically. The aim here was to determine a framework of reporter promoter activation in consonance with such bioprocess-related growth conditions, including oxygen limitation, for genes expressed by $E\sigma^{32}$, especially chaperone genes. *E. coli* JM109 transformed with pQF50KgroE was grown overnight and inoculated into 30mL of LB supplemented with 50 μ g mL⁻¹ of kanamycin and 0.4% glucose. Aerobic and anaerobic cultures were incubated at 37°C before heat shock. Heat-shock was applied for one hour by transferring cultures to 42°C incubators while control cultures remained at 37°C. All work was performed on duplicate cultures, sampled at hourly intervals unless otherwise stated and repeated at least twice. In order to confirm the suitability of the strains for *in vivo* β -galactosidase expression, the activity of untransformed cells was measured during growth resulting in no significant β -galactosidase activity (< 5 nmol of ONPG hydrolyzed per minute per milligram of DCW).

Heat shock during aerobic growth

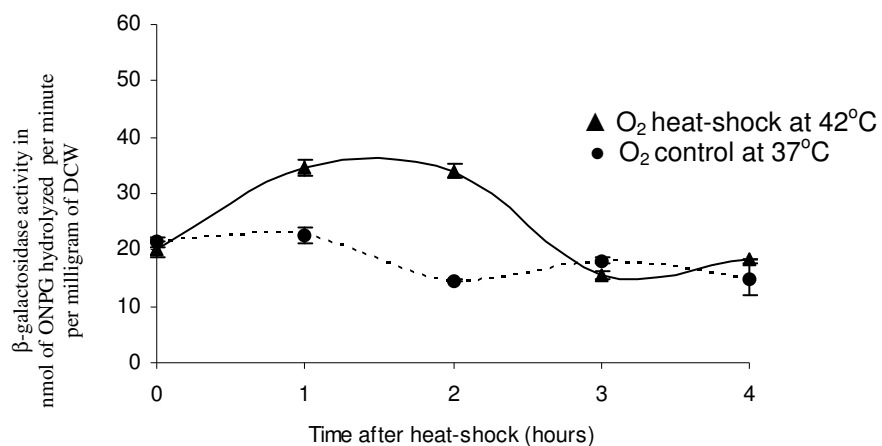
One hour after heat-shock at 42°C, β -galactosidase activity increased in comparison with the

control at 37°C and was maintained higher than the control until the second hour [Fig. 3.1 (a)]. The average activity of the control cultures was maintained during the first hour followed by a decrease in the second hour. This effect is probably due to a high concentration of β -galactosidase in the cytoplasm of the overnight grown cells used to inoculate the culture followed by subsequent dilution during growth. This has been observed frequently in reproductions of this and other experiments. From the third to the fourth hour after shock, cultures at 42°C show a drastic reduction in activity. There are two main hypotheses for this effect: (1) the heat-shock response has counteracted the cause of stress and the cell has returned to a 'rest' state; and (2) there is a growth phase effect as the cultures enter the stationary phase as shown by the growth curves [Fig. 3.1 (b)]. The slight increase of activity shown by the control culture at 3 hours was observed in other reproductions of the experiment and usually appeared towards the end of the exponential phase. This suggests that a growth phase effect cannot be discarded.

Heat shock during anaerobic growth

Data of β -galactosidase activity during anaerobic growth are shown in [Fig. 3.2 (a)]. The activity increased drastically in the heat-shocked cultures after the first hour at 42°C in a similar manner to the aerobic cultures. However, as growth progressed into the second hour, the average activity remained at similar level [Fig. 3.2 (b)]. There is an increase of activity in the third hour which could be attributed to fluctuations in the growth curves or a secondary additive response caused by anaerobic conditions. After the second hour the anaerobic cultures at 37°C exhibited a gradual increase of activity which stabilized in the

a) β -galactosidase activity during aerobic (O_2) growth



b) Growth curves during aerobic growth

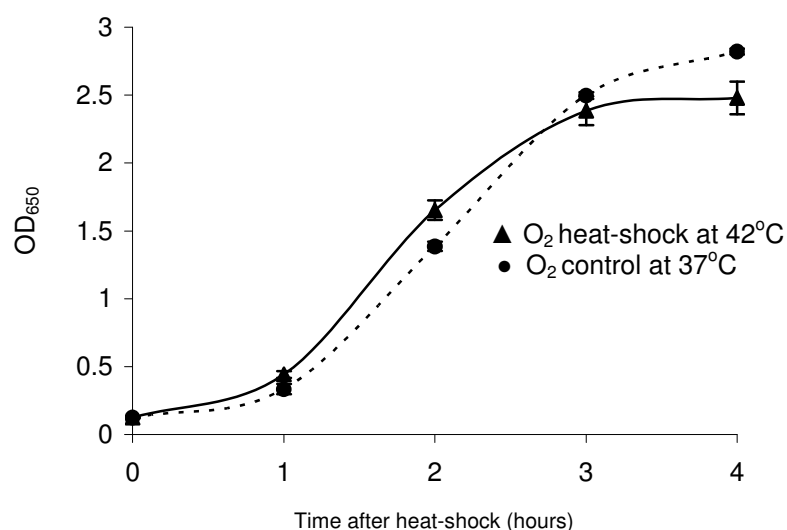


Fig. 3.1 Heat-shock during aerobic growth

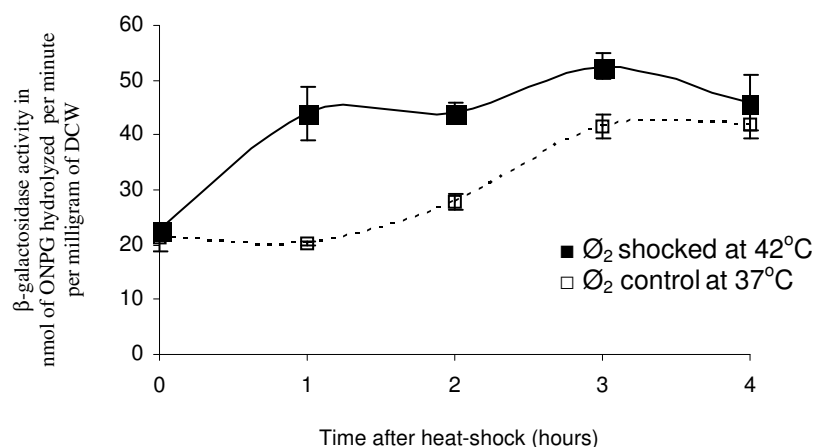
β -galactosidase activities (a) and growth measured by absorbance at 650nm (b) of *E. coli* JM109 pQF50KgroE. All cultures were incubated at 37°C for one hour prior to heat-shock. Heat-shocked cultures, transferred to 42°C , are linked by solid lines while activities measured in the controls at 37°C are linked by dotted lines. Error bars represent the standard deviation of duplicate cultures.

fourth hour at a similar level to the heat-shocked cultures. β -galactosidase activity was higher in anaerobic than aerobic cultures [Fig. 3.1] which could be attributed to faster and more sustained expression of σ^{32} -dependent genes in such conditions as seen by Diaz-Acosta et al. (2006) in combination with growth rate reduction which could have contributed to increased β -galactosidase accumulation within the bacterial cytoplasm. These experiments have demonstrated that the changes in chaperone protein levels shown by Diaz-Acosta et al. (2006) can also be quantified by reporter gene assays, a simpler technique than previously used ($[^{35}\text{S}]$ -methionine radiolabeling and SDS-page).

3.2.1.1 Responsiveness of *groE::lacZ* to consecutive temperature variations during growth

This experiment was designed to observe the influence of temperature variations on β -galactosidase expression in order to test the responsiveness of the reporter system for detecting σ^{32} levels upon temperature fluctuations. *E. coli* JM109 transformed with pQF50KgroE was grown overnight and inoculated into 30ml of LB supplemented with $50\mu\text{g mL}^{-1}$ of kanamycin and 0.4% glucose. Aerobic and anaerobic cultures were incubated at 37°C for one hour after inoculation and then subjected to two subsequent temperature variations: first, from 37°C to 21°C, where cultures stayed for 1.5 hours; and second, from 21°C to 42°C until the end of the experiment. Aerobic and anaerobic control cultures were incubated at 37°C throughout. An additional set of aerobic and anaerobic cultures were subjected to a single variation consisting of a heat shock at 42°C, in order to relate this experiment with prior results. All work was performed on duplicate cultures. Cultures were

a) β -galactosidase activity during anaerobic (O_2) growth



b) Growth curves of anaerobic (O_2) cultures

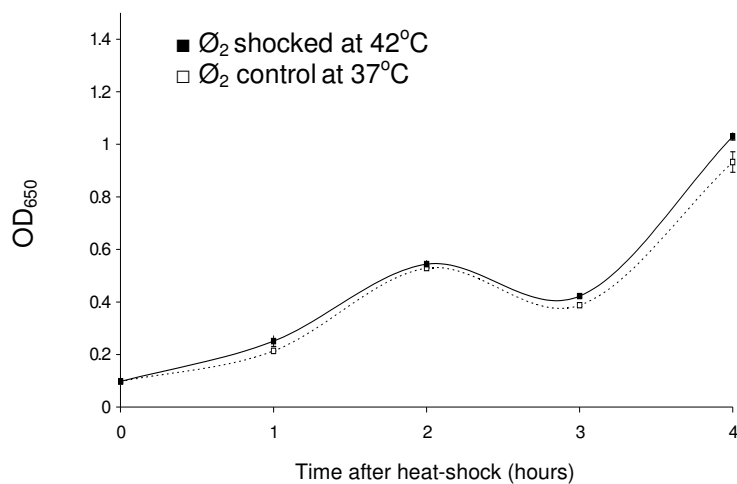


Fig. 3.2 Heat-shock during anaerobic growth

β -galactosidase activities (a) and growth measured by absorbance at 650nm (b) of *E. coli* JM109 pQF50KgroE. All cultures were incubated at 37°C for one hour prior to heat-shock. Heat-shocked cultures, transferred to 42°C, are linked by solid lines while activities measured in the controls at 37°C are linked by dotted lines. Error bars represent the standard deviation of duplicate cultures.

sampled at the time of the first variation, then at 45 minute intervals to enhance resolution in the central stages of growth. The last samples were taken at 1 hour intervals.

Effects of multiple temperature variation observed during aerobic growth

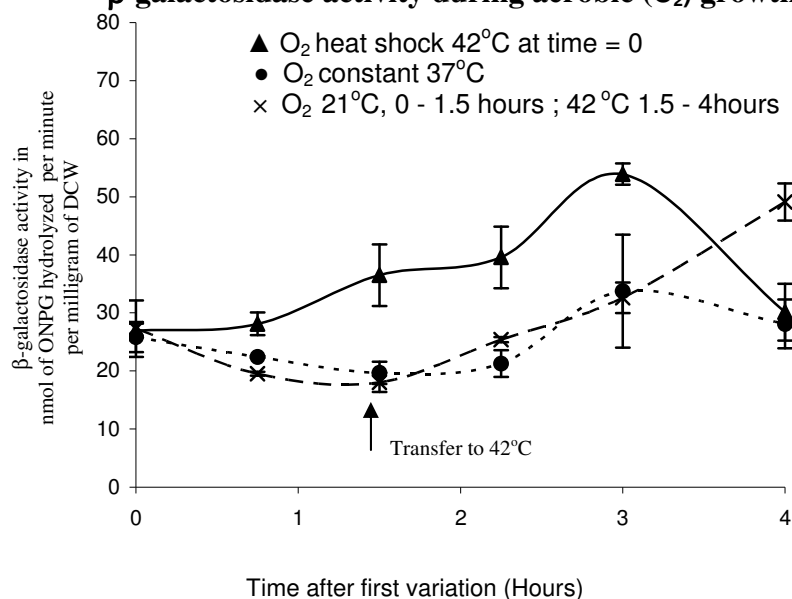
The results [Fig. 3.3] showed similar results compared with previous experiments. For cultures grown at 37°C then at 21°C and 42°C, data can be divided in two phases: the first phase lasted two hours following the temperature decrease to 21°C; continued by a second phase which lasted to the end of the experiment. The activities of the cultures transferred to 21°C (crosses connected with dashed lines) decreased slightly faster than those of the control cultures at 37°C [Fig. 3.3 (a)]. This probably occurs because σ^{32} levels are kept higher in the culture growing at 37°C than the culture transferred at 21°C. Growth curves showed that the cultures at 21°C grew slower than the control at 37°C [Fig. 3.3 (b)]. This downtrend levelled at the end of the 1h and 30 minutes period before the second variation. The second variation, consisting of transfer of cultures to from 21°C to 42°C incubators (indicated by the arrow), triggered two responses in the cultures: (1) reporter activity increased in every sample point exceeding the activities of the control group at 37°C; and (2) an increase in growth rate as seen in [Fig. 3.3 (b)]. The cultures heat-shocked at 42°C (triangles and solid lines) generated similar data as shown previously [Fig. 3.1], although some differences were apparent such as a growth decrease at 3 hours, possibly caused by culture to culture differences. The initial increase of activity did not appear as drastic as in [Fig. 3.1] which can be attributed to a higher initial activity carried from the overnight culture in this experiment [Fig. 3.3 (a)], usually a source of variation. In the final stages of growth there is some variability between duplicates revealed by larger error bars [Fig. 3.3 (b)]. This did not interfere with a decrease of activity towards the end of growth, probably due to acclimatisation to the new conditions as

previously seen. The data from the control cultures at 37°C [Fig. 3.3 (a)] was similar to previous data. There was a rise in activity at hour 3 which generally appeared towards the end of the exponential phase.

Effects of multiple temperature variation observed during anaerobic growth

The data shown in [Fig. 3.4] can be divided in two stages, similar to the aerobic group. The activity of the cultures transferred to 21°C [Fig. 3.4 (a)], shown in crosses connected by dashed lines, reveals a gradual decrease in activity not, as drastic as its aerobic counterpart. The control at 37°C, shown in filled circles, initially followed this decrease before an increase of variability between duplicates. The heat-shocked cultures, transferred to 42°C, exhibited a high increase of activity. The cultures subjected to dual variation were transferred to 42°C after 1.5 hours resulted in an increase of reporter activity and enhanced growth in a very similar manner as revealed by its aerobic counterpart. Towards the fourth hour the activities of the cultures at 42°C followed a downtrend while the activity of the control at 37°C, which was still growing [Fig. 3.4 (b)], seemed to gradually increase with time. On the other hand, the cultures subjected to single variation (heat-shock) exhibited the fastest growth. Compared with aerobic cultures, anaerobic cultures showed higher levels of reporter activity. However, accumulation of β -galactosidase was governed by the temperature increase, especially at heat-shock at 42°C.

a) β -galactosidase activity during aerobic (O_2) growth



b) Growth curves during aerobic growth

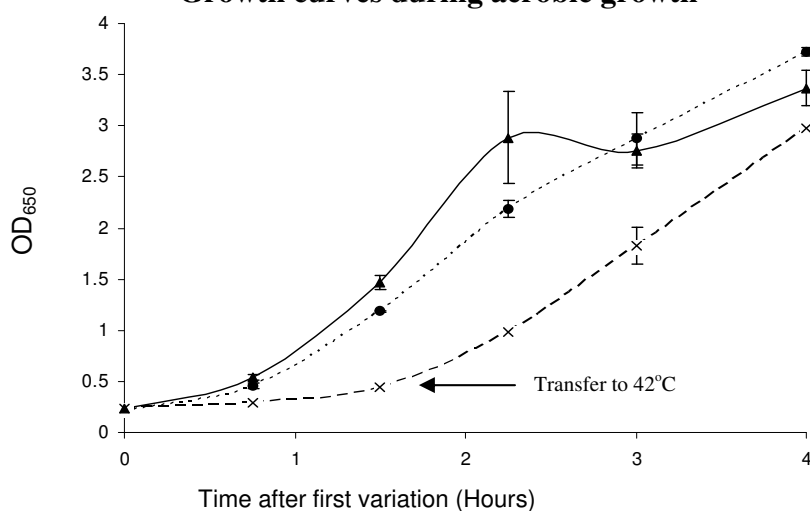
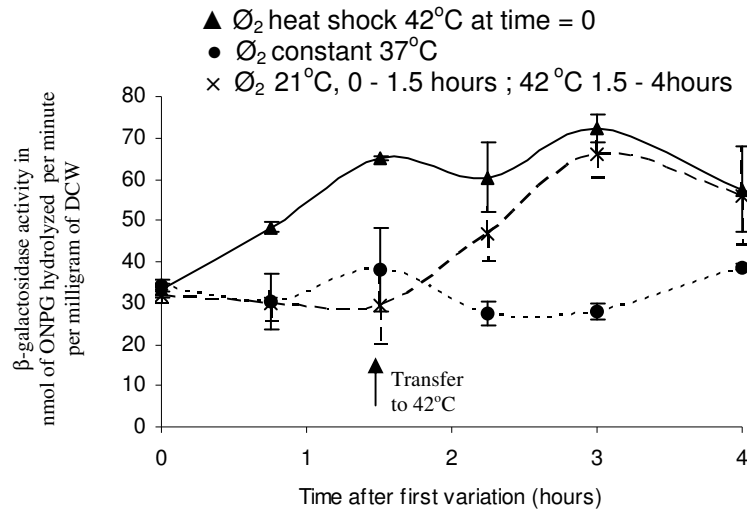


Fig. 3.3 β -galactosidase activity during aerobic growth at variable temperature

β -galactosidase activities (a) and growth measured by absorbance at 650nm (b) of *E. coli* JM109 pQF50KgroE. Solid circles connected by dotted lines represent the data of control cultures incubated at 37°C . Triangles connected by solid lines represent single heat-shock data. Crosses connected by dashed lines indicate data of cultures incubated at 21°C between 0-1.5 hours, and 42°C between 1.5-4 hours. The first variation is from 37°C to 21°C and second variation from 21°C to 42°C , indicated by an arrow.

a) β -galactosidase activity during anaerobic (O_2) growth



b) Growth curves during anaerobic growth

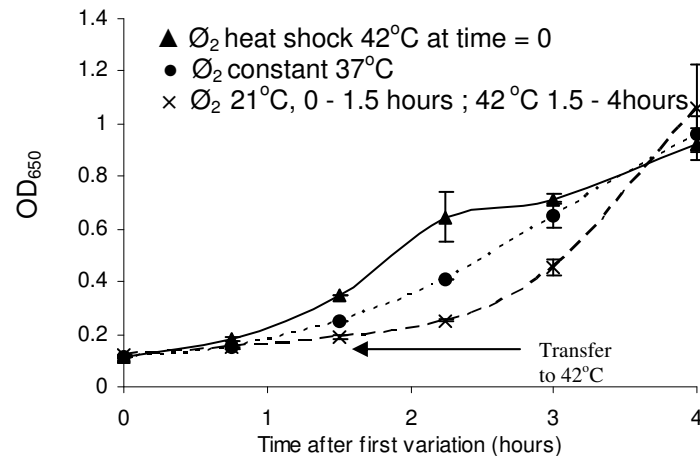


Fig. 3.4 β -galactosidase activity during anaerobic growth at variable temperature

β -galactosidase activities (**a**) and growth measured by absorbance at 650nm (**b**) of *E. coli* JM109 pQF50KgroE. Solid circles connected by dotted lines represent the data of control cultures incubated at constant 37°C . Filled triangles connected by solid lines represent single heat-shock data. Crosses connected by dashed lines indicate data of cultures subjected to double temperature variation. The first variation is from 37°C to 21°C and second variation from 21°C to 42°C , indicated by an arrow.

3.2.1.2 Reduction of activity in the absence of temperature changes

This experiment was performed to test if high β -galactosidase activities in the inoculum could be the cause of the gradual decrease in β -galactosidase activity in the initial stages of growth as seen in the previous experiments. *E. coli* JM109 transformed with pQF50KgroE was grown overnight and inoculated into 30ml of LB supplemented with $50\mu\text{g mL}^{-1}$ of kanamycin and 0.4% glucose. Aerobic and anaerobic control cultures were incubated at 37°C . In addition, a group of aerobic and anaerobic cultures were incubated at to observe possible accentuation of such effect. Due to the reduced growth rates found at 30°C , samples were taken at 1.5h interval for the first three hours and hourly intervals for the final two hours.

The results from the cultures at 37°C (represented by solid lines) [Fig. 3.5 (a)] were comparable to previous experiments with the exception of a higher activity carried forward from the inoculum variable from culture to culture. Aerobic and anaerobic cultures showed a characteristic decrease of activity during the first 1.5 hours, which continued to the 3rd hour. Data delivered similar trends as previously shown, such as an uptrend in anaerobic growth or a spike in activity of aerobic growth at the end of the exponential phase [Fig. 3.5 (b)]. The cultures at 30°C (represented by dotted lines) revealed very pronounced downtrends in activity compared with the 37°C cultures, the decrease in activity, in aerobic cultures, lasting longer and resulting in a much lower activity than that shown by anaerobic cultures. Interestingly, after the first 1.5h [Fig. 3.5 (a)] the average activity at 37°C was lower than that measured in the cultures at 30°C for aerobic and anaerobic cultures. Since the cultures at 37°C grow faster, the residual activity in the overnight could be diluted progressively in the cytoplasm, implying that the expression rate of the reporter gene is lower than the growth rate

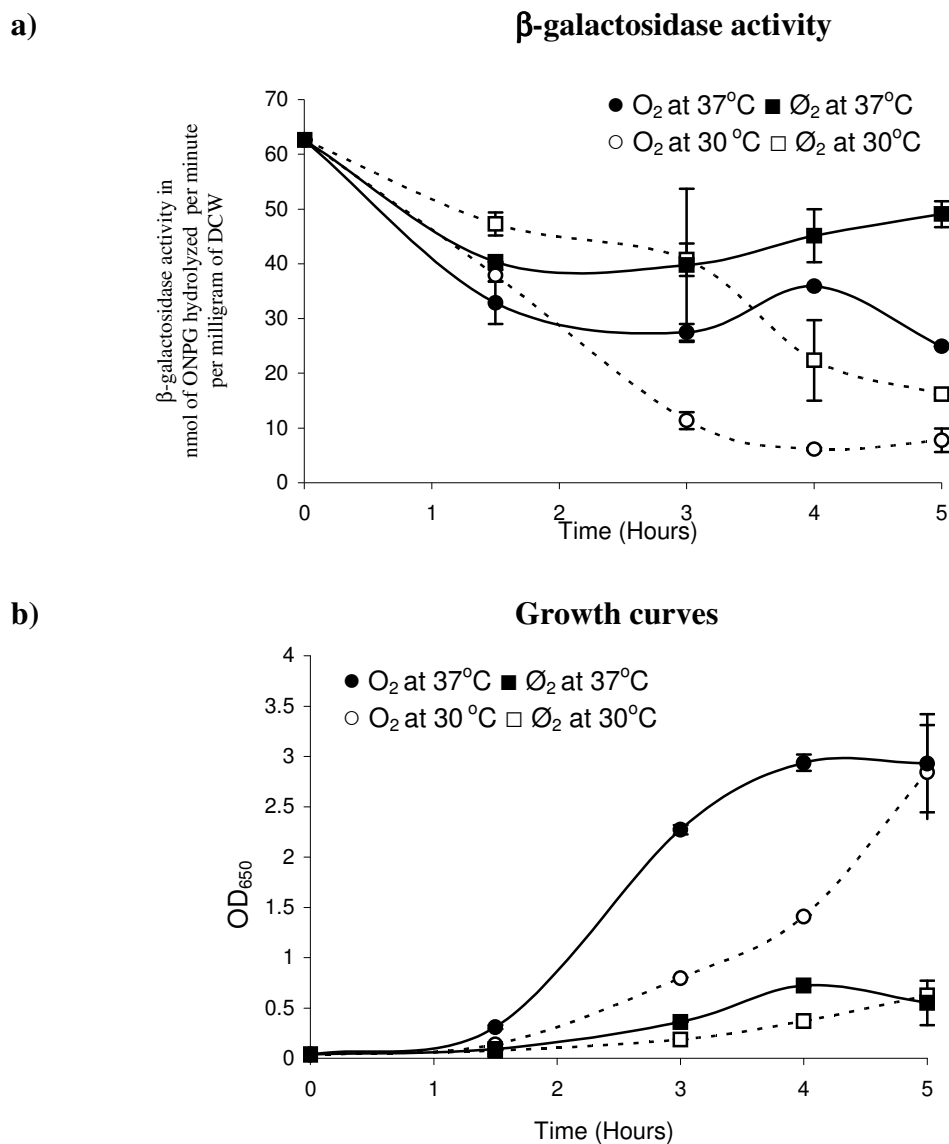


Fig. 3.5 β -galactosidase activity at constant temperature

β -galactosidase activities (**a**) and growth measured by absorbance at 650nm (**b**) during incubation at 30°C or 37°C. Data points of *E. coli* JM109 pQF50KgroE cultures at 37°C are connected by solid lines while data points from cultures at 30°C are connected by dotted lines.

at 37°C. In this sample point, the slow growth in anaerobic cultures was correlated with a slower reduction of β -galactosidase activity. β -galactosidase activity seen in inocula might be not only attributable to excessive accumulation but also to high $E\sigma^{32}$ levels. Thus overnight cultures are likely to be stressed; such possibility is discussed in the appropriate section. From the second hour, aerobic cultures at 37°C began to accumulate activity while the cultures at 30°C continued reducing activity over time. The cultures at 37°C had higher activities than those at 30°C. In addition, the activity of anaerobic cultures exhibited higher activity than the aerobic counterparts.

3.2.2 Response of *groE::lacZ* reporter to β -mercaptoethanol addition during growth

β -mercaptoethanol is a reductant which has been reported to be able to ingress the cytoplasm of *E. coli* (Malpica et al., 2004). Thiol reductants have been utilized before in a technique called “cell conditioning” in which the chemical is utilized to trigger a mild stress response resulting in up-regulation of chaperone genes (Gill et al., 2001). The thiol reductant affected the regulation of the internal environment of the cell triggering a stress response which prepared the cell for recombinant protein expression, thus cell conditioning. While the experiments in this chapter investigated the effect of β -mercaptoethanol on the σ^{32} regulon using the reporter system *groE::lacZ*, in succeeding chapters DTT (dithiothreitol) was also studied.

Expression of lacZ in response to increasing concentrations of β -mercaptoethanol

This experiment was carried out in order to observe the response of *E. coli* to a range of different concentrations of the reductant β -mercaptoethanol. The aim was obtaining a concentration of β -mercaptoethanol sufficient to cause a significant increase of β -galactosidase activity. *E. coli* JM109 transformed with pQF50KgroE was grown overnight and inoculated into 30ml of LB supplemented with $50\mu\text{g mL}^{-1}$ of kanamycin and 0.4% glucose. All the cultures were incubated aerobically and at 30°C in order to minimize interference with temperature triggered *groE::lacZ* response. After 1.5h, β -mercaptoethanol was added to different cultures at final concentrations of 5mM, 10mM and 25mM. All work was performed on duplicate cultures. Control cultures consisted of replicates without β -mercaptoethanol addition. The first sample was taken 1.5 hour after the addition of β -mercaptoethanol; beyond this point extra samples were taken at 1 hour intervals.

The β -galactosidase activities of the stressed cultures is indicated in [Fig. 3.6 (a)] by solid lines. 1.5 hours after β -mercaptoethanol addition, the cultures at 5mM and 10mM decreased their activity in a similar trend to the control. During the final two hours, these cultures remained on activity levels similar to the control (shown by a dotted line). The 10mM culture had higher activity than the 5mM throughout the experiment. The cultures at 25mM had a completely different activity trend after the addition of β -mercaptoethanol. These cultures showed a gradual increase of activity in consecutive sample points. This could be attributed to β -galactosidase accumulation due to a greatly suppressed growth rate [Fig. 3.6 (b)]. Such growth inhibition was shown to be dose dependant as reductant was added. Growth reduction upon addition of 25mM of β -mercaptoethanol may still be attributed to the combination of two factors, firstly, temperature, since *E. coli* JM109 showed slow growth at 30°C ; secondly,

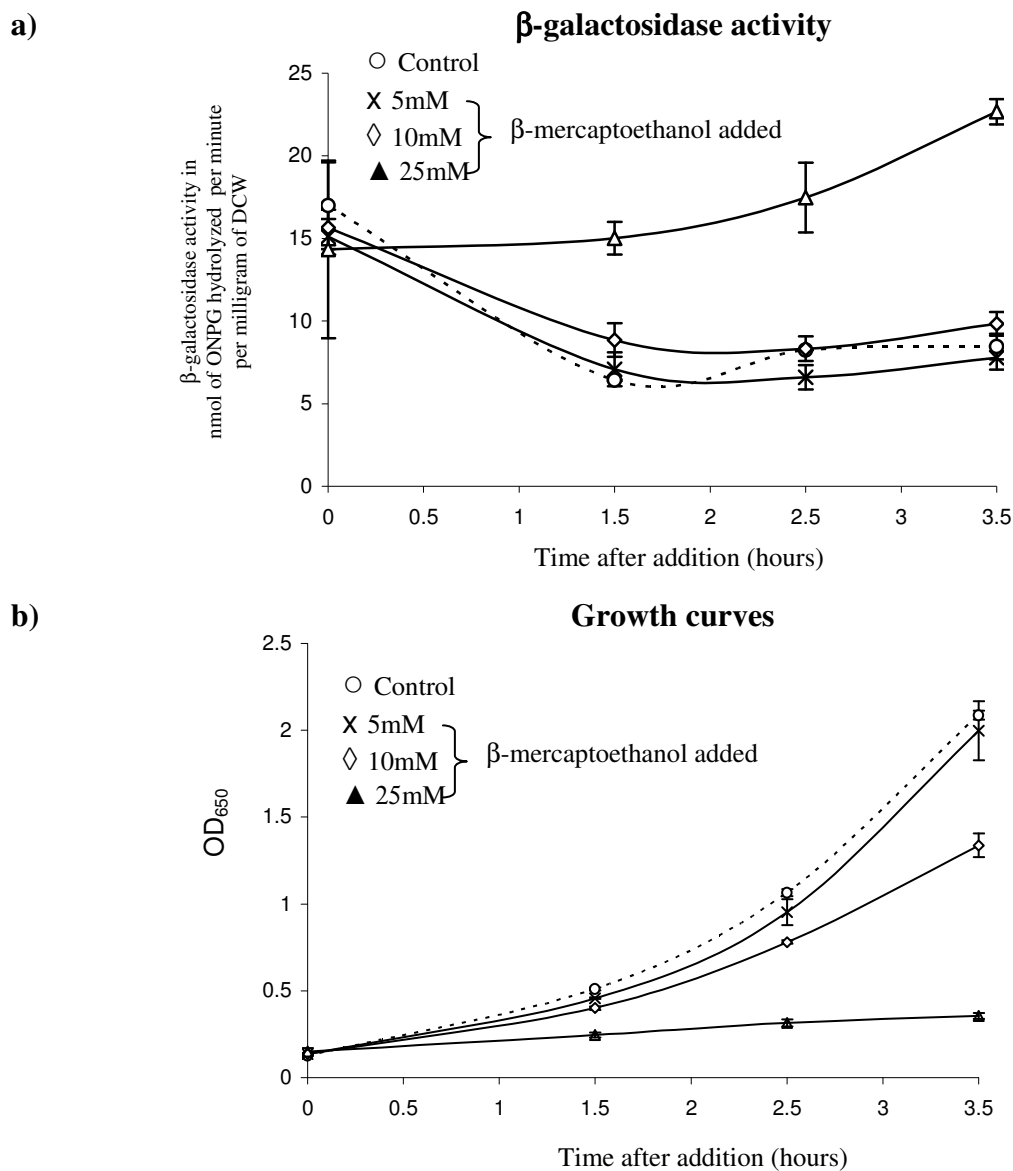


Fig. 3.6 *E. coli* JM109 pQF50KgroE response to β -mercaptoethanol as stress agent

β -galactosidase activities (**a**) and growth measured by absorbance at 650nm (**b**) The data of β -mercaptoethanol cultures is connected by solid lines while the control, with no β -mercaptoethanol, connected with dotted lines. The temperature in this experiment was constant at 30°C for all cultures.

the toxicity of β -mercaptoethanol could be exacerbated by an early addition to the culture. In spite of this, the culture increased in biomass and β -galactosidase was synthesised.

Reporter response to combination of heat-shock and β -mercaptoethanol

The effect previously observed upon addition of β -mercaptoethanol was further tested. The aim was to observe the β -mercaptoethanol effect in different circumstances and if the heat shock response could interfere or merge with the β -mercaptoethanol response. This experiment was performed in aerobic conditions. All cultures were allowed to grow for 1h at 37°C; heat-shock cultures were then transferred to 42°C incubators. The following hour after heat shock, cultures at 37°C and 42°C were supplemented with 25mM β -mercaptoethanol. The experiment incorporated a control growing at constant 37°C and a heat shock only, with no added β -mercaptoethanol. All work was performed in duplicates. The first three samples were taken at 1 hour intervals and every 30 minutes afterwards.

The activities of cultures growing at constant 37°C [Fig. 3.7 (a)], showed by dotted lines, decreased activity during the first hour. Upon β -mercaptoethanol addition (shown by the arrow), the cultures showed increase of activity in the second hour, whereas the control continued to decrease in activity. The heat-shocked cultures, shown by solid lines, did not decrease in β -galactosidase activity immediately after the temperature shift. However, cultures transferred at 42°C accumulated, as expected, more β -galactosidase activity than the cultures at 37°C. The heat-shocked cultures with added β -mercaptoethanol showed faster accumulation rate of β -galactosidase, in an additive manner, than its counterparts without

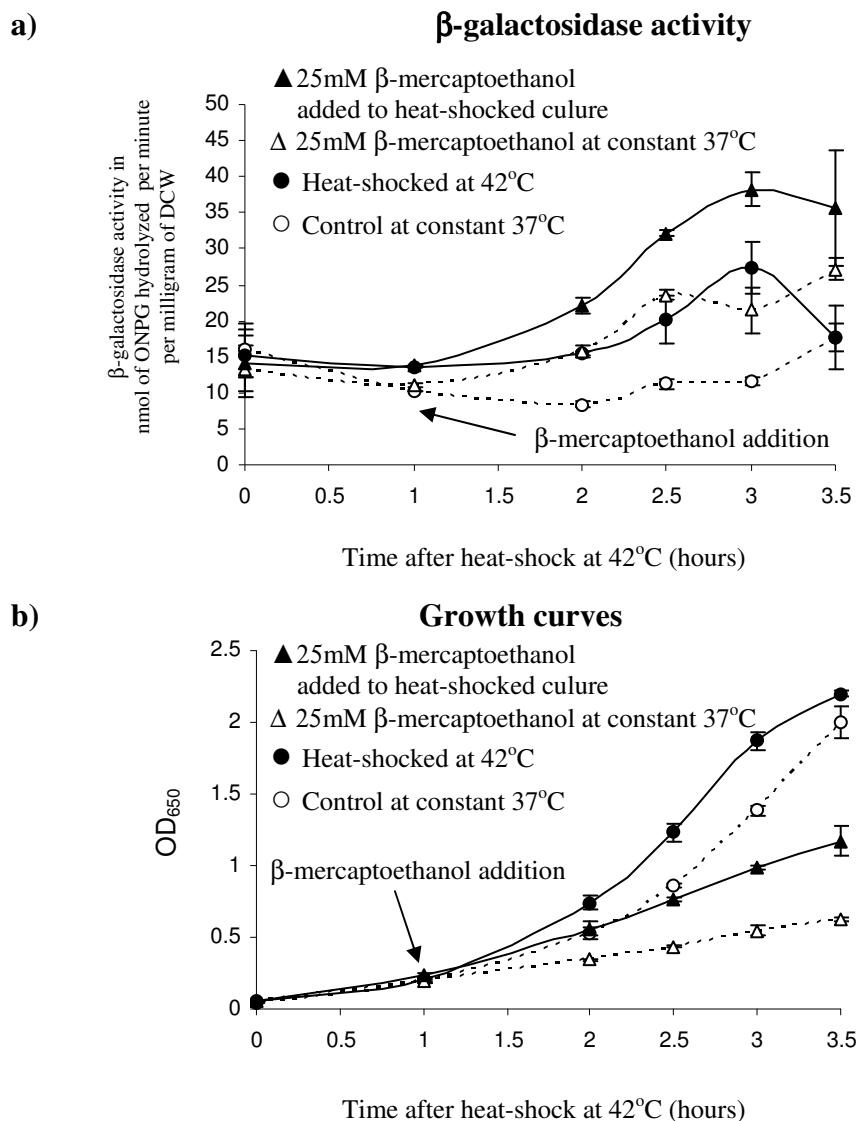


Fig.3.7 Combination of β -mercaptoethanol with heat-shock stress in aerobic cultures

β -galactosidase activities (**a**) and growth curves measured by absorbance at 650nm during aerobic growth (**b**) of *E. coli* JM109 pQF50KgroE. All cultures were allowed to grow for one hour prior to heat-shock at 42°C, heat shocked are represented by solid lines. One hour after heat-shock cultures were stressed with β -mercaptoethanol to 25mM final concentration (indicated with the arrow)

thiol reductant. The cultures at 37°C with β -mercaptoethanol exhibited activities comparable to the heat-shocked cultures, if not greater, at the end of the experiment. The growth curves, shown in [Fig. 3.7 (b)] are typical for the cultures absent of reductant where the heat-shocked culture exhibited a faster growth than the control at 37°C. Cultures challenged with β -mercaptoethanol showed constant growth at reduced rates. By comparing growth data of cultures to which β -mercaptoethanol had been added, cultures at 42°C grew faster than those at constant 37°C. This was probably caused by two circumstances: (1) cells were observed to grow faster at higher temperatures; and (2) heat-shocked cells may be better prepared to respond to β -mercaptoethanol because the σ^{32} regulon is already activated.

Response of groE::lacZ to β -mercaptoethanol addition during anaerobic growth

To observe if the response caused by β -mercaptoethanol was affected by anaerobic conditions, *E. coli* JM109 pQF50KgroE was inoculated into capped boiling tubes, containing 30ml of LB supplemented with 50 μ g mL⁻¹ of kanamycin and 0.4% glucose. All cultures were incubated for 1.5 hours at 37°C prior to transfer of the heat-shock group to an incubator at 42°C. Half an hour later β -mercaptoethanol (25mM final concentration) was added into cultures. Two control groups with no addition were allowed to grow at 37°C and 42°C as an only heat-shock control.

The results for β -galactosidase activity are shown in [Fig. 3.8 (a)]. The group at 37°C, represented by dotted lines, showed the characteristic initial decrease of activity, as seen previously, where the activity in the inoculum was relatively high. During anaerobic growth the 37°C cultures with β -mercaptoethanol (empty triangle series) showed activities slightly lower than the control group (empty circles). This contrasted with the results obtained during

aerobic growth [Fig. 3.7], where β -mercaptoethanol addition increased activities. There are two possible explanations for this: first, greater tolerance of anaerobically growing cells to β -mercaptoethanol (corroborated by the less growth inhibition anaerobically); and second, the higher β -galactosidase activity anaerobically might mask activation by β -mercaptoethanol. It is noteworthy that β -galactosidase activity in anaerobic β -mercaptoethanol cultures at 37°C was comparable to that in control cultures in spite of different growth rate indicating independency between activity and growth rate. In contrast with aerobic cultures [Fig. 3.7], the increase in β -galactosidase activity in response to anaerobic metabolism may have the same underlying trigger as the β -mercaptoethanol response resulting in the merging of both stresses rather than addition.

The heat-shocked cultures, shown by solid lines revealed an increase of activity after transfer to 42°C. This point marked a spike of activity by the single heat-shock group (filled circles) which reduced its activity in the following 1.5 hours. The cultures at 42°C increased activities in synchrony but those with β -mercaptoethanol (filled triangles) resulted in higher activity than the heat-shock alone visible at 3 hours [Fig. 3.8 (a)]. This *lacZ* expression pattern could be attributed: firstly, as consequence of heat-shock; and secondly, maintained due to the presence of β -mercaptoethanol. Cultures absent of the chemical stress proliferated better than those containing the reductant. Compared with aerobic growth [Fig. 3.7 (b)], the anaerobic cultures seemed to tolerate better the toxic effect of β -mercaptoethanol. Notwithstanding this fact, cultures exhibited similar growth rates between 1.5 and 3 hours [Fig. 3.8 (b)] showing that this concentration of β -mercaptoethanol is likely to depress growth rate rather than kill cells. Comparing the β -mercaptoethanol added groups, the cultures shocked at 42°C reached

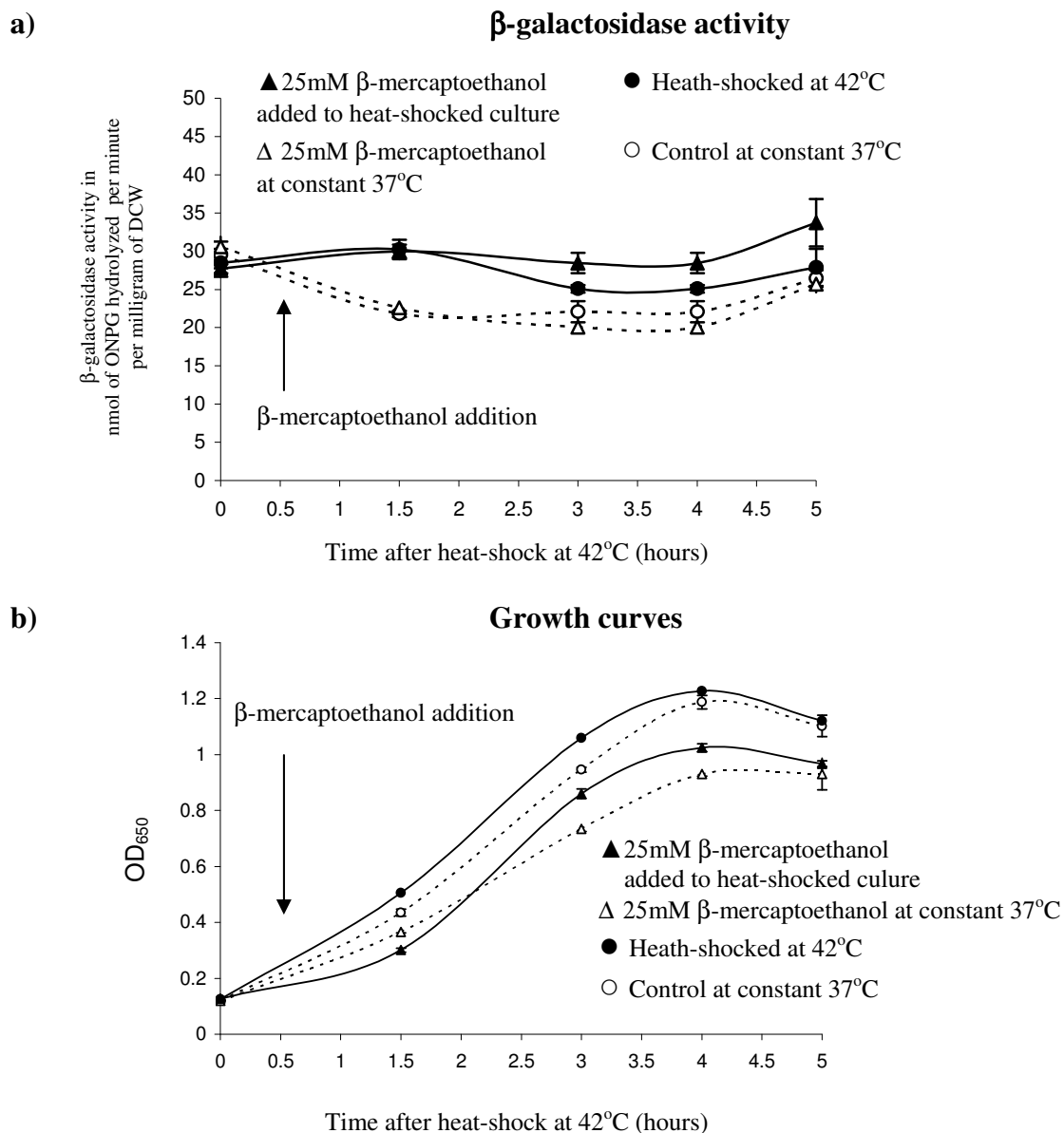


Fig. 3.8 Combination of β -mercaptoethanol with heat-shock stress in anaerobic cultures

β -galactosidase activities (**a**) and the growth curves measured by absorbance at 650nm during aerobic growth (**b**) of *E. coli* JM109 pQF50KgroE. All cultures were allowed to grow at 37°C for 1.5 hours prior to heat-shock. Half hour after heat-shock β -mercaptoethanol was added to a concentration of 25mM, this moment is indicated by the arrow.

higher biomass than the culture at 37°C. This could be attributed to two causes: (1) cultures were observed to grow better at the higher temperature; and (2) heat-shock generated a physiological response that enabled the cells to respond better to the chemical β -mercaptoethanol stress resulting in improved growth.

Expression of $groE::lacZ$ and β -mercaptoethanol influence in late stationary phase

Based on previous experiments, especially in data points towards the end of the growth, it was speculated that $groE::lacZ$ expression could be affected by stationary phase metabolism. Many industrial products are associated with the stationary phase such as enzymes or metabolites bringing interest into monitoring its σ^{32} -associated stress. To observe reporter behaviour in late stationary phase, 30mL of aerobic and anaerobic cultures were inoculated with *E. coli* JM109 pQF50KgroE after inoculation. Cultures were allowed to grow overnight and stressed with 25mM β -mercaptoethanol 3 hours before completion of the 24 hour period. Three samples were taken at 1.5 hour intervals. Aerobic and anaerobic controls, with no β -mercaptoethanol added, were included in the experiment. In addition to β -galactosidase activity assay, samples were tested for colony forming units and plasmid retention.

The aerobic cultures stressed with β -mercaptoethanol (filled triangles) exhibited lower β -galactosidase activities than the unstressed cultures (empty circles) [Fig. 3.9 (a)] before the addition of β -mercaptoethanol which could be attributed to culture to culture variability. β -mercaptoethanol did not increase the activity in the late stationary phase of these cultures. Colony forming units (CFU) were broadly comparable and average plasmid retention, important for industrially used reporter systems, was maintained above 80% for both

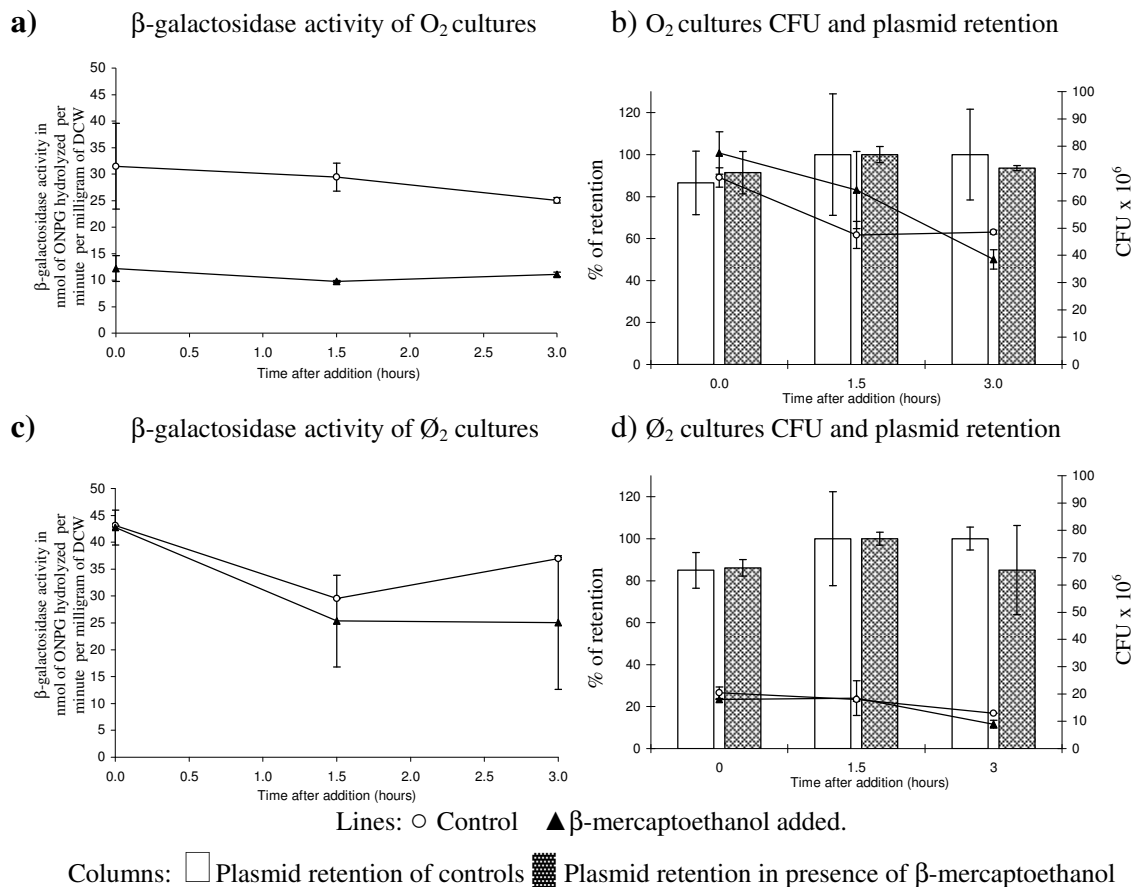


Fig. 3.9 Reporter activity in presence of β -mercaptoethanol in late stationary phase

All cultures of *E. coli* JM109 pQF50KgroE were incubated at constant 37°C. Data of cultures stressed with β -mercaptoethanol (25mM) are indicated with filled triangles.

(a) shows β -galactosidase activity measured from aerobic cultures (O_2) and (b) the colony forming units (CFU) and plasmid retention of aerobic cultures. (c) shows β -galactosidase activity in anaerobic cultures (Ø_2) and (d) CFU and plasmid retention of anaerobic cultures.

conditions [Fig. 3.9 (b)]. The activity observed in stressed anaerobic cultures [Fig. 3.9 (c)] was equivalent to non stressed cultures in the first sample taken suggesting less culture to culture variability. The second and third samples reveal a divergence of average activities but with very wide and overlapping error bars. *E. coli* JM109 sedimented at the bottom of the tubes requiring mixing for sample extraction which could be a source of variability. Broad differences of activities between duplicate complicate the conclusions from this experiment. However, because the reporter plasmid was maintained and CFU resulted in viable colonies it is likely that β -mercaptoethanol was not causing stress on the bulk of the culture. A plausible explanation for this could be that at late stationary phase 25mM of β -mercaptoethanol effect is diluted by higher culture density.

3.2.3 Involvement of σ^E on *groE::lacZ* expression upon β -mercaptoethanol addition

The *rpoH* promoter regulating expression of σ^{32} contains a binding site for σ^E (Kallipolitis and Valentin-Hansen, 1998). Therefore, can be speculated that the previously observed β -mercaptoethanol response could be influenced by an increase of σ^E as a quantitative response to periplasmic stress. The investigation of possible influence of σ^E on the σ^{32} -mediated expression of *groE::lacZ* could show if the heat-shock like stress response caused by β -mercaptoethanol is the result of a multiple challenge to the periplasm and the cytoplasm. To observe the influence of σ^E on reporter expression, data collected from cultures of *E. coli* ECA101 ($\Delta\sigma^E$) was compared with data from the isogenic *E. coli* BW25113 (WT) from the Keio collection (Baba et al., 2006). The knockout mutant $\Delta\sigma^E$ has shown reduced growth in comparison with the wild type upon temperature increases and has shown lack of growth at temperatures above around 30°C (Rouviere et al., 1995).

Expression of groE::lacZ in the strains BW25113 and ECA101 in response to 25mM β -mercaptoethanol

This experiment was performed to observe if the parameters utilized in *E. coli* JM109 can be extrapolated to BW25113 and ECA101 transformed with pQF50KgroE. To study the effect of β -mercaptoethanol, overnight cultures were inoculated into conical flasks containing 30ml of LB, supplemented with 50 μ g mL⁻¹ of kanamycin and 0.4% glucose, and allowed to grow for one hour before β -mercaptoethanol addition to 25mM. Controls with no β -mercaptoethanol added were monitored simultaneously. Samples were taken at 45 minutes intervals except for the last sample which was taken after a one hour interval from the previous. The growth temperature was 37°C throughout. All work was performed on duplicate cultures.

The $\Delta\sigma^E$ culture with β -mercaptoethanol addition showed an increase of activity at 0.75 hours compared with the $\Delta\sigma^E$ control [Fig. 3.10 (a)]. The WT did not show a clear effect. A possible explanation is that these strains are more resistant to β -mercaptoethanol than JM109; in general, the growth inhibition caused by β -mercaptoethanol was less acute than that observed in *E. coli* JM109. The WT control cultures showed higher levels of activity than the $\Delta\sigma^E$ cultures similar to the stressed cultures. It is possible that this was caused by additional expression of *rpoH* through the σ^E -dependent promoter. All cultures reached a similar level of activity towards the stationary phase. The $\Delta\sigma^E$ cultures exhibited a slight decreased growth rate.

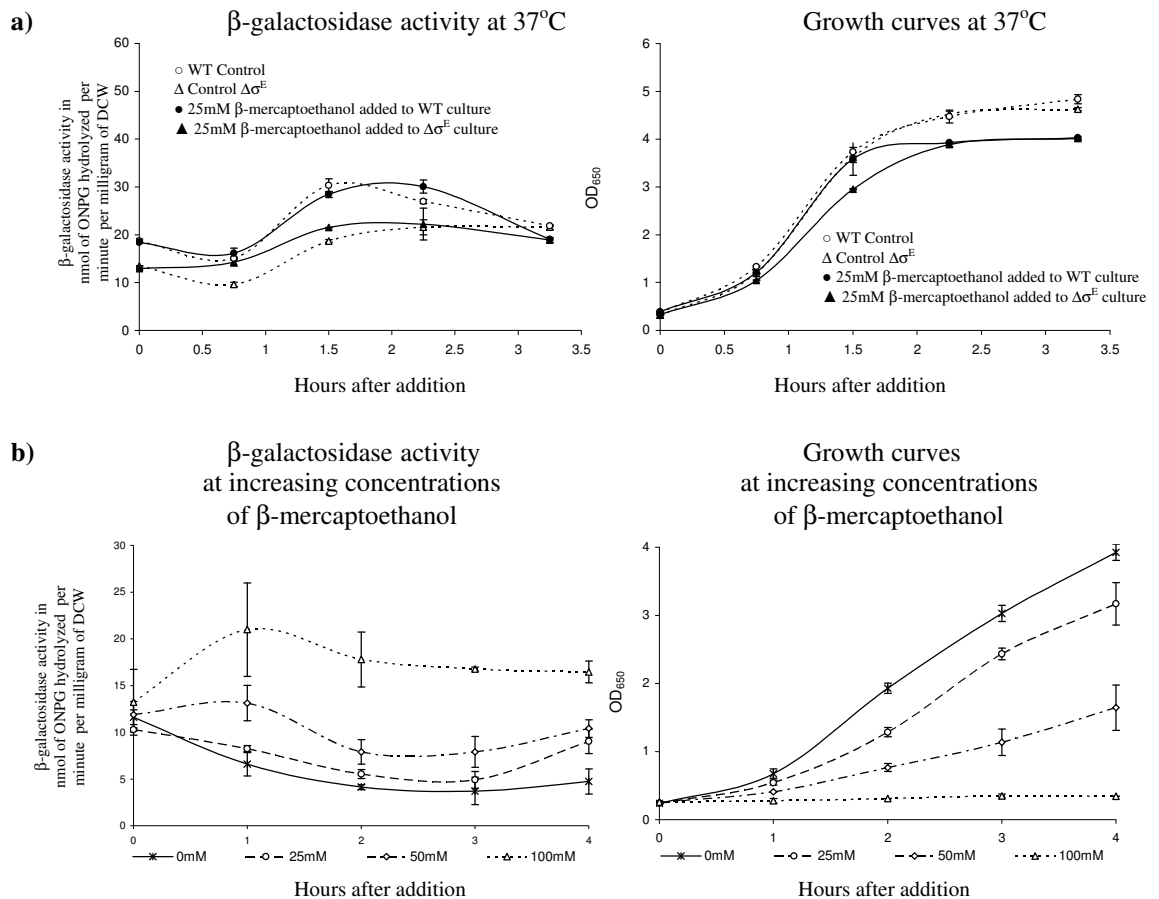


Fig 3.10 Influence of σ^E on β -galactosidase activity.

(a) β -galactosidase activities of *E. coli* BW25113 pQF50KgroE and the σ^E knockout mutant *E. coli* ECA101 harbouring pQF50KgroE (left) and growth curves (right).

Cultures were allowed to grow for 1 hour prior to β -mercaptoethanol addition to a concentration of 25mM. (b) The β -galactosidase activities (left) and growth (right) of σ^E knockout mutant *E. coli* ECA101 with β -mercaptoethanol addition added at different concentrations. The experiment was performed at 30°C with β -mercaptoethanol added one hour after inoculation. While activity seemed to increase accordingly with the β -mercaptoethanol concentration it was observed a strong bacteriostatic effect, especially at 100mM.

Expression of $groE::lacZ$ in $\Delta\sigma^E$ at increasing concentrations of β -mercaptoethanol

In order to find an optimal concentration of β -mercaptoethanol to stimulate expression of $groE::lacZ$; *E. coli* ECA101, as least tolerant strain, was challenged with different concentrations of β -mercaptoethanol. Overnight cultures were inoculated into conical flasks containing 30ml of LB, supplemented with $50\mu\text{g mL}^{-1}$ of kanamycin and 0.4% glucose, and allowed to grow for one hour before addition of β -mercaptoethanol to concentrations of 25mM, 50mM, and 100mM. Control cultures were grown without addition of reductant. This experiment was conducted at constant 30°C in order to prevent interference by the heat-shock response. Samples were taken at hourly intervals. One hour after 100mM β -mercaptoethanol was added cultures showed a sharp increase of activity [**Fig. 3.10 (b)**]. However, shown in the right graph, growth was inhibited in this culture. At 50mM the cultures shows an increase of activity while the cells continue to grow at a reduced rate. Addition at 25mM produced higher activity than the control.

Expression of $groE::lacZ$ in $\Delta\sigma^E$ and WT at 30°C

To compare the response of the $\Delta\sigma^E$ and WT strains to β -mercaptoethanol, overnight cultures were inoculated into conical flasks containing 30ml of LB, supplemented with $50\mu\text{g mL}^{-1}$ of kanamycin and 0.4% glucose, and allowed to grow for one hour before addition of 50mM β -mercaptoethanol. Control cultures were monitored without reductant addition. Cultures were monitored at 45 minutes intervals except for the last sample which was taken after one hour interval. During the first 45 minutes after β -mercaptoethanol addition the activity of all cultures decreased [**Fig. 3.11 (a) left**], similar to observed previously. The activity of WT

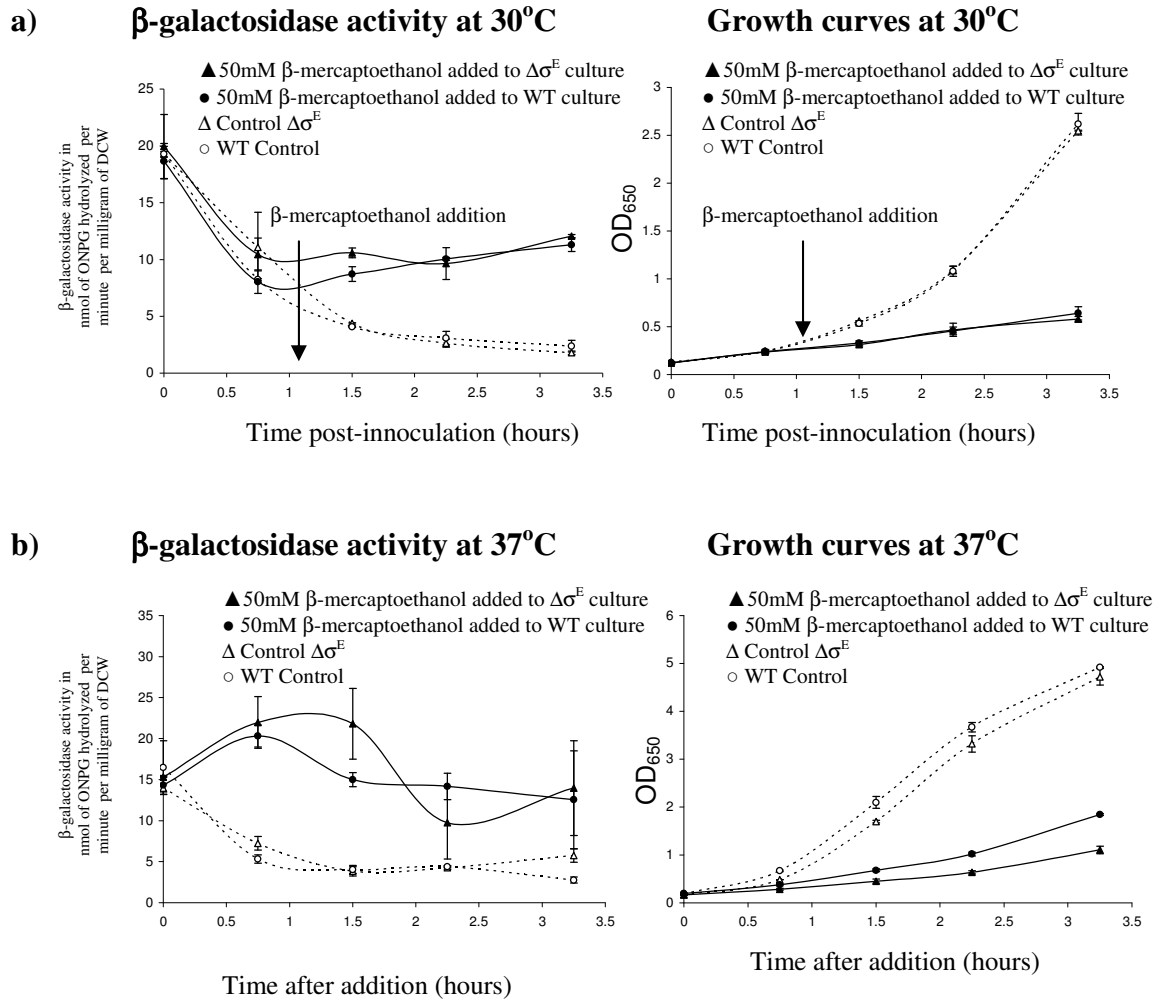


Fig. 3.11 Reporter expression at 30°C and at 37°C of E.coli ECA101 ($\Delta\sigma^E$) strain and BW25113 (WT) harbouring pQF50KgroE upon β -mercaptoethanol addition (50mM)

Control cultures are represented by dotted lines while β -mercaptoethanol with solid lines.

(a) Shows β -galactosidase activity (left) and growth curves measured by absorbance at 650nm (right) of cultures incubated at 30°C with addition of β -mercaptoethanol (arrow).

(b) Shows β -galactosidase activity (left) and growth curves measured by absorbance at 650nm (right) of cultures incubated at 37°C.

cultures seemed to decrease faster than the activity of $\Delta\sigma^E$ in spite of the fact that both cultures exhibited same growth [Fig. 3.11 (a) right]. This could be explained by an increase of $E\sigma^{32}$ in $\Delta\sigma^E$ cultures due to an absence of competition of σ^E for the core RNAP. The effect of β -mercaptoethanol was not significantly different between WT and $\Delta\sigma^E$ in terms of *groE::lacZ* expression or in growth inhibition. Cultures induced with β -mercaptoethanol continued accumulating β -galactosidase for the last three hours in WT and $\Delta\sigma^E$ while control cultures exhibited a gradual decrease in activity.

Expression of groE::lacZ in $\Delta\sigma^E$ and WT at 37°C

This experiment was performed in the same manner as stated previously but the incubation in this case was at 37°C. The reporter activity data [Fig. 3.11 (b) left] revealed a rise in activity in test cultures (shown in solid lines) higher than in the controls (dashed lines), 45 minutes after challenge with β -mercaptoethanol. After this initial increase, the activity of WT cultures reduced the activity while $\Delta\sigma^E$ cultures maintained high activity. After 1.5 hours, the 50mM β -mercaptoethanol $\Delta\sigma^E$ cultures had reduced the activity drastically returning to a similar level as WT in the last hour. The growth curves [Fig. 3.11 (b) right] revealed growth inhibition in the cultures with β -mercaptoethanol where it can be seen that while the increased temperature enhanced the growth rate of the WT cultures, the incubation temperature was not as beneficial to the growth rate of $\Delta\sigma^E$ in the controls and it could have exacerbated the effect of β -mercaptoethanol in comparison with data at 30°C. This is consistent with the observations made by Rouviere et al. (1995) where the $\Delta\sigma^E$ phenotype was shown to grow with difficulties at temperatures higher than 30°C.

Expression of groE::lacZ in $\Delta\sigma^E$ and WT at 42°C

This experiment was performed to test temperature limits of *E. coli* ECA101 and the effect of high growth temperature combined with β -mercaptoethanol. Cultures were prepared as described previously but instead of being heat-shocked, cultures were incubated at 42°C throughout. 50mM β -mercaptoethanol was added after 1h. Control cultures were monitored without reductant addition. Cultures were samples at 45 minutes intervals except for the last sample which was taken after one hour interval.

The activity in $\Delta\sigma^E$ mutants [**Fig. 3.12 (a)**] increased 45 minutes after stress followed by a decrease in activity for 1.5 hours. The final hour cultures exhibited a sharp increase of activity. The growth curves (in the right) show the detrimental effect of temperature on ECA101 biomass. The $\Delta\sigma^E$ control culture grew up to an $OD_{650} = 3.5$ at a comparable rate to that of the WT but, in contrast, it showed loss of biomass during the last hour. Interestingly, β -mercaptoethanol cultures appeared to continue growing. This could be explained by a phenomenon of selection and conditioning of heat-shock resistant individuals caused by β -mercaptoethanol. The WT cultures challenged with β -mercaptoethanol showed an increase of activity [**Fig. 3.12 (b)**], higher than $\Delta\sigma^E$ cultures and WT cultures at 37°C [**Fig. 3.11 (b)**]. The activity of the WT β -mercaptoethanol cultures decreased 1.5 hours after addition [**Fig. 3.12 (b)**] but was maintained above the control cultures throughout the experiment. There was a growth rate reduction [**Fig. 3.12 (b) right**] the β -mercaptoethanol supplemented cultures when compared with the control. However, the increases in β -galactosidase activity were much larger than the decreases in growth rate.

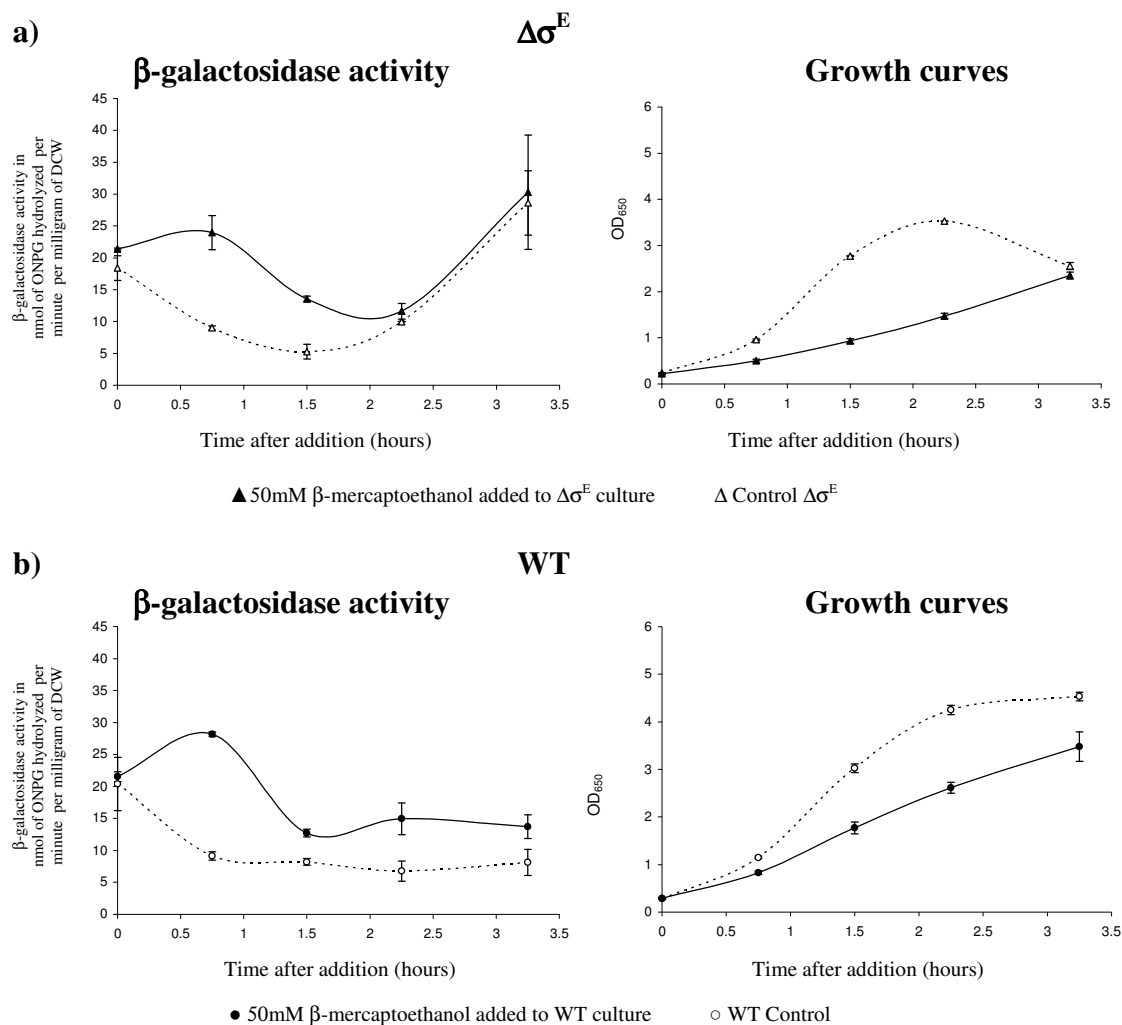


Fig. 3.12 Reporter expression at 42°C of *E.coli* ECA101 ($\Delta\sigma^E$) strain and BW25113 (WT) harbouring pQF50KgroE upon β -mercaptoethanol addition (50mM)

Control cultures are represented by dotted lines while β -mercaptoethanol with solid lines. **(a)** Shows β -galactosidase activity (left) and growth curves measured by absorbance at 650nm (right) of *E.coli* ECA101 ($\Delta\sigma^E$) cultures **(b)** Shows β -galactosidase activity (left) and growth curves measured by absorbance at 650nm (right) of *E.coli* BW25113 (WT) cultures.

3.3 Discussion

3.3.1 Reservoirs of σ^{32} : the effect of temperature

The σ^{32} protein is encoded by the gene *rpoH*. Most of *rpoH* transcription is regulated by three σ^{70} -dependent promoters dedicated to maintaining σ^{32} mRNA levels at a wide range of temperatures (Kallipolitis and Valentin-Hansen, 1998). The *rpoH* expression rate does not increase with temperature (Kamath-Loeb and Gross, 1991). It is the translation rate into σ^{32} that is influenced by temperature variations. This mRNA is folded in a stable secondary structure and constitutes a reservoir of potential σ^{32} . The secondary structure allows some translation into σ^{32} at low temperatures (21-30°C) however the bulk of the transcript accumulated in such form occurs at higher temperatures (37-42°C) by relaxation of the secondary structure (Morita et al., 1999). Once translated, σ^{32} is promptly degraded by the membrane-bound FtsH protease being the main cause of its low half-life (less than one minute) (Tatsuta et al., 1998). This degradation rate constitutes a major sink of σ^{32} . However, the chaperone DnaK and its co-chaperone DnaJ bind σ^{32} reversibly inactivating it which originates a reservoir of latent σ^{32} (Gamer et al., 1996; Tomoyasu et al., 1998).

The reporter plasmid pQF50KgroE needs $E\sigma^{32}$ to express β -galactosidase which implies that active σ^{32} is present and an RNAP core is available (Wang and deHaseth, 2003). Most of the effects observed upon temperature variations in the present study, can be explained by how this variable governs the transference of σ^{32} between these two reservoirs: potential σ^{32} (in form of RNA) and latent σ^{32} -DnaK-DnaJ complex. For example experimental data featuring a heat-shocked aerobic cultures [Fig. 3.1 (a)] revealed an initial spike of activity possibly due to

mass stabilization and translation of *rpoH* mRNA. This is followed by a decrease of $E\sigma^{32}$ presence which could be explained by capture of active σ^{32} by excess of DnaK, produced during the heat-shock response, degradation of idle σ^{32} , or both. Interestingly, heat-shocked anaerobic cultures [Fig. 3.2 (a)], resulted in sustained activity which did not return over time to similar levels as control cultures like aerobic cultures did. Diaz-Acosta et al. (2006) showed by L-[S³⁵] methionine labelling and SDS-PAGE that the response to heat-shock on anaerobic conditions resulted in faster production of chaperone proteins which was sustained for longer periods than in aerobic cultures; this is demonstrated here at the level of transcription regulation for the first time.

The expression of *groE::lacZ* was observed to be constantly higher in anaerobiosis than in aerobiosis. Anaerobic cultures may develop a cytoplasmic environment that requires additional σ^{32} -dependant chaperones at elevated temperatures. Aerobic metabolism produces great amount of reducing potential during the TCA cycle in form of NADH. In the cytoplasm, disulfide bonds are established utilizing the reducing potential of NADH via thiorredoxin or glutaredoxin pathways (Stewart et al., 1998) which could result in a less favourable environment for disulfide bond formation leading to misfolding, hence a need for higher titres of $E\sigma^{32}$. During either aerobic or anaerobic conditions, the β -galactosidase was decreased upon entering in the stationary phase which could be attributed to σ^S interference by weak competition for the RNAP core (Gruber and Gross, 2003) or nutrient starvation (Rosen and Ron, 2002).

Inocula typically had elevated β -galactosidase activities, leading to a decrease in β -galactosidase activity over the first few hours of each experiment. This was especially apparent when cultures were grown at lower temperatures than the overnight (37°C). Experiments showed this feature when cultures were transferred to lower temperatures [Fig. 3.4 (a)] or grown at lower temperature than the inocula [Fig. 3.5 (a)]. The overnight cultures were likely to be stressed due to many factors such as oxygen limitation and starvation, known to be responsible for σ^{32} stress response (Rosen and Ron, 2002). As a result inocula may contain high concentrations of β -galactosidase which can be attributed to $E\sigma^{32}$ available. The concentration of reporter protein per cell decreased either by lack of expression in dividing cells, degradation of β -galactosidase or a combination of both in slowly dividing cells. This effect is most clearly shown in [Fig. 3.5]. After inoculation with an overnight at 37°C, cultures in lag phase (anaerobic and all cultures at 30°C), showed a reduction of activity in the first 1.5 hours possible due to β -galactosidase degradation while, in the same interval, the aerobic culture at 37°C achieved a faster loss of β -galactosidase most likely due to distribution of reporter enzyme during cell division. This effect was balanced by higher *groE::lacZ* activation at 37°C in comparison with 30°C growing cultures which can be seen after 3 hours. This effect was observed consistently in multiple reproductions of the experiments.

3.3.2 β -mercaptoethanol effect triggered a heat-shock like response

When increasing amounts of β -mercaptoethanol were added to aerobic cultures at 30°C it resulted in an increase in the rates of production of β -galactosidase [Fig. 3.6]. At a concentration of 25mM of β -mercaptoethanol, a gradual increase of activity was observed

over time. During growth at such temperature is likely that the β -galactosidase activity was reporting gradual σ^{32} release from DnaK-DnaJ caused by β -mercaptoethanol presence.

As a membrane permeating thiol reductant, β -mercaptoethanol was utilized by Malpica et al. (2004) to successfully reduce ArcB sensors where authors acknowledge a likelihood of causing protein misfolding. It was shown that in mammalian cells, better equipped to form disulphide bonds than *E. coli*, thiol reductants were a cause of protein misfolding (Gelman and Prives, 1996; Lodish and Kong, 1993). In the present study, the effect of β -mercaptoethanol was tested at different temperatures including in combination with heat-shock [Fig. 3.7], and it was shown that the increase of reporter activity caused by β -mercaptoethanol accumulated in an additive manner with the activity generated by a temperature increase. It is probable that heat-shock combined with β -mercaptoethanol resulted in an increase in titre of σ^{32} from its reservoirs, mRNA stabilization and translation upon temperature increase and σ^{32} release from DnaK-DnaJ via GrpE as β -mercaptoethanol disrupts disulphide bond formation in the cytoplasm. It is important to take in account that β -mercaptoethanol is lypophilic (Malpica et al., 2004) and its ability to permeate the membrane may increase with temperature. In addition, DnaK binding of σ^{32} has been shown to be affected by temperature (Chattopadhyay and Roy, 2002). The effect of β -mercaptoethanol was tested on anaerobic cultures [Fig. 3.8] where the impact of β -mercaptoethanol on heat-shocked cultures, based on reporter activity, was minor compared with aerobic cultures.

The cultures when heat-shocked resulted in increase of activity upon heat-shock [Fig. 3.1 (a)]. The decrease of activity that followed could be the result of faster growth rate than β -galactosidase expression rate which leads to less reporter enzyme accumulated per cell. The

decrease of reporter synthesis could be explained by capture of σ^{32} by DnaK-DnaJ produced in excess during heat-shock. Supporting this possibility, elevated temperatures have shown to attenuate DnaK and DnaJ degradation (Arsene et al., 2000). The cultures challenged with β -mercaptoethanol did not show such reduction. A possible explanation is by misfolded protein occupying the DnaK-DnaJ-GrpE cycle leading to higher σ^{32} titres, thus higher reporter activity. In the cultures at 37°C addition of β -mercaptoethanol did not cause an increase of β -galactosidase activity significantly greater than the control [Fig. 3.8] possibly due to the anaerobic - and β -mercaptoethanol- *groE::laxZ* activatory mechanisms overlapping meaning that anaerobic cells are already expressing genes of the σ^{32} regulon masking the action of the β -mercaptoethanol.

The plasmid pQF50KgroE is a low copy number plasmid which was ideal to monitor stress minimizing the burden to the cell caused by plasmid replication. The plasmid conservation was retained 24 hour post-inoculation even after addition of β -mercaptoethanol aerobically and anaerobically [Fig. 3.9]. The β -galactosidase activity observed in late stationary phase was not affected by β -mercaptoethanol [Fig. 3.9 (a)]. This could be attributed to a lower β -mercaptoethanol : cell ratio compared to a early cultures. Viable counts did not reveal great differences between cultures [Fig. 3.9 (b)]. On the other hand, the β -galactosidase activity and viable counts of anaerobic cultures were found to be very similar indicating that anaerobic cells could be more resistant to β -mercaptoethanol. This confirmed previous observations during exponential growth in anaerobic conditions [Fig. 3.8 (b)] where the negative effect of β -mercaptoethanol on growth rate was less accentuated than during aerobic growth [Fig. 3.7 (b)]

3.3.3 The σ^{32} response triggered by β -mercaptoethanol has been shown to be independent of σ^E

Stress responses in the envelope and the periplasm of *E. coli* are governed by σ^E . This alternate sigma factor binds to a specific promoter to transcribe *rpoH*. Due to the suspected mechanism of action of β -mercaptoethanol, it was possible that the heat-shock response was being influenced by envelope stress, as β -mercaptoethanol could be causing protein misfolding in the periplasm. In order to investigate this, the behaviour of a knockout mutant with a deletion of the *rpoE* gene, strain *E. coli* ECA101, from the Keio collection (Baba et al., 2006) was studied in comparison with the isogenic wild type, *E. coli* BW25113. The first considerations when utilizing the new strain was to determine if BW25113 responded to β -mercaptoethanol in a similar manner to *E. coli* JM109. The results indicated that *E. coli* BW25113 needed higher concentrations of β -mercaptoethanol (50mM) to produce a stress response in comparison with *E. coli* JM109 (25mM). This showed that there is strain to strain variability. The reporter activity upon β -mercaptoethanol addition at 30°C, shown in [Fig. 3.11 (a)], revealed a decrease of activity after inoculation followed by an increase of activity upon β -mercaptoethanol addition in the WT and the $\Delta\sigma^E$ cultures indicating that σ^E is not a major factor in σ^{32} accumulation as a consequence of β -mercaptoethanol. However, this does not rule out a possible σ^E response caused by β -mercaptoethanol; it is possible that if a different reporter is used, or a different measurement technique (*e.g.* microarrays), such a response could be detected. In addition, strain ECA101 is extremely sensitive to temperature and other stresses, so comparison of *groE::lacZ* activity with the wild type could be distorted by secondary effects. In fact, $\Delta\sigma^E$ cultures reduced β -galactosidase activity at a slower rate than WT which could indicate that free σ^{32} in σ^{E+} cells is competing with σ^E (Maeda et al.,

2000). It is noteworthy that expression of β -galactosidase in $\Delta\sigma^E$ cells could also be diminished as one of the promoters (P6 is not being utilized, thus resulting in less σ^{32} than that found in WT cultures. When the experiment was performed at 37°C [Fig. 3.11 (b)] both strains responded by increasing β -galactosidase activity to a greater level than at 30°C while control cultures at 37°C β -mercaptoethanol without showed a decrease in β -galactosidase activity. The cultures grown at 42°C resulted in a substantial increase of activity especially in *E. coli* BW25113 [Fig. 3.13 (b)]. As in mutant experiments performed by Rouviere et al. (1995), strain ECA101 exhibited difficulties to survive at 42°C resulting in cell death 4h after inoculation. However, β -mercaptoethanol treated cells appeared to continue growth showing enhanced resistance to the elevated temperature and accumulating β -galactosidase activity. Mutants lacking σ^E have been shown to accumulate compensatory mutations to endure high temperatures (Connolly et al., 1997). Still, the specific nature of such mutation is unknown (Dartigalongue et al., 2001). It is possible that β -mercaptoethanol enabled that selection by allowing survival of the next generation only of those individuals capable of countering the cytoplasmic effect of the thiol reductant.

3.3.4 Cell conditioning

Cell conditioning in fermentation consists of pre-treating the cells by imposing a mild stress situation which results in cells being more resistant to diverse factors. A thiol reductant such as DTT (**dithiothreitol**) was shown to achieve cell conditioning which resulted in enhanced recombinant production titres, attributed to higher basal levels of chaperone concentrations in conditioned cells (Gill et al., 2001). This could relate to the experiments performed where β -mercaptoethanol was added ([Fig. 3.6] to [Fig. 3.12]). In some cases, the elevated β -galactosidase activity in overnight cultures partially masked the stress response expected upon

β -mercaptoethanol addition. On the other hand, β -mercaptoethanol could be utilized as an agent to “condition” the bacteria raising the chaperone expression in order to prepare the cells for recombinant protein production. Alternatively, Hoffmann et al. (2004) presented a way to cause conditioning in cells by pre-culturing the bacteria in complex media before cultivation in chemically defined media which resulted in similar advantages to DTT cell conditioning, further developing this practice in bioprocessing

3.5 Conclusions and future work

For the *groE::lacZ* reporter pQF50KgroE, temperature has been shown to be the major factor governing β -galactosidase activity, hence σ^{32} levels, delivering robust measurable differences upon temperature variation. In agreement with independent observations of chaperone protein concentrations by Diaz-Acosta et al. (2006), the heat-shock response monitored with pQF50KgroE was different between aerobic and anaerobic cultivation.

β -mercaptoethanol is a toxic chemical which limits growth, but in some cases it was shown to result in increases of *groE::lacZ* acting independently of lowered growth rates. The β -mercaptoethanol action was likely to exacerbate the reducing cytoplasm of *E. coli*, leading to disulphide bond disruption, a common cause of peptide misfolding in the cytoplasm. Thus its influence on σ^{32} titre is probably posttranslational. The effect of β -mercaptoethanol resulted in greater β -galactosidase activity when added to heat-shocked aerobic cultures but not in anaerobic cultures. This was possibly due to a merge of the β -mercaptoethanol effect with the sustained β -galactosidase expression, previously demonstrated in anaerobic cultures during

heat-shock. The *groE::lacZ* increment caused by β -mercaptoethanol stress has been shown to be independent of σ^E as it could be observed in a $\Delta\sigma^E$ strain. Previous stress situations, such as high temperature or anaerobic growth, conditioned *E. coli* to produce a more efficient σ^{32} response to β -mercaptoethanol. The reporter pQF50KgroE could be utilized to study the σ^{32} response in diverse culture conditions.

In theory, the β -galactosidase could have been degraded *in vivo* during decreases in intracellular levels of σ^{32} . However, this has to be approached with care as the *in vivo* half-life of the enzyme has not been measured in this work. The biochemical assays revealed *in vivo* β -galactosidase activity but due to the required cell lysis process the results might not be an exact representation of the β -galactosidase in intact cells. There are some limitations for the utilization of the *groE::lacZ* for bioprocess monitoring as it is difficult to isolate the heat-shock response other physiological situations emerging from different growth rates and growth phases. The data obtained with this reporter should be analyzed with care as the heat-shock stress responses monitored by pQF50KgroE were shown to vary in different strains.

The results obtained could be further validated using microarrays or quantitative PCR to compare RNA produced from the *lacZ* gene of pQF50KgroE with RNA of genes of the σ^{32} regulon with special interest on the chaperone genes *dnaK*, *groES/EL* and *ibpA*. Pulse-chase experiments using radiolabelled [^{35}S] could be designed to observe the turnover of these chaperones. Such experiments could also help to resolve in practice the questions regarding the stability of pQF50KgroE β -galactosidase *in vivo*. DnaK binding and release of σ^{32} upon β -mercaptoethanol addition could be confirmed by SDS-PAGE followed by western blotting using anti-DnaK and anti- σ^{32} . In order to apply this reporter system to use in industry it is

recommended to utilize a microplate reader for β -galactosidase activity measurements, or use flow cytometry with Fluorescein-di- β -D-galactopyranoside (FDG) as substrate which enters viable cells and emits green fluorescence on reaction with β -galactosidase (Plovins et al., 1994). Flow injection analysis could be used for on-line monitoring of β -galactosidase in large scale fermentations (Benito et al., 1993).

Chapter Four

**Comparison of bulk measurements from GFP reporter systems
with single cell analysis data for monitoring host response to
reducing environments**

4.1 Introduction

Reporter gene technology has been successfully coupled with fluorimetric techniques thanks to incorporation of autofluorescent proteins in the reporter systems (Vizcaino-Caston et al., 2011). Microbial cells harboring reporter systems accumulate fluorescent protein as a function of promoter activity. The whole organism is generally taken to be the measurement device. Fluorescence can be measured from culture suspensions as bulk amounts by a fluorimeter. The intensity of fluorescence enables comparison between cell suspensions from different conditions. However, the intensities calculated are an average of the fluorescence of the cell suspension per total amount of biomass, frequently inferred from optical density measurements (OD). Flow cytometry is based on the measurement of optical and fluorescent properties of single particles in the sample and has the key advantage of delivering a single cell analysis. In this chapter we compared fluorimetric data from bulk and single cell analysis of bacteria transformed with different reporter systems where GFP expression was driven by different heat shock promoters. In addition, gene reporter technology was utilized to attempt a partial characterization of the *yidQ* promoter (also referred to as *ecfI*) which is addressed independently in this chapter.

The reporter plasmids utilized in these experiments contain full sections of the intergenic regions upstream the gene of interest. The plasmid pAA213 contains the *rpoH* promoter region and pAA212 the *dnaK* promoter region each encompassing 200bp upstream of the each gene of interest. These promoters are cloned upstream of the *gfpmut3* gene in the medium copy number plasmid pFVP25 (20-30 copies per cell) (Aertsen et al., 2004). The plasmids pGROE and pDNAK ([Fig. 2.2] and [Fig. 2.3] respectively) contain promoter regions upstream of a *gfpmut2* gene carried by the very low copy plasmid pUA66 [Fig. 2.1]

(Zaslaver et al., 2006). These plasmids contain a larger promoter insert, spanning into the gene of interest and the immediate gene upstream, thus encompassing the whole intergenic region. The qualities of these reporter plasmids based on their promoters and copy number are contrasted.

The thiol reductant DTT (Dithiothreitol) has been utilized in the past to successfully increase the concentration of the *E. coli* chaperones DnaK and GroE in response to controlled reducing environments (Gill et al., 1998; Gill et al., 2001). The experiments in this chapter aimed to observe the gene regulatory basis for such effects comparing bulk measurements obtained from fluorimetry with data generated by multi-parameter flow cytometry. The effect of β -mercaptoethanol on these full-length promoters was also compared with the σ^{32} -binding site-only promoter of pQF50kgroE (Chapter 3) and with the responses obtained from DTT cultures. The *E. coli* strain BL21* was used in some experiments because of its relevance as a standard industrial host. This strain was not utilized with β -galactosidase reporters because it is *lac*⁺.

4.2 Results

4.2.1 Investigation of the heat-shock effect on pDNAK and pGROE gene expression

These experiments were performed in order to observe the response of the reporter plasmids pDNAK and pGROE to heat-shock. *E. coli* BW25113 transformed either with pDNAK or pGROE were grown in 30mL of LB in conical flasks overnight at 37°C and 0.3mL was inoculated into 250mL conical flasks containing 30mL of LB supplemented with 0.4% glucose and 25 μ g mL⁻¹ of kanamycin. Cultures were incubated aerobically at a constant 30°C.

During sampling, OD_{650nm} and fluorescence at $\lambda_{Ex} = 488\text{nm}$ and $\lambda_{Em} = 509$ were monitored (see [Table 2.3] for luminescence spectrophotometer configuration). Relative fluorescence (RF) is defined as units of fluorescence detected divided by the OD₆₅₀. Single cell data was gathered by flow cytometry (see [Table 2.4] for flow cytometer settings). The green fluorescence histograms represent the number of cells (events) on each channel indicating the level of green fluorescence (pmt2) on a logarithmic scale. The data from these histograms is generally found summarized by the value of its geometric mean (Gm) and its coefficient of variation (CV) as a measure of dispersion. The median (Md) is the value of green fluorescence that divide the population in two halves and is a non-parametric measure of the central tendency of the population. Single cell data were visualized in relation to forward scatter and side scatter measurements. These plots were useful to reveal the heterogeneity in the sample population based on size, shape and cytoplasmic granularity. In this experiment, scatter plots did not reveal any remarkable difference based on reporter plasmid or temperature of incubation.

After 1 hour, cultures were stressed with heat-shock at 42°C. Duplicates were grown at 30°C as controls. Samples were taken at 2 hours intervals after heat-shock. pGROE data is shown in [Fig. 4.1 (a)] The RF of the cultures at 30°C fell and remained constant while heat shocked cultures remained at higher values [Fig. 4.1 (a)]. After four hours at 42°C heat-shock there was a noticeable difference with the cultures at 30°C and 42°C. Growth curves showed marginally better growth at 30°C. Single cell analysis of green fluorescence [Fig. 4.1 (b)] in general correlated with bulk measurements. There is a decrease in coefficient of variation accompanied by an increase in geometric mean (Gm) of green fluorescence two hours post heat-shock. After four hours, the Gm of the green fluorescence histograms showed a decrease that could indicate that the heat-shock response was exhausted.

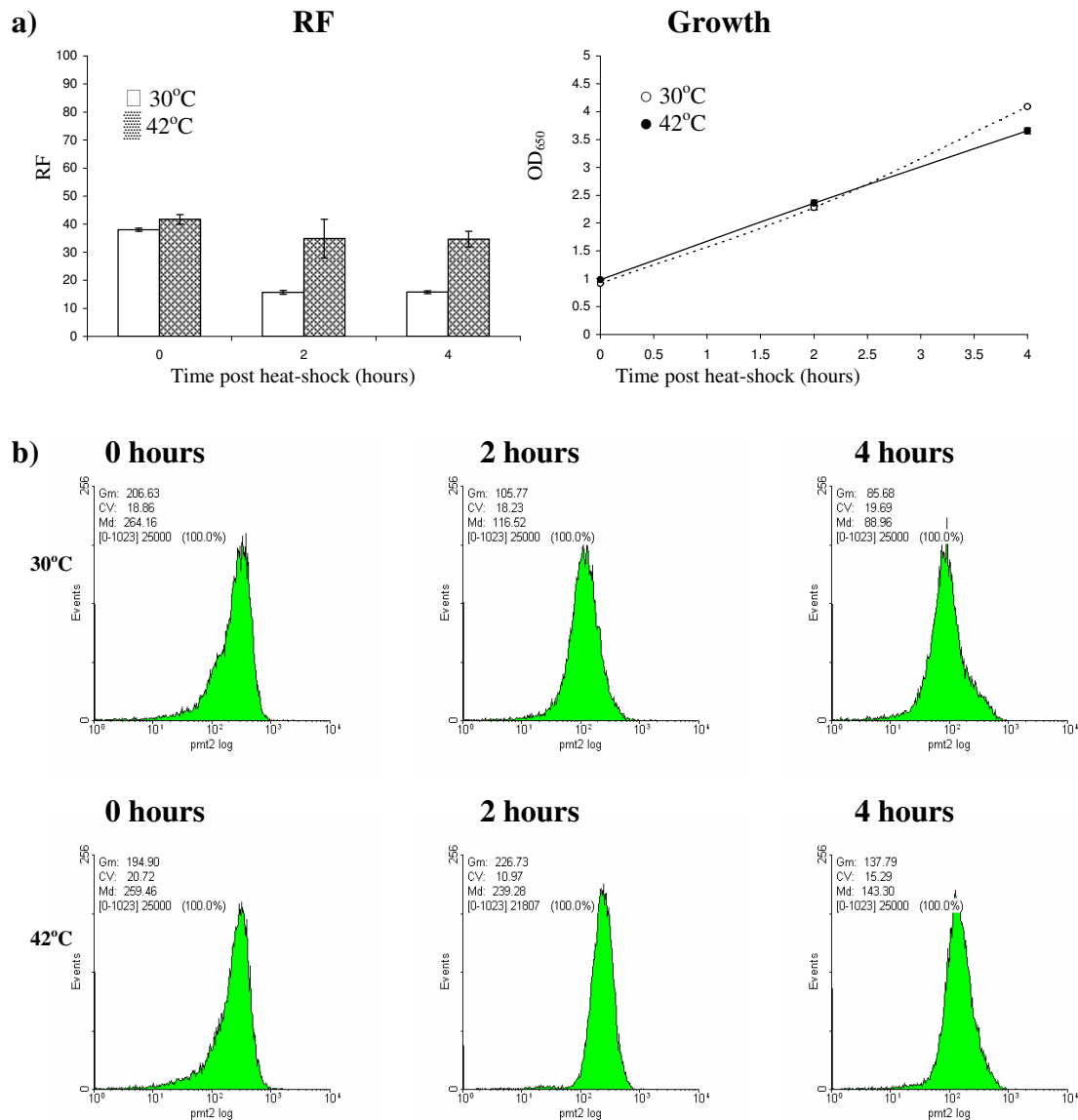


Fig. 4.1 pGROE activity upon heat shock

E. coli BW25113 pGROE were grown in 30ml LB at 30°C and either transferred to 42°C after 1 hours growth or maintained at 30°C. **(a)** Relative fluorescence and OD₆₅₀ **(b)**. single-cell green fluorescence where Gm is the geometric mean, Md is the median and CV is the coefficient of variation were measured every two hours.

pDNAK RF data, [Fig. 4.2 (a)], revealed similar initial high fluorescence as seen in pGROE. These cultures, either at 30°C or heat-shocked cultures decreased in RF 2 hours post-heat shock. This was followed by an increase of RF in the heat-shocked cultures at four hours post heat-shock while the RF of the cultures at 30°C continued decreasing [Fig. 4.2 (a)] showed better growth at 30°C than at 42°C as seen in pGROE. Single cell analysis [Fig. 4.2 (b)] indicated gradual reduction over time of Gm fluorescence in the cultures at 30°C and a gradual but slight increase over time in the heat shocked cultures. It is likely that the bacteria in the inoculum were very stressed or had accumulated GFP from the overnight growth. In spite of this, higher average fluorescence was observed in the cultures shifted to 42°C.

4.2.2 Investigation of DTT effect on the *rpoH* gene expression

To investigate the influence of the thiol reductant DTT on *rpoH* gene expression (encoding σ^{32}), increasing amounts of DTT were added to a culture of *E. coli* BW25113 transformed with the reporter plasmid pAA213 containing an *rpoH::gfpmut3* fusion. The aim of the experiment was to test the effect of increasing amounts of DTT on *rpoH* transcription by fluorimetry.

Aerobic cultures were grown in 250mL conical flasks containing 30mL of LB supplemented with 0.4% glucose and 100 μ g mL^{-1} of ampicillin and inoculated with 0.3mL of overnight grown bacteria. Anaerobic cultures (30mL) were grown in capped tubes. Aerobic and anaerobic cultures were incubated at 30°C and DTT was added at 1, 2 and 3 g L^{-1} when the OD_{650nm} reached approximately 0.5. Cultures free of DTT were incubated as control. All work was performed in duplicates and samples were taken after 1.5, 2.5, 3.5, and 22.5 hours.

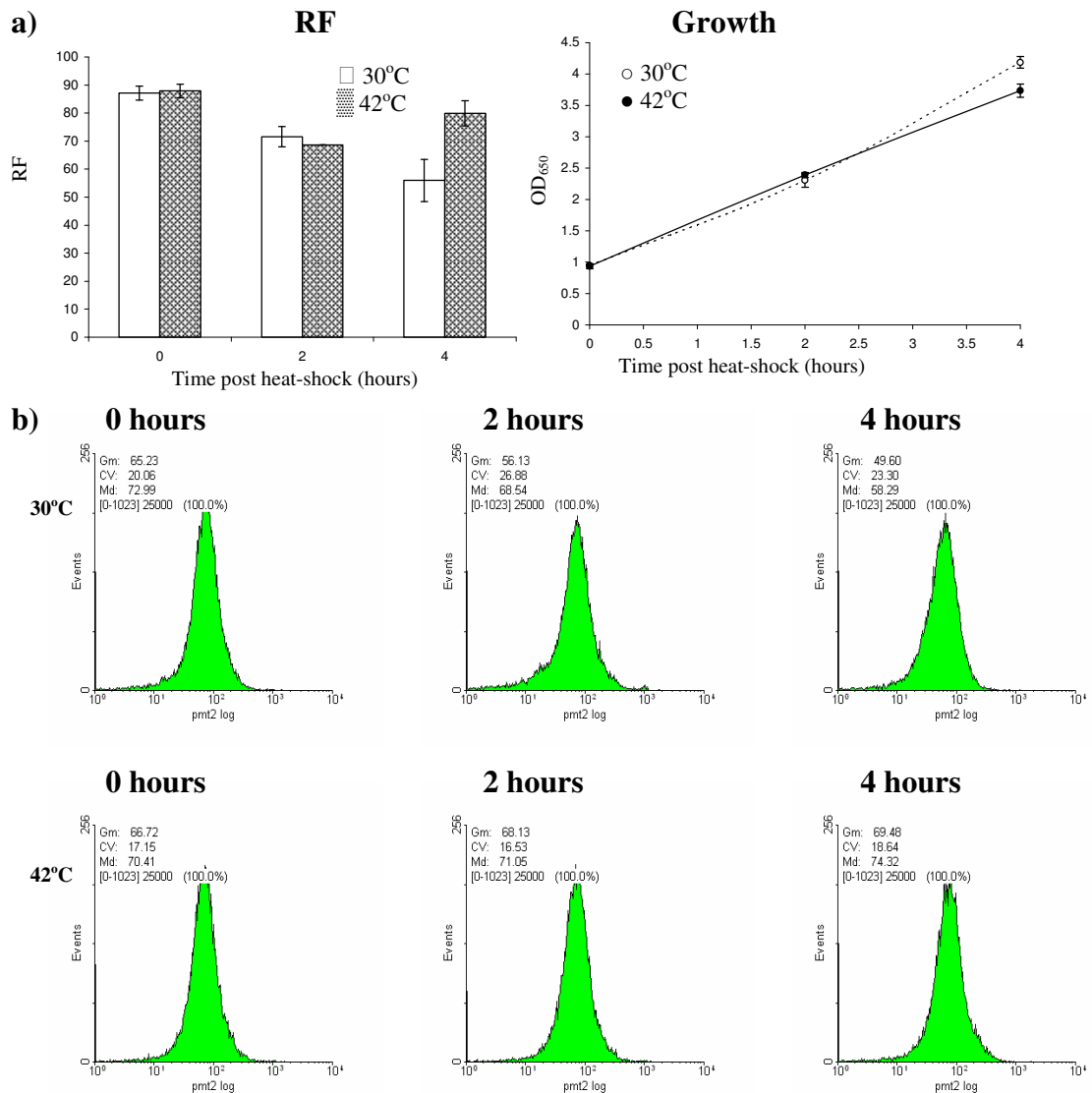


Fig. 4.2 pDNAK activity upon heat shock

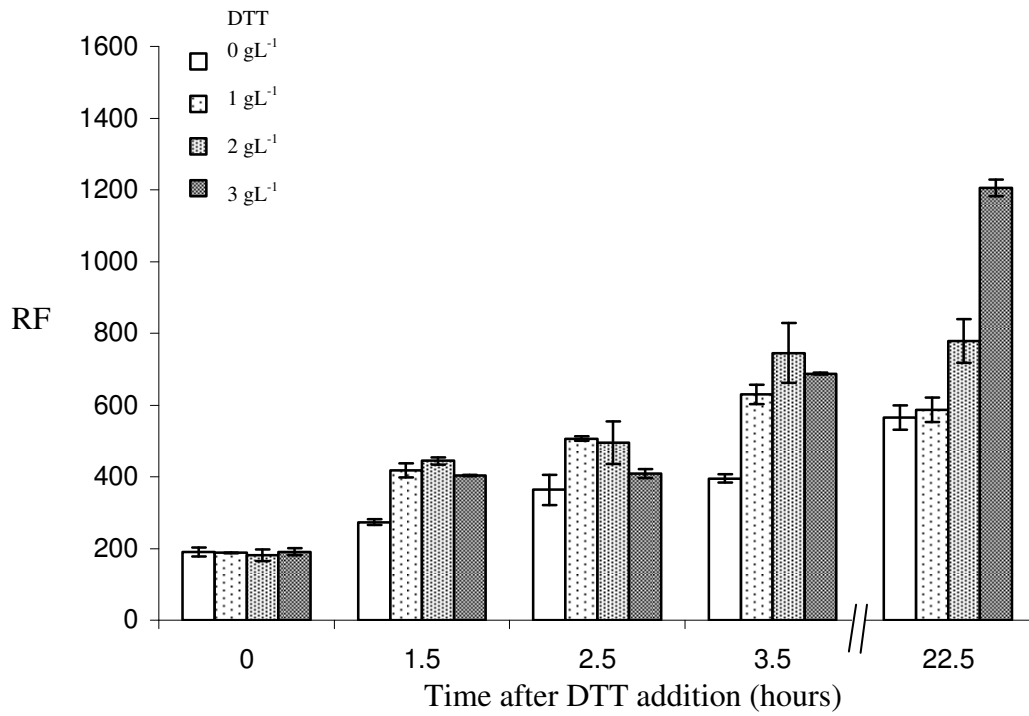
E. coli BW25113 pDNAK were grown in 30ml LB at 30°C and either transferred to 42°C after 1 hours growth or maintained at 30°C. Relative green fluorescence **(a)** and OD₆₅₀. **(b)** Single-cell fluorescence where Gm is the geometric mean, Md is the median and CV is the coefficient of variation were measured every two hours.

Aerobic cultures revealed a general increase of relative fluorescence over time in the controls and the DTT added cultures [Fig. 4.3 (a)]. DTT addition resulted in increased fluorescence 1.5 hours after addition by comparison of the test cultures with the control. Growth inhibition [Fig. 4.3 (b)] was caused by DTT in a dose dependent manner. Despite large differences in growth rate, relative fluorescence was similar in all DTT-added cultures at 1.5 and 3.5 hours, and all DTT-added cultures except for 3gL^{-1} at 2.5 hours. This suggests that DTT addition could trigger a qualitative response rather than a quantitative dose-dependant response.

Anaerobic cultures showed similar behaviour as aerobic cultures accumulating fluorescence over time [Fig 4.4 (a)]. The DTT stressed cultures had greater relative fluorescence accumulation 1.5 hours after DDT addition in comparison with the control. This trend was continued up to 3.5 hours post-addition when cultures with DTT added had 50% higher RF than the control. At this point, the RF observed did not increase in a DTT dose-dependant manner. Overnight incubation revealed higher activity in the 3gL^{-1} culture in comparison with the cultures at 1gL^{-1} and 2gL^{-1} which had similar fluorescence relative to the control. The growth curves [Fig. 4.4 (b)] did not show dose-dependent growth inhibition; the 1gL^{-1} and 2gL^{-1} cultures grew at a similar rate, with the 3gL^{-1} culture growing slightly slower still. Differences of fluorescence accumulation are probably caused by differences in *rpoH* gene expression rate as the growth inhibition found upon increasing amounts of fluorescence is not as marked as in aerobic cultures. In addition, relative fluorescence during anaerobic growth was double than that of aerobic growth with the exception of the overnight sample. Similar observations were made in Chapter Three where β -galactosidase activities measured as a reporter for σ^{32} activity were also greater in anaerobic cultures. These results may indicate that in addition to higher σ^{32} translation rate it is possible that there was also higher *rpoH* transcription rate in anaerobic cultures than in aerobic cultures.

a)

RF (*rpoH::gfpmut3*)



b)

Growth curves

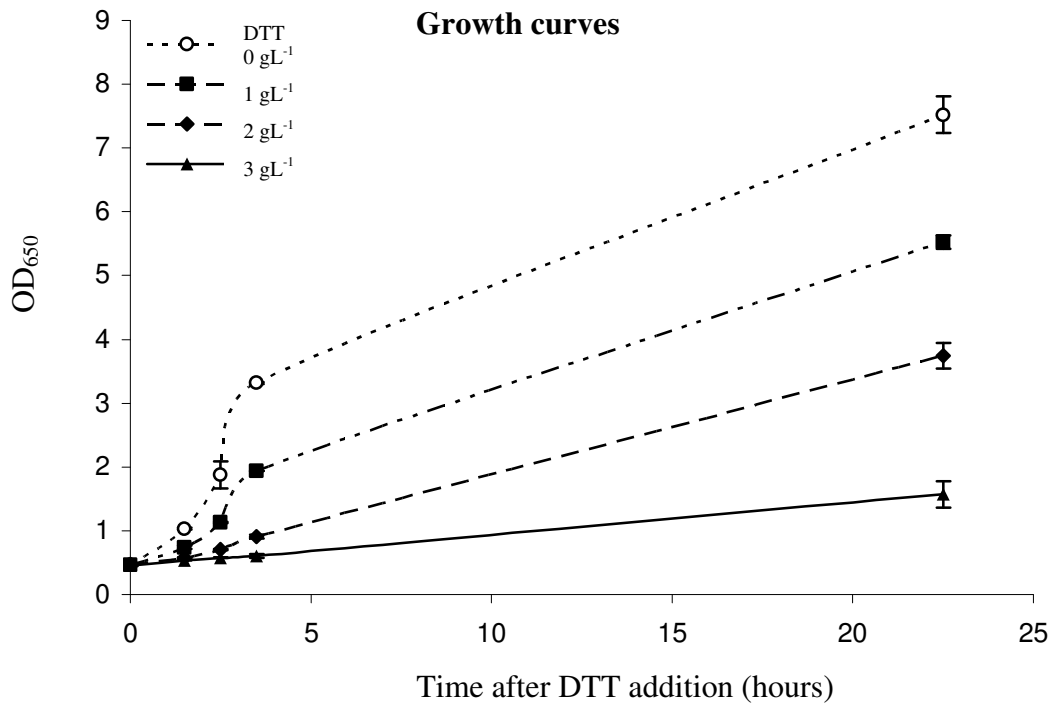


Fig. 4.3 The effect of DTT on pAA213 (*rpoH::gfpmut3*) during aerobic growth

E. coli BW25113 pAA213 was grown aerobically in 30mL LB+0.4% glucose and DTT was added at the indicated concentrations after one hour. Relative fluorescence (a) and OD₆₅₀ (b) were measured at various intervals.

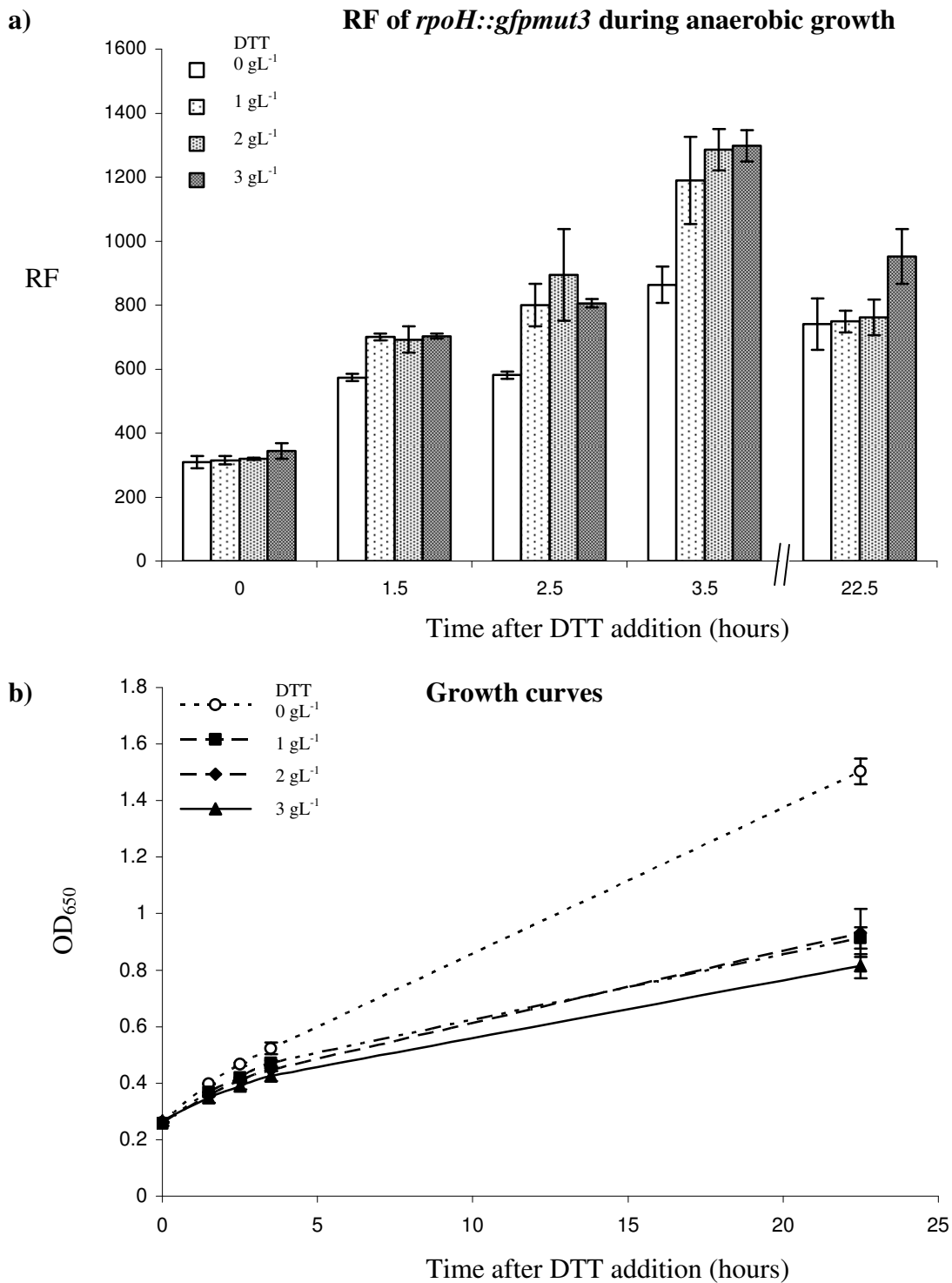


Fig. 4.4 pAA213 (*rpoH::gfpmut3*) expression in *E. coli* BW25113 during anaerobic growth

E. coli BW25113 pAA213 was grown anaerobically in LB+0.4% glucose and DTT was added at the indicated concentrations after one hour. Relative fluorescence (a) and OD₆₅₀ (b) were measured at various intervals.

4.2.3 Comparison of bulk measurements with single cell analysis of *rpoH::gfpmut3* activation in pAA213 in *E. coli* BL21*

As an extension of the previous experiment, fluorimetric (bulk) measurements were compared to single cell measurements (flow cytometry). Bulk and single cell measurements were obtained from *E. coli* BL21* cultures subjected to thiol reductants and anaerobic conditions. *E. coli* BL21* was utilized because of its industrial relevance. Relative fluorescence measurements were expected to reveal overall fluorescence per biomass in response to stress. These results were contrasted with data obtained from flow cytometric analysis allowing single cell analysis of fluorescence as well as light scattering.

Experimental conditions

Conical flasks containing 30mL of LB supplemented with 0.4% glucose and antibiotics as required were inoculated with 0.3mL of *E. coli* BL21*-pAA213 grown overnight. Aerobic and anaerobic test cultures were incubated at a constant 30°C and thiol reductant added, either 3 gL⁻¹ DTT or 80mM β-mercaptoethanol, when the OD_{650nm} reached approximately 1 in aerobic cultures or 0.5 in anaerobic cultures. 80mM of β-mercaptoethanol contains the same molar quantity of -SH groups as 3gL⁻¹ of DTT. Cultures without thiol reductant were used as controls. All work was performed in duplicates. Samples were taken before and two and four hours after reductant addition. During sampling, OD₆₅₀ and fluorescence at λ_{Ex} = 488nm and λ_{Em} = 509 were measured. Samples were analyzed by flow cytometry including propidium iodide (PI) staining in order to observe possible cytotoxic effects of the thiol reductants.

Bulk measurements of rpoH::gfpmut3: Aerobic cultures

Aerobic cultures exhibited results similar to previous data obtained with *E. coli* BW25113. Cultures with the addition of thiol reductants resulted in higher fluorescence 2 hours after addition, which continued to increase 4 hours after addition [Fig 4.5 (a)]. While the average relative fluorescence was higher on addition of DTT than β -mercaptoethanol, the difference was not significant. Cultures showed growth throughout the duration of the experiment [Fig 4.5 (b)]. The cultures stressed with thiol reductants grew at a reduced rate with no significant differences between those with DTT or β -mercaptoethanol added.

Bulk measurements of rpoH::mut3: Anaerobic cultures

The relative fluorescence of anaerobic cultures observed before addition of thiol reductants doubled after two hours even without the addition of DTT or β -mercaptoethanol [Fig 4.5 (c)]. While there were not significant differences between β -mercaptoethanol and control cultures at this point, DTT cultures showed higher fluorescence. Four hours after addition, β -mercaptoethanol cultures exhibited the highest fluorescence while controls and DTT cultures had comparable lower fluorescence. At this point, the DTT cultures showed a drastic reduction of fluorescence in comparison with the aerobic cultures whose RF is nearly 50% higher [Fig 4.5 (a)]. This contrasted with RF at 0 and 2 hours post addition, which increased in all cultures irrespective of reductant addition. The control cultures grew at a constant rate while β -mercaptoethanol and DTT cultures showed growth reduction [Fig 4.5 (d)]. It is possible that a combination of the toxicity of the reductant and the poor natural anaerobic growth of *E. coli* BL21* contributed to this growth stagnation.

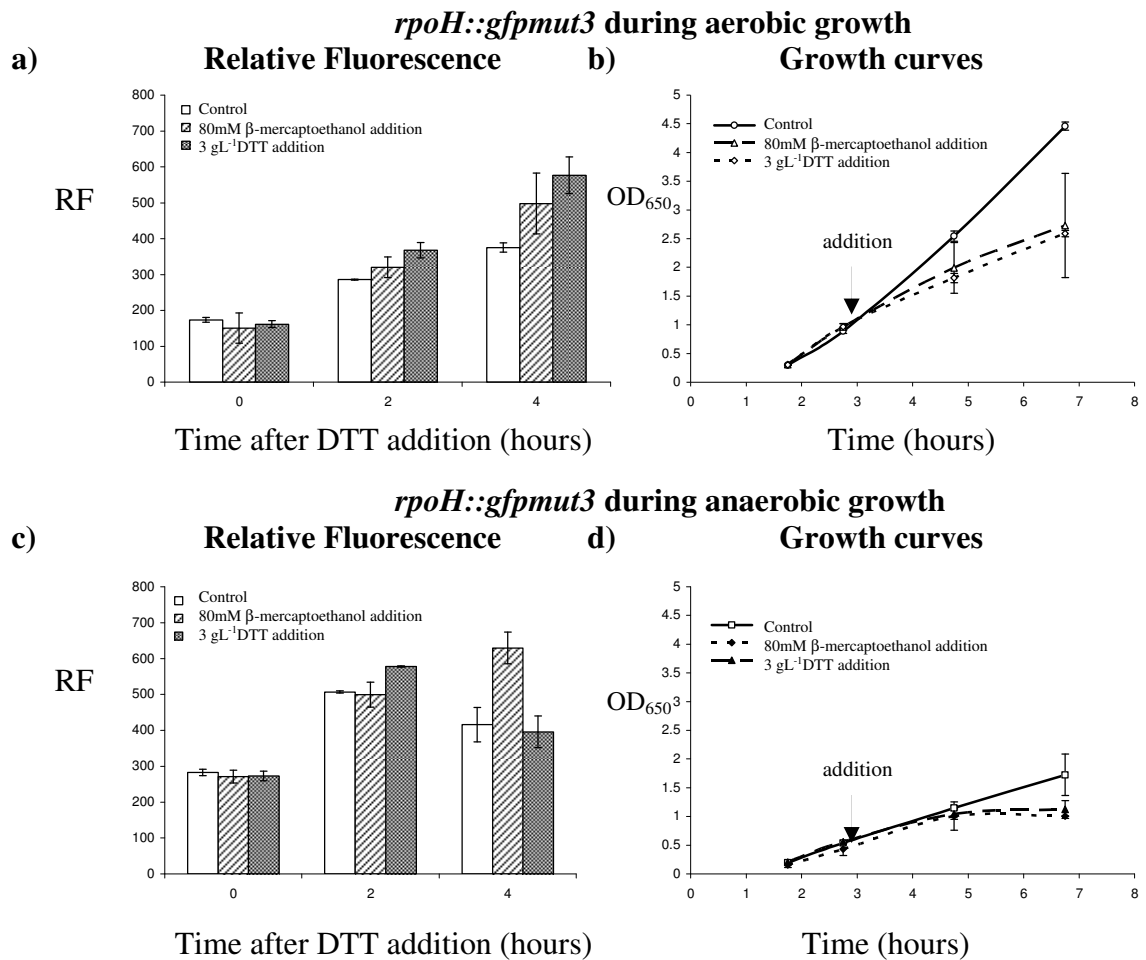


Fig. 4.5 Bulk measurements of expression in *E. coli* BL21* transformed (*rpoH::gfpmut3*) cultures in the presence of reductants

Cultures were grown in 30mL LB at 30°C either aerobically (**a&b**) or anaerobically (**c&d**) until OD₆₅₀ was ~1 in aerobic cultures and ~0.5 in anaerobic cultures. Relative fluorescence (**a&c**) and OD₆₅₀ (**b&d**) were measured to quantify GFP production and growth respectively. The arrow indicates the moment of DTT or β -mercaptoethanol addition.

*Single cell analysis of *rpoH::gfpmut3*: The PI control*

This experiment was performed to assess the performance of PI (Propidium Iodide) staining and to determine the red fluorescence of live and dead cells upon PI addition. Two 1mL aliquots of an overnight culture of *E. coli* BL21* harbouring pAA213 were harvested by centrifugation (5 minutes at 5000 g). The supernatants were discarded and cells resuspended in either PBS, to produce a suspension containing live cells, or 100% ethanol to produce dead cells. Samples were left at room temperature for a minimum of 15 minutes. Cells were harvested by centrifugation (5 minutes at 5000 g) and resuspended in PBS. 2.5µL of each cell sample, were transferred to tubes containing 2mL of PBS and stained with PI.

PI is a stain that only permeates dead cells. It can be seen by red fluorescence produced upon excitation at 488nm when it is interacting with DNA. Live cells did not fluoresce red (shown on the PMT4 axis), while dead cells did [**Fig. 4.6 (a)**]. A mixture of live and dead cells generated a bimodal distribution, with two distinct and non-overlapping populations. In the density plots representing both propidium iodide and GFP fluorescence, two populations were also observed [**Fig. 4.6 (b)**].

The live cells do not take up PI and are therefore do not fluoresce red, but do fluoresce green as they contain GFP. The dead cells take up PI but also have reduced green fluorescence. This is thought to be due to two reasons: first, GFP leakage from the dead cells; and second, energy transfer between GFP and PI, diminishing green fluorescence (Lehtinen et al., 2004). This is an important aspect of the use of GFP as a reporter of metabolically active cells as discussed later in this chapter.

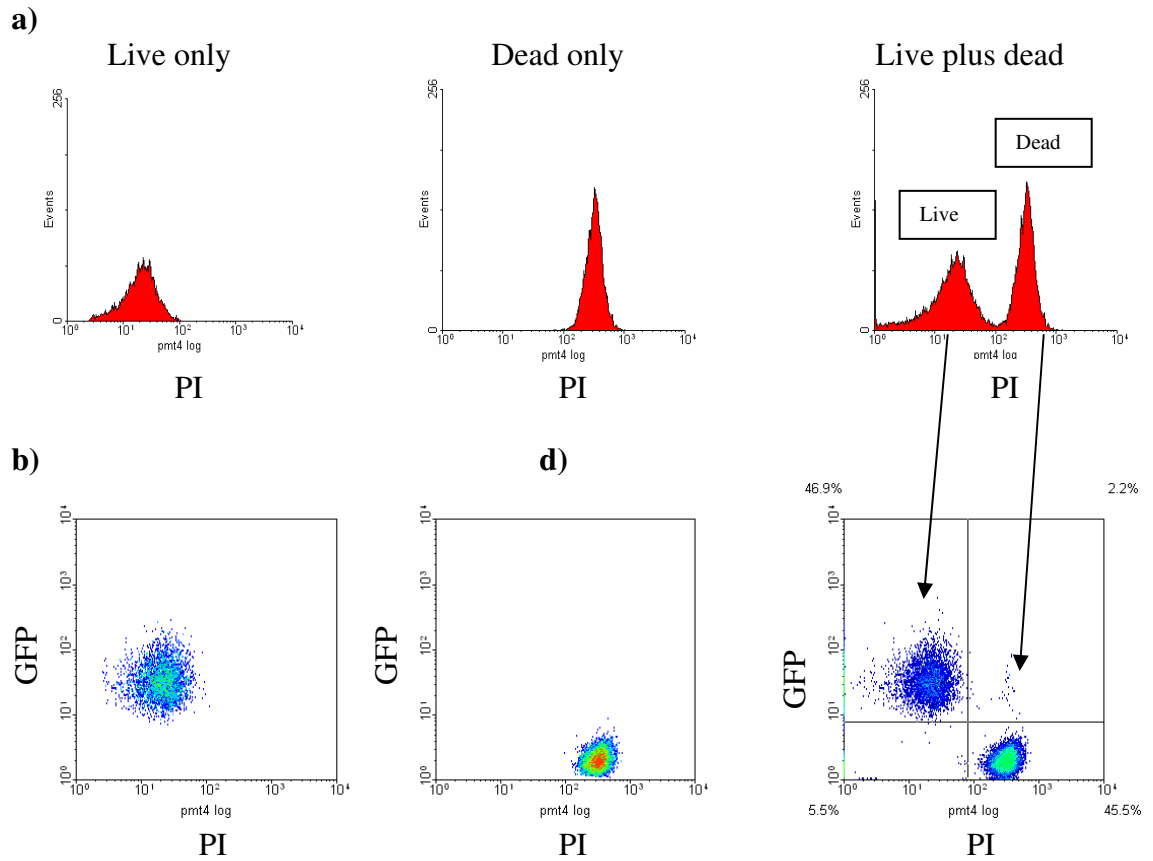


Fig. 4.6 Control experiment to calibrate PI staining of live and dead cells

E. coli BL21* pAA213, either ethanol-killed or live, were stained with propidium iodide and analysed by flow cytometry. **(a)** Red fluorescence (PMT4) histograms representing PI staining. **(b)** Red fluorescence (PMT4; PI) plotted against green fluorescence (PMT2; GFP).

All cultures from [Fig. 4.5] were analyzed with flow cytometry; data shown are representative. The histograms of green fluorescence (pmt2) [Fig. 4.7 (a)] measured before addition of thiol reductants were, as expected, very similar amongst cultures. The plots below these histograms are the respective scatter plots representing the measurements of forward scatter (FS log) in the ordinates and right angle side scatter in the abscises (SS-pmt1) [Fig. 4.7 (b)]. These bi- dimensional graphs are an indication of size and shape of the bacteria. Colour denotes population density. The plots were divided by two lines intersecting with the second decade establishing four quadrants. This was added in order to aid the visualization of movements of cells within the plots, noticeable by the relative position to the quadrant. The density plots pre-addition of thiol reductants did not reveal significant differences between cultures. At the bottom [Fig. 4.7 (c)] shows each corresponding PI histogram. The PI histograms did not reveal any significant dead cell population (in relation to [Fig. 4.6 (a)]).

Aerobic cultivation (rpoH::mut3): Two hours after addition of thiol reductants.

The histograms of green fluorescence [Fig. 4.8 (a)] did not reveal significant differences between control and thiol added cultures except for a slight increase of Gm in the thiol added, which agreed with the relative fluorescence. In the scatter plots, [Fig. 4.8 (b)] FS decreased between pre-addition [Fig. 4.6(b)] and the two hour samples; this could be interpreted as a decreasing cell size. There was no significant difference in FS between the three cultures. PI histograms [Fig. 4.8 (c)] did not reveal significant cell death.

Single cell analysis of *rpoH::gfpmut3* pre-addition

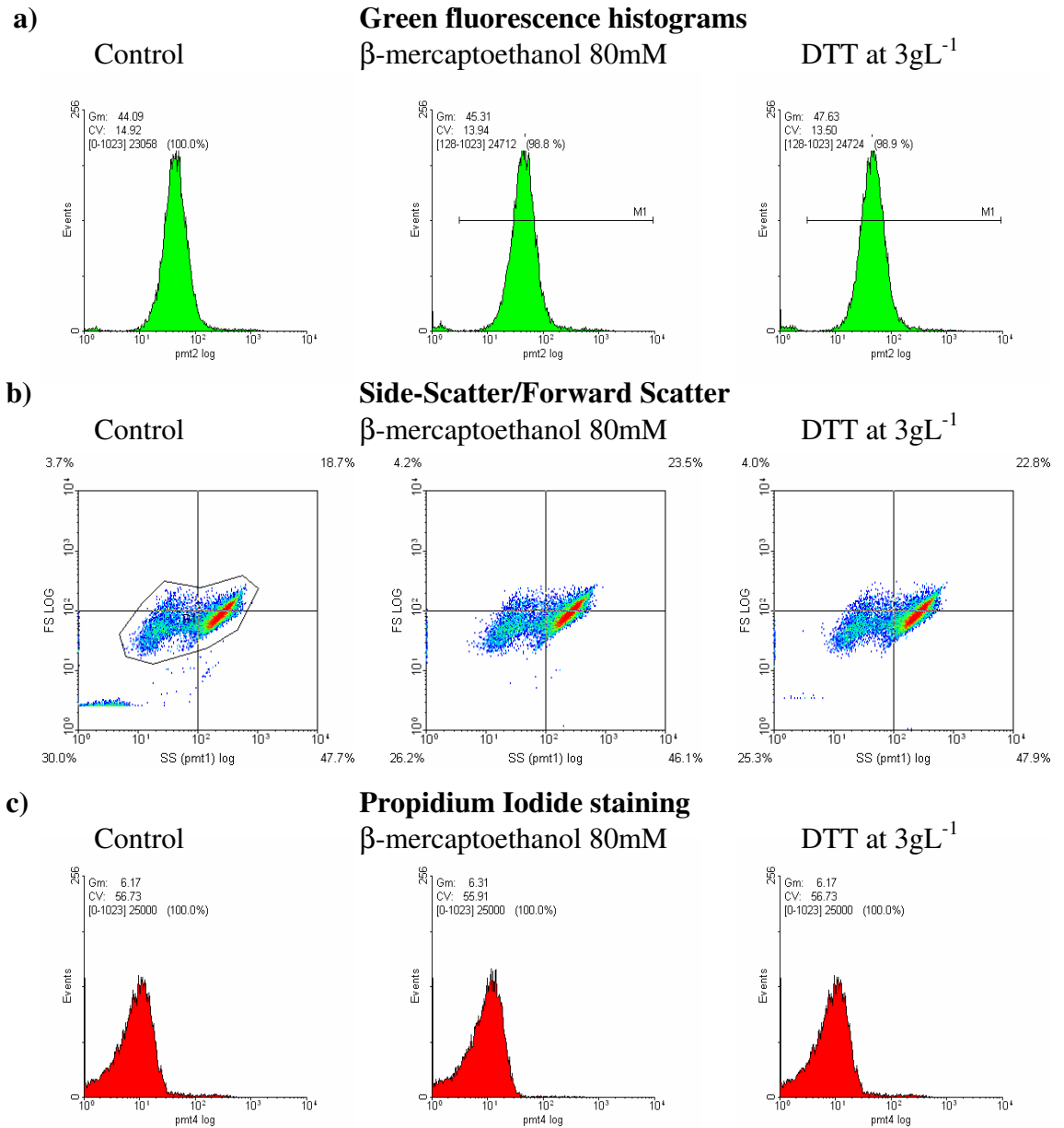


Fig. 4.7 Single cell analysis of *E. coli* BL21* harbouring pAA213 pre-addition of thiols during aerobic growth

Analysis of a sample taken before addition of thiol reductants. It can be observed that the green fluorescence (a) is equivalent between cultures as well as the correlation of forward and side-scatter (b). PI staining (c) did not reveal cell death.

Single cell analysis of *rpoH::gfpmut3* two hours after thiol addition

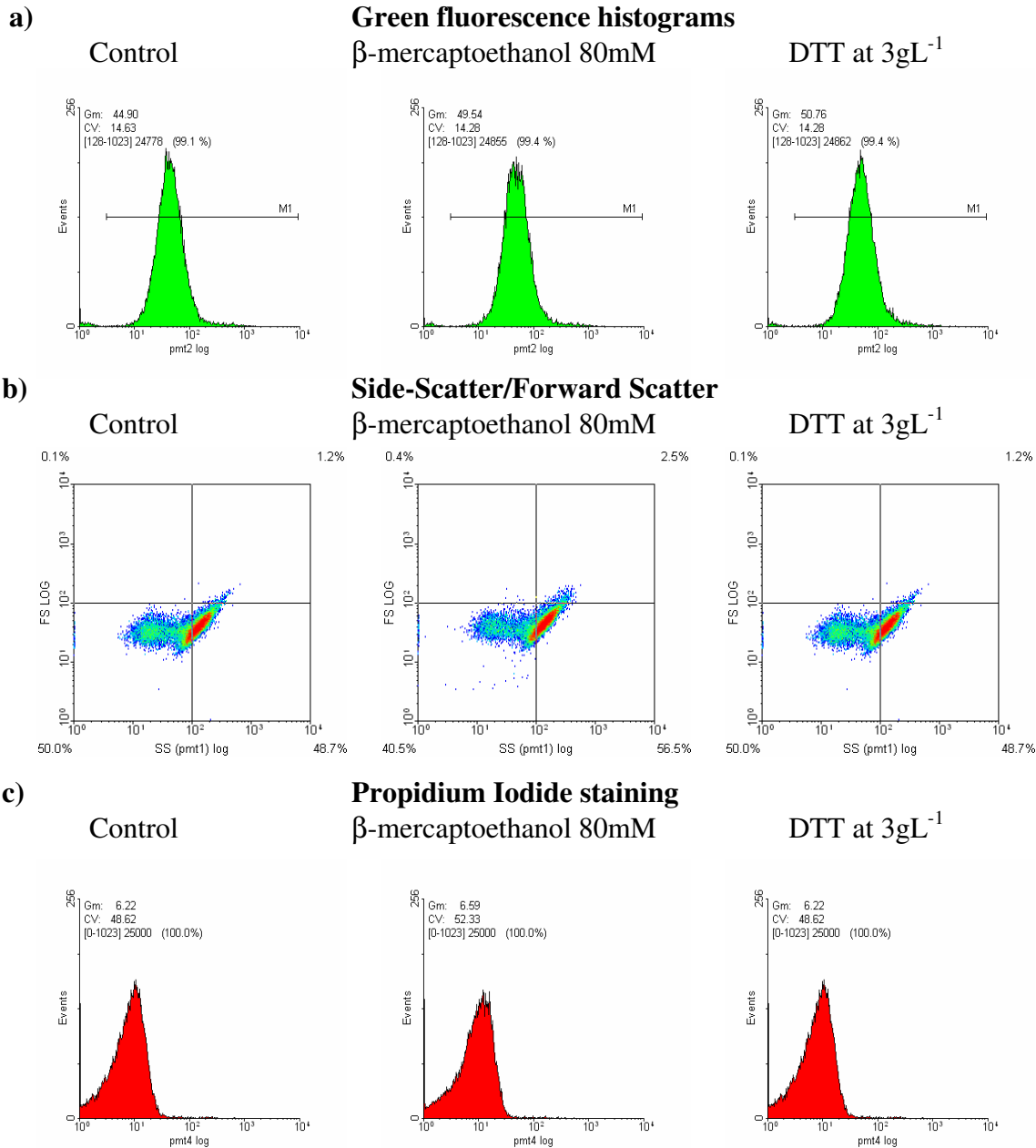


Fig. 4.8 Single cell analysis of *E. coli* BL21* harbouring pAA213 two hours after thiol addition during aerobic growth

(a) Green fluorescence histograms which present averages analogous to the results from fluorimetry. (b) light scatter distribution of the population in values (c) PI staining did not reveal rising dead cell population.

Four hours after addition of thiol reductants

Histograms of green fluorescence contrasted with fluorescence measurements revealing increases of fluorescence in β -mercaptoethanol ($G_m = 80.28$) above DTT added cultures ($G_m = 70.40$) [Fig. 4.9 (a)]. The scatter plots showed similar distribution of FS and SS in the three cultures [Fig. 4.9 (b)]. PI staining [Fig. 4.9 (c)] showed that 4 hours after thiol addition there were no dead cells.

*Anaerobic cultivation (*rpoH::mut3*)*

Again, cultures before the reductant addition revealed highly similar characteristics. Changes occurred after addition of thiols. The samples at two hours post-addition resulted in fluorescence histograms which differed from bulk fluorimetric data as β -mercaptoethanol produced histograms similar to the control [Fig. 4.10 (a)] whereas DTT cultures exhibited a higher fluorescence. The values were much increased in comparison with aerobic cultures at the same sample point ([Fig. 4.5 (c)] versus [Fig. 4.5 (a)] 2 hours post DTT addition). The fluorescence distribution was narrow indicating low intra-population variability. The scatter plots indicated a very similar population distribution for all cultures [Fig. 4.10 (b)]. PI staining did not reveal significant numbers of dead bacteria [Fig. 4.10 (c)].

Single cell analysis of *rpoH::gfpmut3* four hours after thiol addition

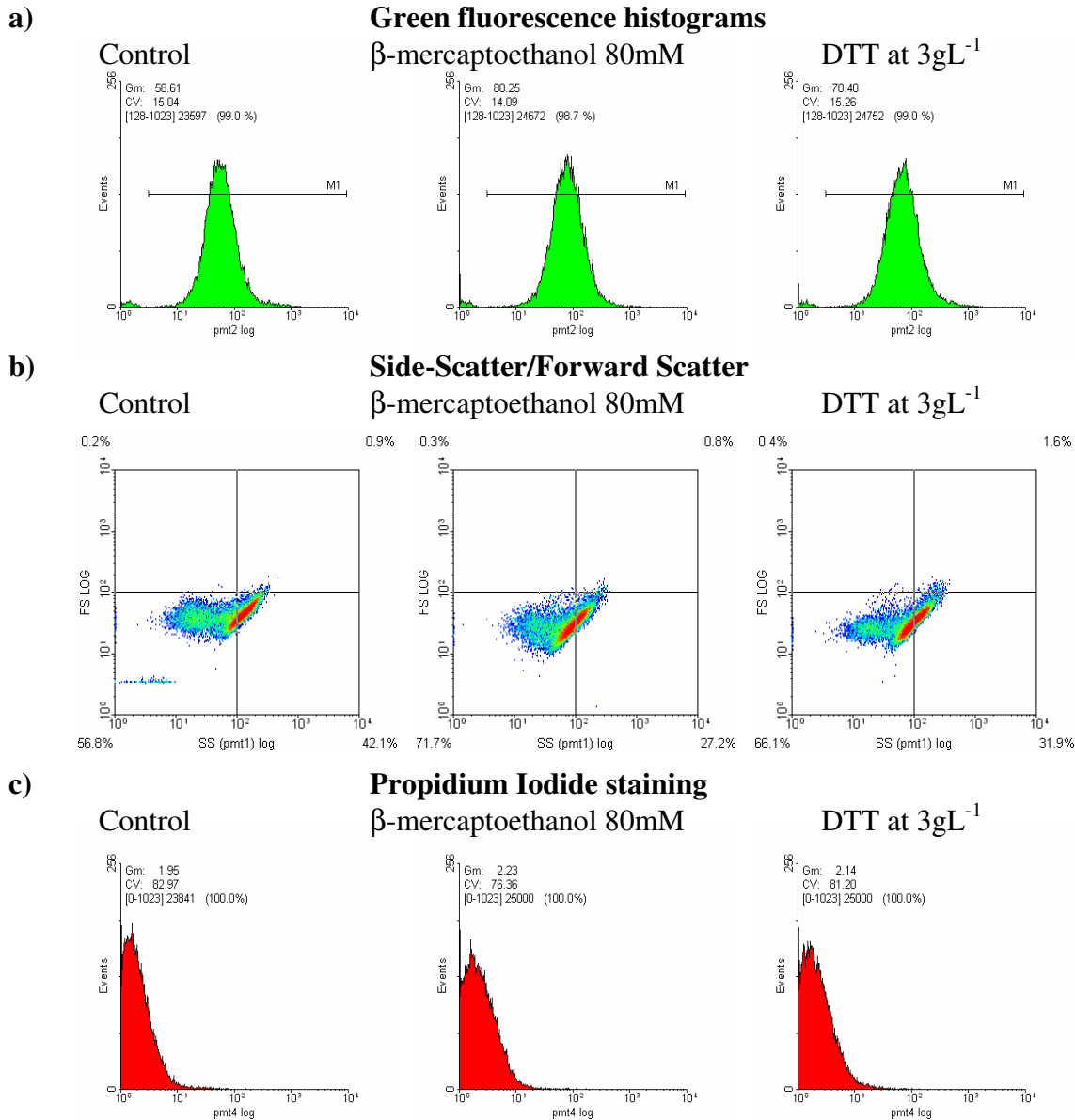


Fig. 4.9 Single cell analysis of *E. coli* BL21* harbouring pAA213 four hours after thiol addition during aerobic growth

Four hours after addition histograms in (a) exhibited an increased of fluorescence in thiol added cultures, especially β -mercaptoethanol. (b) Scatter plots revealed lower FS values on DTT added cultures. PI (c) did not show cell death.

Single cell analysis of *rpoH::gfpmut3* two hours after thiol addition in anaerobiosis

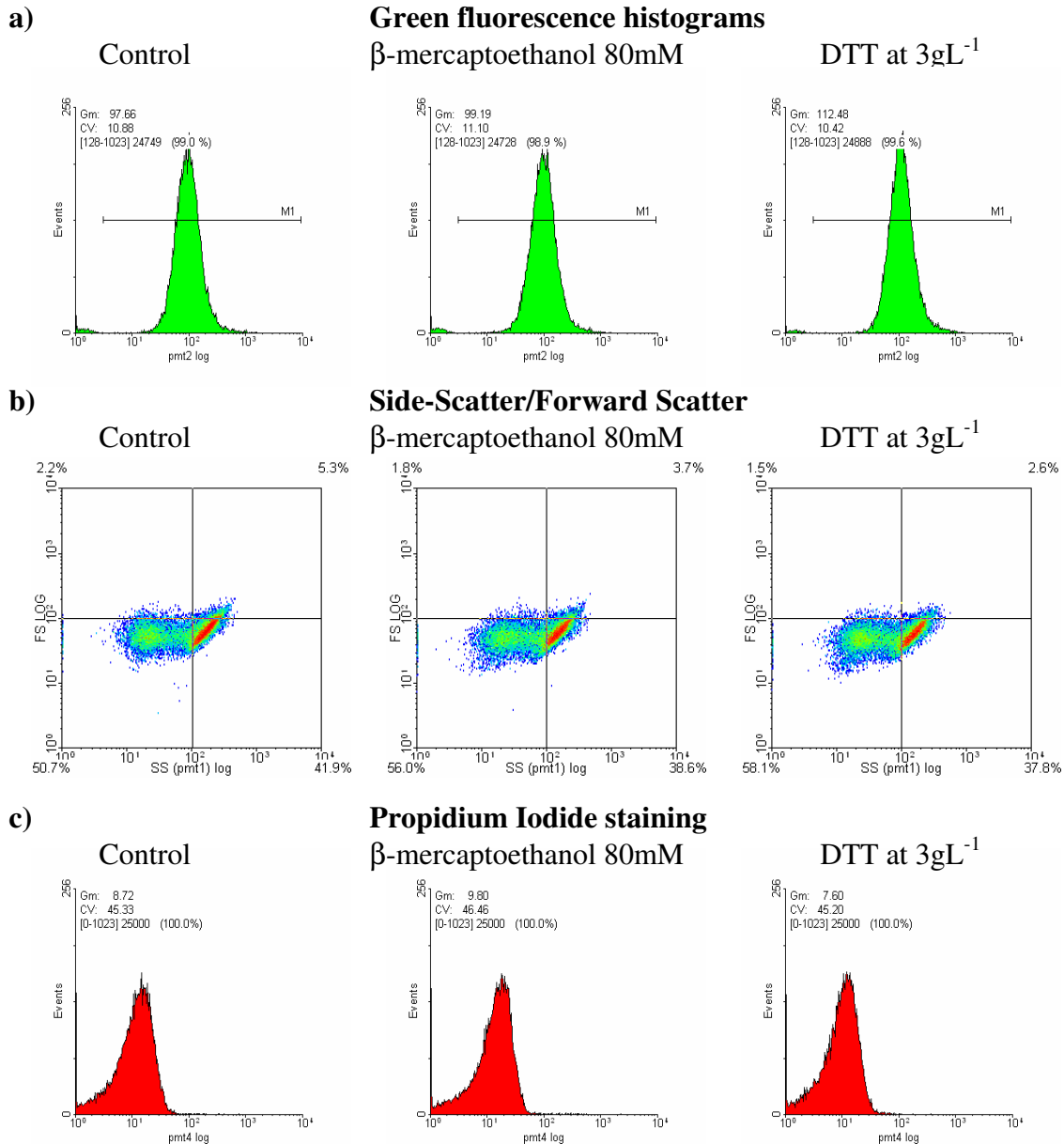


Fig. 4.10 Single cell analysis of *E. coli* BL21* harbouring pAA213 two hours after thiol addition in anaerobiosis

At this sample point green fluorescence histograms (a) were similar except for DTT cultures where Gm was slightly higher. The distribution in FS and SS (b) was similar amongst cultures but showed a greater spread on SS values when compared with aerobic cultures. PI stain did not reveal a noticeable population of dead cells.

Four hours after thiol reductant addition

The histograms at this point showed discrepancy with bulk fluorimetry measurements; the Gm of the control culture was higher than the β -mercaptoethanol and DTT cultures [Fig. 4.11 (a)]. Comparing the scatter plots, the FS and SS values of both cultures treated with reductants decreased, although the reduction in SS values was greater in β -mercaptoethanol than DTT cultures [Fig. 4.11 (b)]. The discrepancies in RF and Gm values may be partially explained through these differences in scatter properties which will be analyzed further subsequently. The PI staining did not reveal a significant population of dead bacteria [Fig. 4.11 (c)], but the β -mercaptoethanol cultures had a different-shaped distribution to untreated and DTT cultures, although within the region expected for live cells [Fig. 4.1 (a)]

4.2.4 Comparative fluorescence study of GFP reporter based on plasmid copy number

The plasmid pAA213 used in previous experiments is a medium copy number plasmid based on pFVP25; it has a ColE1 replicon which on average produces 20-30 copies per cell. The following experiments investigated whether plasmid copy number affects reporter systems.

E. coli BL21* was transformed with pFVP25-based plasmids pAA213 (*rpoH::gfpmut3*) and pAA212 (*dnaK::gfpmut3*) (Aertsen et al., 2004) or the low copy number pUA66-based plasmids pDNAK (*dnaK::gfpmut2*) and pUA66 (promoterless vector). pUA66 features the SC101 origin which controls plasmid replication to < 5 copies per cell (Zaslaver et al., 2006).

Single cell analysis of *rpoH::gfpmut3* four hours after thiol addition in anaerobiosis

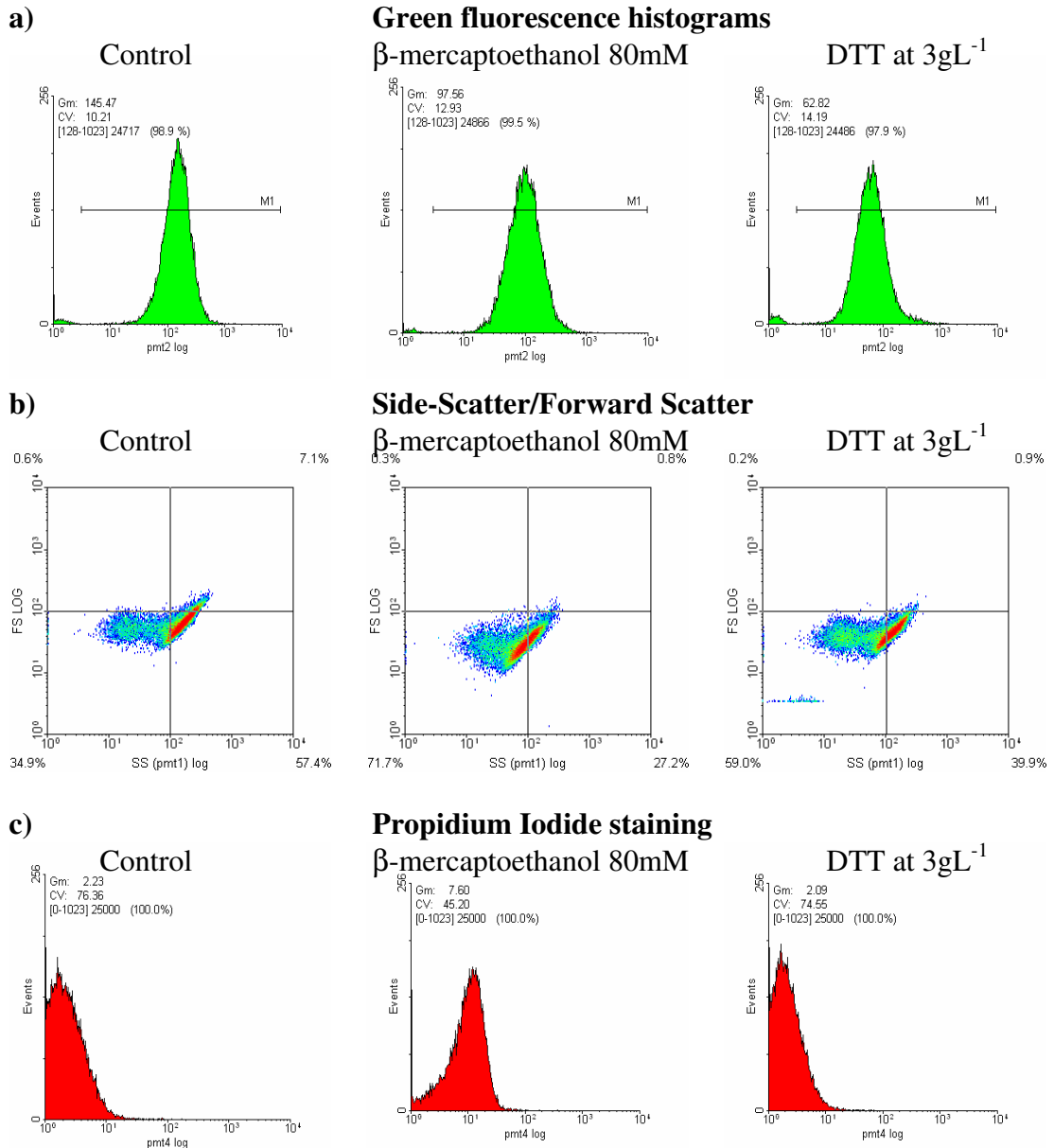


Fig. 4.11 Single cell analysis of *E. coli* BL21* harbouring pAA213 four hours after thiol addition in anaerobiosis

Fluorescence histograms are shown in (a). (b) showed variability in scatter values.. Even if the PI profile looks different in β -mercaptoethanol cultures dead cells were not found in the expected range (c)

Conical flasks containing 30mL of LB supplemented with 0.4% glucose and antibiotics as required were inoculated with 0.3mL of overnight grown bacteria and cultivated aerobically at 30°C. 3g^L⁻¹ DTT was added to cultures four hours after inoculation; cultures free of DTT were incubated as controls. To minimize the impact of DTT on the growth of the cultures were allowed to incubate for a period of 4 hour before addition. Cultures were sampled at regular intervals following DTT addition.

Control cultures containing pFVP25 based vectors grew at similar rates up to the fifth hour post inoculation [Fig. 4.12 (a); (b)]. From this point, the plasmid containing the *rpoH::gfpmut3* fusion grew faster than the culture with the *dnaK::gfpmut3* fusion. Cultures with DTT addition exhibited similar growth inhibition in both reporter systems. However cultures harbouring *dnaK::gfpmut3* seemed to enter stationary phase while the cultures containing *rpoH::gfpmut3* did not show such growth deceleration.

Relative fluorescence increased in cultures with DTT addition [Fig. 4.12 (c); (d)]. Fluorescence readings based on *dnaK::gfpmut3* cultures reached a maximum at 45 minutes post-addition, followed by slight decreases afterwards. The fluorescence generated with the *rpoH::gfpmut3* fusion was, as expected, several fold greater than the fluorescence observed with *dnaK::gfpmut3*.

Unlike cultures containing pFV25 based plasmids, cultures transformed with pUA66 and pDNAK grew at similar rates [Fig. 4.13 (a); (b)]. Both cultures showed growth inhibition

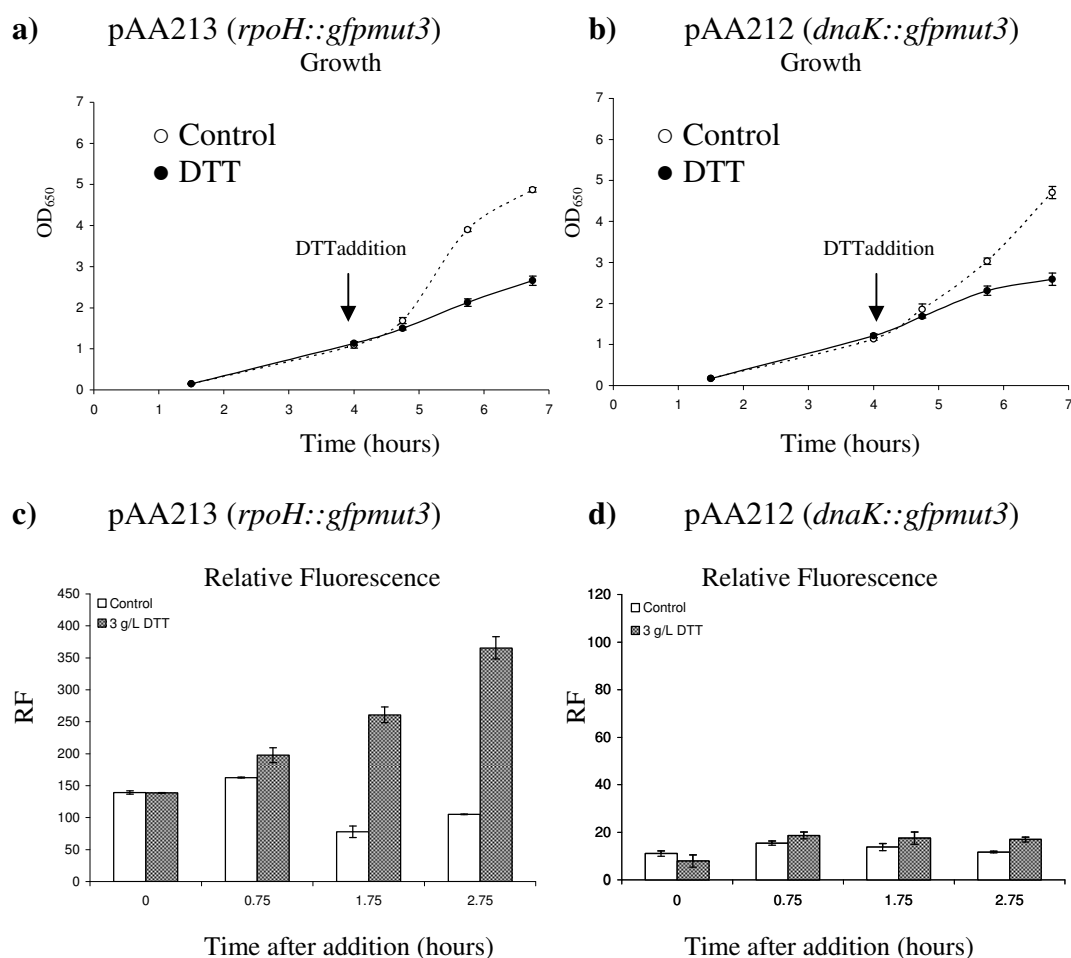


Fig. 4.12 RF and growth of *E. coli* BW25113 transformed with pFVP25 based reporters

Transformants containing either pAA213 (a & c) or pAA212 (b & d) were grown in 30mL of LB for four hours before addition of 3 gL⁻¹ DTT. Growth was monitored by measuring OD₆₅₀ (a & b) and relative fluorescence was measured (c & d).

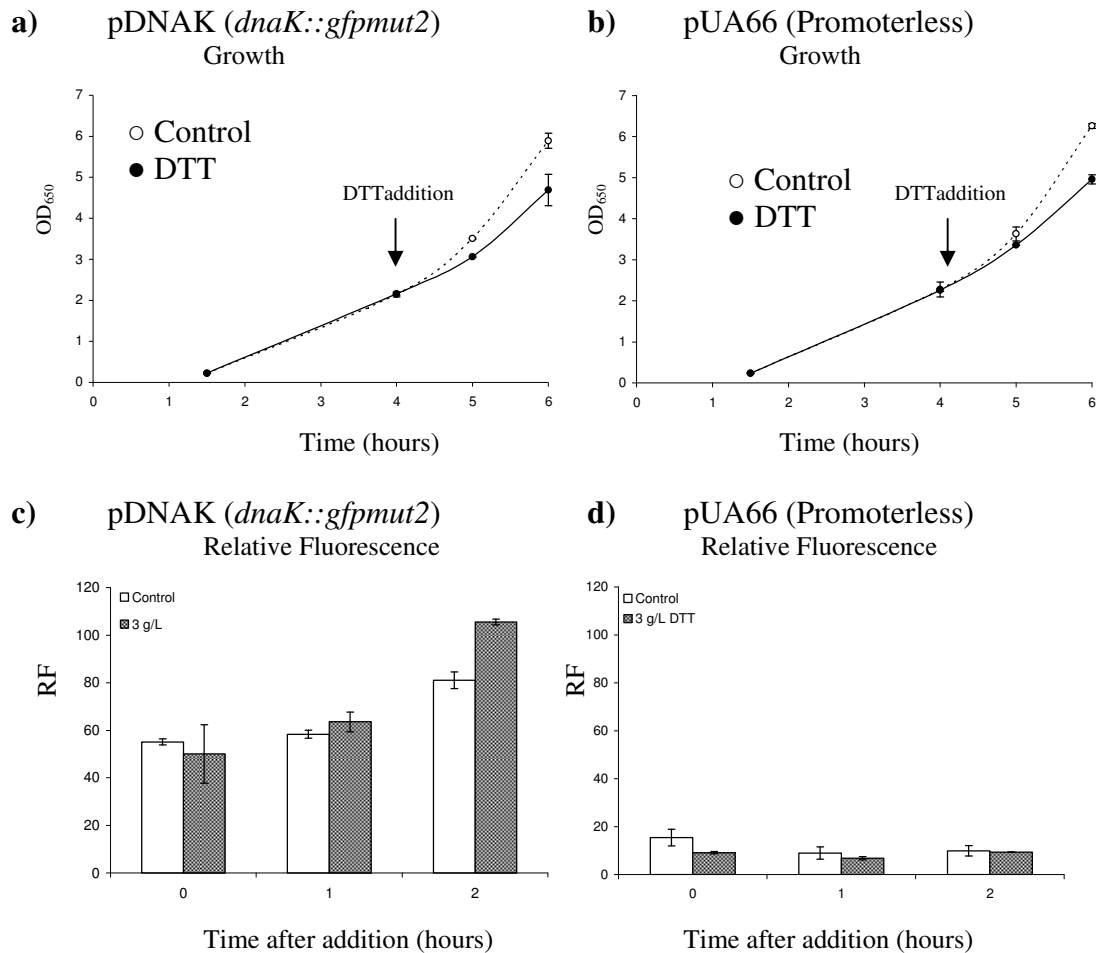


Fig. 4.13 RF and growth of *E. coli* BW25113 transformed with pUA66 based reporters

Transformants containing either pDNAK (a & c) or promoterless pUA66 (b & d) were grown in 30mL of LB for four hours before addition of 3 gL⁻¹ DTT. Growth was monitored by measuring OD₆₅₀ (a & b) and relative fluorescence was measured (c & d).

upon addition of 3gL^{-1} DTT but not as accentuated as that observed with pFV25 plasmids. This could be explained by the lower metabolic burden of low copy number plasmids in comparison with pFV25-based reporters. The cultures carrying promoterless pUA66 grew slightly faster than those with pDNAK. As expected with pUA66, the relative fluorescence was much lower than vectors containing promoters and could represent the basal autofluorescence of the cells [Fig. 4.13 (d)]. The RF of pDNAK cultures increased over time, with the rate of increase being greater in the DTT cultures [Fig. 4.13 (c)]. The fluorescence in general of pDNAK harbouring cells was several-fold greater than the fluorescence observed in cells containing pAA212. This result is counterintuitive as it could be expected that a higher copy number plasmid would generate more reporter protein. The possible causes for such paradox are discussed subsequently.

4.2.5 Comparison of bulk measurements with single cell analysis of *groE::gfpmut2* activation in pGROE

*Bulk measurements of *groE::gfpmut2**

The reporter pGROE was used to monitor the expression of the chaperone GroE in *E. coli* BL21*. This strain is frequently utilized in industry for recombinant protein production, therefore the interest on investigating its *groE* expression. Aerobic and anaerobic cultures of *E. coli* BL21* transformed with pGROE were grown and sampled in the exactly same conditions as described for the experiment shown in [Fig. 4.5]. Stressed cultures were supplemented 80mM β -mercaptoethanol or 3gL^{-1} DTT while aerobic control cultures and cultures in anaerobic conditions were grown free of thiol reductants. Since *E. coli* BL21*

grows poorly anaerobically these cultures were not supplemented with thiols. Fluorimetry measurements revealed an increase in relative fluorescence 2 hours after addition of β -mercaptoethanol while control cultures, DTT and anaerobic cultures showed a decrease in relative fluorescence [**Fig 4.14 (a)**]. Four hours after addition of thiol reductants, stressed cultures displayed higher relative fluorescence than the controls and but the fluorescence of β -mercaptoethanol cultures was twice that of DTT added cultures. The highest relative fluorescence was observed in anaerobic cultures which also displayed the highest rate of increase, nearly 4-fold higher compared to the two hour sample point [**Fig. 4.14 (a)**]. Growth curves [**Fig. 4.14 (b)**] indicated strong growth inhibition in cultures stressed with thiol reductant or grown anaerobically.

Single cell analysis of groE::gfpmut2: Comparison of cultures two hours after thiol addition

At this sample point, the histograms of fluorescence distribution revealed a green fluorescence Gm analogous to the fluorimetry measurements [**Fig 4.15 (a)**]. The β -mercaptoethanol cultures had twice the Gm fluorescence of the control cultures while DTT cultures did not differ significantly from the control. The scatter plots revealed differences in population distribution between control and stressed cultures [**Fig 4.15 (b)**]. A higher density of cells was found in the lower left quadrant in each of these stressed cultures in comparison with the control as a result of a decrease in SS values, which could be indicative of a different morphology. Since FS (indicative of cell diameter) did not varied, it is possible that the cells under these conditions could be showing less cytoplasmic granularity. PI staining resulted in very similar red fluorescence histograms between cultures (not shown) which did not reveal positive staining attributable to cell death.

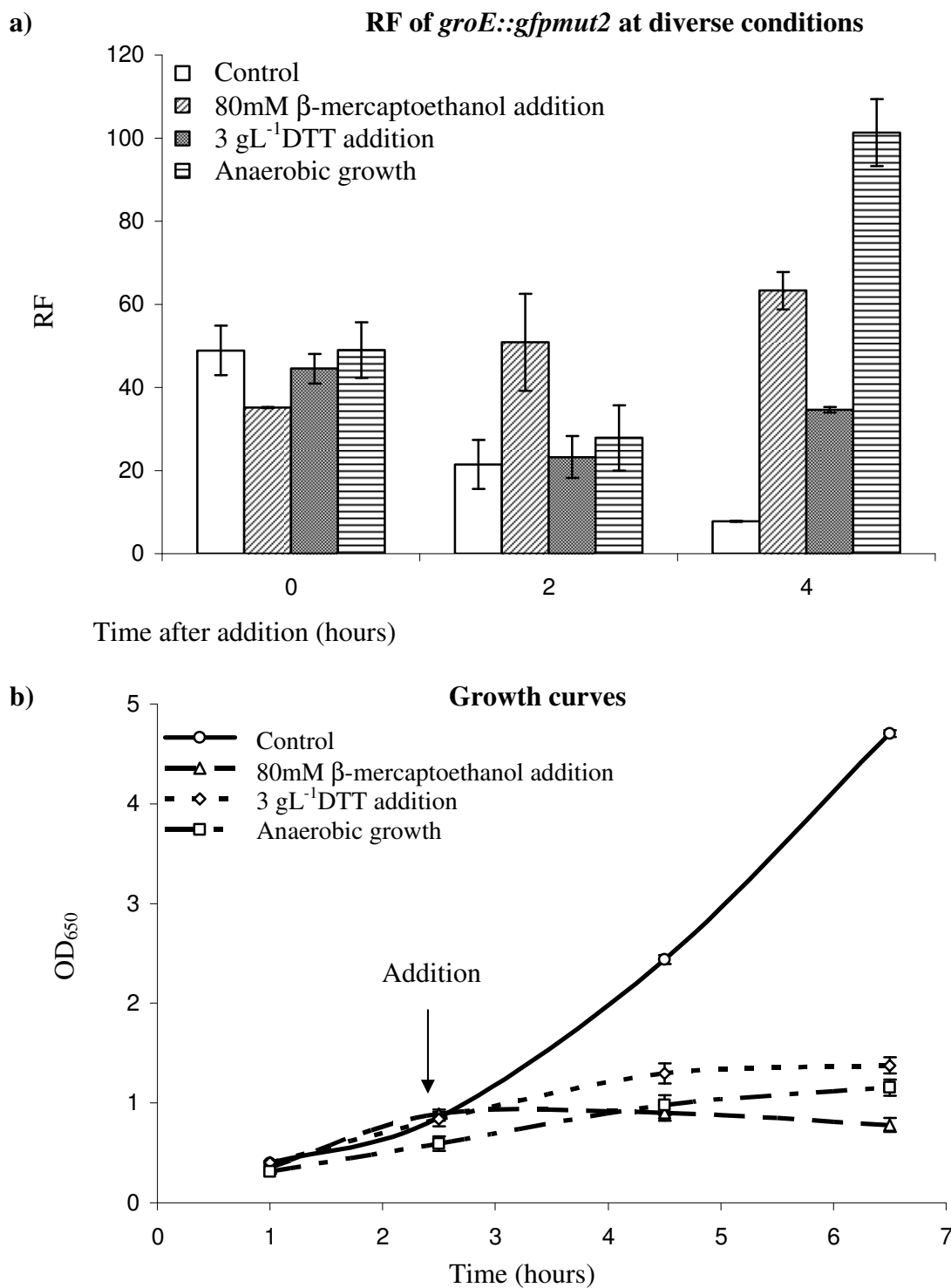


Fig. 4.14 pGROE expression in *E. coli* BL21* in the presence and absence of DTT, β -mercaptoethanol and oxygen

E. coli BL21* transformed with pGROE were grown in 30mL of LB during hours after addition of 3 gL⁻¹ DTT. Relative fluorescence was measured (a) Growth was monitored by measuring OD₆₅₀ (b).

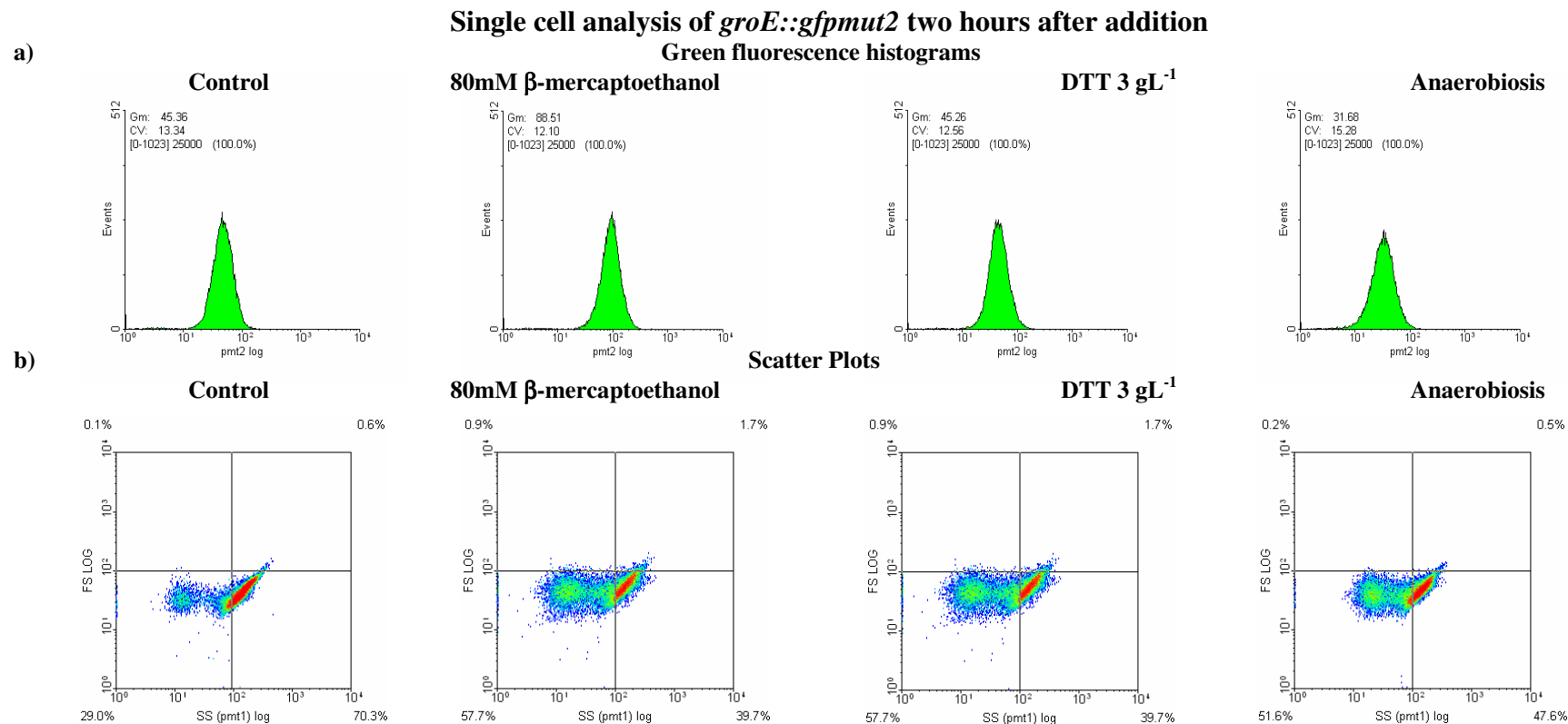


Fig. 4.15 Single cell analysis of pGROE (*groE::gfpmut2*) expression two hours after addition

E. coli cultures BL21* transformed with pGROE were supplemented with 80mM β -mercaptoethanol or 3gL⁻¹ DTT while aerobic control cultures and cultures in anaerobic conditions were grown free of thiol reductants. **(a)** Shows the green fluorescence **(b)** Shows SS/FS spread.

Single cell analysis of groE::gfpmut2: Comparison of cultures four hours after thiol addition (groE::gfpmut2)

The green fluorescence histograms showed that population stressed with β -mercaptoethanol maintained a high fluorescence Gm fluorescence compared with the previous sample while the control and DTT cultures had similar fluorescence Gm [Fig 4.16 (a)]. The fluorescence Gm of anaerobic cultures showed an increase over the control sample but, in contrast with the RF [Fig 4.14 (a)] it did not surpass the Gm β -mercaptoethanol cultures. These cultures, showed the highest in green fluorescence Gm. The scatter plots did not reveal significant FS or SS variations [Fig 4.16 (b)]. The PI staining results did not reveal significant cell death in control and DTT added cultures. A small population positively stained population could be observed in β -mercaptoethanol cultures (1.1% of the total) [Fig 4.16 (c)].

4.2.6 Comparison of bulk measurements with single cell analysis of dnaK::gfpmut2 activation in pDNAK reporter

Bulk measurements of dnaK::gfpmut2

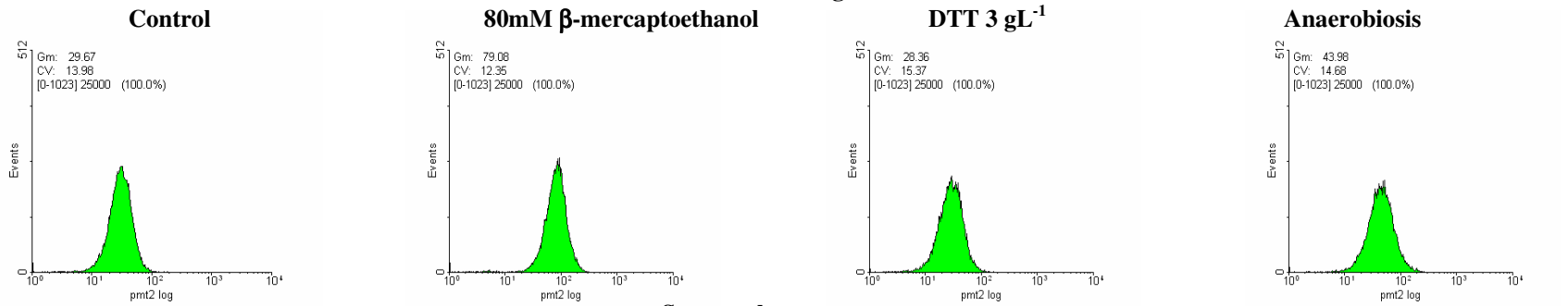
E. coli BL21* transformed with pDNAK were cultured and sampled as the experiments shown in [Fig. 4.5] and [Fig 4.14]. This was done to observe the expression pattern of the chaperone DnaK in *E. coli* BL21* and compare it to that seen in GroE. Pre-addition, cultures had slightly different mean RF values differences were not significant, except in anaerobic cultures which was lower [Fig. 4.17]. Two hours after addition, the control showed higher relative fluorescence than the cultures stressed with thiol reductants. Cultures in anaerobic conditions revealed the lowest fluorescence. Four hours after addition, control cultures exhibited a sharp reduction in relative fluorescence which contrasted with an increase in the RF of β -mercaptoethanol cultures.

Fig 4.16 Single cell analysis of pGROE (*groE::gfpmut2*) expression four hours after addition (next page)

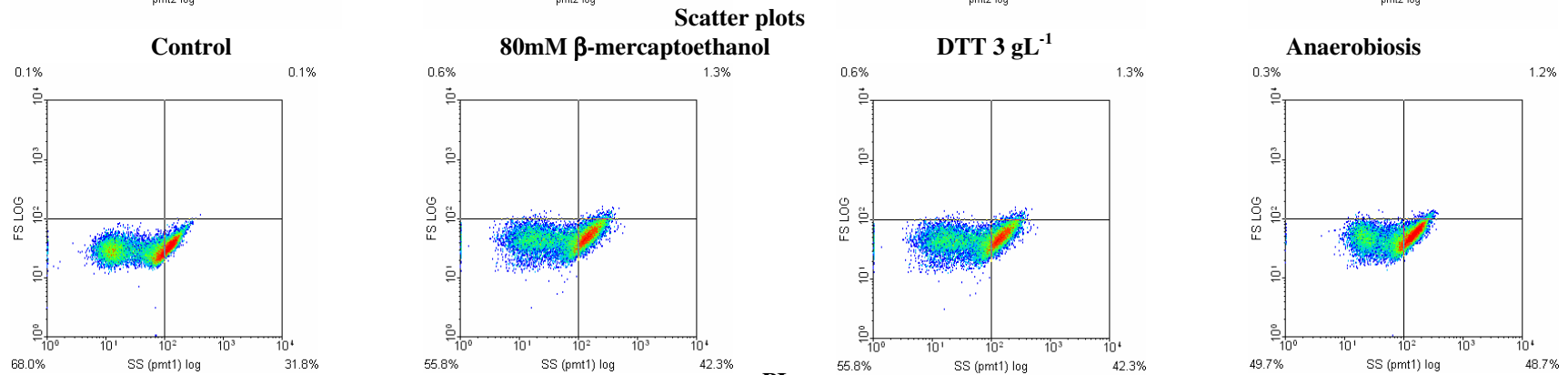
E. coli BL21* cultures transformed with pGROE were supplemented with 80mM β -mercaptoethanol or 3gL⁻¹ DTT while aerobic control cultures and cultures in anaerobic conditions were grown free of thiol reductants. **(a)** Shows the green fluorescence **(b)** Shows SS/FS spread.

a)

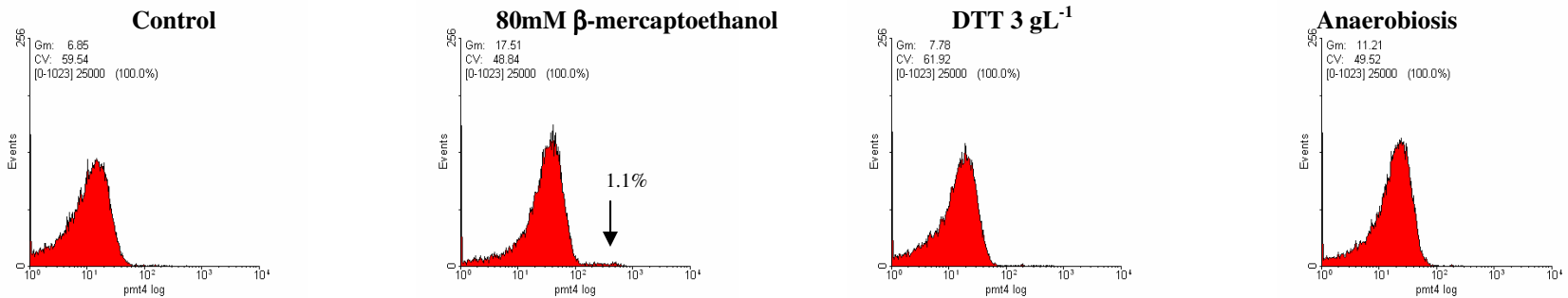
Single cell analysis of *groE::gfpmut2* four hour after addition
Green fluorescence histograms



b)



c)



Cultures stressed with DTT between 2 and 4 hours post-addition did not vary significantly in relative fluorescence [**Fig 17 (a)**]. Anaerobic cultures had decreased reduced their relative fluorescence over time, and RF was lower than other cultures. Growth curves showed control cultures growing at a high rate while the stressed cultures with β -mercaptoethanol, DTT or in anaerobic conditions showed very limited growth [**Fig 17 (b)**].

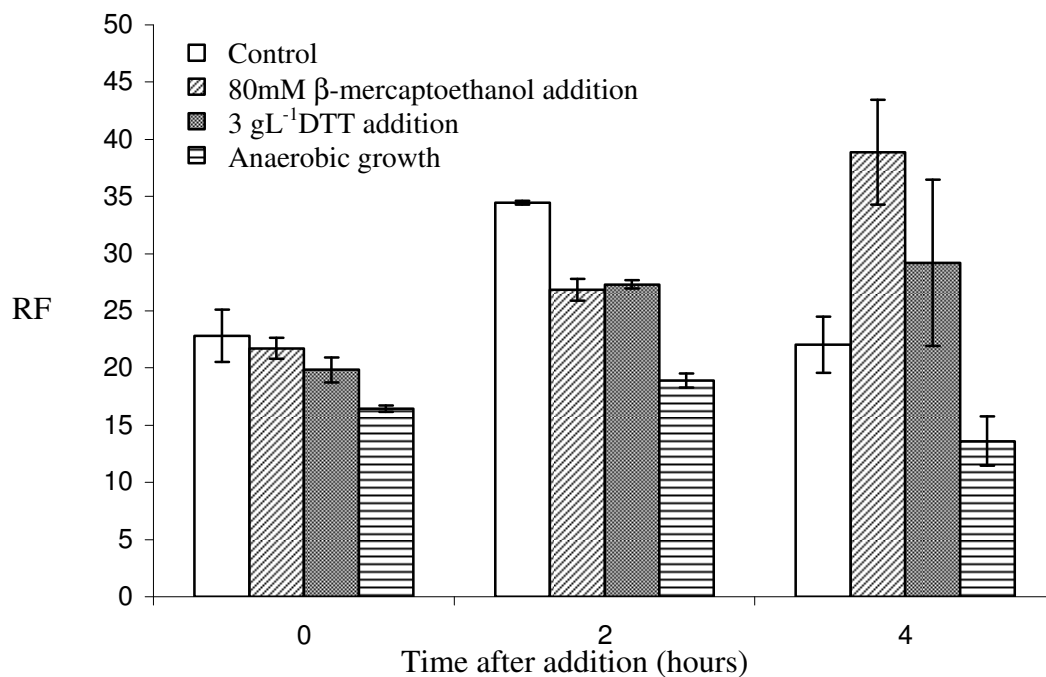
Single cell analysis of dnaK::gfpmut2: Comparison of cultures two hours after thiol addition.

As in the case of *E. coli* BL21* harbouring pGROE, the samples at pre-induction showed very similar characteristics between cultures. Two hours after induction green fluorescence histograms revealed a distribution of fluorescence where Gm values matched RF values from fluorimetry [**Fig 4.18 (a)**]. DTT stressed cultures generated a less fluorescent population compared with the control. The least fluorescent population was registered in anaerobic cultures where the distribution of fluorescence had a larger CV than in the aerobically growing cultures. The light scatter plots revealed a similar population distribution between cultures but different in SS distributions, broader thiol reductant added cultures [**Fig 4.18 (b)**]. PI staining did not show any significant population of dead cells in aerobic cultures although in anaerobic cultures there was a small population (~1.1%) of PI positive cells [**Fig 4.18 (c)**].

Single cell analysis of dnaK::gfpmut2: Comparison of cultures four hours after thiol addition

The green fluorescence histograms [**Fig. 4.19 (a)**] from the control samples possessed a diminished Gm in comparison with the cultures stressed with β -mercaptoethanol. While the control population had a lower Gm than the previous sample (21.03 versus 28.84), the

a) **RF of *dnaK::gfpmut2* at diverse conditions**



b) **Growth curves**

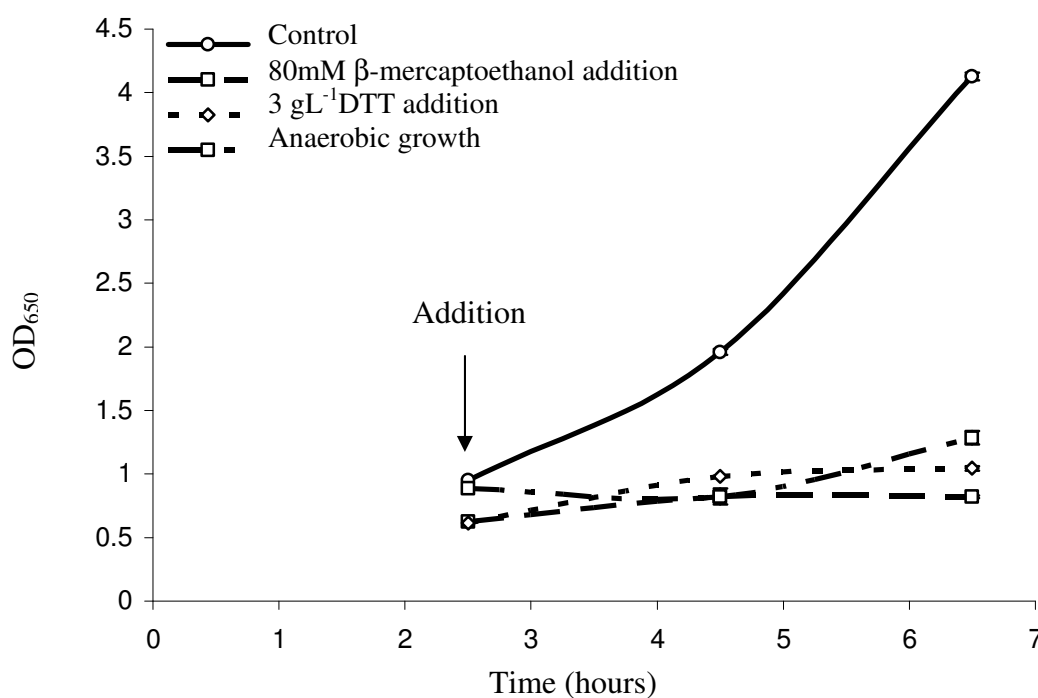


Fig. 4.17 pDNAK expression in *E. coli* BL21* in the presence and absence of DTT, β-mercaptoethanol and oxygen

E. coli BL21* transformed with pDNAK were grown in 30mL of LB during 4 hours after addition of 3 gL⁻¹ DTT. Relative fluorescence was measured (a) Growth was monitored by measuring OD₆₅₀ (b).

Single cell analysis of *dnaK::gfpmut2* two hours after addition
Green fluorescence histograms

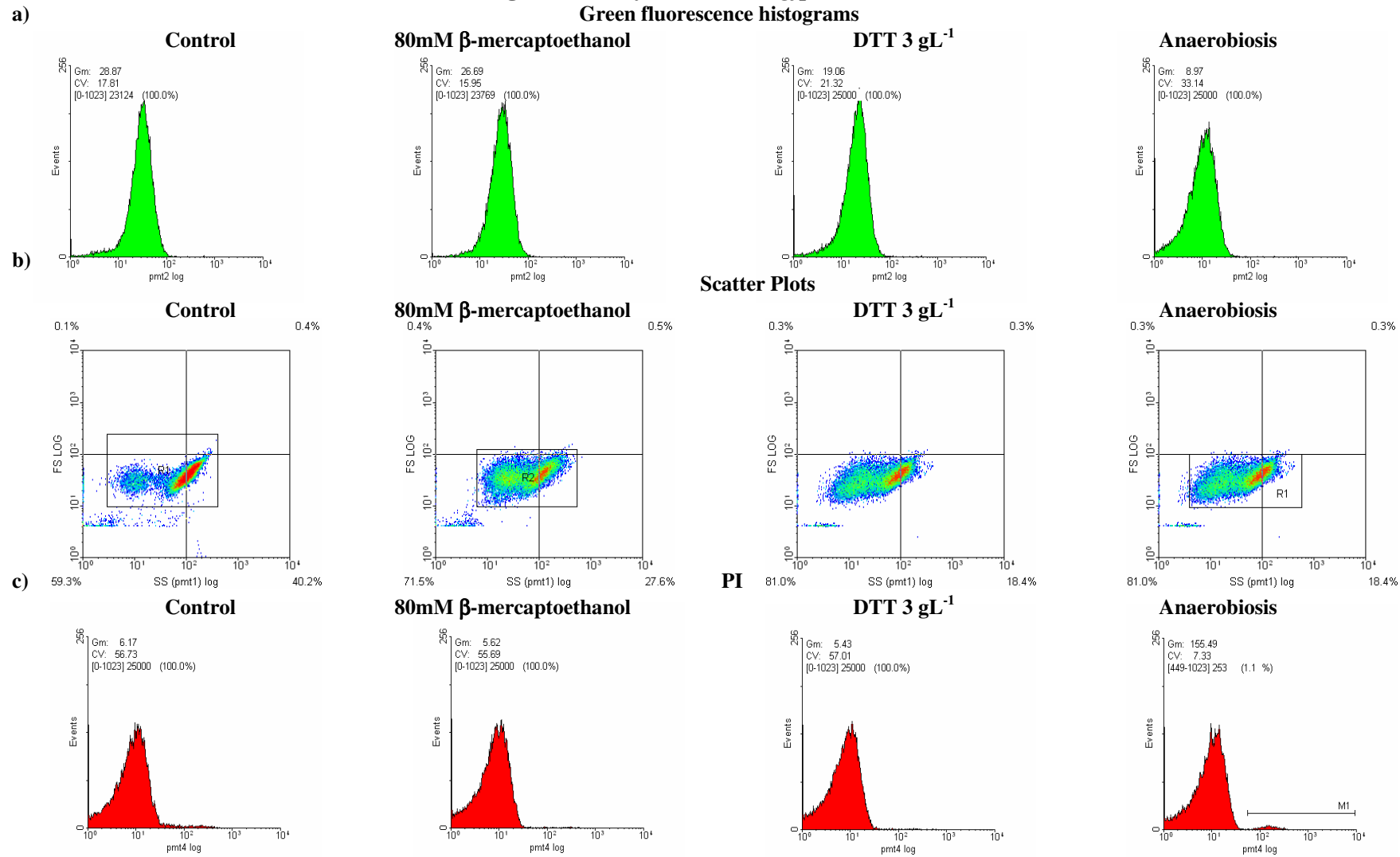


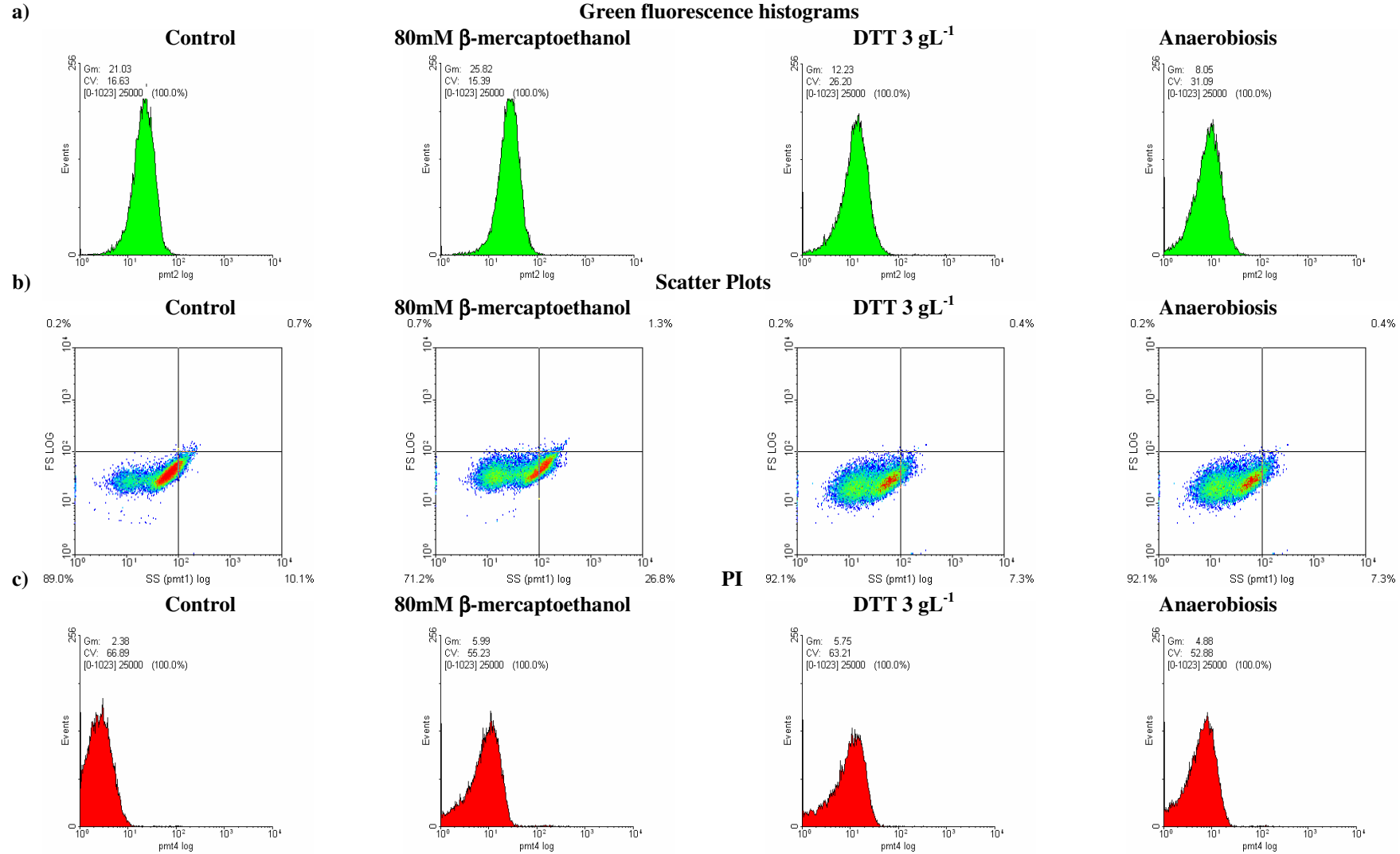
Fig. 4.18 Single cell analysis of *E. coli* BL21*harbouring pDNAK two hours after addition

Cultures *E. coli* BL21* transformed with pDNAK were supplemented 80mM β -mercaptoethanol or 3gL⁻¹ DTT while aerobic control cultures and cultures in anaerobic conditions were grown free of thiol reductants. **(a)** Shows the green fluorescence **(b)** Shows SS/FS spread. **(c)** Shows PI staining.

Fig. 4.19 Single cell analysis of *E. coli* BL21*harbouring pDNAK four hours after addition

Cultures *E. coli* BL21* transformed with pDNAK were supplemented 80mM β -mercaptoethanol or 3gL⁻¹ DTT while aerobic control cultures and cultures in anaerobic conditions were grown free of thiol reductants. **(a)** Shows the green fluorescence **(b)** Shows SS/FS spread. **(c)** Shows PI staining.

Single cell analysis of *dnaK::gfpmut2* four hours after addition
Green fluorescence histograms



Gm of β -mercaptoethanol cultures did not vary with respect to the previous sample [Fig 4.19 (a)]. In cultures stressed with DTT, the fluorescence Gm decreased sharply from the 2 hours sample (19.06 to 12.23). The anaerobic cultures did not show apparent change between 2 and 4 hours except for a slight decrease of FS values [Fig 4.19 (b)]. PI staining [Fig 4.19 (c)] results indicated that there were no significant populations of dead cells.

4.2.7 Partial characterization of *yidQ* regulation using pYIDQ

Origin of pYIDQ

YidQ, also referred to as EcfI (Dartigalongue, Missiakas, and Raina, 2001) (Extracytoplasmic factor I), is annotated in the genome sequence as a putative outer membrane protein of unknown function (Keseler et al., 2011). *yidQ* and *ibpA* are transcribed divergently from overlapping promoters [Fig 2.4]. Originally, in the present study it was planned to clone the *groE*, *dnaK* and *ibpA* promoters into pUA66 to generate custom reporter plasmids. However, cloning was difficult due to the small distance between restriction sites which was only 6 bp, resulting in low efficiency by the second digestion. In addition, the sizes of double digested and single digested vectors were apparently of the same size in agarose gels; therefore it was not possible to isolate pure double digested vector. Following PCR amplification of promoter regions, ligation into pUA66 and transformation, colonies were screened for GFP-producing plasmids by flow cytometry. Plasmids isolated from candidate green fluorescent colonies were sequenced and found to contain the *ibpA* promoter, but in the incorrect orientation, so that the *yidQ* promoter was fused to the *gfp* gene on pUA66. This plasmid was named pYIDQ.

Potential of $yidQ::gfpmut2$ fusion as a reporter

Inspection of the cloned $yidQ$ promoter sequence in pYIDQ revealed a binding site for σ^E (Dartigalongue et al., 2001), making it a potential candidate as a reporter plasmid for periplasmic stress. In addition, there was computational evidence for CpxR-P binding sites in the promoter (Dartigalongue et al., 2001; De Wulf et al., 2002). The following experiments were performed to partially characterize pYIDQ regulation in diverse environments as a potential reporter for bioprocess stress related.

The σ^E promoter was difficult to activate

Factors and chemicals known to activate genes of the σ^E regulon (such as $rpoH$, $degP$ and the $rpoE$ operon) were utilized in an attempt to activate the pYIDQ reporter. *E. coli* BW25113 (WT) and ECA101 ($\Delta\sigma^E$) were transformed with pYIDQ. Overnight cultures were used to inoculate 250mL conical flasks containing 30mL of LB and $25\mu\text{g mL}^{-1}$ of Kanamycin. Cultures were grown aerobically at 30°C. All work was performed on duplicate cultures. In addition, to observe the effect of recombinant protein on $yidQ$ gene expression, *E. coli* BL21* was transformed with pYIDQ and pET-CCP (refer to Chapter Five for further details on this expression vector) were recombinant protein production did not causes a response in pYIDQ.

In successive experiments, cultures were supplemented with 5% ethanol, 175mM CuCl_2 all treatments shown previously to induce σ^E -dependent promoters (Connolly et al., 1997; Hirano et al., 2007). In all cases, WT fluorescence was compared with $\Delta\sigma^E$ during growth; addition of 175mM CuCl_2 or 5% ethanol did not show significant differences in fluorescence between WT and $\Delta\sigma^E$ independently of the carbon source added (either glucose 0.4% (w/v) or glycerol

0.4% (v/v)). However, it was observed that both strains accumulated more fluorescence over time when growing in glycerol than in glucose.

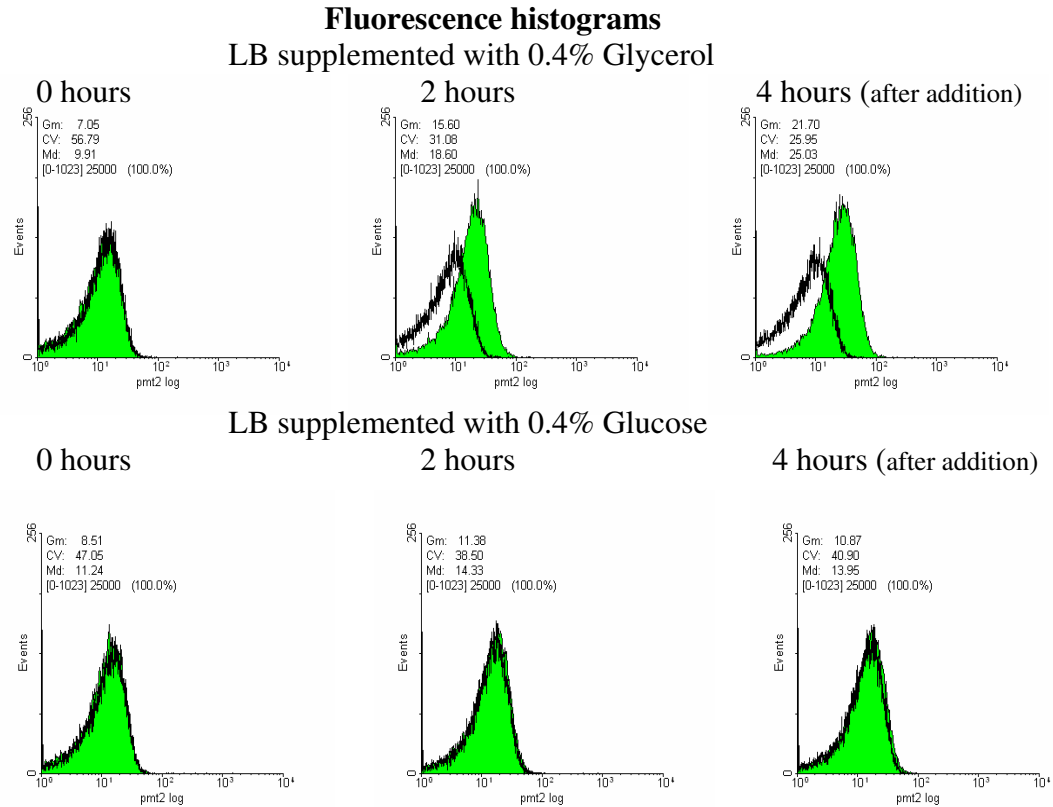
Catabolite repression of pYIDQ

E. coli BW25113 pYIDQ was grown as described above supplemented with glycerol or glucose. BW25113 harboring pUA66 was used as a fluorescence negative control. Cultures were stressed with the addition of ethanol to 5% one hour post-inoculation. Cultures with no ethanol added were grown as a control for ethanol response. Samples were taken at 2 hours intervals.

The green fluorescence histograms [Fig. 4.20 (a)] revealed a gradual increase of green fluorescence (up to two-fold at 4 hours) after ethanol addition only in the cultures supplemented with glycerol. In addition, the toxic effect of ethanol caused greater growth rate reduction in glucose-supplemented cultures, where there was no difference in fluorescence upon addition of ethanol [Fig. 4.20 (b)], while glycerol cultures grew faster and to greater biomass. As growth rate was decreased in the presence of ethanol in both glucose –and glycerol- supplemented culture, independently of green fluorescence increasing, this supports the independence of fluorescence and growth rate.

The toxic effect of ethanol caused greater growth rate reduction in the glucose-supplemented cultures. There was no difference in fluorescence upon addition of ethanol [Fig. 4.20 (b)], but glycerol cultures grew faster and to greater biomass supporting the idea that the accumulation of GFPmut2 is not caused by changes in growth rate. When this experiment was repeated in ECA101 ($\Delta\sigma^E$), there were no significant differences to BW25113 except for a slight

a)



b)

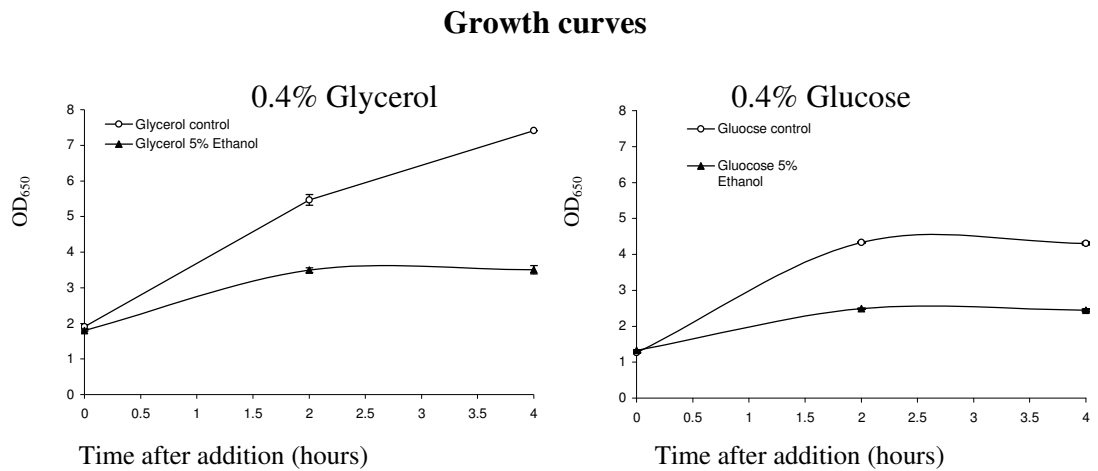


Fig. 4.20 Catabolite repression of pYIDQ

E. coli BW25113 pYIDQ was grown aerobically in 30ml of LB supplemented with either 0.4% glucose or 0.4% glycerol. After one hour, ethanol was added to a concentration of 5% to half of the cultures. Green fluorescence was measured using flow cytometry **(a)**; green traces are plus ethanol, black lines minus ethanol. **(b)** Shows the growth curves.

reduction of fluorescence, therefore the *gfpmut2* is likely to be expressed chiefly from a σ^{70} -dependant promoter.

Intervention of CpxRA

The promoter sequence of pYIDQ contains putative CpxR-P binding sites [Fig. 2.4]. In order to observe the influence of CpxRA on pYIDQ regulation, the *E. coli* mutant strains JW3882 ($\Delta cpxR$) and JW3883 ($\Delta cpxA$) were transformed with pYIDQ. Overnight cultures were used to inoculate 250mL conical flasks containing 30mL of LB and 25mgmL⁻¹ of Kanamycin. Cultures were grown aerobically at 30°C. *E. coli* BW25113 pYIDQ was also cultured to enable comparison with the wild type. Cultures were allowed to grow for 1 hour before addition of ethanol to 5%. Samples were taken at hourly intervals. All work was performed in duplicate. The effect of 5% ethanol on growth was very similar in all cultures [Fig. 4.21 (a)]. Flow cytometric measurement of GFP fluorescence [Fig. 4.21 (b)] revealed that ethanol activated pYIDQ in the wt and $\Delta cpxR$ strains but not the $\Delta cpxA$ strain. Differences in FS/SS distribution were not observed between cultures. At two hours post induction, ethanol stressed $\Delta cpxR$ cultures showed double the average fluorescence of the unstressed. Interestingly, during the first two hours post induction the fluorescence of the stressed $\Delta cpxR$ cultures was even greater than that shown by the WT. The unstressed $\Delta cpxR$ and WT cultures also accumulated fluorescence gradually over time, unlike the $\Delta cpxA$ cultures. Bringing into account previous experiments, this indicated some degree of expression independent of σ^E but dependent on $\Delta cpxA$, CpxA posses a phosphatase domain able to dephosphorylate CpxR-P regulating its activity. It is possible that removal of CpxA implies a diminished control on CpxR-P concentration which could be acting as a *yidQ* repressor.

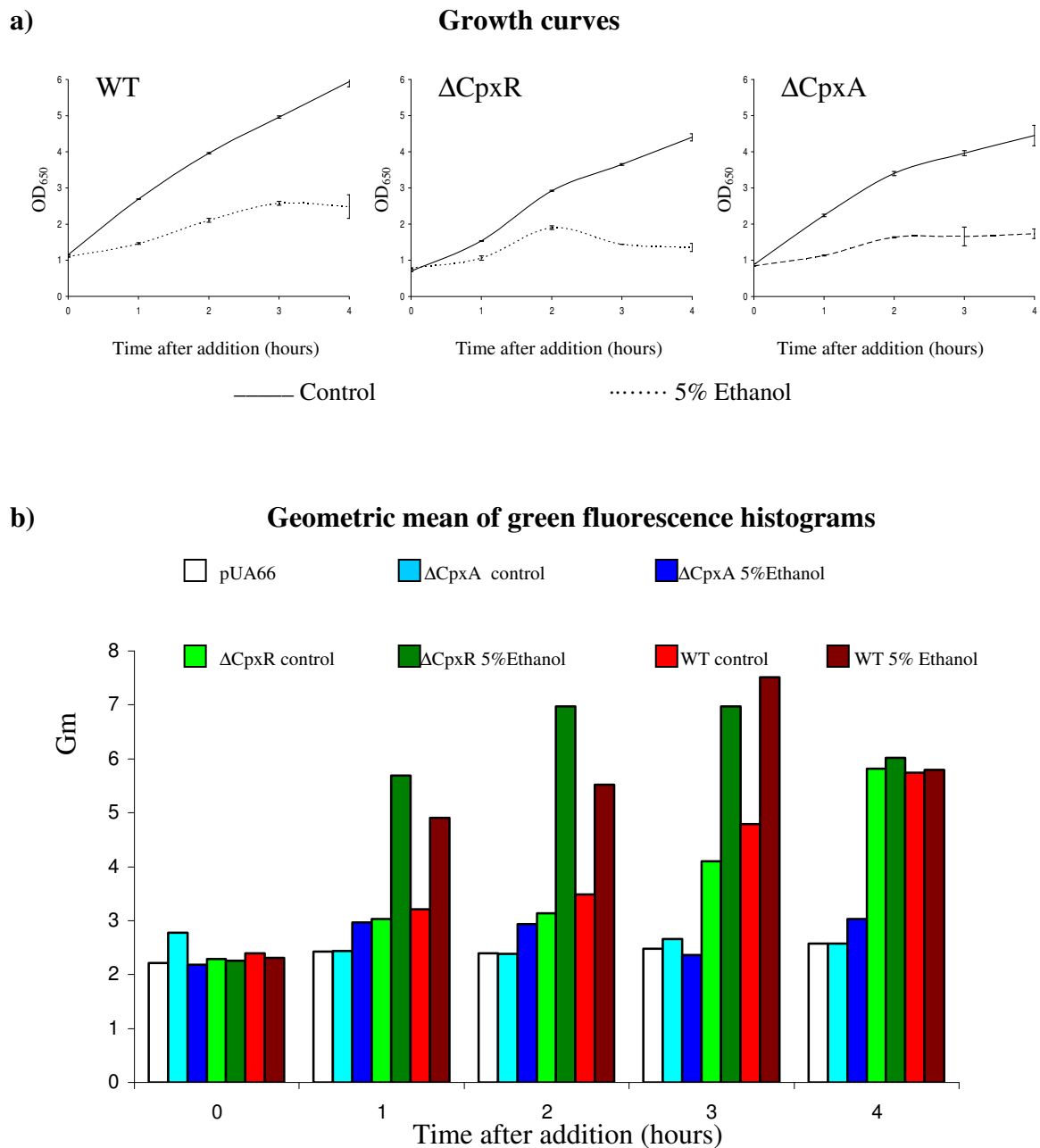


Fig. 4.21 CpxR and CpxA regulation of pYIDQ

E. coli BW25113 (WT), JW3882 (Δ cpxR) and JW3883 (Δ cpxA) were grown at 30°C in LB only. **(a)** Shows growth curves exhibited by different *E. coli* genotype. **(b)** Shows the growth Gm of green fluorescence of representative cultures for each genotype upon 5% Ethanol addition. pUA66 indicates the fluorescence of WT harboring this plasmid as negative control of fluorescence.

4.3 Discussion

Very low copy and medium copy plasmids as vectors for reporter systems

It was observed that bacteria transformed with the very low copy number pUA66 reporter plasmids with a SC101 origin of replication grew better than those transformed with pFVP25vectors carrying the ColE1 replicon (compare [Fig. 4.12] to [Fig. 4.13]). The latter vectors replicate to ~30 copies per cell (Bartolomé et al., 1991) which can result in metabolic burden to the cell. This was reflected in the growth curves. In spite of existing more copies of pAA212 and per cell than in pDNAK transformants and both plasmids containing the same σ^{32} binding sites, pAA212 resulted in less relative fluorescence than the pDNAK reporter based on pUA66. This is likely to be caused because the promoter insert in pAA212, unlike pDNAK, does not include any codon from the *dnaK* gene which is thought to be beneficial for reporter expression.

There are only 20 to 30 copies per cell of active free σ^{32} in the cytoplasm of *E. coli* when incubated at 30°C (Craig and Gross, 1991). These reporter plasmids include ~200bp of DNA upstream of the regulated gene (Aertsen et al., 2004). This region of the *dnaK* promoter includes 3 binding sites for σ^{32} , two of which are σ^{32} -specific while the third may also be bound by σ^{70} (Liberek et al., 1992). This could cause a physiological problem as the σ^{32} pool has to be distributed amongst the σ^{32} binding sites on the genome and the reporter plasmids. Therefore, it is possible that the fluorescence reduction is attributable to available σ^{32} being titrated out by excess σ^{32} binding sites on the medium copy number plasmids. pDNAK, a low copy plasmid, seemed to mitigate this problem as less σ^{32} would be sequestered. Recently some researchers are inclined to integrate the reporter system into the host genome precisely

to reduce such interference (Nemecek et al., 2008). However, utilizing plasmids has other advantages as the reporter system can be mobilized easily into any desired strain as long as it is compatible. As cells containing the medium copy number plasmid pAA213 (*rpoH::gfpmut3*), which does not possess σ^{32} binding sites, grew faster than those containing pAA212 (*dnaK::gfpmut3*), it is possible that the σ^{32} demand from the multiple binding sites created by plasmid replication interfered with the heat-shock response which, as discussed in chapter three, could have a relationship with growth phase. The relative fluorescence measurements obtained using these plasmids were in the range of those obtained by (Aertsen et al., 2004).

Effect of thiol reductants on rpoH expression

Most of the expression of the *rpoH* gene is constitutive, involving several σ^{70} -dependent promoters, a σ^E specific promoter and control by different transcription factors (Kallipolitis and Valentin-Hansen, 1998). In addition, a consensus binding site for σ^{54} (P6) was recently found in the regulatory region (Janaszak et al., 2007) and σ^S may substitute σ^{70} at promoter P1 to ensure transcription during stationary phase (Janaszak et al., 2009). All this regulation seems to ensure that this important gene is expressed in a myriad of stressful situations. Therefore, the effects of thiol reductant on *rpoH::gfpmut3* might be indirect. DNA microarrays have shown that *rpoH* expression may be reduced 2-fold when acetate is present (Oh and Liao, 2000). Therefore, it is possible that these observations are caused by acetate formation as bacteria growing exponentially may ferment the glucose in the media to acetate in batch cultures.

E. coli BL21* has been optimized to reduce acetate formation, although some acetate may be formed during oxygen limitation (Åkesson et al, 2001) possibly explaining why GFP accumulated in exponentially growing *E. coli* BL21* in aerobiosis [Fig 4.5 (a)] while it decreased from two to four hours anaerobically [Fig 4.5 (c)]. The first defence of *E. coli* to thiol reductants has been shown to be in the periplasm where the proteins DsbA and DsbB maintain the oxidizing environment and perform disulphide bond formation, regenerating their redox potential from the respiratory chain (Søballe and Poole, 2000). In aerobic growth it is possible that DsbA and DsbB reduction by thiols resulted in an envelope stress response through σ^E inducing the promoter (P3) of *rpoH::gfpmut3* causing GFP accumulation. Such accumulation is not evident until four hours [Fig 4.11] possibly because the σ^E site on *rpoH* promoter is repressed by CRP becoming active upon glucose depletion (Zheng et al., 2004).

The aerobic cultures with thiol reductants DTT and β -mercaptoethanol, had higher RF values in comparison with the controls, this difference was also visible using single cell analysis ([Fig. 4.8] and [Fig 4.9]). As addition of thiols in general caused growth rate reduction, it becomes difficult to judge if the accumulation observed is a consequence of GFPmut3 synthesis or accumulation over time in non-dividing cells.

Flow cytometry adds a new dimension to the data

Results have shown that in most of the cases the flow cytometric data agreed with the data collected from fluorimetry. However there were some exceptions, particularly in situations when stressed cultures were not showing as much fluorescent increase in single cell analysis as expected from RF; for example, four hours after addition of β -mercaptoethanol to anaerobic cultures, the RF was higher than the corresponding control culture [Fig 4.5 (b)]

whereas the Gm was lower [Fig 4.11 (a)]. Flow cytometry measurements are expected to be more reliable as samples do not require on further steps such as LB removal and dilution reducing analytical error.

Flow cytometry could provide information about heterogeneity in the samples. FS measurements are typically a function of size and therefore biomass as established by a combination of diverse mathematical models with empirical measurements (Robertson, Button, and Koch, 1998). Thus flow cytometry may discriminate cells of different biomass as a function of the forward scatter. As cells replicate plasmids, synthesise proteins or become stressed they may vary their volume and, as consequence, the values in forward scatter. Cells grow and divide during the cell cycle which is likely to produce a distribution of FS values. It is possible that bigger cells are more fluorescent because may contain more GFP. Adding to the recommendations by Leveau and Lindow (2001) about utilization of flow cytometry when samples are from mixed environments or likely to be heterogeneous; it could be also recommended that single cell measurements are used when variability in size or morphology of the cell is expected.

groE::gfpmut2 responded to thiol reductants

The reporter *groE::gfpmut2* has been shown to be activated after addition of β -mercaptoethanol using relative fluorescence and flow cytometry. In general, flow cytometry of very low copy number plasmids correlated better with relative fluorescence measurements than pAA213. The results upon DTT addition presented here did not produce comparable increases in chaperone gene expression to those observed in previous studies, where concentrations between 1gL^{-1} (Gill et al., 2001) and 3gL^{-1} (Gill et al., 1998) produced

significant DnaK and GroE chaperone production. These experiments were performed in larger scale fermentations during longer periods of time, including recombinant protein production which was shown to be a leading factor for chaperone expression. The DnaK and GroEL protein levels were measured in shake flask cultures induced at mid exponential phase (Gill et al., 1998). Their results are hard to compare with those presented here; Gill et al. (1998) pointed out that the chaperone production on DTT addition varied greatly between strains. In addition, they measured protein concentrations rather than gene expression, and the reporter pGROE contains the *groES* promoter, [Fig 2.2 (b)] whereas GroEL production is regulated by both the *groES* and *groEL* promoters. In spite of this, the increase of reporter expression observed here is compatible with their observations. On the other hand, the results obtained here provide evidence of the effect of controlled reducing environment as pGROE became activated upon β -mercaptoethanol addition. The reasons why DTT did not cause similar induction could be attributed to differences in the molecule as β -mercaptoethanol is smaller and could permeate faster to the cytoplasm where is more likely to be the cause of folding stress (Gelman and Prives, 1996; Lodish and Kong, 1993). These results were similar to those obtained in Chapter Three where β -mercaptoethanol was shown to activate a σ^{32} -dependent promoter. However, activity upon heat-shock differed between pQF50kgroE and pGROE, which can be explained by differences in the length of promoter cloned into each plasmid: pQF50kgroE simply comprises the σ^{32} binding site, whereas pGROE contains the full promoter sequence *groE::gfpmut2* including the σ^{70} site [Fig 2.2 (b)] and potential CpxR-P (Lau-Wong et al., 2008) and CRP binding sites as discussed subsequently.

dnaK differed from groE as reporter to thiol reductants

The *dnaK::gfpmut2* reporter responded to thiol reductants differently to the *groE::gfpmut2* reporter. (Gill et al., 1998) reported more expression of DnaK than GroE in samples 20 minutes after addition of DTT. It was argued that, while DnaK is produced faster in response to stress, GroE could accumulate during longer intervals of time. *groE::gfpmut2* data showed agreement with this statement accumulating fluorescence over time consistently above the controls [Fig. 4.14]. *dnaK::gfpmut2* cultures did not produce fluorescence as such early response but it cannot be discarded that such response could develop and reverse in periods of time shorter than 2 hours. Differences amongst strains can affect the comparison. *E. coli* BW25113 harbouring pDNAK showed an increase in RF two hours after DTT addition whereas the same reporter in *E. coli* BL21* showed decreased expression in comparison with the control RF ([Fig. 4.12 (a)] and [Fig. 4.17 (a)] respectively).

While σ^{32} is the major factor directing transcription of both genes, different expression levels may have been originated as consequence of different promoter structures. As these reporters contain a large portion of the promoter region it is possible that transcription factors favour the expression of *groE* upon sustained presence of thiol reductants. Microarray data showed that *dnaK*, and *groES/EL* expression may be partially reduced due to catabolite repression, found to be stronger in *dnaK* than in *groE* (Gosset et al., 2004). However, Crp probably binding sites have not been found at the *dnaK* promoter which may indicate that CRP control could be indirect. An interesting observation of the *dnaK::gfpmut2* fluorescence histograms is the distribution of the fluorescence amongst the population. While *rpoH::gfpmut3* and *groE::gfpmut2* resulted in symmetrical distributions, *dnaK::gfpmut2* fluorescence was distributed asymmetrically. More cell counts were accumulated in the left side of the histogram. This suggests that cell accumulated fluorescence in increasing expression profiles reaching a maximum. Therefore stressed state and non-stressed may differ not only by

variations in the geometric mean but also in the distribution of the data, a characteristic only revealed by single cell analysis.

yidQ regulation is complex

While the reporter pYIDQ does not seem to be relevant to report bioprocess related stress there are some observations which remain interesting. A model for pYIDQ regulation is proposed [Fig. 4.23]. The *yidQ* promoter is an example of how a short DNA sequence may be tightly regulated by σ -factors and DNA binding proteins. There is evidence of catabolite repression [Fig 4.20], and CpxRA intervention [Fig 4.21]. The binding sites for the response regulator CpxR-P were inferred computationally (Dartigalongue et al., 2001; De Wulf et al., 2002). In addition, microarray data (Lau-Wong et al., 2008) revealed strong activation of *ibpA* by CpxR-P, especially in mutants with CpxA* (CpxA altered to perform only phosphatase reactions), lacking the kinase activity, necessary for phosphorylation of CpxR. Since the σ^{32} -dependent promoter of *ibpA* is divergent to and overlapping the putative σ^E -dependent promoter of *yidQ* [Fig. 4.23] it is unlikely that both σ^{32} and sigma σ^E can bind to the divergent promoter region at once. CpxR-P activation of *ibpA* would recruit σ^{32} , thus preventing σ^E binding, thereby repressing *yidQ*. As shown in the diagram [Fig. 4.23], the combination of these factors is perhaps the reason why the *yidQ* promoter on pYIDQ was not activated by stresses that usually activate σ^E regulon genes in WT/ $\Delta\sigma^E$ experiments presented here. The whole *ibpA-yidQ* intergenic region has more potential CpxR-P sites potentially modulating gene regulation (Dartigalongue et al., 2001), but the DNA fragment cloned in pYIDQ only contains two. This may explain why the $\Delta cpxA$ strain had low *yidQ::gfp* activity; removal of the CpxA sensor eliminates the kinase domain but also the phosphatase activity,

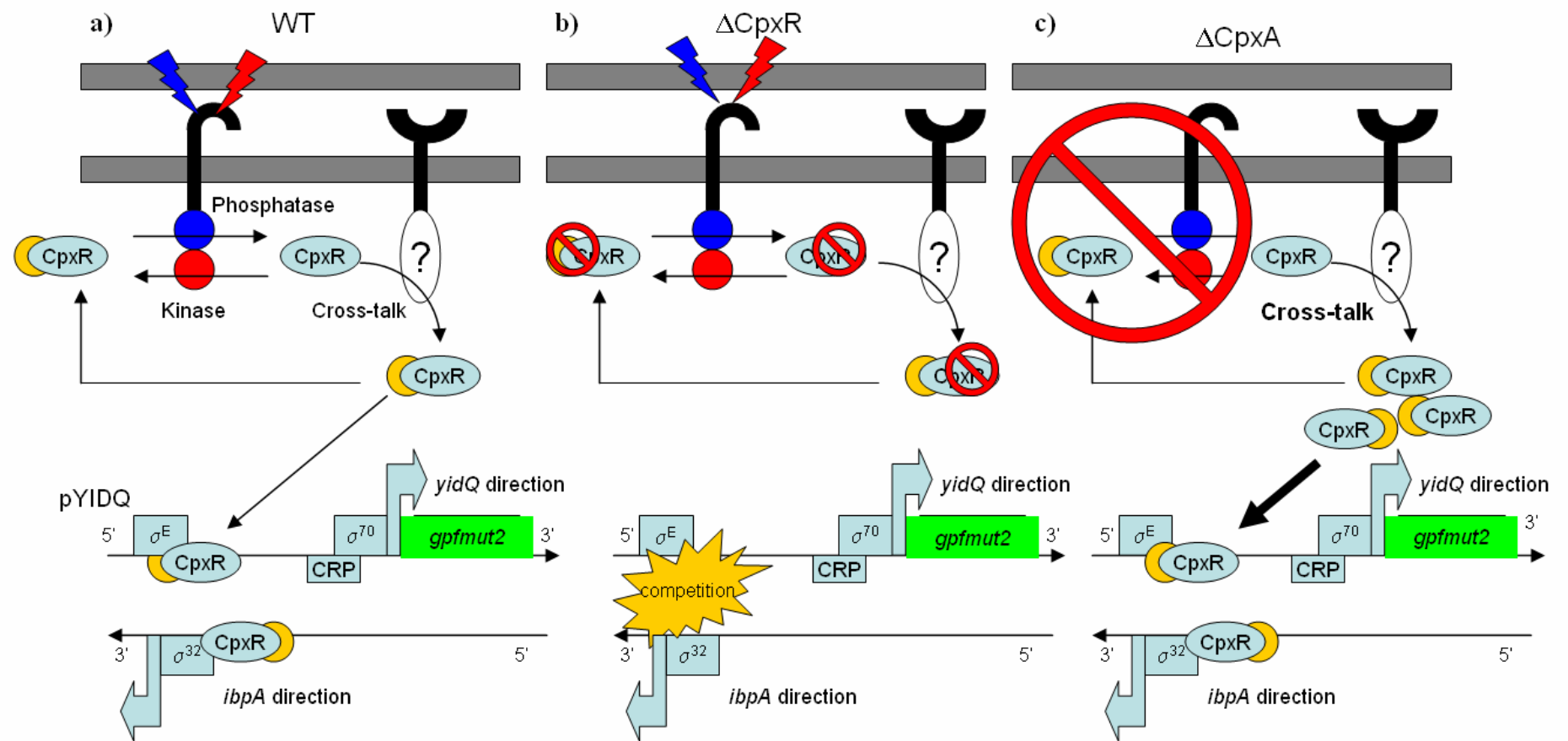


Fig. 4.22 Proposed diagram for pYIDQ regulation by the CpxAR system.

(a) In the case of WT cells. CpxA maintains the levels of CpxR-P (P is represented by a golden circle) via phosphatase (blue) and kinase (red) activities depending on the stimuli sensed. It is possible that an undetermined kinase can phosphorylate CpxR. Below, pYIDQ show how levels of CpxR-P may block the σ^E binding site and favour σ^{32} in agreement with Lau-Wong et al. (2008). (b) In the case of the Δ CpxR mutant, there is no response regulator to be phosphorylated by CpxA. Still, σ^E binding may be by competition of σ^{32} to bind. (c) Total elimination of CpxA results in loss of specific regulation of CpxR-P by phosphatase activity. CpxR may be phosphorylated by other sensor-kinase which was not confirmed by the experiments presented here therefore remains unconfirmed (although EnvZ is a candidate (Siryaporn and Goulian, 2008)). This results in CpxR-P accumulation further repressing promoter. In all cases some σ^{70} -dependent expression modulated by a CRP binding site is expected.

thus potentially allowing CpxR proteins to be phosphorylated by other sensor kinases such as EnvZ system (Siryaporn and Goulian, 2008) to be active for longer. In support of this $\Delta cpxR$ [Fig 4.21] showed high levels of fluorescence upon ethanol addition, even higher than the WT; the lack of CpxR relieving repression. This also suggests that ethanol induction is CpxR-independent, so proceeds via an unknown mechanism, presumably via the σ^{70} -dependant promoter since it also occurred in the $\Delta\sigma E$ strain. The results obtained contrasted with those observed by Dartigalongue et al. (2001), where the *yidQ::lacZ* reporter containing the full length of the promoter was less active in $\Delta cpxR$ than the isogenic strain. This is probably caused because the region which was omitted in pYIDQ contains further CpxR-P with higher affinity.

4.4 Conclusions and future work

The very low copy *dnaK::gfp* reporter plasmid pDNAK outperformed its medium copy number counterpart pFPV25 resulting in enhanced growth and enhanced fluorescence signal. This demonstrates how good reporter system design can deliver better results, including improved measurements but, more importantly; that in order to amplify the signal from the reporter systems multiple copies of the insert might not be the right approach.

Addition of thiol reductants led to an increase in GFP production by all the reporter plasmids tested. This accumulation is thought to be partially due to indirect causes in the case of bacteria harbouring *rpoH:gfpmut3* reporter systems. Flow cytometry supported these observations. PI staining confirmed that thiol reductant addition at the amounts utilized was not lethal in agreement with previous studies (Gill et al., 1998) but bacteriostatic. This difference in growth could have created distortion of the gene expression data by

accumulation of GFP in the stressed cultures because of its elevated half-life *in vivo*. Bulk measurements were reliable in most of the circumstances. However, in specific situations such as heterogeneity of bacterial size or morphology within or between samples, single cell measurements provided useful additional data.

The reporter pYIDQ illustrated the degree of complexity in gene regulation. It would be interesting to further characterize this promoter. The production of a plasmid to overexpress YidQ could reveal its effects on cell morphology and cell physiology if co-transformed with other reporters. Micrographs revealed that aggregates and filaments of cells were more fluorescent than single cells. Induction of biofilm formation in mutant strains and isogenic types could determine if *yidQ* is contact dependant. Construction of a dual reporter plasmid, so that *ibpA* is fused to *gfp* and *yidQ* would allow simultaneous measurements of both promoters. Changes in fluorescence from yellow to green could be determined by multi-parameter flow cytometry, as it enables monitoring of different fluorescence wavelengths at the same time.

Chapter Five

Monitoring *groE* and *dnaK* expression in *E. coli* BL21* during recombinant protein production with gene reporter technology

5.1 Introduction

Recombinant protein production in *E. coli* is known to cause a heat-shock like stress response (Hoffmann and Rinas, 2004). In the following experiments the plasmids pGROE (*groE::gfpmut2*) and pDNAK (*dnaK::gfpmut2*) sourced from (Zaslaver et al., 2006), were tested as reporters of chaperone expression as a measure of stress caused by recombinant protein production rather than external agents such heat-shock or chemicals as described in previous chapters.

CCP was expressed as recombinant protein and protein production strategy

The CCP protein (47kDa) is the cytochrome *c* peroxidase from *Neisseria gonorrhoeae* (Turner et al., 2003). It has been shown to be a lipoprotein localized to the periplasm via a signal peptide but, when overexpressed aerobically in *E. coli*, it accumulates in the cytoplasm as inclusion bodies. The rate limiting steps for protein maturation are the secretion to the periplasm via the Sec pathway followed by the cleavage of the signal peptide. The final maturation step consists on the attachment of an haem group although this occurs during anaerobiosis.

The production strategy is important to guarantee commercially viable titres of recombinant protein which is able to perform its biological function. The process needs to be robust guaranteeing reproducibility in the outcome. The ‘standard’ and ‘improved’ protocols were developed previously to compare the yield of correctly folded recombinant protein when manipulating two variables at a time: growth temperature; and the concentration of the inducer IPTG (Isopropyl β -D-1-thiogalactopyranoside) (Sevastyanovich et al., 2009). In this

chapter, both protocols are compared during CCP expression to observe the suitability of each reporter plasmid, pGROE and pDNAK, for monitoring the physiological stress measured as activation of chaperone genes.

Flow cytometry as a process monitoring tool

Flow cytometry has previously been proposed as a tool for monitoring industrial fermentations (Hewitt and Nebe-Von-Caron, 2001; Hewitt and Nebe-Von-Caron, 2004). Single cell analysis delivers multi-parameter data which can be interpreted to infer diverse aspects of microbial physiology during recombinant protein production such as: inclusion body formation, visible as an increase in side scatter (Lewis et al., 2004); morphological changes, such as filamentation (Wickens et al., 2000); and quantization of recombinant gene expression via fluorescent reporters (Patkar et al., 2002). In this work is shown that the increase in inclusion bodies as a result of recombinant CCP production, and other morphological changes, may interfere with and make analysis of reporter expression difficult. The benefit of coupling flow cytometry with gene reporter technology is that it is possible to elucidate interferences because more information than just fluorescence is collected from each individual cell. This may include characteristics affecting other measurements such as RF (Relative Fluorescence) of CFU (colony forming units).

Following initial experiments in shake flask cultures, the expression of *dnaK::gfpmut2* was studied while CCP was expressed in batch fermentations. Flow cytometry resolved sample heterogeneity and contributed to explain more aspects of the process than bulk measurements. In addition on-line monitoring systems such as DO measurements and mass spectrometry of exhaust gases supported the data generated from off-line samples.

5.2 Results

5.2.1 Synthesis of GFP by pGROE as stress reporter upon pET-CCP activation

To observe *groE::gfpmut2* expression in response to different protocols for recombinant protein production, the reporter plasmid pGROE was co-transformed into *E. coli* BL21* with the expression plasmid pET-CCP. Overnight cultures grown at 30°C were inoculated into 30mL of LB media in 250mL conical flasks supplemented with 0.2% glucose and the appropriate antibiotics. Cultures were incubated accordingly with two different recombinant protein production protocols: standard and improved. The standard protocol consisted of initial incubation at 37°C until the OD₆₅₀ reached ~0.5. Then the cultures were induced with 0.5mM of IPTG and transferred to 25°C incubators. The improved protocol consisted of incubation at 25°C before and after induction with 8μM of IPTG when OD₆₅₀ reached ~0.5. Cultures were shaken constantly at 200rpm. Samples were taken at induction time and every two hours thereafter. All work was performed in duplicate and repeated at least twice. Plasmid retention was observed by replica plating performed at six hours post-induction which confirmed high level of plasmid maintenance of pGROE and pET-CCP at this sample point. *E. coli* BL21* cotransformed with pET-CCP and pUA66 was used as a control.

Bulk measurements

Overall, the standard protocol resulted in higher relative fluorescence in every sample in comparison with the improved protocol [Fig 5.1 (a)]. The RF increased over time under the conditions of both protocols showing a faster rate of increase under the standard conditions than in the improved. This resulted in a widening gap of RF between samples from the two protocols over time. The growth curves showed faster growth rate in the improved protocol

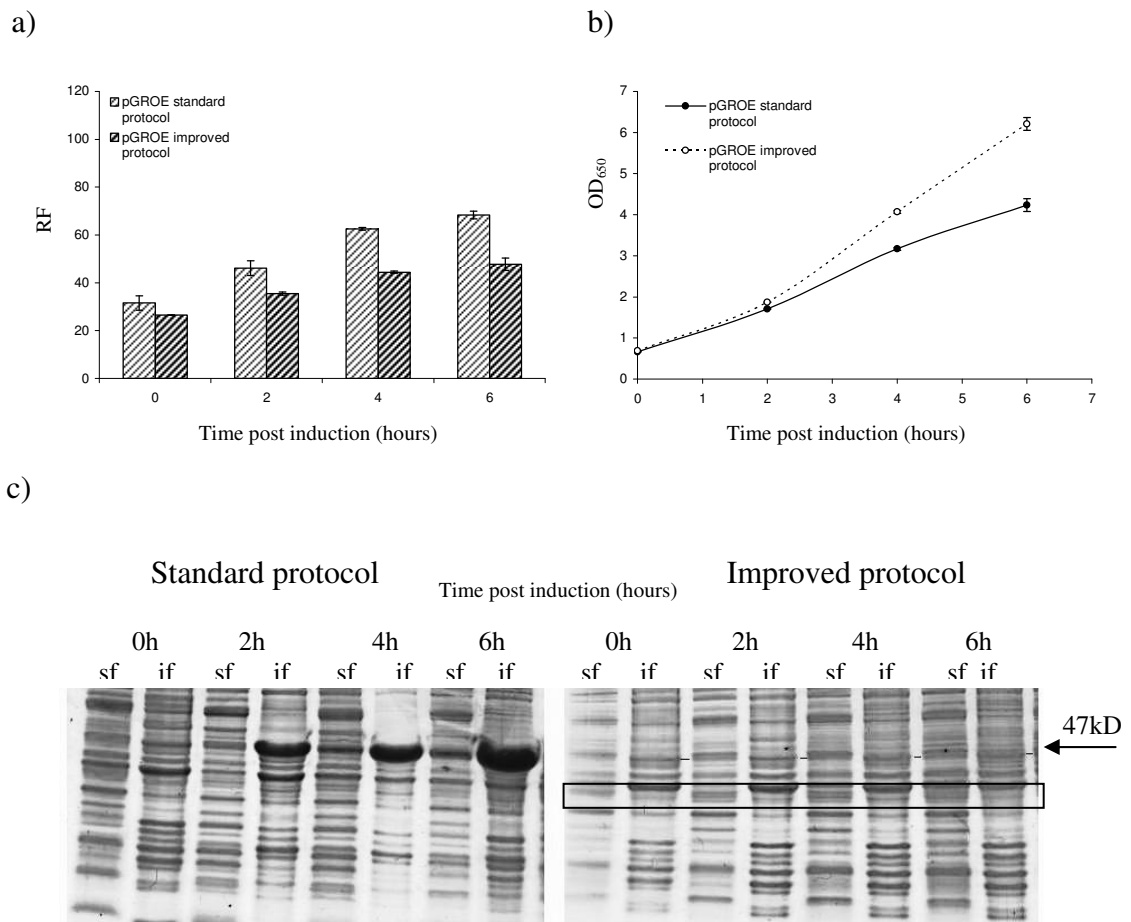


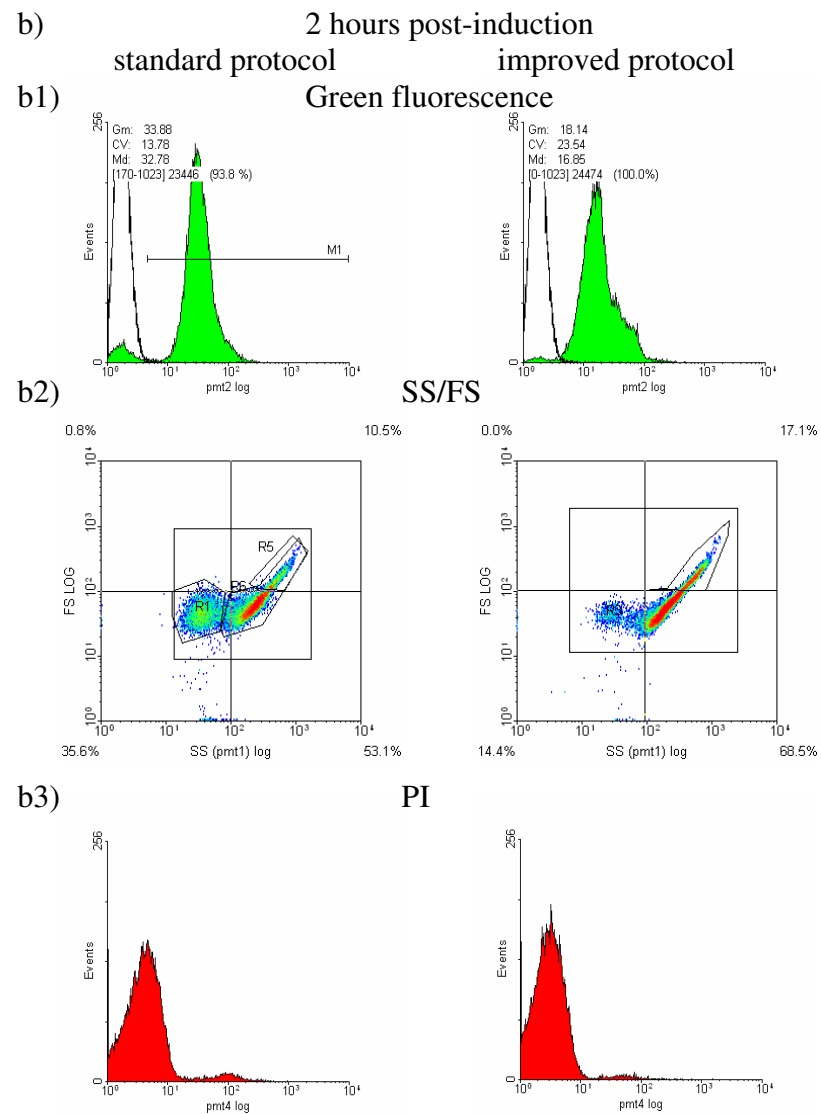
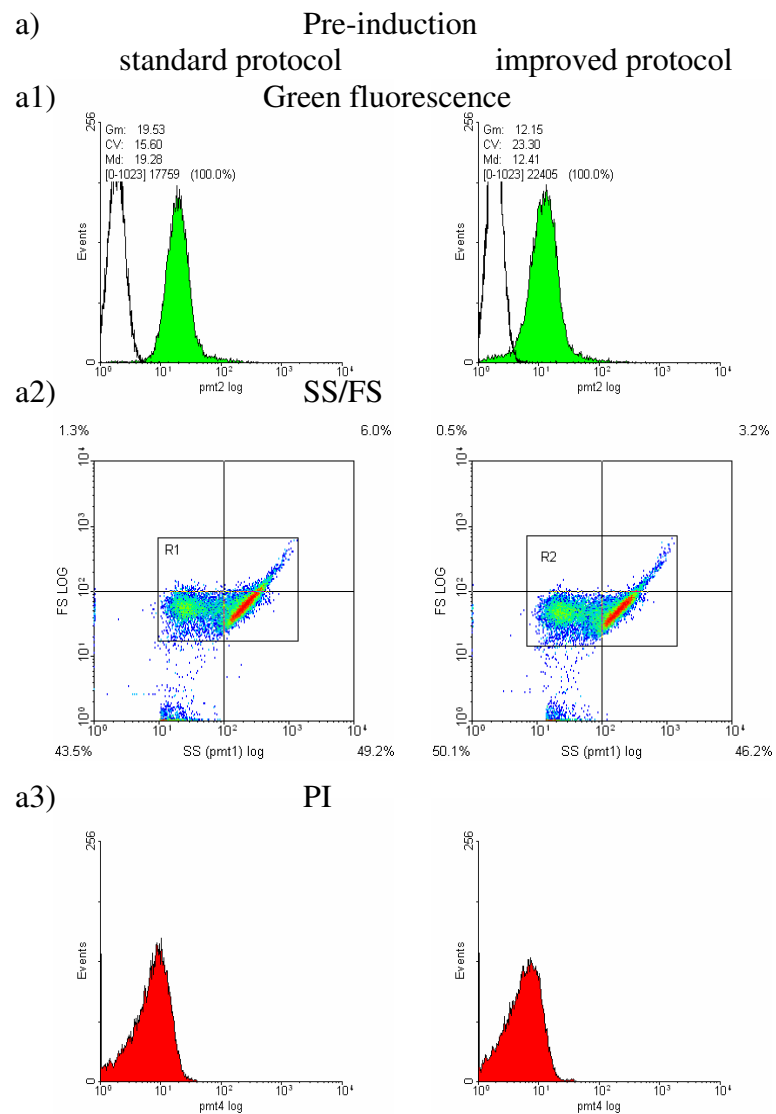
Fig. 5.1 RF, growth curves and SDS-PAGE obtained using pGROE obtained during growth of *E. coli* BL21* pGROE/pET-CCP in standard and improved protocols.

(a) shows the relative fluorescence while (b) shows the growth curves. (c) SDS-PAGE of soluble and insoluble fractions ('sf' and 'if' respectively) extracted with BugBuster. While the CCP overproduction is obvious in the standard protocol, it is highlighted by a black box in the improved protocol.

than in the standard after the second hour post induction [Fig 5.1 (b)]. Cell pellets were treated with BugBuster reagent to separate soluble and insoluble fractions. As shown in SDS-PAGE [Fig 5.1 (c)], CCP appeared mostly in the insoluble fraction which co-precipitates with the cell membrane. This is common when using this reagent for the separation of liposoluble or membrane associated proteins. More CCP was found in insoluble form in the standard protocol than in the improved. Some soluble CCP could be found in the improved protocol confirming the observations from (Sevastyanovich et al., 2009).

Single cell analysis in relation to CCP synthesis: Pre-induction

The single cell analysis of the pre-induction samples resulted in different histograms of green fluorescence [Fig 5.2 (a1)]. The population of cells growing in standard conditions generated a narrow distribution of cells with a geometric mean (Gm) significantly higher than those under the improved conditions. This was probably caused by the temperature difference between protocols as the cells growing in standard conditions were, at this point, at 37°C. The population of cells in improved protocols were more spread in terms of fluorescence, as shown by the coefficient of variation (CV), which could be interpreted as the presence of multiple *groES* expression profiles in the population. The scatter plots [Fig 5.2 (a2)] did not show significant differences in cell size or morphology between protocols and PI [Fig 5.2 (a3)] did not reveal any dead cell populations. SDS-PAGE [Fig 5.1 (c)] revealed a band of CCP protein of similar intensity in both conditions commonly attributed to expression leakage usual with pET vectors.



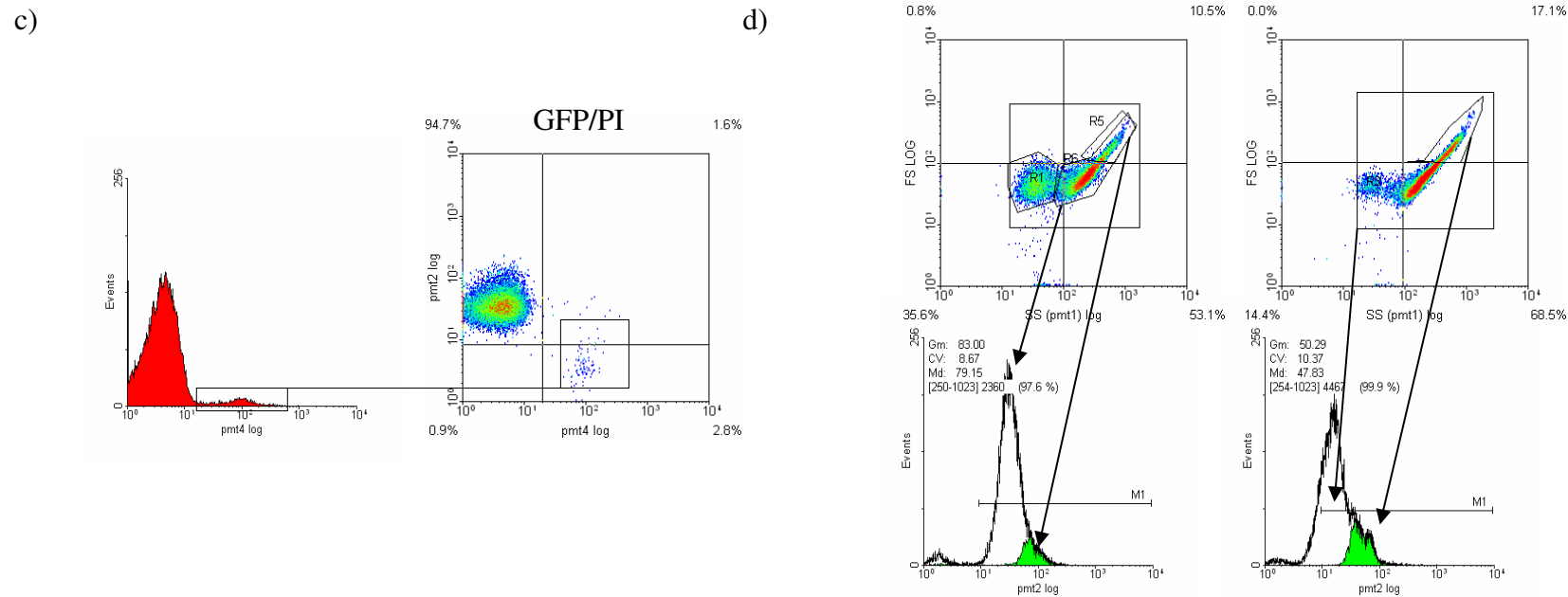


Fig 5.2 Single cell analysis of samples from standard and improved protocols of *E. coli* BL21* pET-CCP/pGROE

(a) shows the results at pre-induction while (b) shows the results 2 hours post-induction, while the numbers indicate as follows (1) green fluorescence histograms. A control culture of *E. coli* BL21* pET-CCP pUA66 is shown overlaid (unfilled black line). These histograms contain information of: Gm, geometric mean; CV, coefficient of variation (spread of data); and Md, median (threshold of PMT2 at which the population is equally divided). *[x-1023]* refers to beginning and end of the marker (M1) followed by the number of events under the marker and percentage of the total. (2) Shows scatter plots and (3) PI histograms. Two hours after induction further gate analysis is shown: (c) analysis of GFP content of the PI stained sample. (d) shows the scatter plot identification of the population in the shoulder of the PMT2 histograms of standard (left) and improved

Single cell analysis in relation to protein synthesis: two hours post-induction

Two hours after IPTG addition, the standard protocol cells were distributed in two populations based on green fluorescence [**Fig 5.2 (b1)**]. One minor population had low fluorescence comparable with control cultures with a promoterless pUA66 reporter vector (shown overlaid as a unfilled line), while the other population exhibited a narrow fluorescence distribution with a more than 50% increase of Gm (from 19.53 pre-induction to 33.88). The improved protocol also showed a small non-fluorescent population. The fluorescent cells were widely distributed with most of the population concentrated around a median (Md) significantly lower than that of the standard protocol. The scatter plots [**Fig 5.2 (b2)**] revealed some differences between protocols. In the improved protocol, a greater proportion of the cells had high FS and SS values (in the top right of the scatter plots) than in the standard protocol. These cells with high FS and SS values could be caused by two factors: first, they could be the result of dividing cells remaining attached; this phenomenon is likely to be stronger in the improved protocol more frequently than in the standard as after 2 hours post induction there is a difference in growth rate between conditions. Second, high SS is indicative of cytoplasm granularity, and could be slightly higher in the standard protocol due to inclusion body formation. This could also be a contributing cause in the improved protocol. The fluorescence of the cells in these sectors of elevated FS/SS is shown in [**Fig 5.2 (d)**] which revealed that cells of this morphology are highly fluorescent. However, as CCP production was greater in the standard protocol, [**Fig 5.1 (c)**], fluorescence in these cells is higher than in improved protocol cells presumably due to increased chaperone expression. PI staining showed that dead cells in the improved protocol were not significant [**Fig 5.2 (b3)**]. In the standard protocol a 5% population of dead cells was revealed which have much lower green

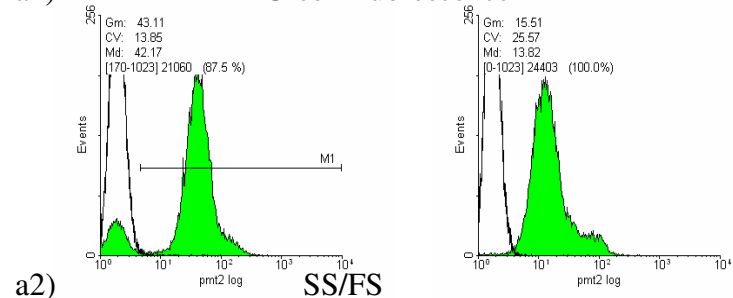
fluorescence than the live cells as shown by the GFP/PI plot [Fig 5.2 (c)] as GFP leaves rapidly the cell upon death.

Single cell analysis in relation to protein synthesis: four hours post-induction

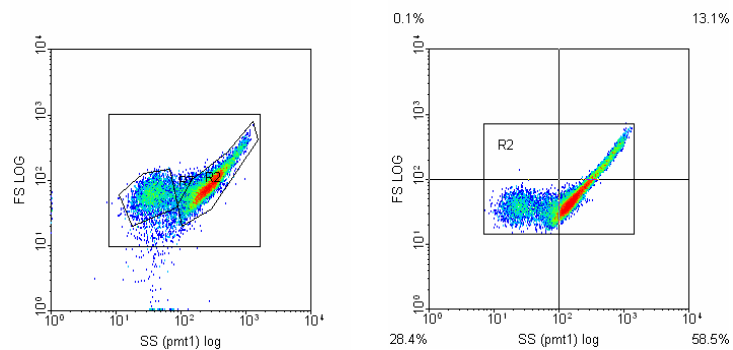
Cultures progressed into the fourth hour post-induction with significant differences in fluorescence behaviour [Fig 5.3 (a1)]. In the cultures under standard conditions, the non-fluorescent population, with respect to PMT2, increased in size becoming nearly 14% of the total population. At the same time, the fluorescent cells increased in brightness ($G_m = 43.11$) in comparison with the 2 hour sample ($G_m = 33.88$). The improved conditions resulted in a wider distribution with a median fluorescence 3-fold lower than the median found in standard conditions, although a sector of the population displayed very high fluorescence. Scatter plots [Fig 5.3 (a2)] revealed a similar distribution of cells between protocols with a subpopulation of cells projecting towards a higher FS/SS region. This population was found within the upper right quadrant of FS/SS, homologous to that described in the previous sample point, and was shown to have high green fluorescence [Fig 5.3 (c)]. The G_m of this subpopulation in the improved protocol was 60.67, more than 3-fold the G_m of the whole population and higher compared with the same subpopulation in the previous sample. The same high FS/SS population in the standard protocol showed a $G_m = 82.25$ which was much higher than that seen in the improved, nearly twice the G_m of the general population and not significantly different than that in the previous sample. In these high FS/SS regions, scatter is similar in cells in both protocols, suggesting similar cells sizes, whereas green fluorescence is far higher in the standard protocol. In the standard protocol, some noise was detected that could be caused by cell debris. PI staining [Fig 5.3 (a3)], shown in the standard protocol by GFP/PI

a) 4 hours post-induction
standard protocol improved protocol

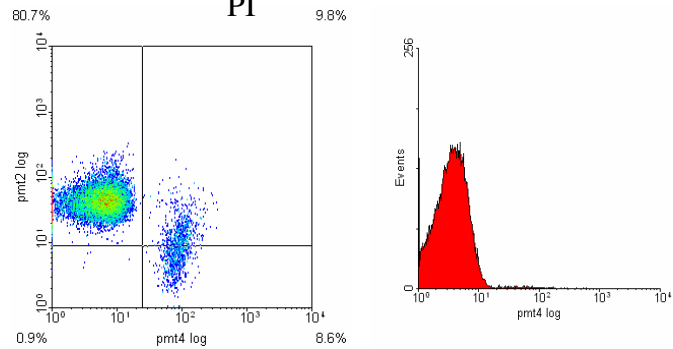
a1) Green fluorescence



a2) SS/FS

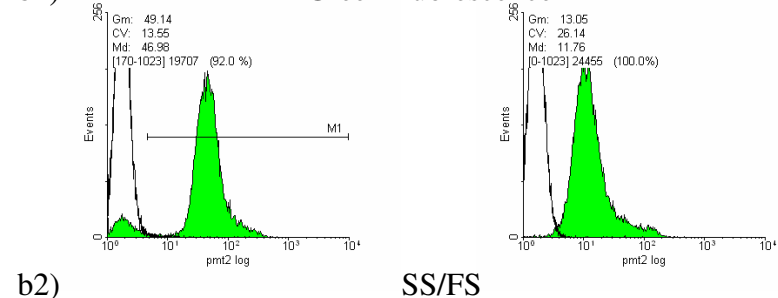


a3) PI

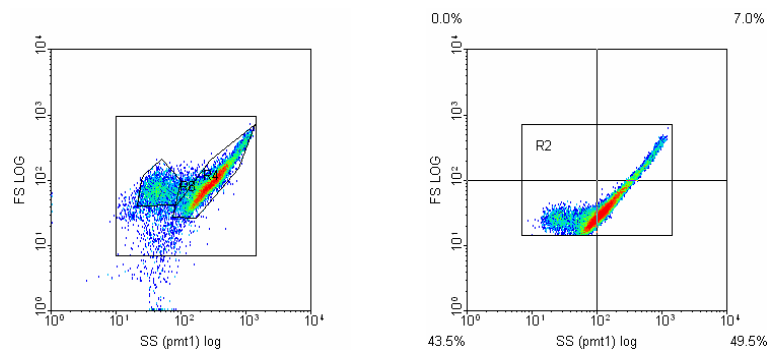


b) 6 hours post-induction
standard protocol improved protocol

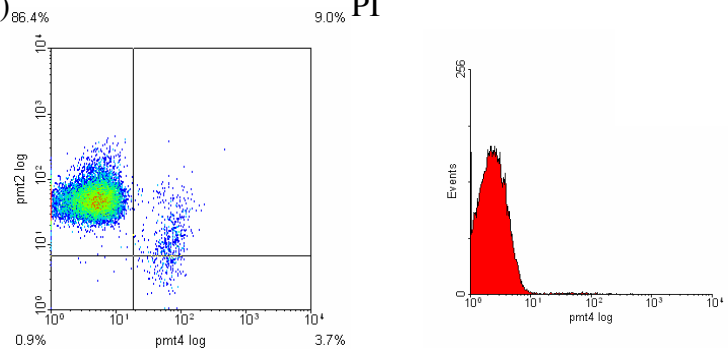
b1) Green fluorescence



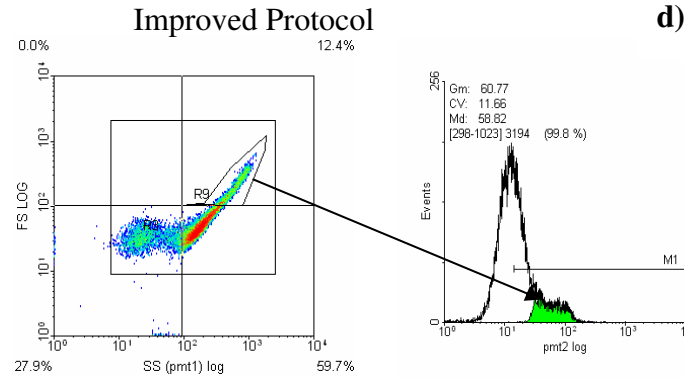
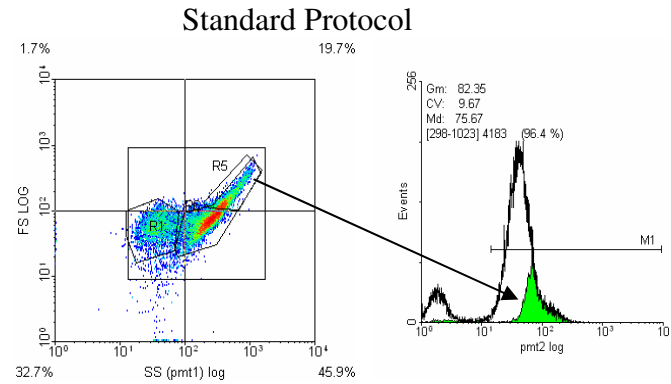
b2) SS/FS



b3) PI



c)



d)

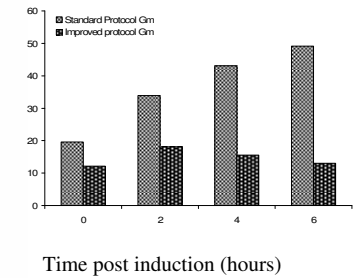


Fig 5.3 Single cell analysis of samples from standard and improved protocols of *E. coli* BL21* pET-CCP/pGROE (continuation)

(a) shows the results at 4 hours post-induction while (b) shows the results six hours post-induction, the numbers indicate as in the previous figure. PI analysis for the standard protocol is displayed in GFP/PI to show that dead cells (PI positive) have low green fluorescence. In the sample four hours after induction further gate analysis is shown: (c) shows a scatter identification of the population sector in the shoulder of the PMT2 histograms of standard (left) and improved (right). (d) shows an histogram of the Gm of green fluorescence obtained over time of the in both protocols.

plots, revealed that the dead cell population increased dramatically in comparison with the previous sample and was again of low green fluorescence. PI histograms did not show significant cell death in the improved protocol.

Single cell analysis in relation to protein synthesis: six hours post-induction

The fluorescence histograms [**Fig 5.3 (b1)**] at this stage revealed two cell populations in the standard protocol, corresponding to non-fluorescent and fluorescent, and one population of fluorescent cells in the improved protocol with a shoulder of higher fluorescence. The fluorescent cells of the standard protocol were distributed narrowly along the channels exhibiting a median more than 4-fold higher than the observed in the improved protocol population. The sample from the improved protocol resulted in the aforementioned shoulder which is the cause of the high coefficient of variation of the distribution. The scatter plots [**Fig 5.2 (b2)**] indicate that this shoulder is caused by cells with high FS and SS, as shown in previous samples. The samples from the standard protocol produced similar scatter plots showing interference of cell debris. PI [**Fig 5.2 (b3)**] confirmed that the non-fluorescent population in PMT2 was formed by dead cells. This population was smaller (~ 12% of the total) than that observed in the 4 hour sample. The SDS-PAGE [**Fig 5.1 (c)**] from cells growing in standard conditions revealed an amount of recombinant protein production much higher than shown by cells in improved conditions. In addition, the amount of recombinant CCP was increased in comparison with the previous sample in standard protocol while the improved protocol showed a constant rate of synthesis. [**Fig. 5.2 (d)**] gives a scope of how the Gm of the samples progressed over time showing accumulation of fluorescence in the standard protocol. Improved protocol samples showed an initial increase at 2 hour post

induction followed by gradual decrease. In comparison with RF [Fig. 5.1 (a)], the fluorescence accumulation in the standard protocol samples was faster and the gap between standard and improved protocols was larger. The leading cause for such differences in measuring methods is that in flow cytometry analysis Gm was obtained for only the live cell, eliminating dead cells which distorted RF measurements.

5.2.2 Synthesis of GFP by pDNAK as stress reporter upon pET-CCP activation

To observe *dnaK::gfpmut2* expression in response to different protocols for recombinant protein production, *E. coli* BL21* was transformed with pET-CCP and pDNAK and grown in the exact same conditions, standard and improved, as previously described for pGROE. Plasmid retention was observed by replica plating performed at six hours post-induction which confirmed high level of plasmid maintenance of both pDNAK and pET-CCP at this sample point. *E. coli* BL21* cotransformed with pET-CCP and pUA66 was used as a control.

Bulk measurements

The relative fluorescence accumulated in a very different manner between the standard and the improved protocols [Fig. 5.4 (a)]. During the standard protocol, the RF increased abruptly two hours after induction followed by a gradual decrease over time, returning to a fluorescence level similar to that observed pre-induction. The RF in the improved conditions accumulated very gradually showing noticeable increases from the fourth to the sixth hour post-induction. The growth curves showed the same behaviour as previously observed in pGROE pET-CCP co-transformants revealing different growth profiles for each protocol. While the standard protocol exhibited a reduction in growth rate from the second hour post-

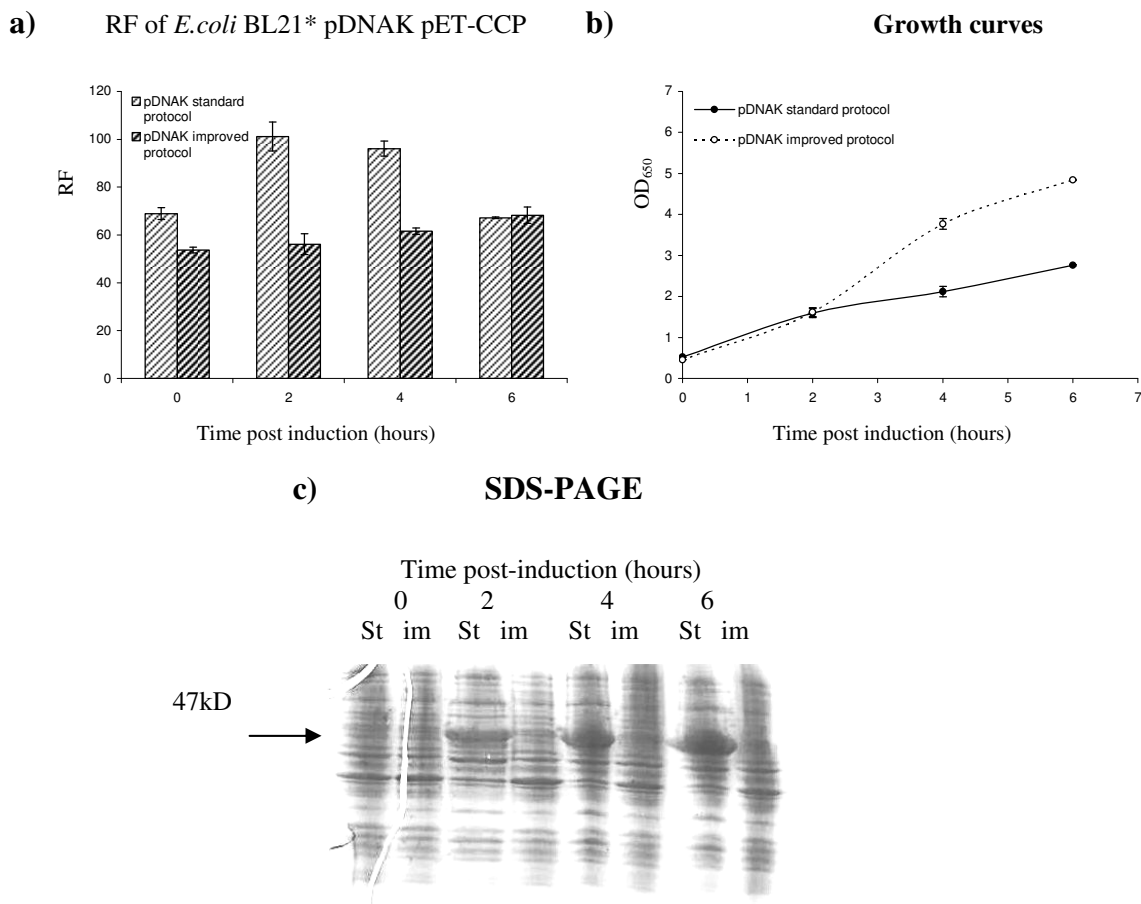


Fig. 5.4 RF, growth curves and SDS-PAGE obtained using pDNAK obtained during growth of *E.coli* BL21* pDNAK/pET-CCP in standard and improved protocols.

(a) shows the relative fluorescence from *dnak::gfpmut2* while (b) shows the growth curves. (c) SDS-PAGE of Standard and Improved insoluble fractions extracted with BugBuster.

induction, the improved protocol delivered a higher growth rate achieving higher biomass [Fig. 5.4 (b)].

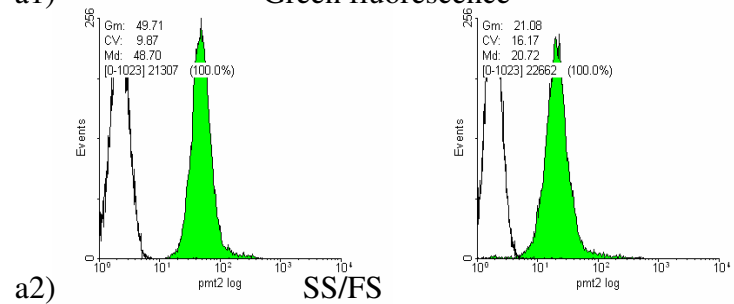
Single cell analysis in relation to protein synthesis: Pre-induction

The difference in green fluorescence between cultivation in standard and improved conditions was more easily observed in the flow cytometric histograms than by RF measurements. The Gm in standard protocol samples was 2.5 fold greater than that observed in the improved protocol [Fig. 5.5 (a1)]. In addition, the data distribution of the former was narrower than the latter (signified by CV) [Fig. 5.5 (a1)]. This difference is likely to be caused by the temperature of incubation and is similar to, but more marked than the effect at the *groES* promoter [Fig. 5.2]. The scatter plots showed very similar population distribution in both protocols [Fig. 5.5 (a2)]. PI staining [Fig. 5.5 (a3)] did not reveal significant cell death and the SDS-PAGE [Fig. 5.4 (d)] showed weak bands corresponding to CCP (47kDa) in both protocols probably attributed to promoter leakage by the pET-CCP vector.

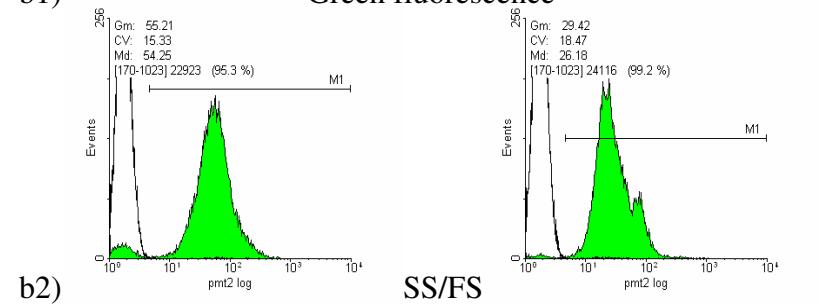
Single cell analysis in relation with protein synthesis: two hours post-induction

In the standard protocol, the fluorescence histograms revealed a minor population of non-fluorescent cells (4.7% of the cells) and a fluorescent population with a wide base [Fig. 5.5 (b1)]. On the other hand, the improved protocol had a shoulder in the distribution of fluorescence. The standard protocol median fluorescence (Md) was 2-fold greater than the improved protocol. The scatter plots [Fig. 5.5 (b2)] showed different spreads of population as the cells cultured in standard conditions were more widely distributed in both axes. In both protocols, as in the pGROE experiment, cells with high FS and SS were present. Analysis of

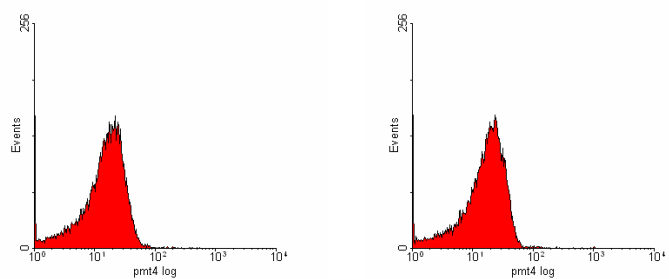
a) Pre-induction
standard protocol improved protocol
a1) Green fluorescence



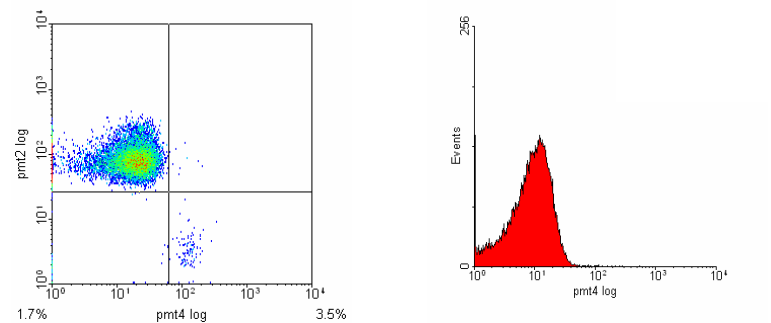
b) 2 hours post-induction
standard protocol improved protocol
b1) Green fluorescence



a3) PI



b3) PI



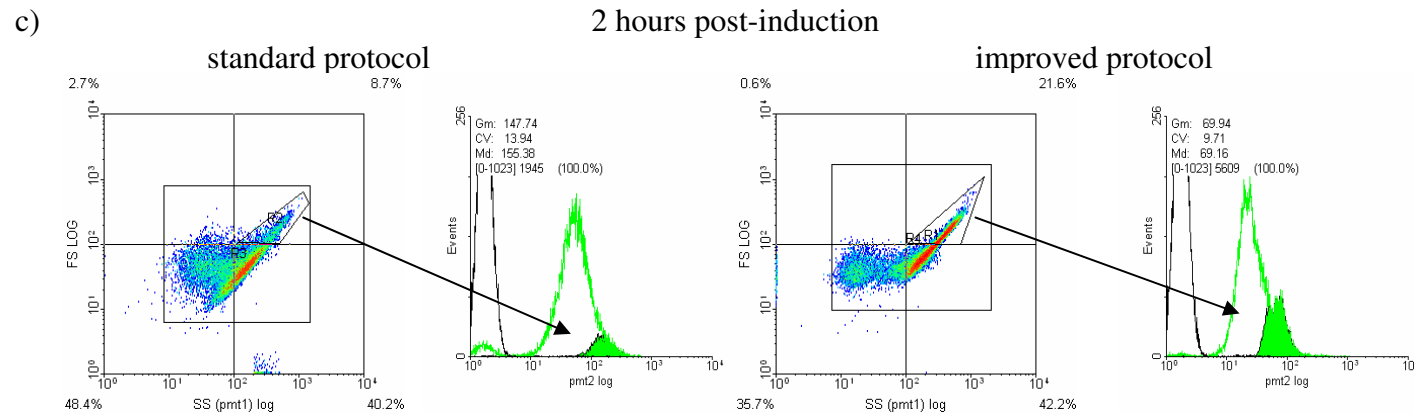


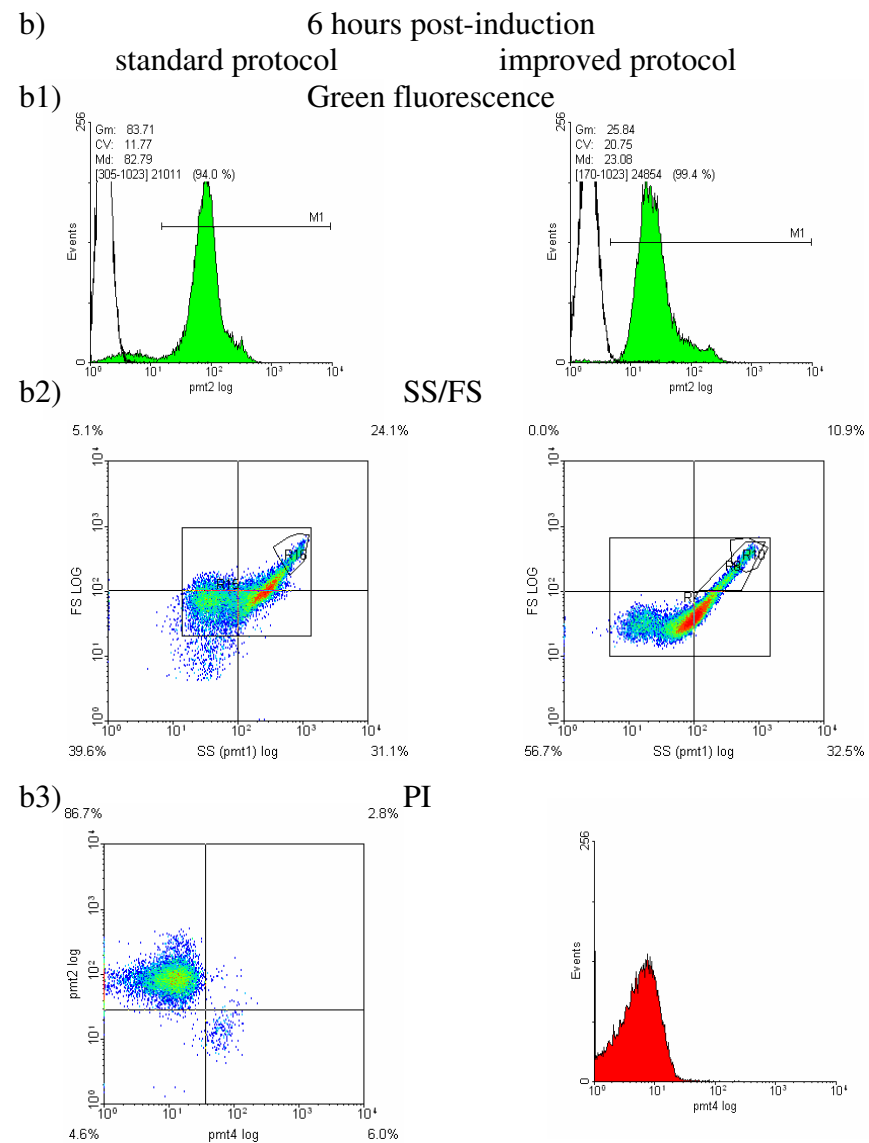
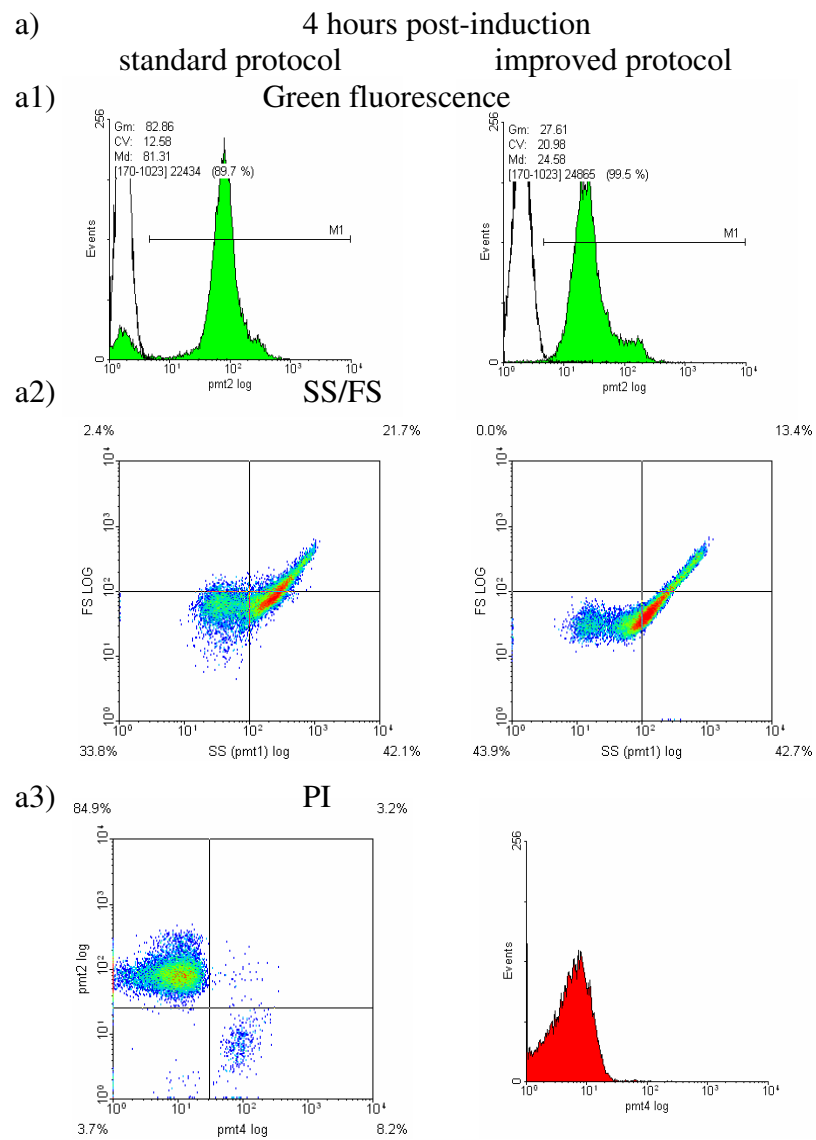
Fig 5.5 Single cell analysis of samples from standard and improved protocols of *E. coli* BL21* pET-CCP/pDNAK

(a) shows the results at pre-induction while (b) shows the results two hours post-induction. (1) shows green fluorescence histograms (for Gm, CV and Md definitions refer to [Fig 5.2]). *E. coli* BL21* pET-CCP/pUA66 are shown overlaid with a black line, (2) Shows scatter plots (3) PI analysis which for standard protocol (b3) is displayed as a GFP/PI plot to show that dead cells have no green fluorescence. For the sample two hours after induction further gate analysis is shown (c) of standard (left) and improved (right) samples; the subpopulations in the upper right quadrant of the scatter plots correspond to the shoulder of the PMT2 histograms (filled green).

fluorescence in this region similarly revealed that these populations are the responsible for high fluorescence histograms, the Gm of which is 2-fold higher in the standard than in the improved protocol [Fig. 5.5 (c)]. Staining with PI revealed that the non-green fluorescent population in the standard protocol [Fig. 5.5 (b3)] were also PI positive, therefore dead. There were no PI positive cells in the improved protocol. SDS-PAGE revealed overproduction of CCP by cells cultured in standard protocol while the cells under improved protocol generated small amounts [Fig. 5.4 (c)].

Single cell analysis in relation to protein synthesis: four hours post-induction

Two hours after the previous sample, the bacteria in the standard conditions nearly doubled the fluorescence Gm in a narrow distribution [Fig. 5.6 (a1)]. The segregated non-green fluorescent population grew in numbers to above 10% of the total population. At this point the median green fluorescence was greater than 3-fold of that observed in the improved conditions. The distribution of the fluorescence in the population of cells in improved conditions revealed a shoulder of high fluorescence comparable to that found in the standard protocol as seen before [Fig. 5.6 (a2)]. Such a distribution was seen in the samples from standard conditions but less accentuated as the FS distribution was much wider. Again, the results from PI staining of standard protocol samples, [Fig. 5.6 (a3)] determined that the non-fluorescent sub-population in PMT2 histograms was formed by dead cells. Such a population was not detected in samples from the improved protocol. SDS-PAGE [Fig. 5.4 (c)] revealed an increase of recombinant CCP overproduction in the standard protocol while the improved protocol showed lower CCP production.



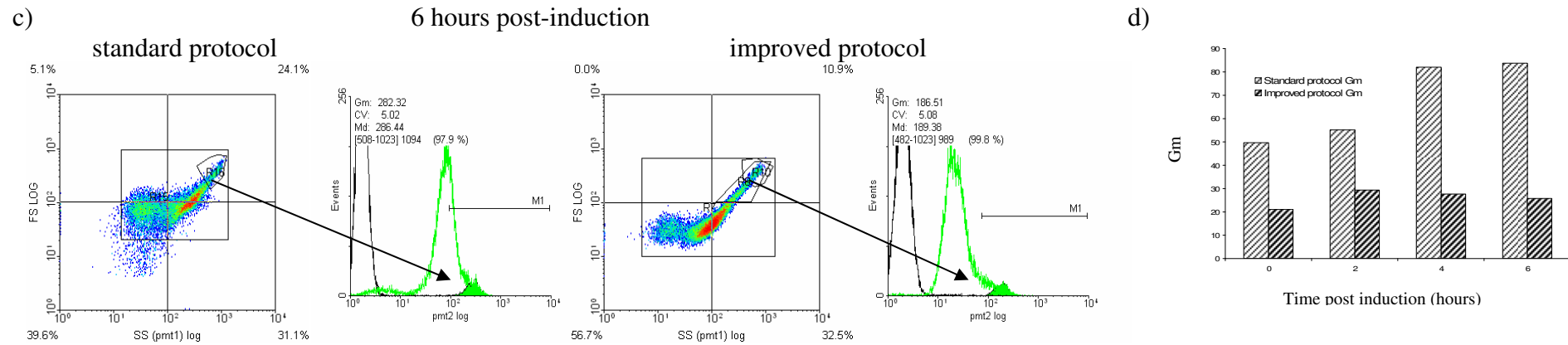


Fig 5.6 Single cell analysis of samples from standard and improved protocols of *E. coli* BL21* pET-CCP/pDNAK (continuation)

(a) shows the results at pre-induction while (b) shows the results four hours post-induction. (1) shows green fluorescence histograms (for Gm, CV and Md definitions refer to [Fig 5.2]). *E. coli* BL21* pET-CCP/pUA66 are shown overlaid with a black line, (2) Shows scatter plots (3) PI analysis which for standard protocol (b3) is displayed as a GFP/PI plot to show that dead cells have no green fluorescence. For the sample two hours after induction further gate analysis is shown (c) of standard (left) and improved (right) samples; the sub-populations in the upper right quadrant of the scatter plots correspond to the shoulder of the PMT2 histograms (filled green). (d) shows an histogram of the Gm of green fluorescence obtained over time of the in both protocols.

Single cell analysis in relation to protein synthesis: six hours post-induction

The fluorescence histograms six hours post-induction revealed a sub-population in the standard protocol with drastically decreased fluorescence [**Fig. 5.6 (b1)**]. The green fluorescent population maintained the fluorescence with respect to the previous sample, the Gm of which in the standard protocol was 3-fold greater in the improved. The scatter plots revealed similar features as highlighted in the previous sample with the exception of an increase of individuals in the subpopulations at high FS and SS in both protocols [**Fig. 5.6 (b2)**]. These subpopulations exhibited high green fluorescence as shown in [**Fig. 5.6 (c)**]. PI staining [**Fig. 5.6 (b3)**] revealed a decrease in the PI positive population, found only under standard conditions, from an 11% at 4 hours to approximately an 8% at six hours. This population was confirmed to be formed mostly of non-green fluorescent cells, although green fluorescence was higher than in the 4 hour sample. This was probably caused by the higher green fluorescence of the live cells, thus meaning that dead cells took longer to lose all GFP. SDS-PAGE [**Fig. 5.4 (c)**] revealed substantial increase, in comparison to the previous sample at 4 hours, in CCP overproduction in the standard protocol while the recombinant protein obtained under the improved protocol just started to be noticeable.

The histogram shown in figure [**Fig. 5.6 (d)**] is an overview of Gm of green fluorescence from single cell analysis. It differed from bulk measurements by a delay in the abrupt activation of *dnaK* in the standard protocol. RF is likely to underestimate the difference in GFPmut2 production between the standard and the improved protocol. This was likely to be caused by two major factors. First, RF was underestimated upon cell death as dead cells are not fluorescent [**Fig. 5.5 (b3)**] but they still affect OD₆₅₀ values, thus decreasing RF. Second,

RF in the improved protocol was overestimated. The eccentric populations in the upper quadrants of the SS/FS plots [Fig. 5.5 (b1)] in both protocols correspond to filamented cells which have an elevated fluorescence [Fig. 5.5 (c)]. These cells may contribute to absolute fluorescence but due to filamentous morphology cells do not act on OD₆₅₀ values resulting in higher RF. Although filamentation occurred in both protocols, while in the improved protocol the filamented population had a Gm 3-fold higher than Gm of the total population, in the improved protocol this was 7-fold higher.

5.2.3 Controls and further growth conditions.

Promoterless reporter vector: pUA66

These controls were performed to demonstrate that GFP production is controlled by the presence of the promoter *dnaK* or *groE*. *E. coli* BL21* transformed with pUA66 and pET-CCP was cultured in the same conditions as described before. The green fluorescence histograms generated from samples were utilized as negative control for fluorescence and overlaid as a black line in the histograms of previous experiments. As shown in [Fig. 5.7 (a)] RF did not significantly increase during growth; the growth curves [Fig. 5.7 (b)] were nearly identical to cultures containing pGROE or pDNAK as reporter plasmids indicating that GFP biosynthesis did not interfere with cell growth. SDS-PAGE [Fig. 5.7 (c)] showed the same CCP overexpression pattern as observed in pGROE/pET-CCP and pDNAK/pET-CCP indicating that the reporter protein production did not interfere with CCP expression. This experiment also confirmed that the growth differences observed between standard and improved protocols are not caused by differences in GFP production by the different reporter

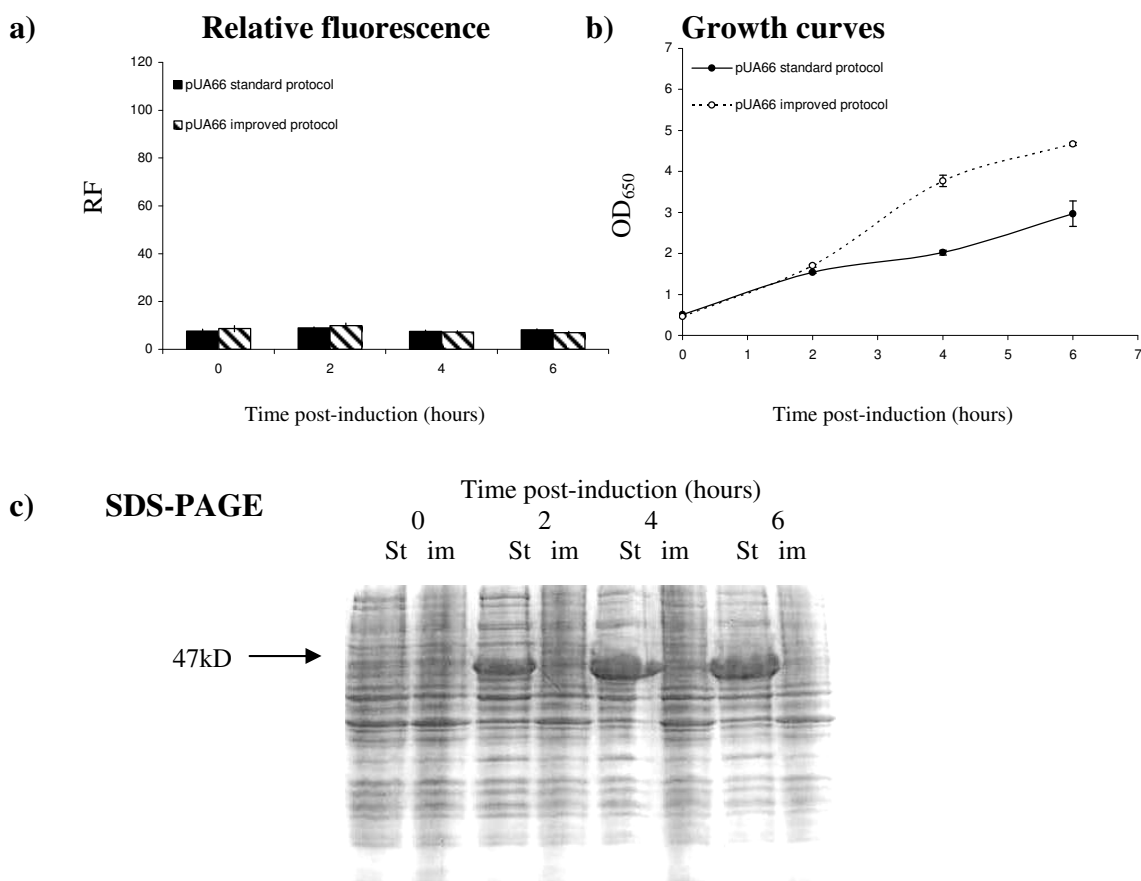


Fig. 5.7 RF and growth curves obtained using the reporter pUA66 in standard and improved protocols

Data shows the RF (a) and growth rates (b) for *E. coli* BL21* transformed with the promotorless GFP reporter pUA66 and pET-CCP. (c) SDS-PAGE which was prepared as described in pDNAK/pET-CCP [Fig 5.4 (d)] confirmed CCP overproduction very similar to that observed in pDNAK/pET-CCP and pGROE/pET-CCP indicating that CCP overexpression occurred independently of GFPmut2 production. (St: Standard protocol; Im: improved protocol)

plasmids. pUA66/pET-CCP revealed the same SS/FS distribution as pGROE/pET-CCP and pDNAK/pET-CCP but no increase of fluorescence in the high SS/FS sub-populations. pUA66/pET-CCP also showed the same results for PI staining, revealing dead cells in the standard but not the improved protocol, confirming that GFP synthesis did not lead to cell death.

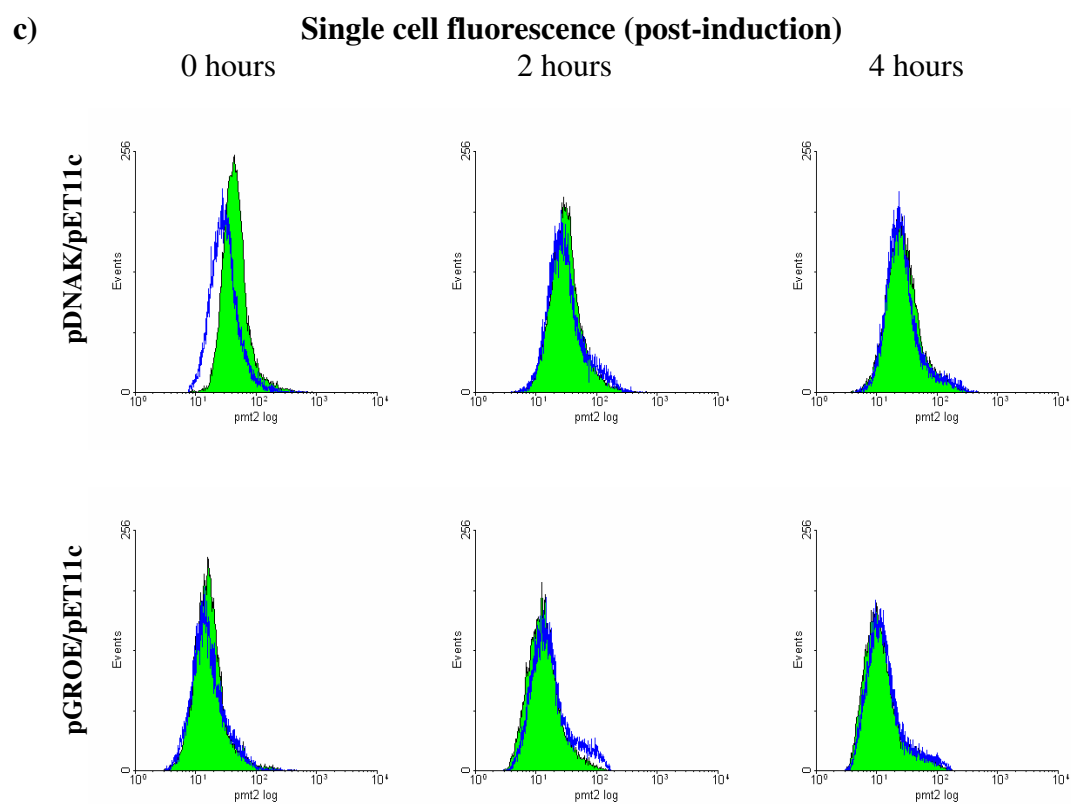
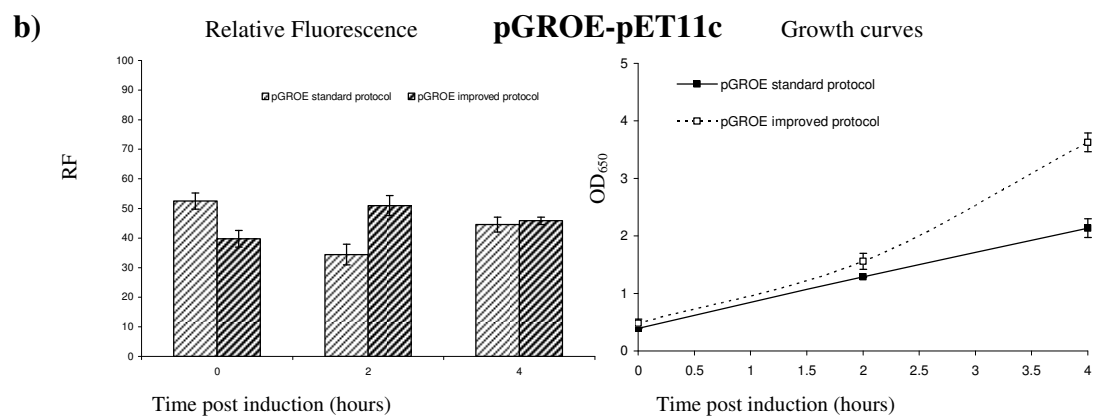
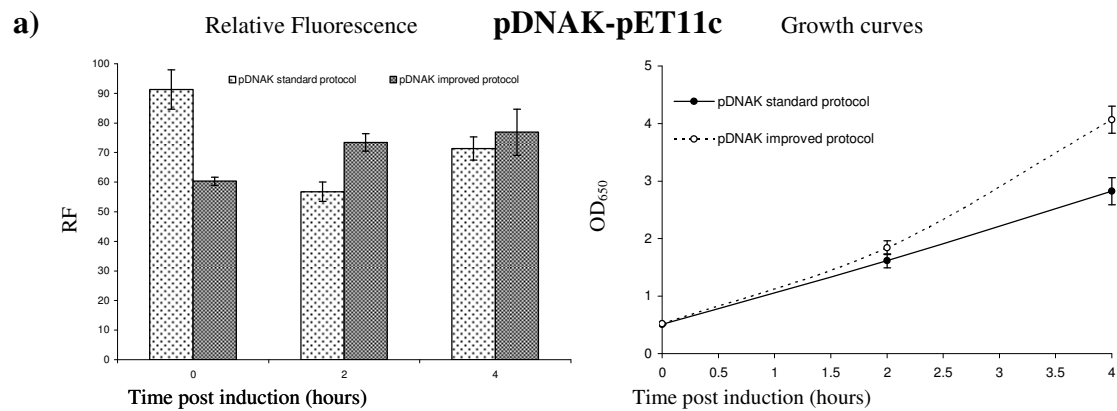
Empty overexpression vector: pET11c

In order to attribute the effects observed in previous experiments to CCP production, double transformants of *E. coli* BL21* harbouring either pGROE/pET11c or pDNAK/pET11c were cultured following standard and improved protocols.

The RF of cultures containing pGROE was higher at pre-induction in the standard protocol than the improved, probably attributed to the increased growth temperature (37°C vs. 25°C) [Fig. 5.8 (a)]. Two hours after IPTG addition, no fluorescence increase was observed in the standard protocol, which had an RF lower than that observed in the improved protocol. The growth curves were similar to those obtained in experiments with pET-CCP/pGROE indicating that it is likely that the growth rate is primarily influenced by the temperature of incubation rather than recombinant protein production. The data collected from pDNAK cultures [Fig. 5.8 (b)] was very similar in RF, as in growth, to that observed during pGROE monitoring which showed that both promoters responded in a very similar manner to the growth conditions. It is interesting that pDNAK/pET11c also showed an initial higher activity in the standard protocol. Following induction (which should not induce a heat shock

Fig 5.8 Results from pET11c vector co-transformed with the reporter plasmids

(a) shows RF (**left**) and growth (**right**) curves for *E. coli* BL21* pDNA/pET11c while (b) shows RF (**left**) and growth (**right**) for *E. coli* BL21* pGROE/pET11c. (c) shows fluorescence histograms, standard protocol is represented filled in green while improved protocol is overlaid in blue.



response, since pET11c does not contain a recombinant gene) and temperature change, the activity in the improved protocol was higher than that in the standard protocol. This could be caused by DnaK capture of σ^{32} at lower temperatures. DnaK could be at higher titres in the standard (37°C) than in the improved (25°C) protocol resulting in quenching of the heat-shock response after temperature change. This experiment also indicated that the differences in growth observed between standard and improved protocols are not caused exclusively by CCP overproduction.

Green fluorescence histograms obtained by flow cytometry [Fig. 5.8 (c)] were similar in standard (filled green) and improved protocols (overlaid in blue). In the standard protocol cultures, pGROE/pET11c and pDNAK/pET11c generated slightly brighter populations (more noticeable in pDNAK/pET11c) prior to induction as the temperature at this point was higher than that in the improved protocol. As time progressed, the fluorescence shown by the histograms was analogous to that observed in RF measurements. The non-green fluorescent populations seen in pGROE/pET-CCP and pDNAK/pET-CCP experiments after induction were not found. Single cell analysis also revealed that cells with pET11c had slightly lower values in SS/FS than pET-CCP but the difference was only noticeable during the standard protocol. This could be attributed to increases in the light scattering due to size and granularity during recombinant protein overproduction. On the other hand, the shoulder in the fluorescence distributions corresponding to high FS/SS populations was also observed in this control, confirming that its occurrence is independent of CCP overexpression. These results reinforced the previous observations in pGROE/pET-CCP and pDNAK/pET-CCP experiments where fluorescence varies as consequence of CCP overexpression.

Further culture conditions: anaerobic growth

As oxygen transfer is a key factor in large scale fermentations and because in the previous experiments reporters were shown to be potentially affected by aeration conditions, anaerobic experiments were performed. The purpose was to observe the reporter activity during recombinant protein production in anaerobic conditions as this could indicate different stress patterns. In addition, anaerobic cultures are likely to have different chaperone expression, thus potentially affecting recombinant protein production. Overnight -grown double transformants of *E. coli* BL21* harbouring either pDNAK/pET-CCP or pGROE/pET-CCP were inoculated into still, capped tubes containing 30mL of LB supplemented with 0.2% glucose and the appropriate antibiotics and were cultured following standard and improved protocols. The RF data in both conditions did not reveal significant changes upon induction [**Fig. 5.9 (a)**]. Interestingly, the values obtained were much greater during anaerobic than aerobic conditions as cells seemed to constantly show values close to the maximum levels of relative fluorescence previously observed for each reporter aerobically, as observed previously in Chapter Three (specially [**Fig. 3.8**]). This could mean that the expression of *dnaK* and *groE* in anaerobic conditions is higher than that triggered by recombinant protein production. Growth curves revealed mainly that *E. coli* BL21* grew poorly anaerobically on glucose as a carbon source [**Fig. 5.9 (b)**]. In addition SDS-PAGE did not reveal significant CCP production.

Glycerol as carbon source

Since glycerol is a carbon source commonly used in industry, overnight grown double transformants of *E. coli* BL21* harbouring pGROE/pET-CCP or pDNAK/pET-CCP were inoculated into 250mL conical flasks containing 30mL LB supplemented with 0.2% glycerol

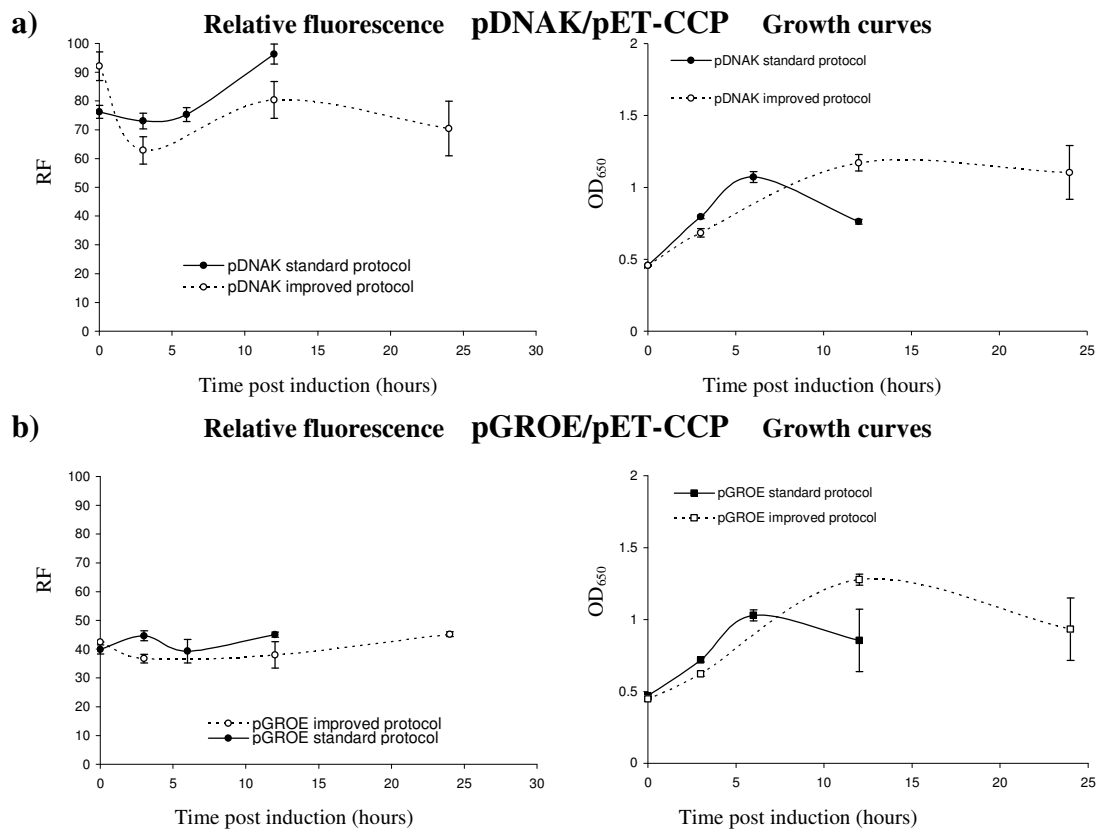


Fig. 5.9 Reporting of stress during recombinant protein production in anaerobic incubation

(a) shows the relative fluorescence of *E. coli* BL21* pDNAK/pET-CCP cultures (left) and the growth curves (right) for each reporter system while (b) shows likewise for *E. coli* BL21* pGROE/pET-CCP cultures.

and the appropriate antibiotics. Bacteria were cultured following standard and improved protocols. In the time frame monitored (0-6 hours), significant differences were not identified, either in cell morphology or fluorescent behaviour, in comparison with glucose cultures. This indicates that these reporter systems could be utilized independently of the carbon source utilized as catabolite repression did not seem to interfere with reporter transcription. However, further experiments spanning more than 6 hours post-inoculation and in larger scale are advised.

dnaK response to recombinant protein production upon increasing concentrations of IPTG

With the aim of testing the response range of *dnaK::gfp* as a reporter for recombinant protein accumulation, 0.3 mL of an overnight culture of *E. coli* BL21* co-transformed with the plasmids pDNAK and pET-CCP was used to inoculate 30mL of LB supplemented with 0.2% glucose and the appropriate antibiotics and grown in constantly shaken conical flasks at 30°C. When the OD₆₅₀ reached 0.5, duplicate cultures were induced with 0.05, 0.5, or 5mM IPTG. A pair of un-induced cultures was included as a control.

Bulk measurements

The cultures produced elevated levels of RF pre-induction followed by a decrease two hours post-induction [Fig 5.10 (a)]. At this point two groups were noticeable; control and 0.05mM cultures showed similar RF values within a lower range than the cultures induced with 0.5 and 5mM IPTG. Four hours post-induction each culture revealed increases of fluorescence according with the amount of IPTG added. Six hours after induction, all cultures showed

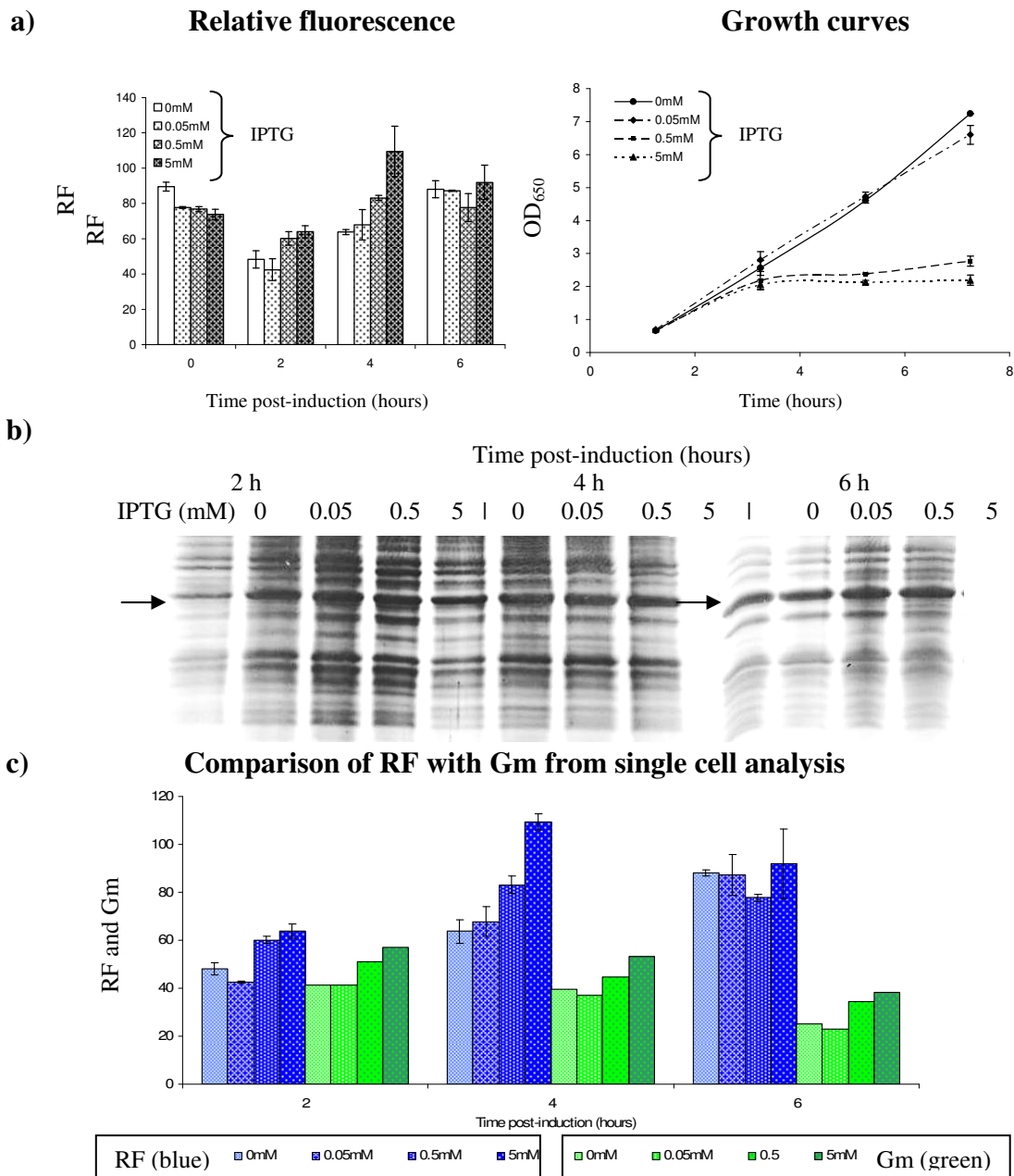


Fig 5.10 *E. coli* BL21* pDNAK/pET-CCP double transformants at different concentrations of IPTG

(a) shows the RF of cultures upon addition of IPTG (left) and growth curves (right)

(b) SDS-PAGE of total cellular protein revealed pET vector leakage (arrow) and gradual increase of CCP over-production with increasing concentration of IPTG. Note that BugBuster reagent was not utilized. (c) Is a comparative histogram of RF and Gm of green fluorescence from flow cytometry

similar RF. The growth curves **[Fig 5.10 (b)]** again revealed two different patterns. While the control cultures and the 0.05mM IPTG induced cultures grew at similar rates, the cultures induced with 0.5 and 5mM of IPTG showed growth stagnation noticeable from the second hour post-induction onwards. This growth stagnation could be causing the intracellular accumulation of GFP in the last two samples. SDS-PAGE **[Fig 5.10 (c)]** revealed that all cultures were producing recombinant protein to a degree as even un-induced cultures have shown a considerable amount of CCP presumably by leakage from the T7 promoter of pET-CCP.

Single cell analysis

At two hours post induction, flow cytometric data was highly comparable with results obtained with bulk measurements. Results could be divided into two groups: (1) control and 0.05mM IPTG; and (2) 0.5 and 5mM IPTG. While the first showed a similar narrow distribution of fluorescence in the population, the second revealed a shifted median towards higher values of fluorescence **[Fig 5.11]**. The cultures induced with 5mM showed just a slight increase of median over those induced with 0.5mM. Interestingly, higher amounts of IPTG rather than produce a shift of Gm resulted on slight increases in the homogeneity of the population interpreted from the CV of fluorescence histograms. As such features were maintained into four hours **[Fig 5.12]**, the differences in fluorescence shown interpreted from bulk measurements were not seen at the single cell level. Six hours post-induction **[Fig 5.13]** the pattern observed in the spread of the fluorescent histograms seemed to have changed and the 0 and 0.05 mM cultures had narrower fluorescence distributions comparison with the 0.5 and 5mM IPTG cultures. PI staining did not reveal dead cell populations throughout the

Post-induction: Two hours

Green fluorescence

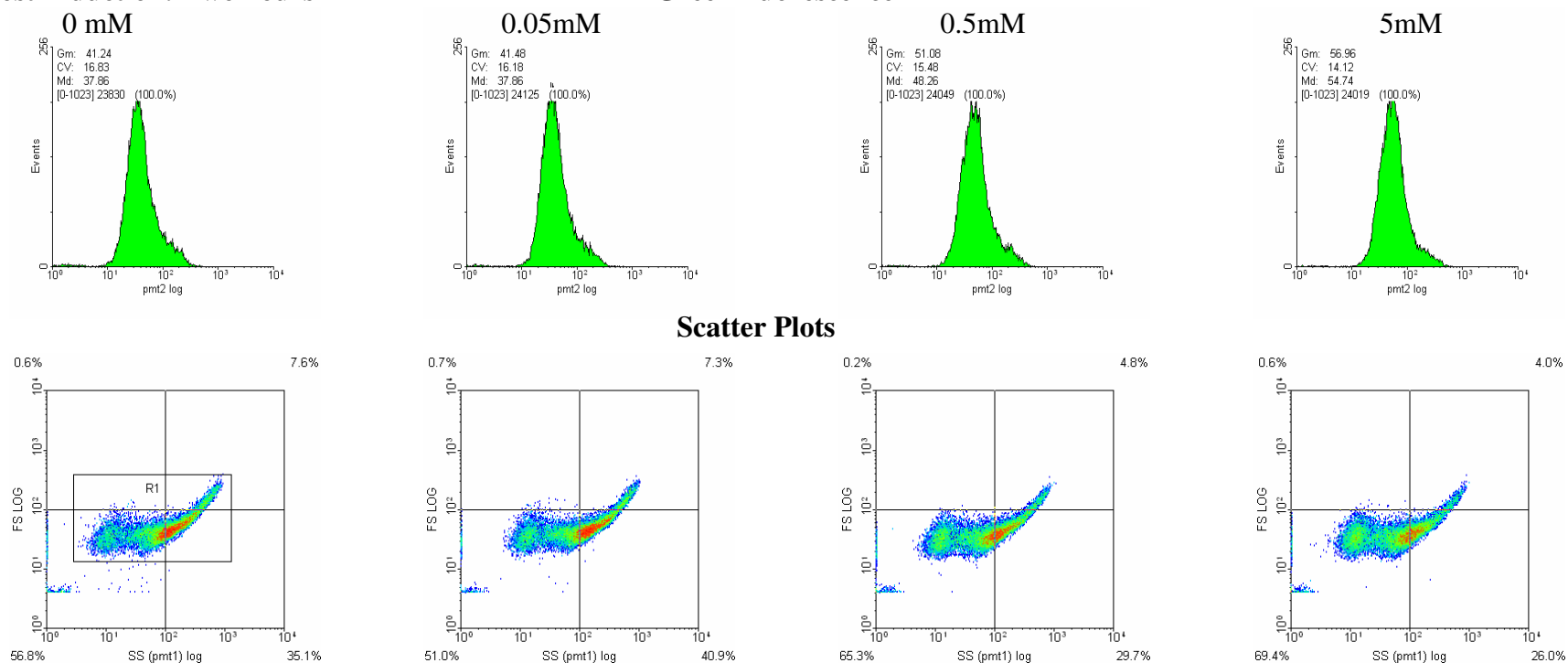


Fig. 5.11 Single cell analysis of *E. coli* BL21* pDNAK/pET-CCP two hours post-induction with 0, 0.05, 0.5 and 5mM IPTG

Green fluorescence histograms are shown in increasing concentration from left to right from representative cultures. The respective scatter plots are shown below.

Post-induction: four hours

Green fluorescence

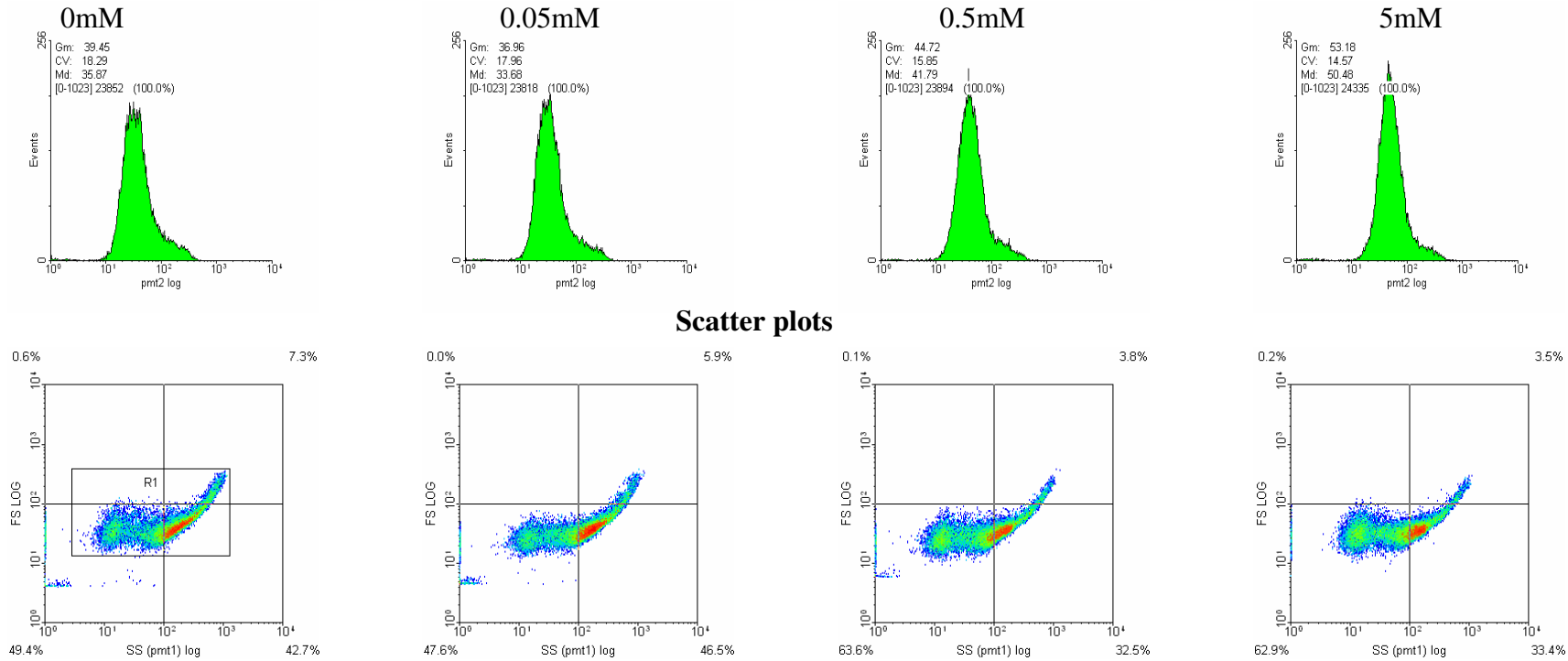


Fig 5.12 Single cell analysis of *E. coli* BL21* pDNAK/pET-CCP two hours post-induction with 0, 0.05, 0.5 and 5mM IPTG

Green fluorescence histograms are shown in increasing concentration from left to right from representative cultures. The respective scatter plots are shown below.

Post-induction: six hours

Green fluorescence

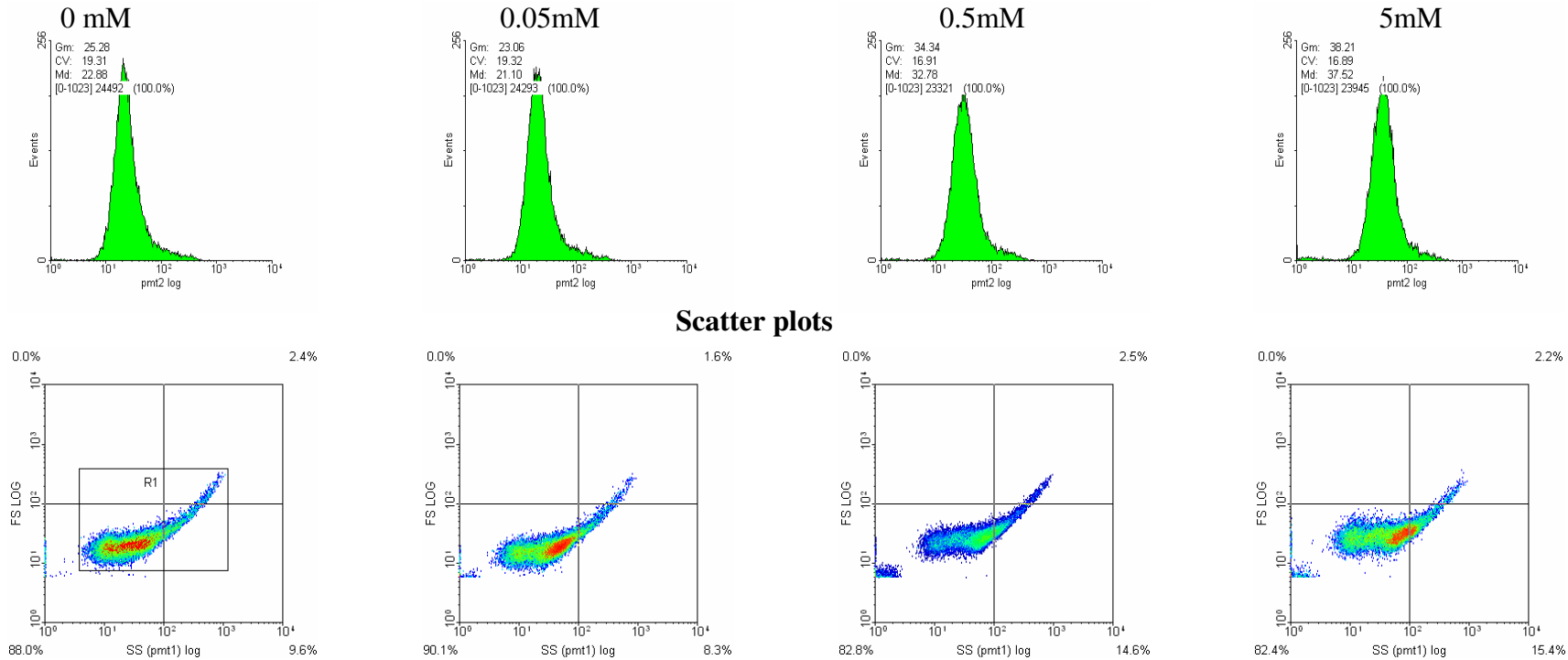


Fig 5.13 Single cell analysis of *E. coli* BL21* pDNAK/pET-CCP two hours post-induction with 0, 0.05, 0.5 and 5mM IPTG

Green fluorescence histograms are shown in increasing concentration from left to right from representative cultures. The respective scatter plots are shown below.

experiment. This could be attributed to the constant growth temperature (30°C); shifts in temperature (37°C to 25°C) could cause additional stress leading to cell death as seen in figures [Fig 5.1] and [Fig 5.4]. The data comparison between RF and Gm [Fig. 10 (c)] revealed that RF and green fluorescence Gm measurements were very similar at 2 hour post-induction. At 4 hour post-induction, the 0.5 and 5mM IPTG were more fluorescence than the 0 and 0.05 mM in RF as shown. This is also shown in green fluorescence Gm but RF measurements showed a much higher difference at 5mM. At 6 hours post-induction, single cell analysis showed more fluorescence in the cultures at 0.5 and 5mM IPTG while RF did not reveal significant differences. These discrepancies are likely to be caused by morphological characteristics of the cell as, in this sample point, single cell analysis showed reduced FS at 0 and 0.05 mM IPTG in comparison with 0.5 and 5mM IPTG induced cultures.

5.2.4 Batch fermentation

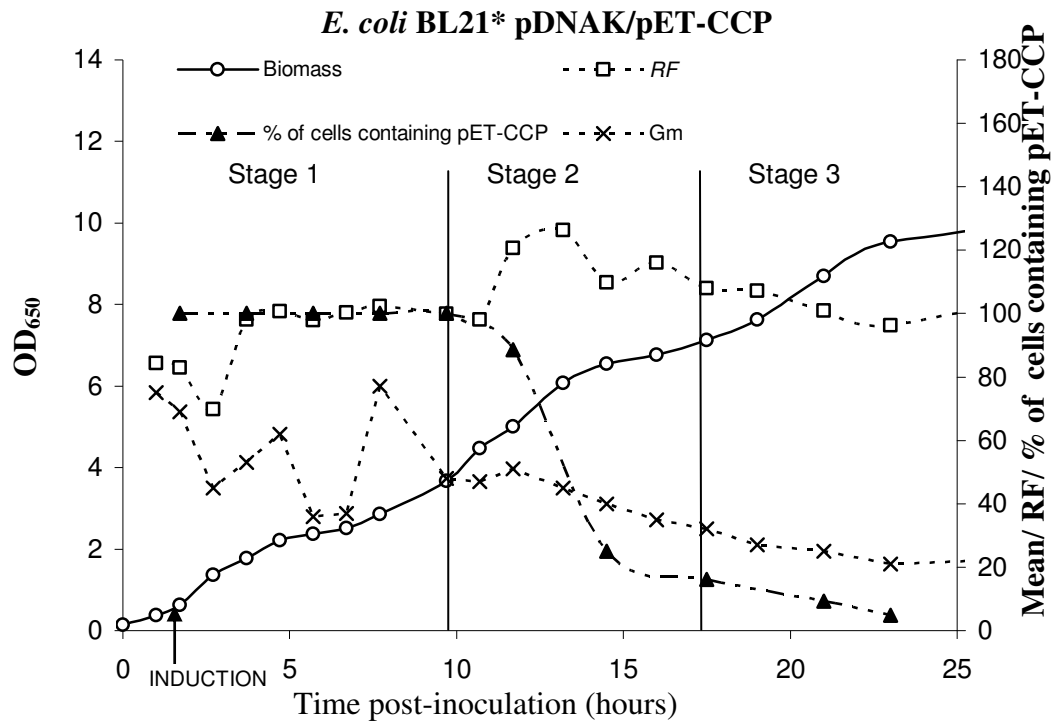
In order to apply the reporter technology in bioprocess, data generated should discriminate normal from stressed states. The reporter plasmid pDNAK was utilized in batch fermentation because, as shown previously [Fig 5.4], pDNAK activation in the standard protocol produced a discrete response upon CCP accumulation while pGROE seemed to accumulate over time as CCP was being made [Fig 5.1]. The sensitivity shown by pDNAK to report stress upon CCP production made it a promising candidate to detect misfolding stress. *E. coli* BL21* co-transformed with pET-CCP and pDNAK was cultivated in a 5L bioreactor (Electrolab) containing 2940mL of LB supplemented with 2% (w/v) glucose and 100µg mL⁻¹ carbenicillin. The medium was inoculated with 56mL (2% of total volume) of seed culture grown in LB at 30°C for 14 hours aerobically. Seed culture was grown in a 1L conical flask containing the

total LB volume required for the bioreactor and supplemented with the appropriate antibiotics. After bioreactor inoculation, bacteria were incubated at 37°C to an OD₆₅₀ of approximately 0.5 at which recombinant protein expression was induced by addition of 0.5mM IPTG. At this point, the incubation temperature was changed to 25°C, following the standard protocol. pH was set at 6.3 and controlled by automated addition of 5% (v/v) HCl and 10% (v/v) NH₃. Aeration was maintained with 1vvm air and stirring speed of 700 rpm. Temperature, pH and dissolved oxygen were monitored on-line by the Electrolab monitoring system. Exhaust gas composition was analyzed on-line by connection to a gas mass spectrometer. Samples taken were analyzed off-line for: biomass, by OD_{650nm}; RF, by fluorimetry; recombinant protein production, by SDS-PAGE; viability and plasmid retention, by CFU and replica plating, and single cell analysis, by flow cytometry. The data shown are representative of two independent repeats.

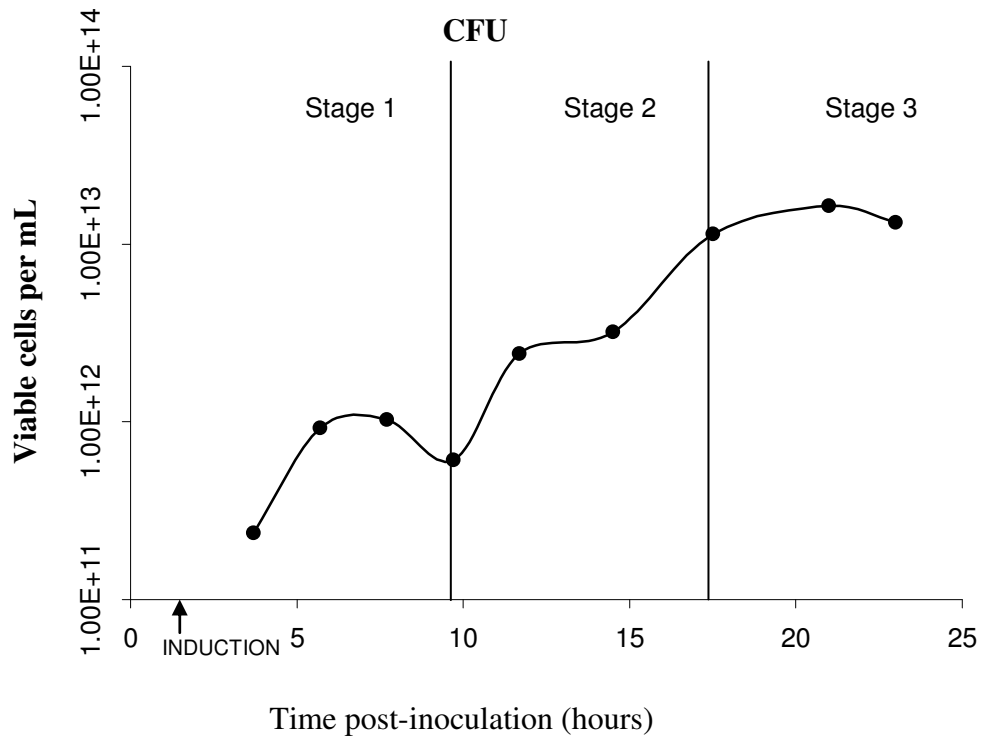
Off-line analysis

Data collected through off-line methods are shown in **Fig 5.14 (a)**. The biomass increased at variable rates resulting in a pattern of a series of 3 consecutive growth and plateau phases. The first stage consisted of recombinant protein production, as shown by SDS-PAGE [**Fig 5.14 (c)**] from the 1.7 to 10.7 hour samples. After induction, the relative green fluorescence (RF) increased and maintained values at approximately 100 RF units from 3.7 to 10.7 hours [**Fig 5.14 (c)**]. At 11.7 hours RF started an uptrend for two sample points up to more than 120 RF units. This was followed by gradual downtrend towards 100 RF units. Single cell analysis, represented by the Gm of green fluorescence histograms increased after induction with the exception of data points at 5.7 and 6.7 hours. Contrasting with RF measurements, Gm decreased sharply at 9.7 hours which was followed by gradual downtrend from 11.7 hours

a)



b)



c)

SDS-PAGE

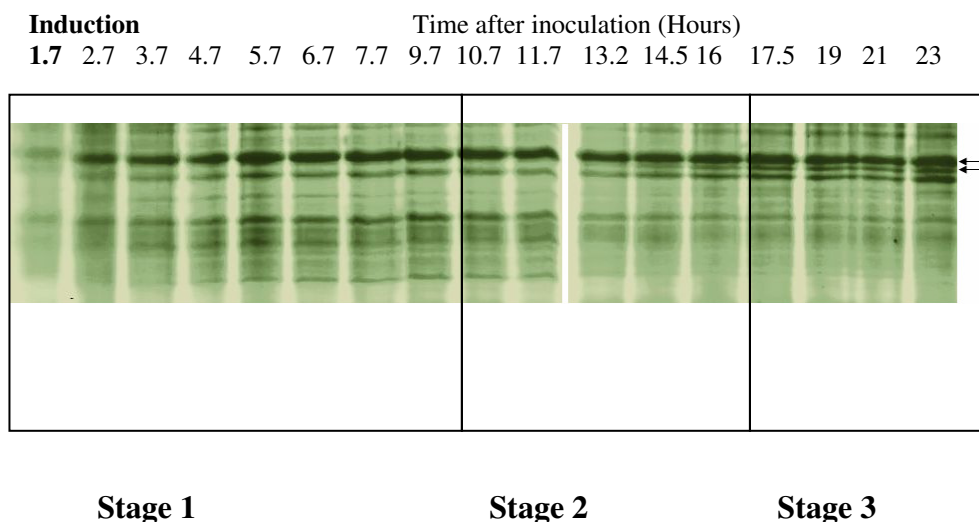


Fig 5.14 Off-line sample analysis during batch fermentation

(a) Analysis of off-line samples. The left Y axis indicates only OD₆₅₀. The other variables are represented against the right Y axis. The values of geometric mean (Gm) of green fluorescence from flow cytometric histograms have been included to facilitate comparison of trends. Readings are divided in 3 stages according to development of significant events during cultivation. (b) shows the viable counts over time and (c) shows the SDS-PAGE where the gel has been divided in the cultures stages. CCP is indicated by the arrow (i). The contiguous band beneath (ii), which becomes most visible in the stage 3, is suspected to be mature CCP following cleavage of the peptide signal (Turner et al., 2003)

until the end of the fermentation. On closer inspection of green fluorescence, data revealed histograms where cell populations shifted to greater values of Gm from one hour post induction to a peak at 4.7 hours [**Fig 5.15 (a)**]. This was followed by a sharp reduction in Gm in samples at 5.7 and 6.7 hours and a sharp increase at 7.5 hours. Independent reproduction of the fermentation showed a similar behaviour characterized by an increase upon induction followed by a decrease in the following sample point. The sharp decline in single cell fluorescence at 9.7 hours post-inoculation marked the beginning of the second stage in the culture.

From forward scatter and side scatter plots, two further analyses were carried out. First, the population of cells with high FS and SS (as previously observed, and thought to be filamentous cells) in respect to the total number of cells is plotted in [**Fig. 5.17 (c)**]. The size of this subpopulation was observed to vary through the fermentation increasing drastically at 4.7 hours and maintaining a constant value up to 7.7 hours. Interestingly, this population increased further at 9.7 hours remaining high until 11.7 hours. This was followed by a sharp decrease and a gradual downtrend. Fluorescence microscopy [**Fig. 5.18 (a)**] was used to observe cells at 4.7 hours and 10.7 hours, confirming that in the later sample, more cells were filamented and the filaments were longer than in the earlier sample. Filaments can be detected by flow cytometry in regions of very elevated FS, SS and green fluorescence. These three values are correlated in such cases. During integration of the signal, the forward scatter detector photodiode of the flow cytometer accumulates the energy received until the cell leaves the laser beam. This means that filaments may be detected because their integrated signal is substantially greater than normal cells.

2.7 hours (1 hour post-induction)

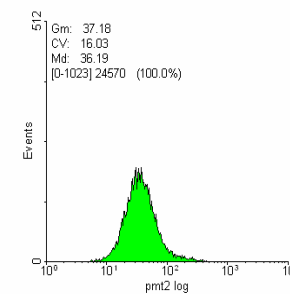
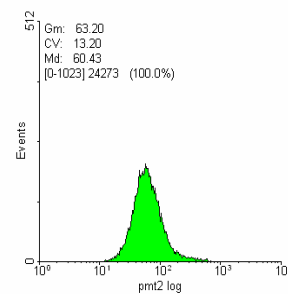
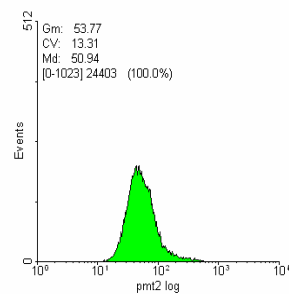
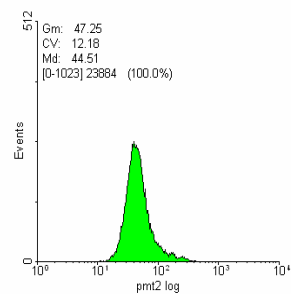
a)

3.7 hours

4.7 hours

5.7 hours

Green fluorescence histograms



b)

Scatter plots

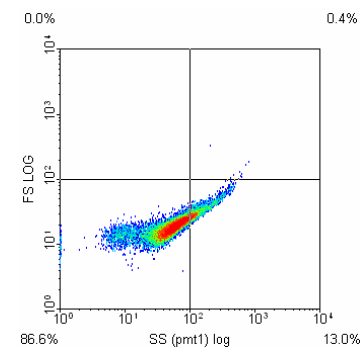
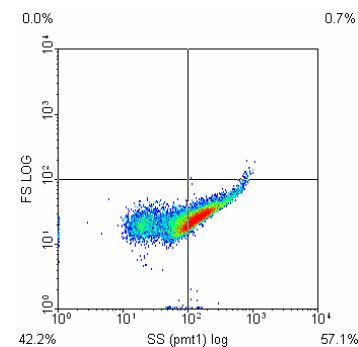
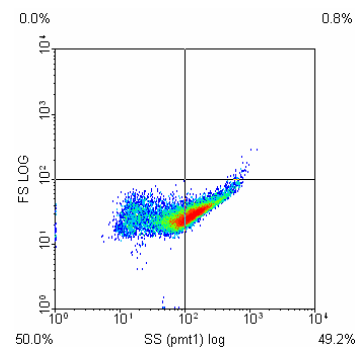
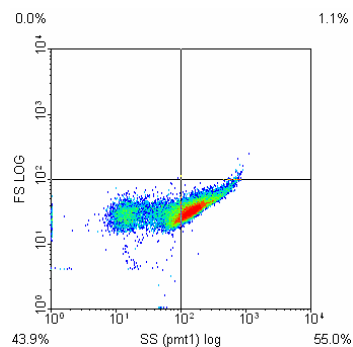


Figure 5.15 Single cell analysis during batch fermentation

(a) shows histograms and (b) scatter plots progressing over time. The graphs shown in this figure correspond to the first burst of GFPmut2 expression. These events are correlated with recombinant protein production and with RQ which is shown subsequently.

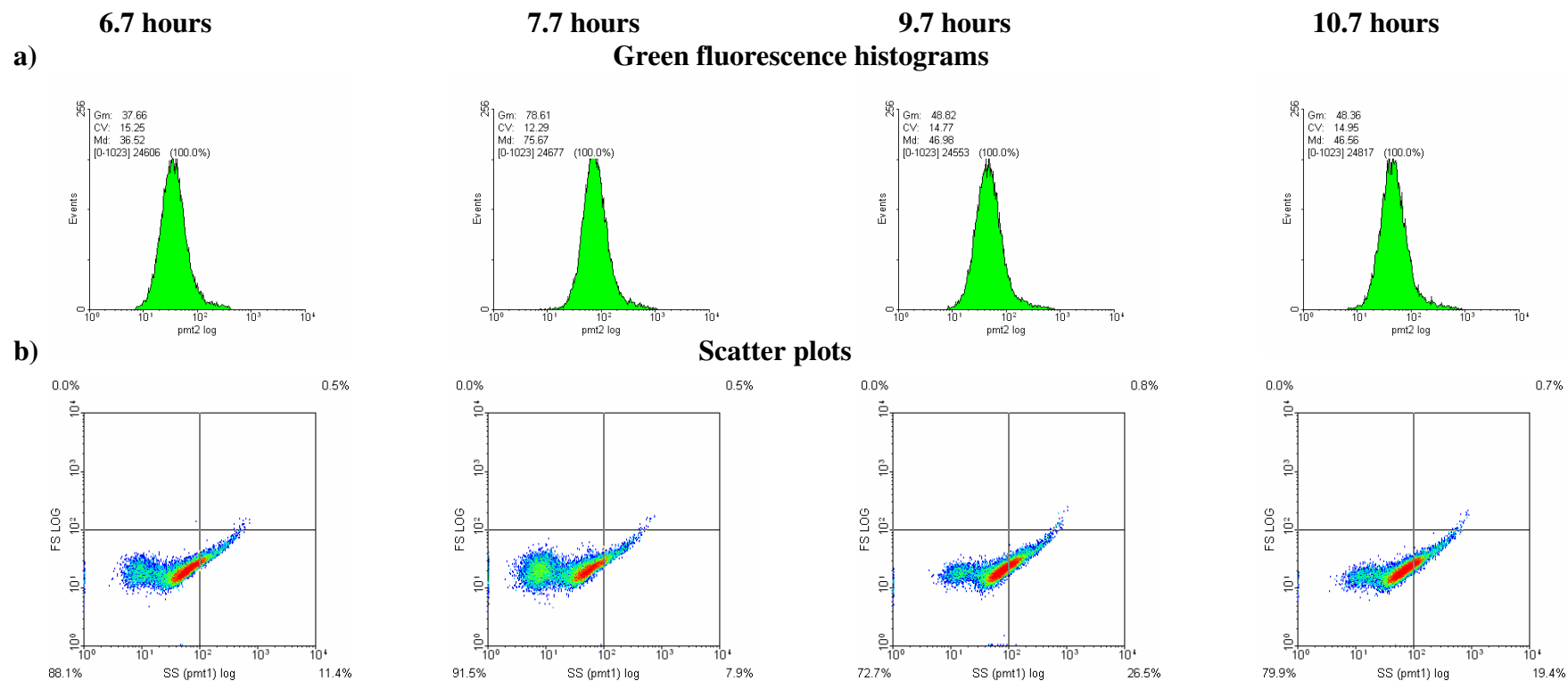
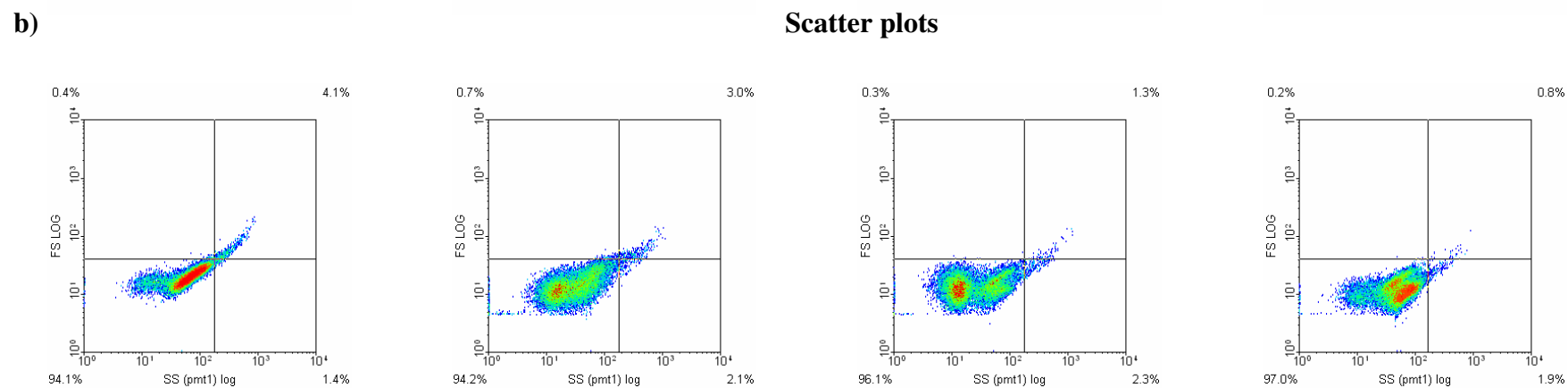
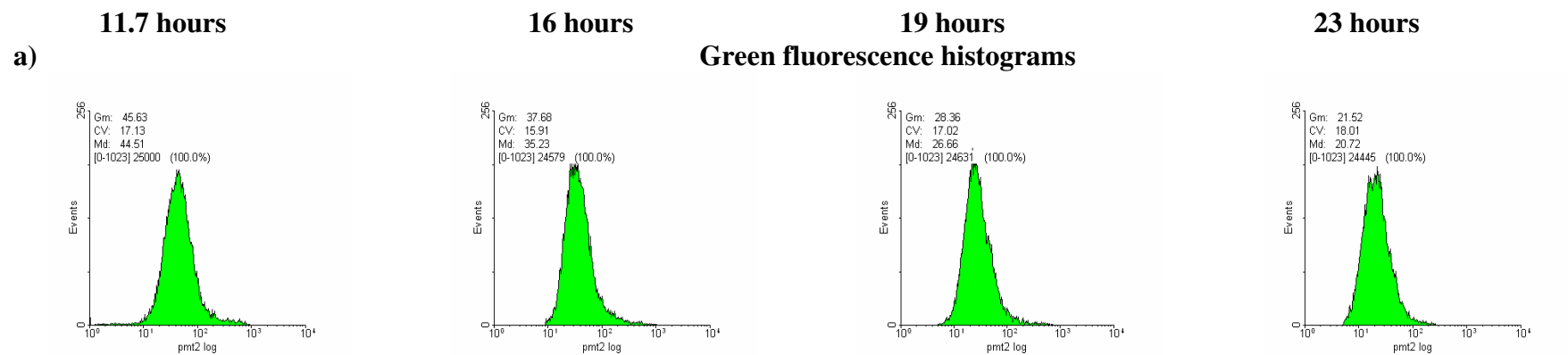


Fig 5.16 Single cell analysis during batch fermentation (continuation)(a) shows histograms and (b) scatter plots progressing over time. The graphs shown in this figure correspond to the second and stronger burst of GFPmut2 expression. These events are also correlated with recombinant protein production and with RQ which is shown subsequently.



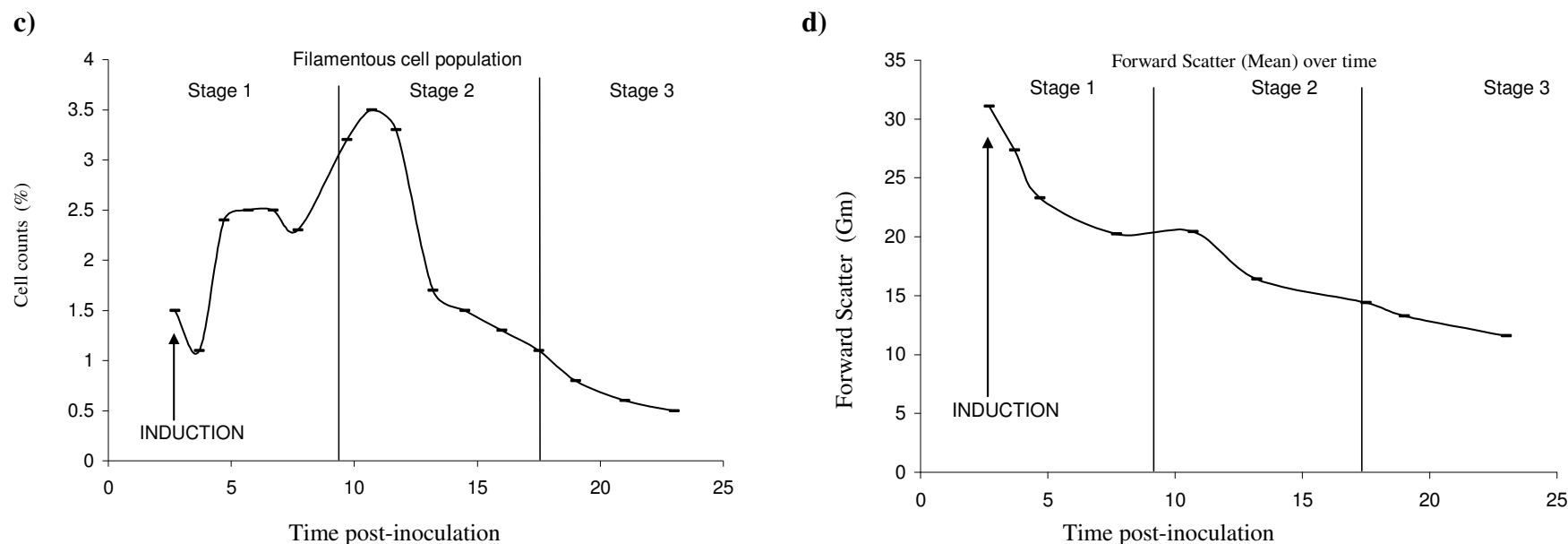


Fig 5.17 Single cell analysis during batch fermentation, last 12 hours, shows decrease in filamentation

(a) shows fluorescence histograms. (b) Shows scatter plots progressing over time. The graph in (c) shows the proportion of populations with high SS/FS (attributed to filamentation) of each sample over time. (d) Shows the Gm of FS histograms from each sample over time. FS is correlated to cell size.

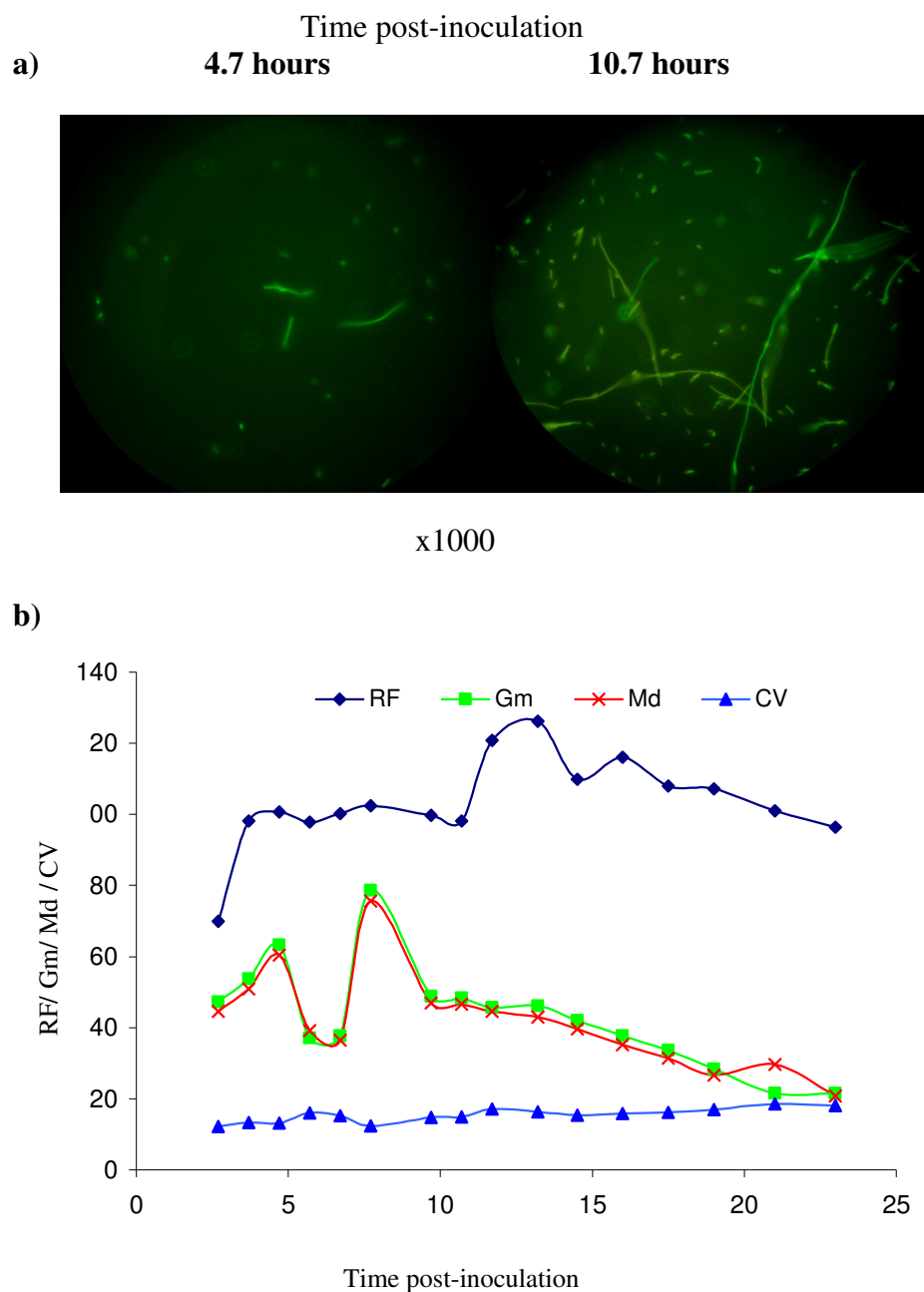


Fig. 5.18 Filamentous morphology of *E. coli* BL21* pDNAK/CCP and RF-Flow cytometry comparison

Filamentation is shown in (a). The graph (b) is a comparison of bulk measurements of fluorescence (RF) with the statistics of single cell analysis Gm (geometric mean), Md (median) and CV (covariance).

Second, the geometric mean of forward scatter is plotted against fermentation time in **[Fig 5.17 (d)]**. The gradual reduction in FS over time could be indicative of plasmid loss as the elimination of a replicating plasmid may reduce the cytoplasmic volume (Gasol et al., 1999)

The second stage of the fermentation was characterized by the decline in plasmid retention noticeable by nearly 20% pET-CCP loss at 11.7 hours increasing to 75% loss at 14.7 hours. pDNAK was retained by at least 90% of the cells throughout. When the plasmid loss entered a slower rate of decline at 14.7 hours, there was a gradual decrease in RF and Gm from single cell analysis. In addition, both the proportion of filamentous cells **[Fig 5.17 (c)]** and the Gm of FS (signifying cell size) **[Fig 5.17 (d)]** decreased throughout the second and third stages.

In the third and final stage, biomass increased was resumed as shown by OD₆₅₀ and viable counts which was followed by a final plateau towards the end of the fermentation. There was no significant pDNAK loss where typical levels above 90% retention were maintained throughout the fermentation. PI staining did not reveal significant populations of dead cells. As shown in previous experiments (**[Fig. 5.1]** and **[Fig. 5.4]**), growth arrest and cell death occurred when CCP was made and the temperature changed. However, when either temperature changed **[Fig. 5.8]**, or CCP was made **[Fig. 5.10]**, dead cells were not detected, only growth rate reduction was observed. Therefore it is possible that the combination of both factors is what was causing cell death. The reason that cell death was not observed here could be because the temperature decrease in the bioreactor is likely to occur more gradually than in the shake flasks, sudden temperature changes being more detrimental to cell viability. The SDS-PAGE **[Fig 5.14 (c)]** revealed additional bands from 14.5 hours onwards which are believed to correspond to maturation of CCP. This process involves the

cleavage of the signal peptide of CCP resulting in a lighter band (Turner et al., 2003) shown in the figure by an arrow (ii).

On-line monitoring

The temperature of the vessel was efficiently maintained. Upon induction, the temperature transition from 37°C to 25°C occurred gradually (~40 minutes). Dissolved oxygen decreased over time correlating with the increase shown in biomass [Fig 5.19 (a)]. The first stage showed a spike in DO occurring simultaneously with cooling to 25°C, probably caused by the temperature change. During the second stage, DO was maintained at approximately 60%. Such measurements contrasted with the biomass results as, in this timeframe, biomass increased at relatively fast rate. On the other hand, most of the pET-CCP loss occurred in this interval of time. The third stage of DO monitoring followed a constant decrease in parallel with biomass increasing.

Mass spectrometry was utilized to analyze and monitor the composition of the exhaust gas from the bioreactor. The consumption of O₂ (OXC) [Fig 5.19 (b)] increased over time accordingly with bacterial growth [Fig 5.14 (a)] in the first stage of the fermentation. In the same time frame, an increase in CO₂ concentration (CDC) was also observed. During the second stage, OXC was constant for four hours followed by further increase from 15 to 20 into the third stage. At this point, (OXC) and (CDC) decreased from 0.4 to 0.3, which was maintained for three hours. The OXC and CDC showed a sharp increase in the last two hours of the fermentation. The respiratory quotient was mostly maintained at 1.1, typical of aerobic fermentation.

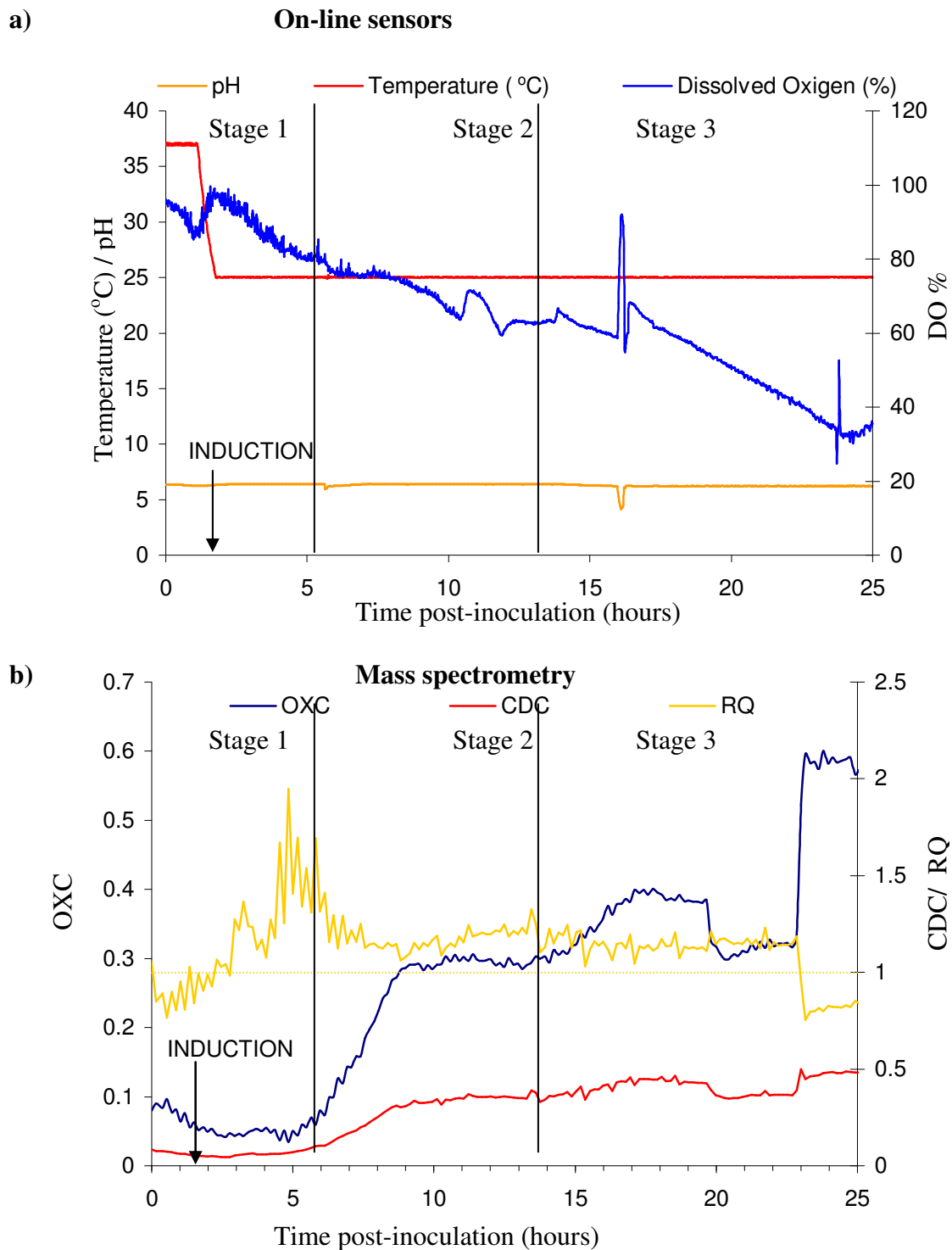


Fig 5.19 On-line monitoring during batch fermentation

(a) Shows the readings of on-line sensors in the bioreactor. (b) Shows the mass spectrometry measurements of gas exhaust from the vessel.

5.3 Discussion

5.3.1 Chaperone expression was greater under standard than under improved protocol

In these experiments it was demonstrated that the fusions *groE::gfpmut2* and *dnaK::gfpmut2* produce measurable amounts of GFP in response to recombinant protein production. This provided evidence that microbial stress (specially the heat-shock response) can be controlled when the improved protocol is used.

One great difference between protocols was the bacterial growth rate. It was found that the major cause of growth reduction in the standard protocol in relation to the improved protocol was the temperature shift as the control culture with empty pET11c [Fig. 5.8] grew in a similar manner to double transformants with pET-CCP instead ([Fig 5.1] and [Fig 5.4]). pET11c control was important because it showed that, while the initial temperature produced an stress response in the standard protocol, the RF and single cell analysis revealed similar levels of GFP expression throughout.

In relation to previous work it is important to emphasize that because the promoter regions are whole intergenic regions (Zaslaver et al., 2006) rather than unique specific promoter (*e.g.* pQF50KgroE, from Chapter Three), interferences in *gfpmut2* expression could be possible from factors derived from growth conditions or recombinant protein production strategy.

pDNAK reported stress in batch fermentation during standard protocol

Stress originated from recombinant protein production was observed in *E. coli* BL21* transformed with pDNAK and pET-CCP. The initial uptrend in the Geometric mean of the fluorescence [Fig 5.14 (a)] was likely to report for DnaK expression upon induction of RPP.

5.3.2 *groE* and *dnaK* promoters reported misfolded protein stress differently

The intracellular roles of GroE and DnaK are complementary (Gragerov et al., 1992). In order to balance the proportion of these chaperones in the cell, they have evolved different promoter structures. Since the reporter plasmids contain the same promoter structure as that on the bacterial genome, they are expected to be activated by physiological demand for chaperones during normal bacterial growth as well as during situations of stress.

In the experiments performed, single cell analysis of pGROE/pET-CCP double transformants [Fig. 5.3], showed accumulation of single cell fluorescence gradually over time. Green fluorescence Gm from the standard protocol were 2, 3 and 4-fold higher at 2, 4 and 6 hours respectively [Fig 5.3 (d)] than that those measured in the improved protocol. Fluorescence histograms can also be compared at 2 hours [Fig 5.2 (b1)], 4 hours [Fig 5.3 (a1)] and 6 hours [Fig 5.3 (b1)]. The histograms also show the median (Md). Median is the value of fluorescence in PMT2 which accumulate half of the total population since it could indicate changes of GFPmut2 expression by the population in the case of selective pressures; while $Md > Gm$ suggest that reporter activity could be increasing, $Md < Gm$ suggests a decrease. Alternatively, pDNAK transformants ([Fig 5.5] and [Fig 5.6]) started at higher fluorescence than pGROE transformants and rather than gradually increase in fluorescence, they responded

with a sharp increase of GFPmut2 expression more than 3-fold greater in the standard protocol compared with the improved protocol [Fig 5.6 (d)]. These results fit with the most supported interpretation of DnaK and GroE cooperation where DnaK is an early respondent to newly synthesised proteins while GroE intervenes at a later stage (Gaitanaris et al., 1994). The RF histograms generated by pGROE [Fig. 5.1 (a)] and pDNAK [Fig. 5.4 (a)] revealed a similar response but it is important to highlight that the relative fluorescence per cell was distorted as dead cells increase OD values but are not fluorescent thus doubly deceasing RF .

Differences in expression patterns can be explained through differences in the structure of these reporter promoters. The *dnaK* promoter region [Fig. 2.3] contains 3 binding sites for σ^{32} establishing three promoters. P1 and P2 account for most of the gene expression as P3 is found partially overlapping P2. The specificity of σ^{32} for these sites is very high as one of the critical aspects of σ^{32} binding is the spacer length which is the base pair distance between the -35 and -10 elements. P1 and P2 possess A/T rich motifs (UP elements) (Nonaka et al., 2006) for α CTD binding. P1 and P2 are sufficiently separated to be able to function simultaneously being able to exploit the available pool of σ^{32} . The *groE* promoter contains a strong σ^{32} binding site which has been exclusively cloned into the reporter plasmid pQF50KgroE (studied in Chapter Three). Studies *in vitro* with this reporter plasmid have shown that there is not significant binding of σ -factors other than σ^{32} to this structure (Wang and deHaseth, 2003). However more recent studies (Wade et al., 2006) indicated a possible overlap of the σ^{32} site with a σ^{70} site as it was found that their *groE* reporter plasmid was able to bind σ^{70} *in vitro*. The reporter plasmids utilized by Wade et al. (2006) study contained a great extension

of the intergenic region similar to that in pGROE [Fig. 2.2] Chapter 2. By analysis of the

promoter region, an UP element and a potential CRP binding site were found (Ebright, Ebright, and Gunasekera, 1989) which could make it susceptible to catabolite repression but experimental evidence is still needed. It is possible that the combination of the A/T rich motif (Nonaka et al., 2006) with the CRP site enable σ^{70} binding to ensure a low level of *groES/EL* operon expression during the absence of σ^{32} -related stress. *E. coli* BL21* pGROE/pET-CCP grown under standard and improved protocols with glycerol or glucose as a carbon source did not show such repression in *groE::gfpmut2* the time frame of 6 hours post-induction, suggesting that catabolite repression was not observed here.

On the other hand, *dnaK* is regulated by σ^{32} and there is no experimental evidence of further modulation. It has been shown that P3 may possess a very low affinity to bind σ^{70} (Wade et al., 2006). This promoter is likely repressed by physical overlap of σ^{32} on P2. As *dnaK* is a first respondent to protein overproduction controlled mainly through σ^{32} it could be better as a reporter for misfolded protein stress during recombinant protein production. Previous results by (Nemecek et al., 2008) agree with this; a low copy number plasmid featuring a *dnaK::gfp* fusion was tested for response to recombinant protein production. While their results agree with those shown in this study, data was collected differently as in this study the measurements shown were based on optical methods and fluorescence rather than ELISA quantification of GFP mass.

The factor σ^{32} is the main coordinator for chaperone expression. The amount of GFPmut2 expression observed comparing cultures transformed with pDNAK or pGROE was roughly 2:1 which is accordance with the number of major σ^{32} promoters contained in each reporter system. The protein CCP is a highly lipophilic protein with a peptide signal for

translocation to the periplasmic side of the inner cell membrane. This could potentially elicit the periplasmic stress response with consequent activation of σ^E and the CpxRA system. The levels of σ^{32} mRNA could be raised by upregulation of *rpoH* gene via σ^E (P6) as shown by (Kallipolitis and Valentin-Hansen, 1998) and CpxR-P has been shown to upregulate *rpoH* expression (Zahrl et al., 2006). However, temperature reduction from 37°C to 25°C decreases the fraction of σ^{32} being translated from the mRNA pool. σ^{32} binding by DnaK also increases as temperature drops making more difficult σ^{32} release by competition with misfolded proteins. The temperature reduction during the standard protocol has shown to be a factor contributing to cell death in shake flask CCP production. When recombinant protein was being made, the shift to 25°C reduction may have impeded the heat-shock response and was likely to be masking an otherwise stronger stress response leading to more chaperone production. The temperature reduction was originally implemented to reduce folding stress in the host, but it seems that it had the unintended consequences of reducing the ability of the cells to act in response to protein overproduction.

5.3.2 The reporter response can be masked

The reporter system has a sensing range defined by the basal expression of GFPmut2 at in the absence of stress and maximum rate of expression for the most extreme cases such as during anaerobic growth [Fig. 5.9]. The reporter plasmids pDNAK and pGROE exhibited high levels of fluorescence after overnight growth as fluorescence accumulated in the inocula. However, if such fluorescence was exclusively caused by GFPmut2 stability, the cellular fluorescence for both reporters would be expected to achieve similar levels as the maximum capacity of *E. coli* to hold GFPmut2 is reached. Instead, bacteria harbouring

pGROE produced significantly less fluorescence than that observed in pDNAK. This effect is clearest in the later stages of anaerobic growth experiments [Fig. 5.9] which shows the regulation of the GFP expression from each plasmid rather than exclusively accumulation over time.

GFPmut2 has been shown to be extremely stable with a half-life *in vivo* > 7 days (Blokpoel et al., 2003). This could cause difficulties to obtain reliable reporter data representative of the levels of short lived proteins such as σ^{32} . However it is possible that GFPmut2 is, in fact, delivering reliable measurements for DnaK protein and GroES/EL subunit presence which are considered to be quite stable. During CCP production experiments, when cells with high chaperone levels (estimated from GFPmut2) were induced, some cell division steps needed to occur until the CCP synthesis overcame the chaperone availability resulting in visible stress reporting.

5.3.3 Fluorimetry versus Flow cytometry for monitoring of reporter activity during bioprocess

Fluorimetry of GFP has proven to be a useful technique for fast monitoring of microbial stress. In addition it has shown a great potential implemented as a fermenter probe for on-line monitoring (Hisiger and Jolicoeur, 2005). However there are limitations that have to be contemplated. First, these measurements underestimated the stress of individual cells as dead cells (PI positive) were shown to rapidly lose green fluorescence (refer to [Fig. 5.2 (c)] and [Fig. 5.5 (b3)]) but still contribute to optical density. Secondly, Relative fluorescence was shown to be affected by heterogeneity in the sample such as filamentation or GFP expression variability. This overestimation can be observed comparing progress the of the

filamented population [Fig. 5.17 (c)] with RF and single cell analysis measurements [Fig. 5.18 (b)] showing that the increase of RF at 11.7 hours could have been caused by an increase in filamentation.

Flow cytometry revealed heterogeneity in cell morphology in batch fermentation

From scatter plots of flow cytometric analysis [Fig. 5.17 (b)], a population of high FS and SS originated. Filamentation can be observed by flow cytometry as group of outliers that projects to a much higher SS/FS region than the main cluster. As the filaments flow in hydrodynamic orientation across the laser beam, the signal is integrated until the filament leaves the beam. This results in high measurements of FS, SS and fluorescence. Fluorescence micrographs confirmed such morphology [Fig. 5.18]. Therefore filamentation can contribute to the overestimation of bulk fluorescence when measured in RF as the OD₆₅₀ does not discriminate filamented cells (Granade, Hehmann, and Artis, 1985) while they are still fluorescent.

After the initial recombinant protein production stage, single cell green fluorescence geometric means decreased whereas [Fig 5.14 (a)] RF measurements increased. A plausible explanation for such differences is the observed variations in cell morphology. The link between cell morphology and heat-shock response is that appropriate DnaK levels are required in the cell for avoiding defects in cell division (Bukau and Walker, 1989). This is the reason why overexpression of DnaK with the intention of obtaining enhanced folding during recombinant protein production results in abnormal septation (Martinez-Alonso et al., 2010), thus filamentous morphology. It has been shown that an inadequate DnaK/GrpE ratio

was observed to cause filamentation demonstrated by GrpE overexpression (Sugimoto et al., 2008).

Interestingly, the non-CCP producing *E. coli* BL21* pET11c produced a similar scatter histogram distribution indicating that filamentation does not occur exclusively because of CCP overproduction [Fig. 5.8 (b)]. It was observed in *rpoH* mutants that plasmid relaxation triggers a heat-shock-like response, independent of *rpoH* (Mizushima et al., 1993), which results in an increase of chaperone expression, especially as DnaK helps to maintain the negative supercoiling of the DNA during heat-shock (Ogata et al., 1996). Therefore its presence and possible alterations of plasmid DNA topology, DNA relaxation or supercoiling, interfere with the demand for DnaK. This could explain the difference of fluorescence in unstressed cultures between single transformants of pDNAK evaluated in the previous chapter. Single transformants fo pDNAK [Fig. 4.17] were less fluorescent than and double transformants pDNAK/pET11c [Fig. 5.8 (a)]. A size effect is discarded as unstressed single transformants pGROE and pGROE/pET11c ([Fig. 4.14] and [Fig. 5.8 (a)]) have not shown such difference. Single transformants of BL21* did not become filamented.

Stress derived from recombinant protein production could result in plasmid loss during batch fermentation

CCP per biomass unit increased during the first stage of the fermentation. The protocol of incubation possibly caused a point at which such recombinant protein production rates were unsustainable for the cell. Studies have shown that misfolded protein is likely to be attracted towards the membrane (Winkler et al., 2010). This is very likely in the case of CCP as it

contains as signal peptide for membrane insertion. It was shown that stress leading to protein misfolding resulted in an asymmetric distribution of aggregates at the poles. As cell division proceeds, some daughter cells are better protected as they inherit newly synthesised poles while others inherit the aggregate-containing poles which results in premature cell senescence (Winkler et al., 2010). This struggle for survival in the fermentor conditions causes the viable cells to lose the production plasmid. It is possible that most of the viable but non-culturable cells still contain the production plasmid. There was great variability in FS during this stage as cells with only reporter plasmid are likely to have smaller cytoplasmic volume than those containing both [5.17 (d)]. Interestingly, as the culture approached the end of the fermentation the number of bacteria with filamented morphology gradually decreased in each sample [5.17 (c)] suggesting that filamentation could be caused by high content of DNA in the cell due to the presence of both plasmids. An interesting observation is that monitoring the *E. coli* BL21* single transformants in the previous chapter, populations of high FS and SS were not observed.

As a general observation, the strain *E. coli* BL21* performed poorly exhibiting low growth rates and low recombinant protein production especially in anaerobiosis. Some experiments could not be performed, such as anaerobic growth with glycerol as a carbon source because *E. coli* BL21 has a mutation in the gene *fnr* (fumarate nitrate reductase) therefore it cannot utilize nitrate as a final acceptor of electrons (Pinske et al., 2011). This implies major defects in anaerobic metabolism; therefore it is recommendable the use of a different host to further study reporter activity in industrial environments.

5.4. Conclusions and future work

While the *groE::gfpmut2* fusion was able to report microbial stress caused by accumulation of recombinant CCP, *dnaK:gfpmut2* produced a more intense response. In the work performed in this study it was shown that, at these scales of cultivation (30-3000 mL) and the timeframes monitored, the single cell analysis revealed monodisperse distributions in the fluorescence histograms. This means that while the cultures revealed some extent of variability there were not observed bimodal distributions of fluorescent cells.

The plasmid pDNAK has been shown to respond to early stages of misfolded protein stress which could make it potentially useful to monitor perturbations in the heat-shock response during cultivation. In the standard protocol, lower temperatures probably resulted in a stronger σ^{32} -DnaK interaction which is accurately regulated by temperature (Chattopadhyay and Roy, 2002). Sudden temperature changes, at small scale during the standard protocol, were shown to reduce stress but could have prevented the heat-shock response developing, leading to cell death. On the other hand, pGROE could be a good reporter of stress caused by misfolded protein which accumulates slowly. In addition, it is important to consider further aspects of microbial physiology that are not being monitored by the reporter plasmid but may interfere with the results. This is because the gene reporter technology has been shown to be influenced by diverse aspects of bacterial cultivation such as growth rate and cell morphology. An advantage of flow cytometry over fluorimetry is the ability to monitor these changes in morphology and relate them to fluorescence and other parameters. For this reason, flow cytometry was very useful. Although the reporter vectors pGROE and pDNAK need further development before they are implemented in industrial bioprocess monitoring,

coupled with flow cytometry they have become a useful system to assist in the design of bioprocessing strategies bringing into account microbial response to misfolding stress.

Further work would be required to determine the suitability of a reporter system to extract the right information. For example, as data revealed, pGROE could be suitable to report stress situations where recombinant protein is accumulated gradually over time. This could be tested in fed-batch fermentation under the improved protocol as recombinant protein is accumulated over longer periods of time. On the other hand, pDNAK appears to be more sensitive than pGROE, but its response is more qualitative (it is either on or off), thus would be better suited to detecting the immediate response to misfolding stress. It is recommended that data from GFP reporter plasmids are validated by comparing the concentration of DnaK and GroE (specifically chaperonin GroES) to that of GFPmut2 in response to misfolding stress with ELISA plates as shown before by (Vostiar et al., 2003) measuring DnaK and (Nemecek et al., 2008) measuring GFP. Alternatively, mRNA from *dnaK* and *groE* genes could be quantified using microarrays (Wade et al., 2006). Such experiments could help to determine the robustness of the reporters during production of other recombinant proteins than CCP. The design and use of reporters other than pDNAK and pGROE is also recommended. Microarray data revealed that the small heat-shock protein IbpA is activated during misfolding stress making its promoter an excellent candidate for a potentially useful reporter system for bioprocess (Smith, 2007).

It is yet to be confirmed whether gradients generated in large-scale fermentations could result in population segregation in terms of chaperone expression during recombinant protein production, similar to that shown in RpoS responses by (Delvigne et al., 2009). Further experiments in scale down reactors models (Hewitt et al., 2007) may deliver interesting data.

Chapter Six

General conclusions

Findings

Reporter gene technology was successfully utilized to monitor the activation of the heat-shock response; either through σ^{32} presence (pQF50kgroE) or activation of the promoters of cellular chaperones DnaK and GroES (via pDNAK and pGROE respectively). The different *E. coli* strains utilized [Table 2.1] showed increased levels of reporter protein in situations of temperature increase, thiol presence, anaerobic conditions, and recombinant protein production.

Measurement of β -galactosidase activity was shown to be a valid method to determine the σ^{32} -mediated stress response but industrial use would benefit from the development of an automated system for β -galactosidase monitoring. Alternatively, the utilization of green fluorescent protein as a reporter protein enabled direct observation of reporter expression in intact cells. Multi-parameter flow cytometry was demonstrated to be advantageous over bulk measurements of relative fluorescence because it enabled visualization of different GFP expression profiles represented in the distribution of green fluorescence histograms. In addition it also delivered useful information about the bacterial population such as cell morphology by light scatter and the presence of dead populations via PI staining.

Limitations of research

Possible differences in GFPmut2 expression during growth could have been obscured due to the high stability of GFPmut2 (>7 days) (Blokpoel et al., 2003). This could cause

interference in the monitoring of expression and repression of the reporter gene (*dnaK::gfpmut2* or *groE::gfpmut2*) as, in the absence of stress, several division cycles are required to observe sufficient reduction of intracellular fluorescence (Roostalu et al., 2008) therefore in non-growing or slow growing cultures, the reporter response could become masked due to already high intracellular levels of GFP. The development and use of new unstable GFP variants, as shown by Blokpoel et al. (2003) and Andersen et al. (1998) could improve the quality of the data of bioprocess monitoring. Another obstacle in utilizing these reporters is the inability of GFP to mature in the absence of oxygen. This situation can be reversed once oxygen is made available to the sample during its analysis. In the case of anaerobic fermentations monitored with on-line fluorimetry probes, GFP use is not advised, in such case other fluorescent proteins (such as the FMN binding fluorescent protein) (Tielker et al., 2009) are recommended.

Although the same general patterns of stress were observed in the different strains used in this study [**Table 2.1**], it was observed variability between strains. This variability has to be considered if an existing reporter plasmid is to be used in new strain as exactly similar fluorescent intensities should not be expected. It was shown that such differences may appear if one of the strains is more resistant, to the stress intended to measure than the other one.

Implications in bioprocess

In order to obtain the best results in terms of efficiency and product quality during bioprocesses, there are three elements that should function together: the microorganism, the bioreactor and the production strategy or protocol. The response of the microorganism to the

bioreactor and production conditions is the result of the integration of a vast number of different physiological variables. Gene reporter technology, as presented in this study, enabled the observation at population level of microbial physiology through one single variable, which is promoter activation of the reporter gene. It is possible that the measurements of this variable could be affected by other events in relation to the cell physiology during growth. For this reason, monitoring the heat-shock response to measure stress caused by recombinant protein production needs further research. Implementation of gene reporter technology as an extended practice in bioprocess remains a long term goal. In the short term it shows great potential for becoming a powerful tool in process development, which could contribute to the design of low stress production strategies ultimately contributing to the achievement of higher yields of good quality recombinant protein product.

Future directions of the gene reporter technology in bioindustry

The future of this technology relies upon the design of improved reporter systems, which coupled with flow cytometry, may offer great opportunities in bioprocess development. These reporters can be engineered to monitor multiple aspects of the stress response simultaneously. Dual reporters have been already proposed to measure genotoxicity and oxidative damage at the same time in single cells (Mitchell and Gu, 2004). Multi-parameter flow cytometry becomes the recommended technique as it can deliver reliable measurements multiplexing of different coloured reporters in addition to many other physiological parameters such as filamentation, cell size and shape or viability aided by fluorogenic stains. Since a flow injector can be successfully coupled with flow cytometry (Zhao et al, 1999) the analysis could be performed on-line generating large amounts of data.

GFP production by low copy number reporter plasmid has shown not to cause interference with recombinant protein production. Therefore, in spite of the difficulty to incorporate this type of monitoring in industry due to the regulatory framework, gene reporter technology coupled with flow cytometry can be of great assistance to monitor microbial physiology in “test-runs” during the stages of process development.

References

- Abbruzzetti, S., Grandi, E., Viappiani, C., Bologna, S., Campanini, B., Raboni, S., Bettati, S., and Mozzarelli, A. (2005). Kinetics of acid-induced spectral changes in the GFPmut2 chromophore. *Journal of the American Chemical Society* **127**(2), 626-635.
- Ades, S. E. (2008). Regulation by destruction: design of the [sigma]E envelope stress response. *Current Opinion in Microbiology* **11**(6), 535-540.
- Ades, S. E., Connolly, L. E., Alba, B. M., and Gross, C. A. (1999). The *Escherichia coli* σ^E -dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-sigma factor. *Genes & Development* **13**(18), 2449-2461.
- Aertsen, A., Vanoirbeek, K., De Spiegeleer, P., Sermon, J., Hauben, K., Farewell, A., Nystrom, T., and Michiels, C. W. (2004). Heat shock protein-mediated resistance to high hydrostatic pressure in *Escherichia coli*. *Applied Environ Microbiol* **70**(5), 2660-6.
- Åkesson, M., Hagander, P., and Axelsson, J. P. (2001). Avoiding acetate accumulation in *Escherichia coli* cultures using feedback control of glucose feeding. *Biotechnology and Bioengineering* **73**(3), 223-230.
- Albano, C. R., Randers-Eichhorn, L., Bentley, W. E., and Rao, G. (1998). Green fluorescent protein as a real time quantitative reporter of heterologous protein production. *Biotechnology Progress* **14**(2), 351-354.
- Alexeeva, S., Hellingwerf, K. J., and de Mattos, M. J. T. (2003). Requirement of ArcA for redox regulation in *Escherichia coli* under microaerobic but not anaerobic or aerobic conditions. *Journal of Bacteriology* **185**(1), 204-209.
- Allen, S. P., Polazzi, J. O., Gierse, J. K., and Easton, A. M. (1992). Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli*. *Journal of Bacteriology* **174**(21), 6938-47.
- Alm, E., and Baker, D. (1999). Matching theory and experiment in protein folding. *Current Opinion in Structural Biology* **9**(2), 189-96.
- Alphen, W. V., and Lugtenberg, B. (1977). Influence of osmolarity of the growth medium on the outer membrane protein pattern of *Escherichia coli*. *Journal of Bacteriology* **131**(2), 623-30.
- Andersen, D. C., and Reilly, D. E. (2004). Production technologies for monoclonal antibodies and their fragments. *Current Opinion in Biotechnology* **15**(5), 456-462.
- Andersen, J. B., Sternberg, C., Poulsen, L. K., Bjorn, S. P., Givskov, M., and Molin, S. (1998). New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Applied Environmental Microbiology* **64**(6), 2240-6.
- Andrews, B. T., Schoenfish, A. R., Roy, M., Waldo, G., and Jennings, P. A. (2007). The Rough Energy Landscape of Superfolder GFP Is Linked to the Chromophore. *Journal of Molecular Biology* **373**(2), 476-490.
- Anfinsen, C. B. (1972). The formation and stabilization of protein structure. *Biochemical Journal* **128**(4), 737-49.
- Anfinsen, C. B. (1973). Principles That Govern Folding of Protein Chains. *Science* **181**(4096), 223-230.
- Anfinsen, C. B., Haber, E., Sela, M., and White, F. H., Jr. (1961). The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proceedings of the National Academy of Sciences of the United States of America* **47**, 1309-14.
- Apetri, A. C., and Horwich, A. L. (2008). Chaperonin chamber accelerates protein folding through passive action of preventing aggregation. *Proceedings of the National Academy of Sciences of the United States of America* **105**(45), 17351-17355.
- Arsene, F., Tomoyasu, T., and Bukau, B. (2000). The heat shock response of *Escherichia coli*. *International Journal of Food Microbiol* **55**(1-3), 3-9.

- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular and Systems Biology* **2**, 2006 0008.
- Bahl, M. I., Sorensen, S. J., and Hestbjerg Hansen, L. (2004). Quantification of plasmid loss in *Escherichia coli* cells by use of flow cytometry. *FEMS Microbiology Letters* **232**(1), 45-9.
- Baker, D. (2000). A surprising simplicity to protein folding. *Nature* **405**(6782), 39-42.
- Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology* **10**(5), 411-21.
- Baneyx, F., and Mujacic, M. (2004). Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology* **22**(11), 1399-408.
- Barondeau, D. P., Tainer, J. A., and Getzoff, E. D. (2006). Structural evidence for an enolate intermediate in GFP fluorophore biosynthesis. *Journal of the American Chemical Society* **128**(10), 3166-8.
- Bartolomé, B., Jubete, Y., Martínez, E., and de la Cruz, F. (1991). Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene* **102**(1), 75-78.
- Batchelor, E., Walthers, D., Kenney, L. J., and Goulian, M. (2005). The *Escherichia coli* CpxA-CpxR envelope stress response system regulates expression of the porins OmpF and OmpC. *Journal of Bacteriology* **187**(16), 5723-31.
- Becker, G. W., and Hsiung, H. M. (1986). Expression, secretion and folding of human growth hormone in *Escherichia coli*. Purification and characterization. *Febs Letters* **204**(1), 145-50.
- Benito, A., Valero, F., Lafuente, J., Vidal, M., Cairo, J., Sola, C., and Villaverde, A. (1993). Uses of beta-galactosidase tag in on-line monitoring production of fusion proteins and gene expression in *Escherichia coli*. *Enzyme Microb Technol* **15**(1), 66-71.
- Bentley, W. E., Mirjalili, N., Andersen, D. C., Davis, R. H., and Kompala, D. S. (1990). Plasmid-Encoded Protein - the Principal Factor in the Metabolic Burden Associated with Recombinant Bacteria. *Biotechnology and Bioengineering* **35**(7), 668-681.
- Benz, I., and Schmidt, M. A. (2002). Never say never again: protein glycosylation in pathogenic bacteria. *Molecular Microbiology* **45**(2), 267-276.
- Bertelsen, E. B., Chang, L., Gestwicki, J. E., and Zuiderweg, E. R. P. (2009). Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proceedings of the National Academy of Sciences of the United States of America* **106**(21), 8471-8476.
- Bessette, P. H., Aslund, F., Beckwith, J., and Georgiou, G. (1999). Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc Natl Acad Sci U S A* **96**(24), 13703-8.
- Blackwell, J. R., and Horgan, R. (1991). A novel strategy for production of a highly expressed recombinant protein in an active form. *Febs Letters* **295**(1-3), 10-2.
- Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., ColladoVides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**(5331), 1453-62.
- Blokpoel, M. C., O'Toole, R., Smeulders, M. J., and Williams, H. D. (2003). Development and application of unstable GFP variants to kinetic studies of mycobacterial gene expression. *Journal of Microbiol Methods* **54**(2), 203-11.
- Boor, K. J. (2006). Bacterial Stress Responses: What Doesn't Kill Them Can Make Them Stronger. *PLoS Biol* **4**(1), e23.

- Botterman, J., and Zabeau, M. (1985). High-Level Production of the EcorI Endonuclease under the Control of the P1 Promoter of Bacteriophage-Lambda. *Gene* **37**(1-3), 229-239.
- Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., and Sigler, P. B. (1994). The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* **371**(6498), 578-86.
- Brockwell, D. J., and Radford, S. E. (2007). Intermediates: ubiquitous species on folding energy landscapes? *Current Opinion in Structural Biology* **17**(1), 30-37.
- Bronstein, I., Martin, C. S., Fortin, J. J., Olesen, C. E., and Voyta, J. C. (1996). Chemiluminescence: sensitive detection technology for reporter gene assays. *Clinical Chemistry* **42**(9), 1542-6.
- Bukau, B., and Walker, G. C. (1989). Cellular defects caused by deletion of the *Escherichia coli* dnaK gene indicate roles for heat shock protein in normal metabolism. *Journal of Bacteriology* **171**(5), 2337-46.
- Busby, S., and Ebright, R. H. (1994). Promoter Structure, Promoter Recognition, and Transcription Activation in Prokaryotes. *Cell* **79**(5), 743-746.
- Busby, S., and Ebright, R. H. (1999). Transcription activation by catabolite activator protein (CAP). *Journal of Molecular Biology* **293**(2), 199-213.
- Busby, S. J. W., and Savery, N. J. (2001). Transcription Activation at Bacterial Promoters. In "eLS". John Wiley & Sons, Ltd.
- Cabilly, S. (1989). Growth at sub-optimal temperatures allows the production of functional, antigen-binding Fab fragments in *Escherichia coli*. *Gene* **85**(2), 553-7.
- Cai, S. J., and Inouye, M. (2002). EnvZ-OmpR interaction and osmoregulation in *Escherichia coli*. *Journal of Biological Chemistry* **277**(27), 24155-61.
- Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., Trester-Zedlitz, M. L., and Darst, S. A. (2002a). Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Molecular Cell* **9**(3), 527-39.
- Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2002). A monomeric red fluorescent protein. *Proceedings of the National Academy of Science U S A* **99**(12), 7877-82.
- Carey, K. (2011). Biosimilars encircle Rituxan, US debates innovator exclusivity. *Nature Biotechnology* **29**(3), 177-8.
- Carrier, T. A., and Keasling, J. D. (1999). Library of synthetic 5' secondary structures to manipulate mRNA stability in *Escherichia coli*. *Biotechnology Progress* **15**(1), 58-64.
- Carrio, M. M., and Villaverde, A. (2001). Protein aggregation as bacterial inclusion bodies is reversible. *Febs Letters* **489**(1), 29-33.
- Casadaban, M. J., and Cohen, S. N. (1980). Analysis of Gene-Control Signals by DNA-Fusion and Cloning in *Escherichia-Coli*. *Journal of Molecular Biology* **138**(2), 179-207.
- Cha, H. J., Srivastava, R., Vakharia, V. M., Rao, G., and Bentley, W. E. (1999). Green fluorescent protein as a noninvasive stress probe in resting *Escherichia coli* cells. *Applied and Environmental Microbiology* **65**(2), 409-414.
- Chaba, R., Grigorova, I. L., Flynn, J. M., Baker, T. A., and Gross, C. A. (2007). Design principles of the proteolytic cascade governing the sigma(E)-mediated envelope stress response in *Escherichia coli*: keys to graded, buffered, and rapid signal transduction. *Genes & Development* **21**(1), 124-136.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**(5148), 802-5.

Chalova, V. I., Kim, W. K., Woodward, C. L., and Ricke, S. C. (2007). Quantification of total and bioavailable lysine in feed protein sources by a whole-cell green fluorescent protein growth-based *Escherichia coli* biosensor. *Applied Microbiology and Biotechnology* **76**(1), 91-9.

Chandrasekhar, G. N., Tilly, K., Woolford, C., Hendrix, R., and Georgopoulos, C. (1986). Purification and properties of the groES morphogenetic protein of *Escherichia coli*. *Journal of Biological Chemistry* **261**(26), 12414-9.

Chattopadhyay, R., and Roy, S. (2002). DnaK-Sigma 32 Interaction Is Temperature-dependent. *Journal of Biological Chemistry* **277**(37), 33641-33647.

Chaudhuri, R. R., Loman, N. J., Snyder, L. A. S., Bailey, C. M., Stekel, D. J., and Pallen, M. J. (2008). xBASE2: a comprehensive resource for comparative bacterial genomics. *Nucleic Acids Research* **36**(suppl 1), D543-D546.

Chilcott, G. S., and Hughes, K. T. (2000). Coupling of flagellar gene expression to flagellar assembly in *Salmonella enterica* serovar typhimurium and *Escherichia coli*. *Microbiol Mol Biol Rev* **64**(4), 694-708.

Chirico, G., Cannone, F., Beretta, S., Diaspro, A., Campanini, B., Bettati, S., Ruotolo, R., and Mozzarelli, A. (2002). Dynamics of green fluorescent protein mutant2 in solution, on spin-coated glasses, and encapsulated in wet silica gels. *Protein Science* **11**(5), 1152-61.

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.

Choi, J. H., and Lee, S. Y. (2004). Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Applied Microbiology and Biotechnology* **64**(5), 625-635.

Chou, C. (2007). Engineering cell physiology to enhance recombinant protein production in *Escherichia coli*. *Applied Microbiology and Biotechnology* **76**(3), 521-532.

Clarke, D. J., Calder, M. R., Carr, R. J. G., Blake-Coleman, B. C., Moody, S. C., and Collinge, T. A. (1985). The development and application of biosensing devices for bioreactor monitoring and control. *Biosensors* **1**(3), 213-320.

Clementsich, F., and Bayer, K. (2006). Improvement of bioprocess monitoring: development of novel concepts. *Microbial Cell Factories* **5**(19), 1-11

Cody, C. W., Prasher, D. C., Westler, W. M., Prendergast, F. G., and Ward, W. W. (1993). Chemical structure of the hexapeptide chromophore of the Aequorea green-fluorescent protein. *Biochemistry* **32**(5), 1212-8.

Connolly, L., Peñas, A. D. L., Alba, B. M., and Gross, C. A. (1997). The response to extracytoplasmic stress in *Escherichia coli* is controlled by partially overlapping pathways. *Genes & Development* **11**(15), 2012-2021.

Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**(1), 33-8.

Coucheney, F., Gal, L., Beney, L., Lherminier, J., Gervais, P., and Guzzo, J. (2005). A small HSP, Lo18, interacts with the cell membrane and modulates lipid physical state under heat shock conditions in a lactic acid bacterium. *Biochimica et Biophysica Acta* **1720**(1-2), 92-8.

Cousens, L. S., Shuster, J. R., Gallegos, C., Ku, L. L., Stempien, M. M., Urdea, M. S., Sanchez-Pescador, R., Taylor, A., and Tekamp-Olson, P. (1987). High level expression of proinsulin in the yeast, *Saccharomyces cerevisiae*. *Gene* **61**(3), 265-75.

Craig, E. A., and Gross, C. A. (1991). Is Hsp70 the cellular thermometer? *Trends in Biochemical Sciences* **16**, 135-140.

Creighton, T. E. (1988). Toward a Better Understanding of Protein Folding Pathways. *Proceedings of the National Academy of Sciences of the United States of America* **85**(14), 5082-5086.

- Crommelin, D. J. A., Storm, G., Verrijk, R., de Leede, L., Jiskoot, W., and Hennink, W. E. (2003). Shifting paradigms: biopharmaceuticals versus low molecular weight drugs. *International Journal of Pharmaceutics* **266**(1-2), 3-16.
- Danese, P. N., and Silhavy, T. J. (1998). CpxP, a stress-combative member of the Cpx regulon. *Journal of Bacteriology* **180**(4), 831-9.
- Dartigalongue, C., Missiakas, D., and Raina, S. (2001). Characterization of the *Escherichia coli* sigma(E) regulon. *Journal of Biological Chemistry* **276**(24), 20866-20875.
- Davey, H. M., and Kell, D. B. (1996). Flow cytometry and cell sorting of heterogeneous microbial populations: the importance of single-cell analyses. *Microbiol Reviews* **60**(4), 641-96.
- Davey, H. M., and Winson, M. K. (2003). Using flow cytometry to quantify microbial heterogeneity. *Current Issues in Molecular Biology* **5**(1), 9-15.
- De Groot, A. S., and Scott, D. W. (2007). Immunogenicity of protein therapeutics. *Trends in Immunology* **28**(11), 482-490.
- de Lorenzo, V., and Danchin, A. (2008). Synthetic biology: discovering new worlds and new words. *EMBO Reports* **9**(9), 822-7.
- de Marco, A., and De Marco, V. (2004). Bacteria co-transformed with recombinant proteins and chaperones cloned in independent plasmids are suitable for expression tuning. *Journal of Biotechnology* **109**(1-2), 45-52.
- De Wulf, P., McGuire, A. M., Liu, X., and Lin, E. C. C. (2002). Genome-wide Profiling of Promoter Recognition by the Two-component Response Regulator CpxR-P in *Escherichia coli*. *Journal of Biological Chemistry* **277**(29), 26652-26661.
- Dean, P. N., Bagwell, C. B., Lindmo, T., Murphy, R. F., and Salzman, G. C. (1990). Introduction to flow cytometry data file standard. *Cytometry* **11**(3), 321-2.
- Deepthike, H. U., Tecon, R., Van Kooten, G., Van der Meer, J. R., Harms, H., Wells, M., and Short, J. (2009). Unlike PAHs from Exxon Valdez crude oil, PAHs from Gulf of Alaska coals are not readily bioavailable. *Environmental Science & Technology* **43**(15), 5864-70.
- del Solar, G., Giraldo, R., Ruiz-Echevarria, M. J., Espinosa, M., and Diaz-Orejas, R. (1998). Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**(2), 434-64.
- DeLasPenas, A., Connolly, L., and Gross, C. A. (1997). The sigma(E)-mediated response to extracytoplasmic stress in *Escherichia coli* is transduced by RseA and RseB, two negative regulators of sigma(E). *Molecular Microbiology* **24**(2), 373-385.
- DeLisa, M. P., Li, J. C., Rao, G., Weigand, W. A., and Bentley, W. E. (1999). Monitoring GFP-operon fusion protein expression during high cell density cultivation of *Escherichia coli* using an on-line optical sensor. *Biotechnology and Bioengineering* **65**(1), 54-64.
- Delvigne, F., Boxus, M., Ingels, S., and Thonart, P. (2009). Bioreactor mixing efficiency modulates the activity of a prpO::GFP reporter gene in *E. coli*. *Microbial Cell Factories* **8**, 15.
- Deng, L., Vysotski, E. S., Markova, S. V., Liu, Z.-J., Lee, J., Rose, J., and Wang, B.-C. (2005). All three Ca²⁺-binding loops of photoproteins bind calcium ions: The crystal structures of calcium-loaded apo-aequorin and apo-obelin. *Protein Science* **14**(3), 663-675.
- Diamant, S., Ben-Zvi, A. P., Bukau, B., and Goloubinoff, P. (2000). Size-dependent disaggregation of stable protein aggregates by the DnaK chaperone machinery. *Journal of Biological Chemistry* **275**(28), 21107-21113.
- Diamond, L. W., Nathwani, B. N., and Rappaport, H. (1982). Flow cytometry in the diagnosis and classification of malignant lymphoma and leukemia. *Cancer* **50**(6), 1122-35.

- Diaz-Acosta, A., Sandoval, M. L., Delgado-Olivares, L., and Membrillo-Hernandez, J. (2006). Effect of anaerobic and stationary phase growth conditions on the heat shock and oxidative stress responses in *Escherichia coli* K-12. *Archives of Microbiology* **185**(6), 429-38.
- Díaz, M., Herrero, M., García, L. A., and Quirós, C. (2010). Application of flow cytometry to industrial microbial bioprocesses. *Biochemical Engineering Journal* **48**(3), 385-407.
- Dietz, H. (2004). Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proceedings of the National Academy of Sciences* **101**(46), 16192-16197.
- DiGiuseppe, P. A., and Silhavy, T. J. (2003). Signal detection and target gene induction by the CpxRA two-component system. *Journal Bacteriol* **185**(8), 2432-40.
- Dill, K. A., and Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nature Structural Biology* **4**(1), 10-19.
- Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M., and Karplus, M. (2000). Understanding protein folding via free-energy surfaces from theory and experiment. *Trends in Biochemical Sciences* **25**(7), 331-339.
- Dombroski, A. J., Johnson, B. D., Lonetto, M., and Gross, C. A. (1996). The sigma subunit of *Escherichia coli* RNA polymerase senses promoter spacing. *Proceedings of the National Academy of Sciences of the United States of America* **93**(17), 8858-8862.
- Dombroski, A. J., Walter, W. A., and Gross, C. A. (1993). Amino-Terminal Amino-Acids Modulate Sigma-Factor DNA-Binding Activity. *Genes & Development* **7**(12A), 2446-2455.
- Doyle, S. M., and Wickner, S. (2009). Hsp104 and ClpB: protein disaggregating machines. *Trends in Biochemical Sciences* **34**(1), 40-48.
- Drepper, T., Huber, R., Heck, A., Circolone, F., Hillmer, A. K., Buchs, J., and Jaeger, K. E. (2010). Flavin mononucleotide-based fluorescent reporter proteins outperform green fluorescent protein-like proteins as quantitative in vivo real-time reporters. *Applied Environ Microbiol* **76**(17), 5990-4.
- Durr, C., Nothaft, H., Lizak, C., Glockshuber, R., and Aepli, M. (2010). The *Escherichia coli* glycoprotein display system. *Glycobiology* **20**(11), 1366-1372.
- Ebright, R. H. (2000). RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *Journal of Molecular Biology* **304**(5), 687-98.
- Ebright, R. H., Ebright, Y. W., and Gunasekera, A. (1989). Consensus DNA site for the *Escherichia coli* catabolite gene activator protein (CAP): CAP exhibits a 450-fold higher affinity for the consensus DNA site than for the *E. coli* lac DNA site. *Nucleic Acids Research* **17**(24), 10295-305.
- Ellis, R. J. (1993). The general concept of molecular chaperones. *Philosophical Transactions of the Royal Society B: Biological Sciences* **339**(1289), 257-61.
- Ellis, R. J. (2006). Molecular chaperones: assisting assembly in addition to folding. *Trends in Biochemical Sciences* **31**(7), 395-401.
- Ellis, R. J. (2007). Protein misassembly: macromolecular crowding and molecular chaperones. *Advances in Experimental Medicine and Biology* **594**, 1-13.
- Ellis, R. J., and Minton, A. P. (2003). Cell biology - Join the crowd. *Nature* **425**(6953), 27-28.
- Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic Gene Expression in a Single Cell. *Science* **297**(5584), 1183-1186.
- Elowitz, M. B., Surette, M. G., Wolf, P. E., Stock, J., and Leibler, S. (1997). Photoactivation turns green fluorescent protein red. *Current Biology* **7**(10), 809-12.

- Elowitz, M. B., Surette, M. G., Wolf, P. E., Stock, J. B., and Leibler, S. (1999). Protein mobility in the cytoplasm of *Escherichia coli*. *Journal of Bacteriology* **181**(1), 197-203.
- Enz, S., Braun, V., and Crosa, J. H. (1995). Transcription of the region encoding the ferric dicitrate-transport system in *Escherichia coli*: similarity between promoters for fecA and for extracytoplasmic function sigma factors. *Gene* **163**(1), 13-18.
- Erickson, J. W., and Gross, C. A. (1989). Identification of the sigma E subunit of *Escherichia coli* RNA polymerase: a second alternate sigma factor involved in high-temperature gene expression. *Genes & Development* **3**(9), 1462-71.
- Escher, A., O'Kane, D. J., Lee, J., and Szalay, A. A. (1989). Bacterial luciferase alpha beta fusion protein is fully active as a monomer and highly sensitive in vivo to elevated temperature. *Proceedings of the National Academy of Sciences of the United States of America* **86**(17), 6528-32.
- Eyles, S. J., and Gierasch, L. M. (2010). Nature's molecular sponges: Small heat shock proteins grow into their chaperone roles. *Proceedings of the National Academy of Sciences of the United States of America* **107**(7), 2727-2728.
- Ezzell, C. (1987a). Genentech lose Protropin case. *Nature* **329**(6138), 379.
- Ezzell, C. (1987b). Protropin status questioned by FDA decision. *Nature* **326**(6110), 231.
- Farewell, A., Kvint, K., and Nystrom, T. (1998). Negative regulation by RpoS: a case of sigma factor competition. *Molecular Microbiology* **29**(4), 1039-51.
- Farinha, M. A., and Kropinski, A. M. (1990). Construction of broad-host-range plasmid vectors for easy visible selection and analysis of promoters. *Journal Bacteriol* **172**(6), 3496-9.
- Fedorov, A. N., and Baldwin, T. O. (1997). Cotranslational protein folding. *Journal of Biological Chemistry* **272**(52), 32715-32718.
- Ferrer-Miralles, N., Domingo-Espin, J., Corchero, J. L., Vazquez, E., and Villaverde, A. (2009). Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories* **8**, 17.
- Fink, A. L. (1998). Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Folding & Design* **3**(1), R9-R23.
- Fischer, B., Sumner, I., and Goodenough, P. (1993). Isolation, Renaturation, and Formation of Disulfide Bonds of Eukaryotic Proteins Expressed in Escherichia-Coli as Inclusion-Bodies. *Biotechnology and Bioengineering* **41**(1), 3-13.
- Fujita, N., and Ishihama, A. (1987). Heat-shock induction of RNA polymerase sigma-32 synthesis in *Escherichia coli*: Transcriptional control and a multiple promoter system. *Molecular and General Genetics MGG* **210**(1), 10-15.
- Gaal, T., Ross, W., Estrem, S. T., Nguyen, L. H., Burgess, R. R., and Gourse, R. L. (2001). Promoter recognition and discrimination by EsigmaS RNA polymerase. *Molecular Microbiology* **42**(4), 939-54.
- Gaitanaris, G. A., Vysokanov, A., Hung, S. C., Gottesman, M. E., and Gragerov, A. (1994). Successive action of *Escherichia coli* chaperones in vivo. *Molecular Microbiology* **14**(5), 861-9.
- Gamer, J., Multhaup, G., Tomoyasu, T., McCarty, J. S., Rudiger, S., Schonfeld, H. J., Schirra, C., Bujard, H., and Bukau, B. (1996). A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the *Escherichia coli* heat shock transcription factor sigma32. *EMBO Journal* **15**(3), 607-17.
- Gao, R., and Stock, A. M. (2009). Biological Insights from Structures of Two-Component Proteins. *Annual Review of Microbiology* **63**(1), 133-154.

- Garcia, J. R., Cha, H. J., Rao, G., Marten, M. R., and Bentley, W. E. (2009). Microbial nar-GFP cell sensors reveal oxygen limitations in highly agitated and aerated laboratory-scale fermentors. *Microbial Cell Factories* **8**, 6.
- Gasol, J. M., Zweifel, U. L., Peters, F., Fuhrman, J. A., and Hagstrom, A. (1999). Significance of size and nucleic acid content heterogeneity as measured by flow cytometry in natural planktonic bacteria. *Applied Environmental Microbiology* **65**(10), 4475-83.
- Gelman, M. S., and Prives, J. M. (1996). Arrest of subunit folding and assembly of nicotinic acetylcholine receptors in cultured muscle cells by dithiothreitol. *Journal of Biological Chemistry* **271**(18), 10709-14.
- Georgellis, D., Kwon, O., and Lin, E. C. C. (2001). Quinones as the redox signal for the Arc two-component system of bacteria. *Science* **292**(5525), 2314-2316.
- Georgopoulos, C. P., Hendrix, R. W., Kaiser, A. D., and Wood, W. B. (1972). Role of the host cell in bacteriophage morphogenesis: effects of a bacterial mutation on T4 head assembly. *Nature New Biology* **239**(89), 38-41.
- Gill, R. T., Cha, H. J., Jain, A., Rao, G., and Bentley, W. E. (1998). Generating controlled reducing environments in aerobic recombinant *Escherichia coli* fermentations: effects on cell growth, oxygen uptake, heat shock protein expression, and in vivo CAT activity. *Biotechnology and Bioengineering* **59**(2), 248-59.
- Gill, R. T., DeLisa, M. P., Valdes, J. J., and Bentley, W. E. (2001). Genomic analysis of high-cell-density recombinant *Escherichia coli* fermentation and "cell conditioning" for improved recombinant protein yield. *Biotechnology and Bioengineering* **72**(1), 85-95.
- Givan, A. L. (1992) Flow Cytometry: First Principles (Second Edition). New York: Wiley Interscience.
- Gosset, G., Zhang, Z., Nayyar, S., Cuevas, W. A., and Saier, M. H., Jr. (2004). Transcriptome analysis of Crp-dependent catabolite control of gene expression in *Escherichia coli*. *Journal of Bacteriol* **186**(11), 3516-24.
- Gragerov, A., Nudler, E., Komissarova, N., Gaitanaris, G. A., Gottesman, M. E., and Nikiforov, V. (1992). Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **89**(21), 10341-4.
- Granade, T. C., Hehmann, M. F., and Artis, W. M. (1985). Monitoring of filamentous fungal growth by in situ microspectrophotometry, fragmented mycelium absorbance density, and ¹⁴C incorporation: alternatives to mycelial dry weight. *Applied Environmental Microbiology* **49**(1), 101-8.
- Green, J., Scott, C., and Guest, J. R. (2001). Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP. *Advances in Microbial Physiology* **44**, 1-34.
- Gruber, T. M., and Gross, C. A. (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual Review of Microbiology* **57**, 441-466.
- Guex, N., and Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**(15), 2714-23.
- Guisbert, E., Herman, C., Lu, C. Z., and Gross, C. A. (2004). A chaperone network controls the heat shock response in *E. coli*. *Genes & Development* **18**(22), 2812-21.
- Gupta, S. D., Lee, B. T., Camakaris, J., and Wu, H. C. (1995). Identification of cutC and cutF (nlpE) genes involved in copper tolerance in *Escherichia coli*. *Journal of Bacteriol* **177**(15), 4207-15.
- Hakkila, K., Maksimow, M., Karp, M., and Virta, M. (2002). Reporter Genes lucFF, luxCDABE, gfp, and dsred Have Different Characteristics in Whole-Cell Bacterial Sensors. *Analytical Biochemistry* **301**(2), 235-242.
- Hampsey, M. (2001). Omega meets its match. *Trends in Genetics* **17**(4), 190-191.

- Han, W., and Christen, P. (2003). Mechanism of the targeting action of DnaJ in the DnaK molecular chaperone system. *Journal of Biological Chemistry* **278**(21), 19038-43.
- Harris, B. Z., and Lim, W. A. (2001). Mechanism and role of PDZ domains in signaling complex assembly. *J Journal of Cell Science* **114**(Pt 18), 3219-31.
- Hartl, F. U., and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. *Nature Structural Molecular Biology* **16**(6), 574-581.
- Hartley, J. L. (2006). Cloning technologies for protein expression and purification. *Curr Opin Biotechnol* **17**(4), 359-66.
- Harvey, B. R., Georgiou, G., Hayhurst, A., Jeong, K. J., Iverson, B. L., and Rogers, G. K. (2004). Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from *Escherichia coli*-expressed libraries. *Proceedings of the National Academy of Sciences of the United States of America* **101**(25), 9193-9198.
- Hasegawa, J.-Y., Fujimoto, K., Swerts, B., Miyahara, T., and Nakatsuji, H. (2007). Excited states of GFP chromophore and active site studied by the SAC-CI method: Effect of protein-environment and mutations. *Journal of Computational Chemistry* **28**(15), 2443-2452.
- Hastings, J. W. (1983). Biological diversity, chemical mechanisms, and the evolutionary origins of bioluminescent systems. *Journal of Molecular Evolution* **19**(5), 309-21.
- Hastings, J. W., and Morin, J. G. (1969). Calcium-Triggered Light Emission in Renilla . A Unitary Biochemical Scheme for Coelenterate Bioluminescence. *Biochemical and Biophysical Research Communications* **37**(3), 493-&.
- Hawley, D. K., and McClure, W. R. (1983). Compilation and Analysis of Escherichia-Coli Promoter DNA-Sequences. *Nucleic Acids Research* **11**(8), 2237-2255.
- Head, J. F., Inouye, S., Teranishi, K., and Shimomura, O. (2000). The crystal structure of the photoprotein aequorin at 2.3 angstrom resolution. *Nature* **405**(6784), 372-376.
- Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995). Improved green fluorescence. *Nature* **373**(6516), 663-4.
- Heim, R., Prasher, D. C., and Tsien, R. Y. (1994). Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **91**(26), 12501-4.
- 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), 178-82.
- Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W., and Ellis, R. J. (1988). Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* **333**(6171), 330-4.
- Hengge-Aronis, R. (2002a). Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase. *Microbiology and Molecular Biology Reviews* **66**(3), 373-95, table of contents.
- Hengge-Aronis, R. (2002b). Stationary phase gene regulation: what makes an *Escherichia coli* promoter sigma(s)-selective? *Current Opinion in Microbiology* **5**(6), 591-595.
- Heumann, D., and Roger, T. (2002). Initial responses to endotoxins and Gram-negative bacteria. *Clinica Chimica Acta* **323**(1-2), 59-72.
- Hewitt, C. J., and Nebe-Von-Caron, G. (2001). An industrial application of multiparameter flow cytometry: assessment of cell physiological state and its application to the study of microbial fermentations. *Cytometry* **44**(3), 179-87.

- Hewitt, C. J., and Nebe-Von-Caron, G. (2004). The application of multi-parameter flow cytometry to monitor individual microbial cell physiological state. *Advances in Biochemical Engineering/Biotechnology* **89**, 197-223.
- Hewitt, C. J., Onyeaka, H., Lewis, G., Taylor, I. W., and Nienow, A. W. (2007). A comparison of high cell density fed-batch fermentations involving both induced and non-induced recombinant *Escherichia coli* under well-mixed small-scale and simulated poorly mixed large-scale conditions. *Biotechnology and Bioengineering* **96**(3), 495-505.
- Heyde, M., and Portalier, R. (1987). Regulation of major outer membrane porin proteins of *Escherichia coli* K 12 by pH. *Molecular and General Genetics* **208**(3), 511-7.
- Hirakawa, H., Inazumi, Y., Masaki, T., Hirata, T., and Yamaguchi, A. (2005). Indole induces the expression of multidrug exporter genes in *Escherichia coli*. *Molecular Microbiology* **55**(4), 1113-1126.
- Hirano, Y., Hossain, M. M., Takeda, K., Tokuda, H., and Miki, K. (2007). Structural studies of the Cpx pathway activator NlpE on the outer membrane of *Escherichia coli*. *Structure* **15**(8), 963-76.
- Hisiger, S., and Jolicoeur, M. (2005). A multiwavelength fluorescence probe: Is one probe capable for on-line monitoring of recombinant protein production and biomass activity? *Journal of Biotechnology* **117**(4), 325-336.
- Ho, B. K., Thomas, A., and Brasseur, R. (2003). Revisiting the Ramachandran plot: Hard-sphere repulsion, electrostatics, and H-bonding in the alpha-helix. *Protein Science* **12**(11), 2508-2522.
- Hoffmann, F., Heuvel, J. v. d., Zidek, N., and Rinas, U. (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.
- Hoffmann, F., and Rinas, U. (2001). On-line estimation of the metabolic burden resulting from the synthesis of plasmid-encoded and heat-shock proteins by monitoring respiratory energy generation. *Biotechnology and Bioengineering* **76**(4), 333-40.
- Hoffmann, F., and Rinas, U. (2004). Roles of heat-shock chaperones in the production of recombinant proteins in *Escherichia coli*. *Advances in Biochemical Engineering and Biotechnology* **89**, 143-61.
- Houry, W. A., Frishman, D., Eckerskorn, C., Lottspeich, F., and Hartl, F. U. (1999). Identification of in vivo substrates of the chaperonin GroEL. *Nature* **402**(6758), 147-54.
- Humphreys, D. P., Carrington, B., Bowering, L. C., Ganesh, R., Sehdev, M., Smith, B. J., King, L. M., Reeks, D. G., Lawson, A., and Popplewell, A. G. (2002). A plasmid system for optimization of Fab' production in *Escherichia coli*: importance of balance of heavy chain and light chain synthesis. *Protein Expression and Purification* **26**(2), 309-20.
- Inouye, S., Noguchi, M., Sakaki, Y., Takagi, Y., Miyata, T., Iwanaga, S., and Tsuji, F. I. (1985). Cloning and sequence analysis of cDNA for the luminescent protein aequorin. *Proceedings of the National Academy of Sciences of the United States of America* **82**(10), 3154-8.
- Isaac, D. D., Pinkner, J. S., Hultgren, S. J., and Silhavy, T. J. (2005). The extracytoplasmic adaptor protein CpxP is degraded with substrate by DegP. *Proceedings of the National Academy of Sciences of the United States of America* **102**(49), 17775-17779.
- Ishihama, A. (2000). Functional modulation of *Escherichia coli* RNA polymerase. *Annual Review of Microbiology* **54**, 499-518.
- Ishizuka, H., Hanamura, A., Inada, T., and Aiba, H. (1994). Mechanism of the down-Regulation of Camp Receptor Protein by Glucose in *Escherichia-Coli* - Role of Autoregulation of the Crp Gene. *EMBOJournal* **13**(13), 3077-3082.

- Jackson, S. E., Craggs, T. D., and Huang, J. R. (2006). Understanding the folding of GFP using biophysical techniques. *Expert Review of Proteomics* **3**(5), 545-59.
- Jaeger, J. A., Turner, D. H., and Zuker, M. (1990). Predicting optimal and suboptimal secondary structure for RNA. *Methods in Enzymology* **183**, 281-306.
- Jakobs, S., Schauss, A. C., and Hell, S. W. (2003). Photoconversion of matrix targeted GFP enables analysis of continuity and intermixing of the mitochondrial lumen. *FEBS Letters* **554**(1-2), 194-200.
- Janaszak, A., Majczak, W., Nadratowska, B., Szalewska-Palasz, A., Konopa, G., and Taylor, A. (2007a). A sigma54-dependent promoter in the regulatory region of the *Escherichia coli* rpoH gene. *Microbiology* **153**(Pt 1), 111-23.
- Janaszak, A., Majczak, W., Nadratowska, B., Szalewska-Palasz, A., Konopa, G., and Taylor, A. (2007b). A sigma54-dependent promoter in the regulatory region of the *Escherichia coli* rpoH gene. *Microbiology* **153**(1), 111-123.
- Janaszak, A., Nadratowska-Wesołowska, B., Konopa, G., and Taylor, A. (2009). The P1 promoter of the *Escherichia coli* rpoH gene is utilized by sigma70-RNAP or sigmaS-RNAP depending on growth phase. *FEMS Microbiology Letters* **291**(1), 65-72.
- Janetopoulos, C., Jin, T., and Devreotes, P. (2001). Receptor-mediated activation of heterotrimeric G-proteins in living cells. *Science* **291**(5512), 2408-11.
- Jayaraman, P. S., Peakman, T. C., Busby, S. J. W., Quincey, R. V., and Cole, J. A. (1987). Location and sequence of the promoter of the gene for the NADH-dependent nitrite reductase of *Escherichia coli* and its regulation by oxygen, the Fnr protein and nitrite. *Journal of Molecular Biology* **196**(4), 781-788.
- Jefferis, R. (2009). Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action. *Trends in Pharmacological Science United States of America* **30**(7), 356-62.
- Jefferson, R. A., Burgess, S. M., and Hirsh, D. (1986). beta-Glucuronidase from *Escherichia coli* as a gene-fusion marker. *Proceedings of the National Academy of Sciences United States of America* **83**(22), 8447-8451.
- Jishage, M., and Ishihama, A. (1998). A stationary phase protein in *Escherichia coli* with binding activity to the major sigma subunit of RNA polymerase. *Proceedings of the National Academy of Sciences United States of America* **95**(9), 4953-8.
- Jishage, M., Iwata, A., Ueda, S., and Ishihama, A. (1996). Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: Intracellular levels of four species of sigma subunit under various growth conditions. *Journal of Bacteriology* **178**(18), 5447-5451.
- Johnson, I. S. (1983). Human insulin from recombinant DNA technology. *Science* **219**(4585), 632-7.
- Jones, C. H., Danese, P. N., Pinkner, J. S., Silhavy, T. J., and Hultgren, S. J. (1997). The chaperone-assisted membrane release and folding pathway is sensed by two signal transduction systems. *EMBO Journal* **16**(21), 6394-6406.
- Jubelin, G., Vianney, A., Beloin, C., Ghigo, J. M., Lazzaroni, J. C., Lejeune, P., and Dorel, C. (2005). CpxR/OmpR interplay regulates curli gene expression in response to osmolarity in *Escherichia coli*. *Journal of Bacteriology* **187**(6), 2038-2049.
- Jung, W. K., Koo, H. C., Kim, K. W., Shin, S., Kim, S. H., and Park, Y. H. (2008). Antibacterial activity and mechanism of action of the silver ion in *Staphylococcus aureus* and *Escherichia coli*. *Appl Environ Microbiol* **74**(7), 2171-8.
- Kadokura, H., Katzen, F., and Beckwith, J. (2003). Protein disulfide bond formation in prokaryotes. *Annual Review of Biochemistry* **72**, 111-135.

- Kallipolitis, B. H., and Valentin-Hansen, P. (1998). Transcription of *rpoH*, encoding the *Escherichia coli* heat-shock regulator sigma(32), is negatively controlled by the cAMP-CRP/CytR nucleoprotein complex. *Molecular Microbiology* **29**(4), 1091-1099.
- Kamath-Loeb, A. S., and Gross, C. A. (1991). Translational regulation of sigma 32 synthesis: requirement for an internal control element. *Journal of Bacteriology* **173**(12), 3904-6.
- Kane, J. F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Current Opinion in Biotechnology* **6**(5), 494-500.
- Kanehara, K., Ito, K., and Akiyama, Y. (2002). YaeL (EcfE) activates the sigma(E) pathway of stress response through a site-2 cleavage of anti-sigma(E), RseA. *Genes & Development* **16**(16), 2147-2155.
- Karplus, M. (1997). The Levinthal paradox: yesterday and today. *Folding & Design* **2**(4), S69-S75.
- Keiler, K. C., Waller, P. R., and Sauer, R. T. (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**(5251), 990-3.
- Kenney, L. J. (2010). How important is the phosphatase activity of sensor kinases? *Current Opinion in Microbiology* **13**(2), 168-176.
- Kerner, M. J., Naylor, D. J., Ishihama, Y., Maier, T., Chang, H. C., Stines, A. P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M., and Hartl, F. U. (2005). Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* **122**(2), 209-20.
- Keseler, I. M., Collado-Vides, J., Gama-Castro, S., Ingraham, J., Paley, S., Paulsen, I. T., Peralta-Gill, M., and Karp, P. D. (2005). EcoCyc: a comprehensive database resource for *Escherichia coli*. *Nucleic Acids Research* **33**, D334-D337.
- Keseler, I. M., Collado-Vides, J., Santos-Zavaleta, A., Peralta-Gil, M., Gama-Castro, S., Muniz-Rascado, L., Bonavides-Martinez, C., Paley, S., Krummenacker, M., Altman, T., Kaipa, P., Spaulding, A., Pacheco, J., Latendresse, M., Fulcher, C., Sarker, M., Shearer, A. G., Mackie, A., Paulsen, I., Gunsalus, R. P., and Karp, P. D. (2011). EcoCyc: a comprehensive database of *Escherichia coli* biology. *Nucleic Acids Research* **39**(Database issue), D583-90.
- Khoroshilova, N., Popescu, C., Munck, E., Beinert, H., and Kiley, P. J. (1997). Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O₂: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proceedings of the National Academy of Sciences United States of America* **94**(12), 6087-92.
- Kim, D. E., Gu, H. D., and Baker, D. (1998). The sequences of small proteins are not extensively optimized for rapid folding by natural selection. *Proceedings of the National Academy of Sciences of the United States of America* **95**(9), 4982-4986.
- Kipriyanov, S. M., Moldenhauer, G., and Little, M. (1997). High level production of soluble single chain antibodies in small-scale *Escherichia coli* cultures. *Journal of Immunological Methods* **200**(1-2), 69-77.
- Kitagawa, M., Matsumura, Y., and Tsuchido, T. (2000). Small heat shock proteins, IbpA and IbpB, are involved in resistances to heat and superoxide stresses in *Escherichia coli*. *FEMS Microbiol Lett* **184**(2), 165-71.
- Kleywegt, G. J., and Jones, T. A. (1996). Phi/psi-chology: Ramachandran revisited. *Structure* **4**(12), 1395-1400.
- Klinkert, B., and Narberhaus, F. (2009). Microbial thermosensors. *Cellular and Molecular Life Sciences* **66**(16), 2661-2676.
- Knapp, W., Strobl, H., and Majdic, O. (1994). Flow cytometric analysis of cell-surface and intracellular antigens in leukemia diagnosis. *Cytometry* **18**(4), 187-98.

- Konz, J. O., King, J., and Cooney, C. L. (1998). Effects of oxygen on recombinant protein expression. *Biotechnology Progress* **14**(3), 393-409.
- Kotelnikova, E. A., Makeev, V. J., and Gelfand, M. S. (2005). Evolution of transcription factor DNA binding sites. *Gene* **347**(2), 255-63.
- Kowarik, M., Numao, S., Feldman, M. F., Schulz, B. L., Callewaert, N., Kiermaier, E., Catrein, I., and Aebi, M. (2006). N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science* **314**(5802), 1148-1150.
- Krell, T., Lacal, J., Busch, A., Silva-Jimenez, H., Guazzaroni, M. E., and Ramos, J. L. (2010). Bacterial sensor kinases: diversity in the recognition of environmental signals. *Annual Review of Microbiology* **64**, 539-59.
- Kresse, G. B. (2009). Biosimilars - Science, status, and strategic perspective. *European Journal of Pharmaceutics and Biopharmaceutics* **72**(3), 479-486.
- Krojer, T., Sawa, J., Schafer, E., Saibil, H. R., Ehrmann, M., and Clausen, T. (2008). Structural basis for the regulated protease and chaperone function of DegP. *Nature* **453**(7197), 885-90.
- Kuczynska-Wisnik, D., Kedzierska, S., Matuszewska, E., Lund, P., Taylor, A., Lipinska, B., and Laskowska, E. (2002). The *Escherichia coli* small heat-shock proteins IbpA and IbpB prevent the aggregation of endogenous proteins denatured in vivo during extreme heat shock. *Microbiology-Society for General Microbiology* **148**, 1757-1765.
- Lange, R., and Hengge-Aronis, R. (1991). Identification of a central regulator of stationary-phase gene expression in *Escherichia coli*. *Molecular Microbiology* **5**(1), 49-59.
- Lange, R., and Hengge-Aronis, R. (1994a). The Cellular Concentration of the Sigma(S) Subunit of RNA-Polymerase in *Escherichia coli* Is Controlled at the Levels of Transcription Translation, and Protein Stability. *Genes & Development* **8**(13), 1600-1612.
- Lange, R., and Hengge-Aronis, R. (1994b). The NlpD Gene Is Located in an Operon with Rpos on the *Escherichia coli* Chromosome and Encodes a Novel Lipoprotein with a Potential Function in Cell-Wall Formation. *Molecular Microbiology* **13**(4), 733-743.
- Langley, K. E., Villarejo, M. R., Fowler, A. V., Zamenhof, P. J., and Zabin, I. (1975). Molecular basis of beta-galactosidase alpha-complementation. *Proceedings of the National Academy of Sciences United States of America* **72**(4), 1254-7.
- Laskowska, E., Wawrzynow, A., and Taylor, A. (1996). IbpA and IbpB, the new heat-shock proteins, bind to endogenous *Escherichia coli* proteins aggregated intracellularly by heat shock. *Biochimie* **78**(2), 117-122.
- Lau-Wong, I. C., Locke, T., Ellison, M. J., Raivio, T. L., and Frost, L. S. (2008). Activation of the Cpx regulon destabilizes the F plasmid transfer activator, TraJ, via the HslVU protease in *Escherichia coli*. *Molecular Microbiology* **67**(3), 516-527.
- Lederberg, J. (1950). The beta-d-galactosidase of *Escherichia coli*, strain K-12. *J Bacteriol* **60**(4), 381-92.
- Lederberg, J. (2004). *E. coli* K-12. *Microbiology Today* **31**, 116.
- Lee, S., Sowa, M. E., Watanabe, Y.-h., Sigler, P. B., Chiu, W., Yoshida, M., and Tsai, F. T. F. (2003a). The Structure of ClpB: A Molecular Chaperone that Rescues Proteins from an Aggregated State. *Cell* **115**(2), 229-240.
- Lee, S., Sowa, M. E., Watanabe, Y. H., Sigler, P. B., Chiu, W., Yoshida, M., and Tsai, F. T. F. (2003b). The structure of clpB: A molecular chaperone that rescues proteins from an aggregated state. *Cell* **115**(2), 229-240.
- Lehmann, K., Hoffmann, S., Neudecker, P., Suhr, M., Becker, W. M., and Rosch, P. (2003). High-yield expression in *Escherichia coli*, purification, and characterization of properly folded major peanut allergen Ara h 2. *Protein Expression and Purification* **31**(2), 250-9.

- 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* **60A**(2), 165-172.
- Leong, S. R., DeForge, L., Presta, L., Gonzalez, T., Fan, A., Reichert, M., Chuntharapai, A., Kim, K. J., Tumas, D. B., Lee, W. P., Gribbling, P., Snedecor, B., Chen, H., Hsei, V., Schoenhoff, M., Hale, V., Deveney, J., Koumenis, I., Shahrokh, Z., McKay, P., Galan, W., Wagner, B., Narindray, D., Hebert, C., and Zapata, G. (2001). Adapting pharmacokinetic properties of a humanized anti-interleukin-8 antibody for therapeutic applications using site-specific pegylation. *Cytokine* **16**(3), 106-119.
- Levchenko, I., Smith, C. K., Walsh, N. P., Sauer, R. T., and Baker, T. A. (1997). PDZ-like domains mediate binding specificity in the Clp/Hsp100 family of chaperones and protease regulatory subunits. *Cell* **91**(7), 939-47.
- Leveau, J. H., and Lindow, S. E. (2001). Predictive and interpretive simulation of green fluorescent protein expression in reporter bacteria. *J Bacteriol* **183**(23), 6752-62.
- Levinthal, C. (1968). Are There Pathways for Protein Folding. *Journal De Chimie Physique Et De Physico-Chimie Biologique* **65**(1), 44-45.
- Lewis, G., Taylor, I. W., Nienow, A. W., and Hewitt, C. J. (2004). The application of multi-parameter flow cytometry to the study of recombinant *Escherichia coli* batch fermentation processes. *Journal of Industrial Microbiology and Biotechnology* **31**(7), 311-22.
- Liberek, K., Galitski, T. P., Zylicz, M., and Georgopoulos, C. (1992). The DnaK chaperone modulates the heat shock response of *Escherichia coli* by binding to the sigma 32 transcription factor. *Proceedings of the National Academy of Sciences* **89**(8), 3516-3520.
- Lipinska, B., Fayet, O., Baird, L., and Georgopoulos, C. (1989). Identification, Characterization, and Mapping of the *Escherichia-Coli* Htra Gene, Whose Product Is Essential for Bacterial-Growth Only at Elevated-Temperatures. *Journal of Bacteriology* **171**(3), 1574-1584.
- Liu, X., and Ferenci, T. (1998). Regulation of porin-mediated outer membrane permeability by nutrient limitation in *Escherichia coli*. *Journal of Bacteriology* **180**(15), 3917-22.
- Lodish, H. F., and Kong, N. (1993). The secretory pathway is normal in dithiothreitol-treated cells, but disulfide-bonded proteins are reduced and reversibly retained in the endoplasmic reticulum. *Journal of Biological Chemistry* **268**(27), 20598-605.
- Lonetto, M., Gribskov, M., and Gross, C. A. (1992). The Sigma-70 Family - Sequence Conservation and Evolutionary Relationships. *Journal of Bacteriology* **174**(12), 3843-3849.
- Lopez-Amoros, R., Comas, J., and Vives-Rego, J. (1995). Flow cytometric assessment of *Escherichia coli* and *Salmonella typhimurium* starvation-survival in seawater using rhodamine 123, propidium iodide, and oxonol. *Applied and Environmental Microbiology* **61**(7), 2521-6.
- Lorenz, W. W., Cormier, M. J., O'Kane, D. J., Hua, D., Escher, A. A., and Szalay, A. A. (1996). Expression of the *Renilla reniformis* luciferase gene in mammalian cells. *Journal of Bioluminescence and Chemiluminescence* **11**(1), 31-7.
- Luo, Z.-H., and Hua, Z.-C. (1998). Increased solubility of glutathione S-transferase-p16 (GST-p16) fusion protein by co-expression of chaperones GroES and GroEL in *Escherichia coli*. *IUBMB Life* **46**(3), 471-477.
- Mackie, G. A. (1998). Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* **395**(6703), 720-723.
- Madigan, M. T. M., J M; Dunlap, P V; Clark, D P (2009). "Biology of Microorganisms." 12 ed. Brock Pearson Benjamin Cummings, San Francisco.
- Maeda, H., Fujita, N., and Ishihama, A. (2000). Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Research* **28**(18), 3497-3503.

- Makino, Y., Amada, K., Taguchi, H., and Yoshida, M. (1997). Chaperonin-mediated folding of green fluorescent protein. *Journal of Biological Chemistry* **272**(19), 12468-74.
- Malpica, R., Franco, B., Rodriguez, C., Kwon, O., and Georgellis, D. (2004). Identification of a quinone-sensitive redox switch in the ArcB sensor kinase. *Proceedings of the National Academy of Sciences of the United States of America* **101**(36), 13318-13323.
- Marincs, F. (2000). On-line monitoring of growth of *Escherichia coli* in batch cultures by bioluminescence. *Appl Microbiol Biotechnol* **53**(5), 536-41.
- Marston, F. A. O. (1986). The Purification of Eukaryotic Polypeptides Synthesized in *Escherichia-Coli*. *Biochemical Journal* **240**(1), 1-12.
- Martin, J. (1998). Protein folding assisted by the GroEL/GroES chaperonin system. *Biochemistry (Mosc)* **63**(4), 374-81.
- Martinez-Alonso, M., Garcia-Fruitos, E., Ferrer-Miralles, N., Rinas, U., and Villaverde, A. (2010). Side effects of chaperone gene co-expression in recombinant protein production. *Microbial Cell Factories* **9**, 64.
- Martinez-Antonio, A., and Collado-Vides, J. (2003). Identifying global regulators in transcriptional regulatory networks in bacteria. *Current Opinion in Microbiology* **6**(5), 482-9.
- Matin, A. (1990). Molecular analysis of the starvation stress in *Escherichia coli*. *FEMS Microbiology Ecology* **7**(2-3), 185-195.
- Matuszewska, E., Kwiatkowska, J., Kuczynska-Wisnik, D., and Laskowska, E. (2008). *Escherichia coli* heat-shock proteins IbpA/B are involved in resistance to oxidative stress induced by copper. *Microbiology* **154**(Pt 6), 1739-47.
- Matuszewska, M., Kuczynska-Wisnik, D., Laskowska, E., and Liberek, K. (2005). The small heat shock protein IbpA of *Escherichia coli* cooperates with IbpB in stabilization of thermally aggregated proteins in a disaggregation competent state. *Journal of Biological Chemistry* **280**(13), 12292-8.
- Mayerhofer, R., Langridge, W. H. R., Cormier, M. J., and Szalay, A. A. (1995). Expression of Recombinant Renilla Luciferase in Transgenic Plants Results in High-Levels of Light-Emission. *Plant Journal* **7**(6), 1031-1038.
- Mazor, Y., Van Blarcom, T., Iverson, B. L., and Georgiou, G. (2008). E-clonal antibodies: selection of full-length IgG antibodies using bacterial periplasmic display. *Nature Protocols* **3**(11), 1766-1777.
- McCarty, J. S., Buchberger, A., Reinstein, J., and Bukau, B. (1995). The role of ATP in the functional cycle of the DnaK chaperone system. *J Mol Biol* **249**(1), 126-37.
- McEwen, J., and Silverman, P. (1980). Chromosomal Mutations of *Escherichia-Coli* That Alter Expression of Conjugative Plasmid Functions. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **77**(1), 513-517.
- Meighen, E. A. (1988). Enzymes and Genes from the Lux Operons of Bioluminescent Bacteria. *Annual Review of Microbiology* **42**, 151-176.
- Mika, F., and Hengge, R. (2005). A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of sigma(S) (RpoS) in E-coli. *Genes & Development* **19**(22), 2770-2781.
- Mileykovskaya, E., and Dowhan, W. (1997). The Cpx two-component signal transduction pathway is activated in *Escherichia coli* mutant strains lacking phosphatidylethanolamine. *Journal of Bacteriology* **179**(4), 1029-1034.
- Miroux, B., and Walker, J. E. (1996). Over-production of Proteins in *Escherichia coli* : Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *Journal of Molecular Biology* **260**(3), 289-298.

- Mishin, A. S., Subach, F. V., Yampolsky, I. V., King, W., Lukyanov, K. A., and Verkhusha, V. V. (2008). The first mutant of the *Aequorea victoria* green fluorescent protein that forms a red chromophore. *Biochemistry* **47**(16), 4666-4673.
- Missiakas, D., Mayer, M. P., Lemaire, M., Georgopoulos, C., and Raina, S. (1997). Modulation of the *Escherichia coli* sigmaE (RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins. *Molecular Microbiology* **24**(2), 355-71.
- Mitchell, R. J., and Gu, M. B. (2004). An *Escherichia coli* biosensor capable of detecting both genotoxic and oxidative damage. *Applied Microbiology Biotechnology* **64**(1), 46-52.
- Miyawaki, A., Llopis, J., Heim, R., McCaffery, J. M., Adams, J. A., Ikura, M., and Tsien, R. Y. (1997). Fluorescent indicators for Ca²⁺ based on green fluorescent proteins and calmodulin. *Nature* **388**(6645), 882-7.
- Mizuno, T. (1997). Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of *Escherichia coli*. *DNA Research* **4**(2), 161-8.
- Mizushima, T., Natori, S., and Sekimizu, K. (1993). Relaxation of supercoiled DNA associated with induction of heat shock proteins in *Escherichia coli*. *Molecular and General Genetics* **238**(1-2), 1-5.
- Mogk, A., Tomoyasu, T., Goloubinoff, P., Rudiger, S., Roder, D., Langen, H., and Bukau, B. (1999). Identification of thermolabile *Escherichia coli* proteins: prevention and reversion of aggregation by DnaK and ClpB. *EMBO J* **18**(24), 6934-49.
- Moore, L. J., Mettert, E. L., and Kiley, P. J. (2006). Regulation of FNR dimerization by subunit charge repulsion. *Journal of Biological Chemistry* **281**(44), 33268-75.
- Morise, H., Shimomura, O., Johnson, F. H., and Winant, J. (1974). Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry* **13**(12), 2656-62.
- Morita, M., Kanemori, M., Yanagi, H., and Yura, T. (1999). Heat-induced synthesis of sigma(32) in *Escherichia coli*: Structural and functional dissection of rpoH mRNA secondary structure. *Journal of Bacteriology* **181**(2), 401-410.
- Morschhäuser, J., Michel, S., and Hacker, J. (1998). Expression of a chromosomally integrated, single-copy GFP gene in *Candida albicans*, and its use as a reporter of gene regulation. *Molecular and General Genetics MGG* **257**(4), 412-420.
- Muller, S., and Davey, H. (2009). Recent advances in the analysis of individual microbial cells. *Cytometry A* **75**(2), 83-5.
- Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., and Darst, S. A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**(5571), 1285-90.
- Murphy, R. F., and Chused, T. M. (1984). A proposal for a flow cytometric data file standard. *Cytometry* **5**(5), 553-5.
- Myasnikov, A. G., Simonetti, A., Marzi, S., and Klaholz, B. P. (2009). Structure-function insights into prokaryotic and eukaryotic translation initiation. *Current Opinion in Structural Biology* **19**(3), 300-309.
- Nakajima-Shimada, J., Iida, H., Tsuji, F. I., and Anraku, Y. (1991). Monitoring of intracellular calcium in *Saccharomyces cerevisiae* with an apoaquorin cDNA expression system. *Proceedings of the National Academy of Sciences of the United States of America* **88**(15), 6878-82.
- Nakamoto, H., and Vigh, L. (2007). The small heat shock proteins and their clients. *Cellular and Molecular Life Sciences* **64**(3), 294-306.
- Nakayama, S., and Watanabe, H. (1995). Involvement of cpxA, a sensor of a two-component regulatory system, in the pH-dependent regulation of expression of *Shigella sonnei* virF gene. *Journal of Bacteriology* **177**(17), 5062-9.

- Narayanan, N., and Chou, C. P. (2008). Periplasmic chaperone FkpA reduces extracytoplasmic stress response and improves cell-surface display on *Escherichia coli*. *Enzyme and Microbial Technology* **42**(6), 506-513.
- Naylor, L. H. (1999). Reporter gene technology: the future looks bright. *Biochemical Pharmacology* **58**(5), 749-757.
- Nemecek, S., Marisch, K., Juric, R., and Bayer, K. (2008a). Design of transcriptional fusions of stress sensitive promoters and GFP to monitor the overburden of *Escherichia coli* hosts during recombinant protein production. *Bioprocess and Biosystems Engineering* **31**(1), 47-53.
- Nishimura, A., Morita, M., Nishimura, Y., and Sugino, Y. (1990). A rapid and highly efficient method for preparation of competent *Escherichia coli* cells. *Nucleic Acids Research* **18**(20), 6169.
- Niswender, K. D., Blackman, S. M., Rohde, L., Magnuson, M. A., and Piston, D. W. (1995). Quantitative Imaging of Green Fluorescent Protein in Cultured-Cells - Comparison of Microscopic Techniques, Use in Fusion Proteins and Detection Limits. *Journal of Microscopy-Oxford* **180**, 109-116.
- Nonaka, G., Blankschien, M., Herman, C., Gross, C. A., and Rhodius, V. A. (2006). Regulon and promoter analysis of the *E. coli* heat-shock factor, sigma32, reveals a multifaceted cellular response to heat stress. *Genes & Development* **20**(13), 1776-89.
- Novick, R. P. (1987). Plasmid Incompatibility. *Microbiological Reviews* **51**(4), 381-395.
- Novo, D. J., Perlmutter, N. G., Hunt, R. H., and Shapiro, H. M. (2000). Multiparameter flow cytometric analysis of antibiotic effects on membrane potential, membrane permeability, and bacterial counts of *Staphylococcus aureus* and *Micrococcus luteus*. *Antimicrobial Agents and Chemotherapy* **44**(4), 827-34.
- Ogawa, T., Yamada, Y., Kuroda, T., Kishi, T., and Moriya, S. (2002). The *datA* locus predominantly contributes to the initiator titration mechanism in the control of replication initiation in *Escherichia coli*. *Molecular Microbiology* **44**(5), 1367-75.
- Ogata, Y., Mizushima, T., Kataoka, K., Kita, K., Miki, T., and Sekimizu, K. (1996). DnaK heat shock protein of *Escherichia coli* maintains the negative supercoiling of DNA against thermal stress. *Journal of Biological Chemistry* **271**(46), 29407-14.
- Oh, M. K., and Liao, J. C. (2000). Gene expression profiling by DNA microarrays and metabolic fluxes in *Escherichia coli*. *Biotechnology Progress* **16**(2), 278-86.
- Olson, K. C., Fenno, J., Lin, N., Harkins, R. N., Snider, C., Kohr, W. H., Ross, M. J., Fodge, D., Prender, G., and Stebbing, N. (1981). Purified Human Growth-Hormone from *Escherichia-Coli* Is Biologically-Active. *Nature* **293**(5831), 408-411.
- Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., and Remington, S. J. (1996). Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**(5280), 1392-5.
- Patkar, A., Vijayasankaran, N., Urry, D. W., and Srienc, F. (2002). Flow cytometry as a useful tool for process development: rapid evaluation of expression systems. *Journal of Biotechnology* **93**(3), 217-229.
- Patra, A. K., Mukhopadhyay, R., Mukhija, R., Krishnan, A., Garg, L. C., and Panda, A. K. (2000). Optimization of inclusion body solubilization and renaturation of recombinant human growth hormone from *Escherichia coli*. *Protein Expression and Purification* **18**(2), 182-92.
- Peterkofsky, A., and Gazdar, C. (1974). Glucose inhibition of adenylate cyclase in intact cells of *Escherichia coli* B. *Proceedings of the National Academy of Sciences of the United States of America* **71**(6), 2324-8.
- Petsch, D., and Anspach, F. B. (2000). Endotoxin removal from protein solutions. *Journal of Biotechnology* **76**(2-3), 97-119.
- Pickett, M. J., and Goodman, R. E. (1966). Beta-Galactosidase for Distinguishing between *Citrobacter* and *Salmonella*. *Applied Microbiology* **14**(2), 178-&.

- Pinto, E., Queiroz, M.-J. R. P., Vale-Silva, L. A., Oliveira, J. F., Begouin, A., Begouin, J.-M., and Kirsch, G. (2008). Antifungal activity of synthetic di(hetero)arylamines based on the benzo[b]thiophene moiety. *Bioorganic & Medicinal Chemistry* **16**(17), 8172-8177.
- Pinske, C., Bonn, M., Kruger, S., Lindenstrauss, U., and Sawers, R. G. (2011). Metabolic deficiencies revealed in the biotechnologically important model bacterium *Escherichia coli* BL21(DE3). *PLoS One* **6**(8), e22830.
- Plovins, A., Alvarez, A. M., Ibanez, M., Molina, M., and Nombela, C. (1994). Use of fluorescein-di-beta-D-galactopyranoside (FDG) and C12-FDG as substrates for beta-galactosidase detection by flow cytometry in animal, bacterial, and yeast cells. *Applied and Environmental Microbiology* **60**(12), 4638-41.
- Pocsi, I., Taylor, S. A., Richardson, A. C., Smith, B. V., and Price, R. G. (1993). Comparison of Several New Chromogenic Galactosides as Substrates for Various Beta-D-Galactosidases. *Biochimica Et Biophysica Acta* **1163**(1), 54-60.
- Pogliano, J., Lynch, A. S., Belin, D., Lin, E. C. C., and Beckwith, J. (1997). Regulation of *Escherichia coli* cell envelope proteins involved in protein folding and degradation by the Cpx two-component system. *Genes & Development* **11**(9), 1169-1182.
- Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G., and Cormier, M. J. (1992). Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* **111**(2), 229-33.
- Pribnow, D. (1975). Nucleotide-Sequence of an Rna-Polymerase Binding-Site at an Early T7 Promoter. *Proceedings of the National Academy of Sciences of the United States of America* **72**(3), 784-788.
- Rader, R. A. (2008). (Re)defining biopharmaceutical. *Nature Biotechnology* **26**(7), 743-751.
- Raivio, T. L., and Silhavy, T. J. (1999). The sigmaE and Cpx regulatory pathways: overlapping but distinct envelope stress responses. *Current Opinion in Microbiology* **2**(2), 159-65.
- Raivio, T. L., and Silhavy, T. J. (2001). Periplasmic stress and ECF sigma factors. *Annual Review of Microbiology* **55**, 591-624.
- Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963). Stereochemistry of Polypeptide Chain Configurations. *Journal of Molecular Biology* **7**(1), 95-&.
- Redelman, D. (2004). CytometryML. *Cytometry A* **62**(1), 70-3.
- Regot, S., Macia, J., Conde, N., Furukawa, K., Kjellen, J., Peeters, T., Hohmann, S., de Nadal, E., Posas, F., and Sole, R. (2011). Distributed biological computation with multicellular engineered networks. *Nature* **469**(7329), 207-11.
- Reid, B. G., and Flynn, G. C. (1997). Chromophore formation in green fluorescent protein. *Biochemistry* **36**(22), 6786-91.
- Rinas, U., Hoffmann, F., Betiku, E., Estape, D., and Marten, S. (2007). Inclusion body anatomy and functioning of chaperone-mediated in vivo inclusion body disassembly during high-level recombinant protein production in *Escherichia coli*. *Journal of Biotechnology* **127**(2), 244-257.
- Robertson, B. R., Button, D. K., and Koch, A. L. (1998). Determination of the Biomasses of Small Bacteria at Low Concentrations in a Mixture of Species with Forward Light Scatter Measurements by Flow Cytometry. *Applied and Environmental Microbiology*. **64**(10), 3900-3909.
- Rodriguez, F., Arsene-Ploetze, F., Rist, W., Rudiger, S., Schneider-Mergener, J., Mayer, M. P., and Bukau, B. (2008). Molecular basis for regulation of the heat shock transcription factor sigma 32 by the DnaK and DnaJ chaperones. *Molecular Cell* **32**(3), 347-58.
- Roostalu, J., Joers, A., Luidalepp, H., Kaldalu, N., and Tenson, T. (2008). Cell division in *Escherichia coli* cultures monitored at single cell resolution. *BMC Microbiology* **8**, 68.

- Rosen, R., and Ron, E. Z. (2002). Proteome analysis in the study of the bacterial heat-shock response. *Mass Spectrometry Reviews* **21**(4), 244-65.
- Rosenow, M. A., Huffman, H. A., Phail, M. E., and Wachter, R. M. (2004). The crystal structure of the Y66L variant of green fluorescent protein supports a cyclization-oxidation-dehydration mechanism for chromophore maturation. *Biochemistry* **43**(15), 4464-72.
- Rouviere, P. E., De Las Penas, A., Meccas, J., Lu, C. Z., Rudd, K. E., and Gross, C. A. (1995). rpoE, the gene encoding the second heat-shock sigma factor, sigma E, in *Escherichia coli*. *EMBO Journal* **14**(5), 1032-42.
- Rudd, K. E. (2000). EcoGene: a genome sequence database for *Escherichia coli* K-12. *Nucleic Acids Research* **28**(1), 60-64.
- Rudiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO Journal* **16**(7), 1501-7.
- Ruther, U. (1980). Construction and properties of a new cloning vehicle, allowing direct screening for recombinant plasmids. *Molecular and General Genetics* **178**(2), 475-7.
- Salmon, K. A., Hung, S. P., Steffen, N. R., Krupp, R., Baldi, P., Hatfield, G. W., and Gunsalus, R. P. (2005). Global gene expression profiling in *Escherichia coli* K12: effects of oxygen availability and ArcA. *Journal of Biological Chemistry* **280**(15), 15084-96.
- Sambrook J., Fritsch E.F., and T., M. (1989). "Molecular Cloning: A Laboratory Manual 2nd Ed." Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sawers, G. (1999). The aerobic/anaerobic interface. *Current Opinion in Microbiology* **2**(2), 181-7.
- Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**(5235), 467-70.
- Schlieker, C., Bukau, B., and Mogk, A. (2002). Prevention and reversion of protein aggregation by molecular chaperones in the E. coli cytosol: implications for their applicability in biotechnology. *Journal of Biotechnology* **96**(1), 13-21.
- Schweder, T., Lee, K. H., Lomovskaya, O., and Matin, A. (1996). Regulation of *Escherichia coli* starvation sigma factor σ^S by ClpXP protease. *Journal of Bacteriology* **178**(2), 470-476.
- Seamer, L. C., Bagwell, C. B., Barden, L., Redelman, D., Salzman, G. C., Wood, J. C., and Murphy, R. F. (1997). Proposed new data file standard for flow cytometry, version FCS 3.0. *Cytometry* **28**(2), 118-22.
- Sevastyanovich, Y., Alfasi, S., and Cole, J. (2009). Recombinant protein production: a comparative view on host physiology. *New Biotechnology* **25**(4), 175-180.
- Sevastyanovich, Y., Alfasi, S., Overton, T., Hall, R., Jones, J., Hewitt, C., and Cole, J. (2009). 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.
- Shakhnovich, E. (2006). Protein folding thermodynamics and dynamics: Where physics, chemistry, and biology meet. *Chemical Reviews* **106**(5), 1559-1588.
- Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. *Nature Methods* **2**(12), 905-909.
- Shapiro AL, Viñuela E, Maizel JV Jr. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels.(1967) *Biochemical and Biophysics Research Communications*. **28**(5):815-20.
- Sherman, M. Y., and Goldberg, A. L. (1992). Involvement of the Chaperonin DnaK in the Rapid Degradation of a Mutant Protein in *Escherichia-Coli*. *EMBO Journal* **11**(1), 71-77.

- Shimomura, O. (1979). Structure of the Chromophore of Aequorea Green Fluorescent Protein. *FEBS Letters* **104**(2), 220-222.
- Shimomura, O. (2005). The discovery of aequorin and green fluorescent protein. *Journal of Microscopy* **217**(Pt 1), 1-15.
- Shimomura, O., Johnson, F. H., and Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *Journal of Comparative Physiology* **59**, 223-39.
- Sidhu, S. S. (2007). Full-length antibodies on display. *Nature Biotechnology* **25**(5), 537-538.
- Siegenthaler, R. K., Grimshaw, J. P., and Christen, P. (2004). Immediate response of the DnaK molecular chaperone system to heat shock. *Febs Letters* **562**(1-3), 105-10.
- Siemering, K. R., Golbik, R., Sever, R., and Haseloff, J. (1996). Mutations that suppress the thermosensitivity of green fluorescent protein. *Current Biology* **6**(12), 1653-63.
- Simmons, L. C., Reilly, D., Klimowski, L., Raju, T. S., Meng, G., Sims, P., Hong, K., Shields, R. L., Damico, L. A., Rancatore, P., and Yansura, D. G. (2002). Expression of full-length immunoglobulins in *Escherichia coli*: rapid and efficient production of aglycosylated antibodies. *Journal of Immunological Methods* **263**(1-2), 133-147.
- Singh, S. M., and Panda, A. K. (2005). Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering* **99**(4), 303-310.
- Singh, S. S., Typas, A., Hengge, R., and Grainger, D. C. (2011). *Escherichia coli* σ^{70} senses sequence and conformation of the promoter spacer region. *Nucleic Acids Research* **39**(12), 5109-18.
- Sinha, G. (2011). EU mAb biosimilars path. *Nature Biotechnology: News* **29**(10), 1.
- Siryaporn, A., and Goulian, M. (2008). Cross-talk suppression between the CpxA-CpxR and EnvZ-OmpR two-component systems in E-coli. *Molecular Microbiology* **70**(2), 494-506.
- Sklar, J. G., Wu, T., Kahne, D., and Silhavy, T. J. (2007). Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes & Development* **21**(19), 2473-84.
- Slepenkov, S. V., and Witt, S. N. (2002). The unfolding story of the *Escherichia coli* Hsp70 DnaK: is DnaK a holdase or an unfoldase? *Molecular Microbiology* **45**(5), 1197-1206.
- Smith, H. E. (2007). The transcriptional response of *Escherichia coli* to recombinant protein insolubility. *Journal of Structural and Functional Genomics* **8**(1), 27-35.
- Søballe, B., and Poole, R. K. (2000). Ubiquinone limits oxidative stress in *Escherichia coli*. *Microbiology* **146**(4), 787-796.
- Sola, R. J., and Griebenow, K. (2009). Effects of Glycosylation on the Stability of Protein Pharmaceuticals. *Journal of Pharmaceutical Sciences* **98**(4), 1223-1245.
- Sonoda, H., Kumada, Y., Katsuda, T., and Yamaji, H. (2011). Cytoplasmic production of soluble and functional single-chain Fv-Fc fusion protein in *Escherichia coli*. *Biochemical Engineering Journal* **53**(3), 253-259.
- Sorensen, H. P., and Mortensen, K. K. (2005). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology* **115**(2), 113-28.
- Sorensen, S. J., Sorensen, A. H., Hansen, L. H., Oregaard, G., and Veal, D. (2003). Direct detection and quantification of horizontal gene transfer by using flow cytometry and *gfp* as a reporter gene. *Current Microbiology* **47**(2), 129-33.

- Speed, M. A., Wang, D. I. C., and King, J. (1996). Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion body composition. *Nature Biotechnology* **14**(10), 1283-1287.
- Spidlen, J., Gentleman, R. C., Haaland, P. D., Langille, M., Le Meur, N., Ochs, M. F., Schmitt, C., Smith, C. A., Treister, A. S., and Brinkman, R. R. (2006). Data standards for flow cytometry. *OMICS* **10**(2), 209-14.
- Spiess, C., Beil, A., and Ehrmann, M. (1999). A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell* **97**(3), 339-47.
- Spiro, S., and Guest, J. R. (1990). FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. *FEMS Microbiology Reviews* **6**(4), 399-428.
- Steitz, T. A. (2008). A structural understanding of the dynamic ribosome machine. *Nature Reviews Molecular Cell Biology* **9**(3), 242-253.
- Stewart, E. J., Aslund, F., and Beckwith, J. (1998). Disulfide bond formation in the *Escherichia coli* cytoplasm: an in vivo role reversal for the thioredoxins. *EMBO Journal* **17**(19), 5543-50.
- Stewart, V., Bledsoe, P. J., and Williams, S. B. (2003). Dual overlapping promoters control napF (periplasmic nitrate reductase) operon expression in *Escherichia coli* K-12. *Journal of Bacteriology* **185**(19), 5862-70.
- Stock, A. M., Robinson, V. L., and Goudreau, P. N. (2000). Two-component signal transduction. *Annual Review of Biochemistry* **69**, 183-215.
- Stricker, J., Cookson, S., Bennett, M. R., Mather, W. H., Tsimring, L. S., and Hasty, J. (2008). A fast, robust and tunable synthetic gene oscillator. *Nature* **456**(7221), 516-U39.
- Stryer, L., and Haugland, R. P. (1967). Energy transfer: a spectroscopic ruler. *Proceedings of the National Academy of Sciences of the United States of America* **58**(2), 719-26.
- Sugimoto, S., Saruwatari, K., Higashi, C., and Sonomoto, K. (2008). The proper ratio of GrpE to DnaK is important for protein quality control by the DnaK–DnaJ–GrpE chaperone system and for cell division. *Microbiology* **154**(7), 1876-1885.
- Swartz, J. R. (2001). Advances in *Escherichia coli* production of therapeutic proteins. *Current Opinion in Biotechnology* **12**(2), 195-201.
- Takano, T., and Kakefuda, T. (1972). Involvement of a bacterial factor in morphogenesis of bacteriophage capsid. *Nature New Biology* **239**(89), 34-7.
- Tatsuta, T., Tomoyasu, T., Bukau, B., Kitagawa, M., Mori, H., Karata, K., and Ogura, T. (1998). Heat shock regulation in the ftsH null mutant of *Escherichia coli*: dissection of stability and activity control mechanisms of sigma(32) in vivo. *Molecular Microbiology* **30**(3), 583-593.
- Tecon, R., Wells, M., and van der Meer, J. R. (2006). A new green fluorescent protein-based bacterial biosensor for analysing phenanthrene fluxes. *Environmental Microbiology* **8**(4), 697-708.
- Terranova, U., and Nifosi, R. (2010). A role for molecular compression in the post-translational formation of the Green Fluorescent Protein chromophore. *Chemical Physics* **371**(1-3), 76-83.
- Thattai, M., and van Oudenaarden, A. (2004). Stochastic gene expression in fluctuating environments. *Genetics* **167**(1), 523-30.
- Thomas, C. P., Booth, T. F., and Roy, P. (1990). Synthesis of bluetongue virus-encoded phosphoprotein and formation of inclusion bodies by recombinant baculovirus in insect cells: it binds the single-stranded RNA species. *Journal of General Virology* **71** (Pt 9), 2073-83.
- Tielker, D., Eichhof, I., Jaeger, K. E., and Ernst, J. F. (2009). Flavin Mononucleotide-Based Fluorescent Protein as an Oxygen-Independent Reporter in *Candida albicans* and *Saccharomyces cerevisiae*. *Eukaryotic Cell* **8**(6), 913-915.

- Tomoyasu, T., Ogura, T., Tatsuta, T., and Bukau, B. (1998). Levels of DnaK and DnaJ provide tight control of heat shock gene expression and protein repair in *Escherichia coli*. *Molecular Microbiol* **30**(3), 567-81.
- Travers, A. A., and Burgess, R. R. (1969). Cyclic Re-Use of Rna Polymerase Sigma Factor. *Nature* **222**(5193), 537-&.
- Trill, J. J., Shatzman, A. R., and Ganguly, S. (1995). Production of monoclonal antibodies in COS and CHO cells. *Current Opinion in Biotechnology* **6**(5), 553-60.
- Tsien, R. Y. (1998). The green fluorescent protein. *Annual Review of Biochemistry* **67**, 509-44.
- Turner, S., Reid, E., Smith, H., and Cole, J. (2003). A novel cytochrome c peroxidase from *Neisseria gonorrhoeae*: a lipoprotein from a Gram-negative bacterium. *Biochem J* **373**(Pt 3), 865-73.
- Tyagi, N. K., Fenton, W. A., and Horwich, A. L. (2009). GroEL/GroES cycling: ATP binds to an open ring before substrate protein favoring protein binding and production of the native state. *Proceedings of the National Academy of Sciences of the United States of America* **106**(48), 20264-20269.
- Tyagi, N. K., Fenton, W. A., and Horwich, A. L. (2010). ATP-triggered ADP release from the asymmetric chaperonin GroEL/GroES/ADP(7) is not the rate-limiting step of the GroEL/GroES reaction cycle. *Febs Letters* **584**(5), 951-953.
- Ugrinov, K. G., and Clark, P. L. (2010). Cotranslational folding increases GFP folding yield. *Biophysical Journal* **98**(7), 1312-20.
- Uندن, G., and Bongaerts, J. (1997). Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors. *Biochim et Biophysica Acta* **1320**(3), 217-34.
- van der Meer, J. R., and Belkin, S. (2010). Where microbiology meets microengineering: design and applications of reporter bacteria. *Nature Reviews Microbiology* **8**(7), 511-522.
- Veal, D. A., Deere, D., Ferrari, B., Piper, J., and Attfield, P. V. (2000). Fluorescence staining and flow cytometry for monitoring microbial cells. *Journal of Immunological Methods* **243**(1-2), 191-210.
- Visca, P., Leoni, L., Wilson, M. J., and Lamont, I. L. (2002). Iron transport and regulation, cell signalling and genomics: lessons from *Escherichia coli* and *Pseudomonas*. *Molecular Microbiology* **45**(5), 1177-90.
- Vizcaino-Caston, I., Wyre, C., and Overton, T. W. (2011). Fluorescent proteins in microbial biotechnology-new proteins and new applications. *Biotechnol Letters*.
- Vostiar, I., Tkac, J., and Mandenius, C. F. (2003). Monitoring of the heat-shock response in *Escherichia coli* using an optical biosensor. *Analytical Biochemistry* **322**(2), 156-63.
- Wade, J. T., Roa, D. C., Grainger, D. C., Hurd, D., Busby, S. J., Struhl, K., and Nudler, E. (2006). Extensive functional overlap between sigma factors in *Escherichia coli*. *Nature Structural and Molecular Biology* **13**(9), 806-14.
- Waldo, G. S., Standish, B. M., Berendzen, J., and Terwilliger, T. C. (1999). Rapid protein-folding assay using green fluorescent protein. *Nature Biotechnology* **17**(7), 691-695.
- Wallberg, F., Sundstrom, H., Ledung, E., Hewitt, C. J., and Enfors, S. O. (2005). Monitoring and quantification of inclusion body formation in *Escherichia coli* by multi-parameter flow cytometry. *Biotechnol Letters* **27**(13), 919-26.
- Walsh, G. (2006). Biopharmaceutical benchmarks 2006. *Nature Biotechnology* **24**(7), 769-76.
- Walsh, N. P., Alba, B. M., Bose, B., Gross, C. A., and Sauer, R. T. (2003). OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. *Cell* **113**(1), 61-71.

- Wang, Y., and deHaseth, P. L. (2003). Sigma 32-dependent promoter activity in vivo: sequence determinants of the groE promoter. *Journal of Bacteriology* **185**(19), 5800-6.
- Ward, W. W., and Cormier, M. J. (1975). Extraction of Renilla-type luciferin from the calcium-activated photoproteins aequorin, mnemiopsin, and berovin. *Proceedings of the National Academy of Sciences of the United States of America* **72**(7), 2530-4.
- Weisblum, B., Graham, M. Y., Gryczan, T., and Dubnau, D. (1979). Plasmid copy number control: isolation and characterization of high-copy-number mutants of plasmid pE194. *Journal of Bacteriology* **137**(1), 635-43.
- Wickens, H. J., Pinney, R. J., Mason, D. J., and Gant, V. A. (2000). Flow cytometric investigation of filamentation, membrane patency, and membrane potential in *Escherichia coli* following ciprofloxacin exposure. *Antimicrobial Agents and Chemotherapy* **44**(3), 682-7.
- Wickner, S., Gottesman, S., Skowyra, D., Hoskins, J., Mckenney, K., and Maurizi, M. R. (1994). A Molecular Chaperone, ClpA, Functions Like DnaK and DnaJ. *Proceedings of the National Academy of Sciences of the United States of America* **91**(25), 12218-12222.
- Wickner, S., Maurizi, M. R., and Gottesman, S. (1999). Posttranslational quality control: Folding, refolding, and degrading proteins. *Science* **286**(5446), 1888-1893.
- Wilken, C., Kitzing, K., Kurzbauer, R., Ehrmann, M., and Clausen, T. (2004). Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease. *Cell* **117**(4), 483-494.
- Williams, D. C., Van Frank, R. M., Muth, W. L., and Burnett, J. P. (1982). Cytoplasmic inclusion bodies in *Escherichia coli* producing biosynthetic human insulin proteins. *Science* **215**(4533), 687-9.
- Wilson, T., and Hastings, J. W. (1998). Bioluminescence. *Annual Review of Cell and Developmental Biology* **14**, 197-230.
- Winkler, J., Seybert, A., Konig, L., Pruggnaller, S., Haselmann, U., Sourjik, V., Weiss, M., Frangakis, A. S., Mogk, A., and Bukau, B. (2010). Quantitative and spatio-temporal features of protein aggregation in *Escherichia coli* and consequences on protein quality control and cellular ageing. *EMBO Journal* **29**(5), 910-23.
- Winter, J., and Jakob, U. (2004). Beyond Transcription—New Mechanisms for the Regulation of Molecular Chaperones. *Critical Reviews in Biochemistry and Molecular Biology* **39**(5-6), 297-317.
- Wolfe, A. J., Parikh, N., Lima, B. P., and Zemaitaitis, B. (2008). Signal integration by the two-component signal transduction response regulator CpxR. *Journal of Bacteriology* **190**(7), 2314-2322.
- Wollmann, P., and Zeth, K. (2007). The structure of RseB: A sensor in periplasmic stress response of *E. coli*. *Journal of Molecular Biology* **372**(4), 927-941.
- Wösten, M. M. S. M. (1998). Eubacterial sigma-factors. *FEMS Microbiology Reviews* **22**(3), 127-150.
- Wu, J., and Newton, A. (1997). The Caulobacter heat shock sigma factor gene rpoH is positively autoregulated from a sigma32-dependent promoter. *Journal of Bacteriology* **179**(2), 514-521.
- Wu, X., Jornvall, H., Berndt, K. D., and Oppermann, U. (2004). Codon optimization reveals critical factors for high level expression of two rare codon genes in *Escherichia coli*: RNA stability and secondary structure but not tRNA abundance. *Biochemical and Biophysical Research Communications* **313**(1), 89-96.
- Xu, Z., Horwich, A. L., and Sigler, P. B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. *Nature* **388**(6644), 741-50.
- Yamamoto, K., and Ishihama, A. (2006). Characterization of copper-inducible promoters regulated by CpxA/CpxR in *Escherichia coli*. *Biosciences Biotechnology and Biochemistry* **70**(7), 1688-95.

- Yang, F., Moss, L. G., and Phillips, G. N. (1996). The molecular structure of green fluorescent protein. *Nature Biotechnology* **14**(10), 1246-1251.
- Yura, T., and Nakahigashi, K. (1999). Regulation of the heat-shock response. *Current Opinion in Microbiology* **2**(2), 153-8.
- Zacharias, D. A., Violin, J. D., Newton, A. C., and Tsien, R. Y. (2002). Partitioning of Lipid-Modified Monomeric GFPs into Membrane Microdomains of Live Cells. *Science* **296**(5569), 913-916.
- Zahrl, D., Wagner, M., Bischof, K., and Koraimann, G. (2006). Expression and assembly of a functional type IV secretion system elicit extracytoplasmic and cytoplasmic stress responses in *Escherichia coli*. *Journal of Bacteriol* **188**(18), 6611-21.
- Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M. G., and Alon, U. (2006). A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nature Methods* **3**(8), 623-8.
- Zhang, L., Patel, H. N., Lappe, J. W., and Wachter, R. M. (2006). Reaction progress of chromophore biogenesis in green fluorescent protein. *Journal of the American Chemical Society* **128**(14), 4766-72.
- Zhao, R., Natarajan, A., and Srienc, F. (1999). A flow injection flow cytometry system for on-line monitoring of bioreactors. *Biotechnology and Bioengineering* **62**(5), 609-17.
- Zheng, D., Constantinidou, C., Hobman, J. L., and Minchin, S. D. (2004). Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Research* **32**(19), 5874-5893.
- Zheng, M., Wang, X., Templeton, L. J., Smulski, D. R., LaRossa, R. A., and Storz, G. (2001). DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. *Journal of Bacteriology* **183**(15), 4562-70.
- Zhou, X. H., Keller, R., Volkmer, R., Krauss, N., Scheerer, P., and Hunke, S. (2011). Structural Basis for Two-component System Inhibition and Pilus Sensing by the Auxiliary CpxP Protein. *Journal of Biological Chemistry* **286**(11), 9805-9814.
- Zimmer, M. (2002). Green fluorescent protein (GFP): Applications, structure, and related photophysical behavior. *Chemical Reviews* **102**(3), 759-781.
- Zwanzig, R. (1997). Two-state models of protein folding kinetics. *Proceedings of the National Academy of Sciences of the United States of America* **94**(1), 148-150.