**Universidade do Minho**
Escola de Ciências

Mário Jorge Faria Barroca

# Optimisation of the fed-batch production of a silk-elastin-like protein: overcoming acetate accumulation and plasmid instability

outubro de 2013

Optimisation of the fed-batch production of a silk-elastin-like protein: overcoming acetate accumulation and plasmid instability
Mário Jorge Faria Barroca
UMinho | 2013

**Universidade do Minho**
Escola de Ciências

Mário Jorge Faria Barroca

# Optimisation of the fed-batch production of a silk-elastin-like protein: overcoming acetate accumulation and plasmid instability

Dissertação de Mestrado
Mestrado em Biofísica e Bionanossistemas

Trabalho realizado sob a orientação do
**James Anthony Collins**
(Investigador Auxiliar, Universidade do Minho)
e da
**Margarida Casal**
(Prof. Catedrática, Universidade do Minho)

outubro de 2013

# DECLARAÇÃO

**Nome**

Mário Jorge Faria Barroca

**Endereço eletrónico:** mariobarroc@gmail.com

**Número do Bilhete de Identidade:** 13741239

**Título da Dissertação**

Optimisation of the fed-batch production of a silk-elastin-like protein: overcoming acetate accumulation and plasmid instability

**Orientadores**

James Anthony Collins (Investigador Auxiliar, Universidade do Minho) e Margarida Casal (Prof. Catedrática, Universidade do Minho)

**Ano de conclusão** 2013

**Designação do Mestrado:**

Mestrado em Biofísica e Bionanossistemas



Universidade do Minho, ___/___/_____

Assinatura: _____

# Acknowledgments/Agradecimentos

À generalidade dos professores do Mestrado em Biofísica e Bionanossistemas; pelos saberes e competências que em mim desenvolveram, pelas reflexões que me proporcionaram.

Em particular, ao Doutor Tony Collins e à Professora Doutora Margarida Casal, por me terem aceite como orientando, pelos sábios ensinamentos científicos e pedagógicos durante o desenvolvimento do projecto de investigação e escrita desta dissertação, mas também pelo estímulo, compreensão e amizade demonstradas. Obrigado!!

Um agradecimento muito especial aos meus companheiros do LBM, Raul, Margarida, João, Joana, André, Telma, Pedro, por todo tempo dedicado, paciência, sabedoria, respeito, vícios, e sobretudo a grande amizade que fizemos ao longo deste ano. Tornaram tudo muito mais fácil.

Um muito Obrigado!!!

A todos os meus colegas de mestrado, em especial, Rogério, David, Ana Luísa, Pedro, Diana, Filipa, Filipa Gonçalves, um agradecimento pelo companheirismo e força que proporcionaram.

Agradeço a todos os técnicos do departamento que foram sempre profissionais e ajudaram bastante na realização do trabalho, principalmente a D. Isabel, a Manuela Rodrigues e a Magda Graça por animarem me atenderem sempre que precisava, mesmo quando não precisavam de o fazer. Um sincero Obrigado!!

Agradeço a todos os meus amigos que me aturaram durante este ano, principalmente a minha grandes amigas Nádia e Rosa Maria, que me deram forças e paciência para a realização desta tese, assim como pelos momentos divertidos que me fizeram viver, como boas amigas que são

Por último, mas sempre em primeiro, um agradecimento muito especial e muito sentido ao Serafim, à Manuela e ao Gonçalo pelas atenções que lhes devia e que este trabalho lhes tomou, assim como todo o amor e dedicação que deram a este filho. Sem eles nada era possível. Obrigado do fundo do coração!!!!!

# Abstract

Silk-elastin-like proteins (SELPs) are protein based polymers composed of repetitive amino acid sequence motifs, or variants thereof, found in silk fibroin (GAGAGS) and mammalian elastin (VPGVG). These polymers have a high potential for use in the pharmaceutical, regenerative medicine and materials fields. The successful employment of these polymers in these different areas requires that they can be simply prepared in large quantities on an industrial scale. The present study attempted to increase the production levels of a novel recently described SELP (SELP-59-A) in *E. coli*.

Prior to the present study, the highest reported production level for SELPs was 500 mg/L in *E. coli* BL21(DE3) with an optimised batch fermentation in shake flasks approach[1]. It was shown that SELP production was limited with this approach by the accumulation of acetic acid to toxic levels as well as by plasmid loss on induction. In an attempt to overcome these limitations, the use of a fed-batch production approach and of various plasmid stabilisation systems were investigated in the present study. A fed-batch approach was developed and optimised for the successful high cell density production of SELP-59-A. Pre- (0.4 $h^{-1}$) and post-induction (0.1 $h^{-1}$) growth rates, dissolved oxygen concentration (15 %), dry cell weight (DCW) on induction (75 g/L) and IPTG concentration for induction (1 mM IPTG) were optimised and allowed for an almost 9-fold increase in SELP-59-A production. This allowed for a reduced cost process with a fed-batch phase of approximately 8 hours. Limiting factors identified were plasmid loss on induction, limitations in the oxygen transfer efficiency of the fermentation instrumentation used and an increased metabolic stress on the host cells following induction. Investigation of ampicillin resistance, kanamycin resistance and a Type II toxin/antitoxin post segregational suicide system (ccdB/ccdA) for reduced plasmid loss indicated the latter to allow for improved plasmid stability and a 50 % increase in SELP-59-A production under the conditions used. Removal or inactivation of the ampicillin resistance marker from the expression vector did not affect production but allows for the production of SELPs for use in biomedical applications. The SELP production levels reported here are the highest to date and represent a 10-fold increase on that previously reported.

# Resumo

As proteínas do tipo seda-elastina (SELPs) são uma nova classe de polímeros recombinantes de origem proteica, cuja composição consiste na combinação das sequências repetitivas de aminoácidos comummente encontradas nas proteínas naturais, fibroína (GAGAGS) e elastina (VPGVG). Estes polímeros têm um elevado potencial de aplicação nas áreas de medicina regenerativa, farmacêutica e materiais. Todavia, para que a sua aplicação seja bem sucedida, surge a necessidade de serem facilmente produzidos em larga escala. Como objetivo principal deste estudo, pretendeu-se aumentar os níveis de produção de uma SELP em *E. coli*. Até à data, o nível de produtividade volumétrica mais elevado descrito para SELPs foi de 500 mg/L, obtido em *E. coli*, e em condições otimizadas de fermentação em matrazes. Foi demonstrado que a produção de SELP era limitada não só devido à acumulação de ácido acético, conduzindo a níveis tóxicos para a célula, mas também devido à perda de plasmídeo pela indução. No âmbito desta tese e numa tentativa de ultrapassar estas limitações, recorreu-se a uma abordagem de produção em fed-batch, assim como a vários sistemas de estabilização de plasmídeo. Assim, foi desenvolvido e otimizado um processo de *fed-batch*, o qual demonstrou resultar na produção de SELP-59-A em culturas de elevada densidade celular. Com parâmetros otimizados tais como taxas de crescimento de pré- (0.4 h-1) e pós- indução (0.1h-1), concentração de oxigénio dissolvido (15%) e, indução com IPTG (1 mM IPTG, a um peso seco celular de 75 g/L), permitiu um aumento de quase 9 vezes em relação à produção volumétrica previamente descrita. Isto permitiu obter um processo de custo reduzido com uma fase de *fed-batch* de cerca de 8 horas. Como fatores limitantes, foram identificados: a perda de plasmídeo na indução, limitações na eficiência da transferência de oxigénio pelo equipamento utilizado e um aumento do stress metabólico nas células hospedeiras após indução. Os estudos de resistência à ampicilina, canamicina e um sistema toxina/antitoxina do tipo II (ccdB/ccdA), foram utilizados para reduzir a perda de plasmídeo, resultando num aumento da estabilidade do plasmídeo e por conseguinte num aumento de 50 % na produção de SELP-59-A. Os níveis de produção SELP relatados neste estudo são os mais altos descritos até hoje, representando um aumento de cerca de 10 vezes maior do que descrito na literatura.

# Index

# Abbreviations list

Δ*bla* – Ampicillin resistance gene

Δ*blap* – Ampicillin resistance promoter

$amp^R$ – Ampicillin resistance

kDa – Kilo Dalton

DCW – Dry Cell Weight

DNA – Deoxyribonucleic acid

$dO_2$ – Dissolved oxygen

g – Gram

h – Hour

HPLC – High performance liquid chromatography

IPTG – Isopropyl β-D-1-thiogalactopyranoside

ITT – Inverse temperature transition

$kan^R$ – Kanamycin resistance

L – Litre

LB – Lysogeny broth

M – Molar

min – Minute

mRNA – Messenger ribonucleic acid

NBS – New Brunswick Scientific

nm – Nanometers

ºC – Degrees Celsius

OD – Optical density

PBP - Protein Based Polymer

PCR – Polymerase chain reaction

qPCR – quantitative PCR

RNA – Ribonucleic acid

rpm – Revolutions per minute

SDS – Sodium dodecyl sulfate

SELP – Silk-Elastin-like Polymer

TCA – Tricarboxylic acid

w – Weight

DNA – Deoxyribonucleic acid

# List of figures

# List of tables

# 1. Introduction

## 1.1. Polymers

Polymers (poly – many, meros – parts) are naturally occurring or synthetic compounds made up of a large number of linked repeating structural subunits. The relatively simple basic subunits, known as monomers, are joined together (typically by covalent bonds) to form oligomers of varying sizes (dimers, trimers etc.) and eventually, through a large number of repeats, give rise to polymers. Polymers typically have molecular weights greater than 5000 g/mole, thereby allowing for chain entanglement, and can consist of up to millions of repeated units per polymer chain[2.3].

### 1.1.1. Protein-Based Polymers

Protein-based polymers (PBPs), composed of repeating units of amino acids, have recently emerged as a promising new class of bioinspired materials. The repetitive amino acid motifs of these biopolymers are typically based on naturally occurring proteins and in particular on repetitive motifs found in naturally occurring fibrous or elastomeric proteins such as silk, elastin and resilin. These natural proteins serve functional, protective and structural purposes in nature and are typically characterized by remarkable physical, mechanical and biological properties with, in particular, a high strength and stability being often noted[3.4.5]. PBPs, designed to display similar or improved properties to these, are of much interest from both a fundamental and applied point of view, and find potential application in biomedicine, pharmaceuticals, nanotechnology and as materials. They can be processed into fibres, films, nanoparticles and/or hydrogels and offer many advantages over conventional petroleum-based polymers, including being more environmentally friendly, renewable and biocompatible. Furthermore, they can be prepared by use of recombinant DNA technology which, in contrast to conventional synthetic polymers that tend to be produced as a polydispersed population, allows for the production of monodispersed PBPs in which the sequence, composition and length is strictly controlled. Here, the gene encoding

the protein based polymer of interest is introduced into a suitable host whereupon the protein producing machinery (i.e. transcription, translation and folding machinery) of this host is utilised to produce the polymer[6.7.8].

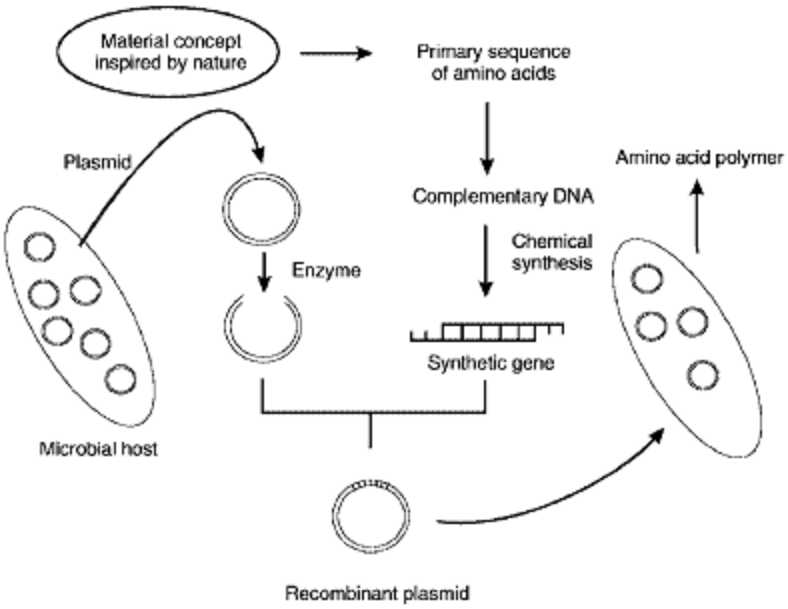



Figure 1. An overview of the process for protein based polymer synthesis via protein engineering[5]**.**

The basic principle of PBP synthesis via recombinant DNA technology is illustrated in Figure 1. Amino acid sequences are designed to create specific folding patterns and desired new material properties, with the design being generally based on naturally occurring fibrous proteins found in nature but with novel designs being also possible. Indeed, use of protein engineering techniques to modify natural amino acid sequences might allow for novel materials with novel improved properties. The primary amino acid sequence is then reverse-translated into its corresponding nucleotide sequence and the desired DNA fragment synthesized by molecular biology approaches. Indeed, as many fibrous proteins are characterized by repetitive amino acid sequences, it is often possible to multimerise a smaller synthesized oligonucleotide sequence to prepare an artificial gene that codes for proteins of high molecular weight[8]. The gene is then incorporated into circular plasmid DNA which can be transformed to an appropriate host, or directly integrated into host genome. Many hosts are currently available, including bacteria, yeast, fungi, plants and animals[3.4.5] with

the bacterial host *E. coli* being the most commonly used as it displays a high growth rate, is relatively cheap and easy to cultivate and has been extensively studied.

## 1.2. Fibrous Proteins

Fibrous proteins such as elastin, silk, resilin, abductin or wheat gluten are self-assembling proteins that display properties of extensibility, elastic recoil and/or high strength and stability. They are present in a variety of tissues, have precise biological functions and are critical for survival. Examples of their exploitation in nature include: insect wings that depend on resilin for flight by releasing stored energy to power the return stroke in response to each wing beat; human arteries that contain elastin in the walls to allow for contraction and the pumping of blood throughout the body; and spider silk that can be tougher than steel and allows for the entrapment of prey [9.10]. Although evolutionarily unrelated, they all share a common sequence design involving short highly repetitive sequences that impart elasticity, and often also strength, interspersed with elements capable of forming cross-links that help stabilise the polymer structures [11] (Table 1). The presence of regularly repeated sequences implies the formation of a regular structure and although the direct determination of the structures of elastic proteins has proved problematic, the limited information that is available indicates that the repetitive sequences do form regular structures and that these may be important in the elastic mechanisms[9.12].

Table 1. Repeat motifs and crosslinking mechanisms for elastomeric proteins [9].

| Group | Protein | Elastic repeat motif[a] | Nature of crosslinks |
|---|---|---|---|
| 1 | Abductin | GGFGGMGGGX | Unknown, disulphide (?) and/or via tyrosine residues (?) |
| 2 | Elastin | VPGG<br>VPGVG<br>APGVGV | Via lysine residues in the alanine-rich regions |
| | Byssus | GPGGG | Unknown, metal complexation with histidine-rich regions (?) and/or via tyrosine residues (?) |
| | Flagelliform silk | GPGGx | Unknown, noncovalent interactions in spacer regions (?) |
| 3 | Dragline silk | GPGQQ<br>GPGGY<br>GGYGPGS | Noncovalent interactions between alanine-rich domains |
| | HMW subunit | PGQGQQ<br>GYYPTSPQQ<br>GQQ | Disulphide between N- and C-terminal domains |
| 4 | Titin (I-band) | PPAKVPEVPKKPVPEEKVPVPVP<br>KKPEA | In Z-disc and M-line, mechanism unknown |
| 5 | Resilin | GGRPSDSYGAPGGGN<br>GYSGGRPGGQDLG | N/A |

(a) Group 1: glycine-rich; group 2: glycine- and proline-rich; group 3: glycine-, proline- and glutamine-rich; and group 4: proline-, glutamic acid-, valine- and lysine-rich. x is any amino acid.

All of the elastomeric proteins have incredible elasticity, allowing for high deformation without rupture when submitted to stress and a recovery of their original form when relaxed. This property is known as resilience and is dependent on both the lengths and properties of the elastic domains and the degree of cross-linking. It appears to result from a predominantly entropic mechanism, with a decrease in conformational entropy occurring on stretching due to the restricted number of low energy conformations that the extended polypeptide chains can adopt. On removal of the stretching force, there is an increase in the number of conformations the polypeptide chains can adopt, thus providing the free energy for elastic recoil. Furthermore, many of these proteins are characterized by extremes of strength and toughness with, for example, spider dragline silk displaying a toughness higher than Kevlar and even steel (Table 2). This is attributed to the close packing of chains allowed by the small side-chained amino acids in the repetitive domains as well as to extensive cross linking.

Table 2. Mechanical properties of natural and synthetic fibres[13.14].

| Material | Density (g/cm) | Strength (GPa) | Stiffness (GPa) | Extensibility (%) | Toughness (MJ/m³) |
|---|---|---|---|---|---|
| *Araneus diadematus* (spider) dragline silk | 1.3 | 1.1 | 10 | 27 | 180 |
| *Bombyx mori* (silkworm) silk | 1.3 | 0.6 | 7 | 18 | 70 |
| Elastin | 1.3 | 0.002 | 0.001 | 150 | 2 |
| Nylon 6.6 | 1.1 | 0.95 | 5 | 18 | 80 |
| Kevlar 49 | 1.4 | 3.6 | 130 | 2.7 | 50 |
| Steel | 7.8 | 1.5 | 200 | 0.8 | 6 |
| Wool (at 100% relative humidity) | 1.3 | 0.2 | 0.5 | 5 | 60 |
| Carbon fiber | 1.8 | 4 | 300 | 1.3 | 25 |
| Rubber | - | 0.05 | - | 850 | 80 |

Although fibrous proteins are found in many animals, only a few are known in detail, with the most studied including: abductin from arthropods, byssus from mussel, resilin from insects, silks from spiders and elastin from vertebrates[3.9.13.14].

In this thesis PBPs based on elastin and silk will be investigated and both of these fibrous proteins will now be discussed.

### 1.2.1. Elastin

The resilience of some tissues, such as blood vessels, lungs, tendons and skin of higher vertebrates, is due to elastin, a connective tissue protein that provides a combination of strength and flexibility to the extracellular matrix. Elastins are composed of simple repeating sequences such as the pentapeptide, VPGVG, the hexapeptide, APGVGV, the nonapeptide, VPGFGVGAG and the tetrapeptide VPGG (V = valine, P = proline, G = glycine, A = alanine)[12.15]. Of these, the most thoroughly investigated is the pentapeptide sequence VPGVG[3.8.9]. This

sequence is the most abundant in animal elastin and adapts a β-helix structure in the protein.

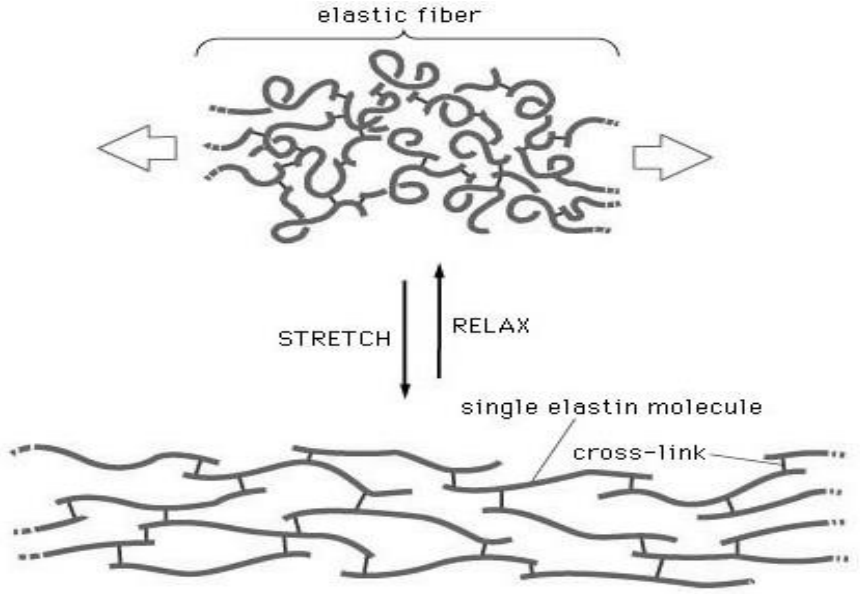



Figure 2. Illustration of stretch and relax state of elastin[15].

Elastin displays its mechanical properties only in the condensed, partially hydrated phase that is formed above the critical aggregation temperature. It is generally believed that the elasticity of elastin is an entropy-based property resulting from the hydrophobic nature of the protein[16].

### 1.2.2. Elastin-like Polymers

Elastin-like polymers (ELPs) are a new type of protein-based polymer derived from the repetitive amino acid sequence motifs of natural elastin. Here the repetitive elastin motifs are modified and/or combined in varying proportions to give rise to this new set of polymers with potentially novel mechanical and biological properties.

All functional ELPs show the so-called “inverse temperature transition” (ITT). This is a phase transitional behaviour in which, below a certain critical temperature (*Tt*), and in the presence of water, the polymer chains remain disordered. On the

contrary, above *Tt*, the polymer chain hydrophobically folds and assembles to form a phase separated state in which the chains adopt a dynamic, regular, non-random structure. In this folded and associated state, the chain loses essentially all of the ordered water structures of hydrophobic hydration[17.18].

All ELPs behave as thermo-responsible smart polymers. However, adequate substitution of the monomer amino acids by other, can be used to modify the properties and expand the responsive nature of this family of polymers. Indeed many different polymers have been designed which show the capacity to perform more than 15 pairwise free energy transductions involving the intensive variables of mechanical force, temperature, pressure, chemical potential, electrochemical potential and electromagnetic radiation[17.18.19.20.21].

### 1.2.3. Silk

Silk is the most thoroughly studied of the fibrous proteins. There are many forms of silk, of which that from *Bombyx mori* (Chinese silkworm silk) and dragline spider silk from *Nephila clavipes* (the golden orb weaver) have drawn most attention[22]. Spiders produce a large variety of silk types, each meant for a specific purpose, such as for the trapping of prey or for suspension. Indeed these different types of spider silk are examples of nature's high performance fibres with remarkable combinations of strength and toughness.

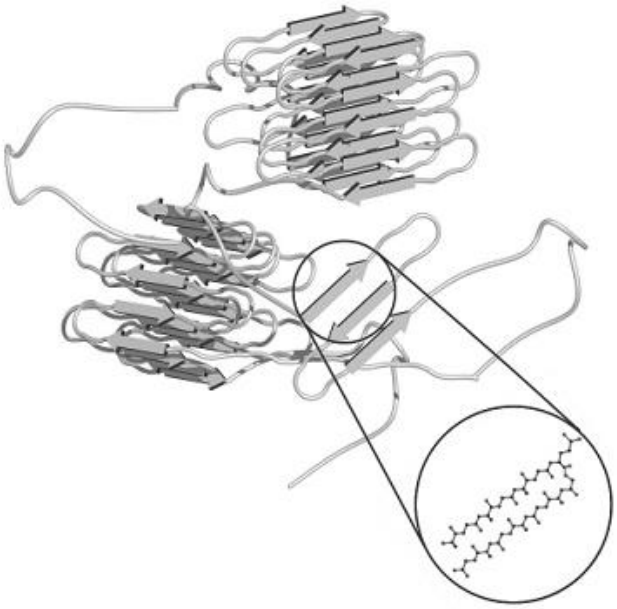


Figure 3. Illustration of Silk structure.

Silks from both silkworms and spiders contain repetitive sequences of crystalline (e.g. GAGAGS) and amorphous domains (e.g. GPGGx, GPGQQ, GPGGY) (Table 1 and Figure 3). The crystalline domains are β−sheet regular structures which are believed to be responsible for the strength of the material, whereas the amorphous protein matrix introduces flexibility and increases the energy to break while also allowing the crystalline domains to orient under strain and thereby allow for increased strength.

### 1.2.4. Silk-Elastin-like Proteins

Silk-elastin-like proteins (SELPs) are a family of biopolymers based on the highly repetitive amino acid sequence blocks of the naturally occurring fibrous proteins silk and elastin. They are diblock copolymers which combine the physicochemical and biological properties of the high tensile strength silk with highly resilient elastin, allowing for the fabrication of diverse materials with a high potential for use in the pharmaceutical, regenerative medicine and materials fields. Indeed, by varying the content and/or ratio of silk and elastin, SELPs can be fine tuned to give various polymers of diverse mechanical and biological properties. In fact, by

this means one can vary the strength, flexibility, solubility, biodegradability and/or biocompatibility of the material produced[1]. For example, a high content of silk β-sheet crystals gives rise to a polymer of high thermal and chemical stability but with reduced aqueous solubility. On the other hand, the periodic inclusion of the elastomeric sequence would reduce the overall crystallinity and increases its flexibility, aqueous solubility and biocompatibility. Hence one can fine tune the SELP with the desired properties to accomplish a particular function[23.24.25].

Some examples of potential applications for SELPs include: controlled release systems for intracellular, intratumoral and gastrointestinal drug and gene delivery, in tissues regeneration as a support for the creation and repair of new tissues, in contact lens and synthetic corneas, and also in the automobile industry as environmentally friendly and biodegradable plastics[25.26.27]. While SELPs and indeed PBPs in general have a high potential for application and offer many advantages over currently used petroleum based polymers, they are currently characterised by a high production cost which constitutes a major obstacle to the commercial viability of these. Indeed, the principal problem with the production of SELPs is their low production levels. They are produced by heterologous production with, up to now, *E. coli* being the most commonly used production host and with only mg/L productivities being reported[1.24].

Hence a current major area of interest in the study of PBPs is in the improvement of the production levels and reduction of the production costs of these. This present project is focused on an attempt to optimise the production levels of a novel recently developed SELP[28.29].

# 1.3. Background

A number of novel SELPs consisting of multiple blocks of the silkworm silk consensus sequence GAGAGS in various combinations with a variant (VPAVG) of the natural mammalian elastin repetitive sequence block VPGVG have recently been synthesised and produced by the team where this work was developed[24]. The single amino acid change (G to A) in the elastin motif has been found to have an effect on its mechanical properties, including a change in its mechanical response from elastic to plastic deformation, a two orders of magnitude increase in its young modulus value and the revelation of a hysteresis behaviour accompanying its reversible transition on heating. Hence, these novel SELPs incorporating this unique elastin-like repetitive motif (VPAVG) with that of silk (GAGAGS) should give rise to a new set of polymers with expanded and potentially enhanced properties and applicability[1.24].

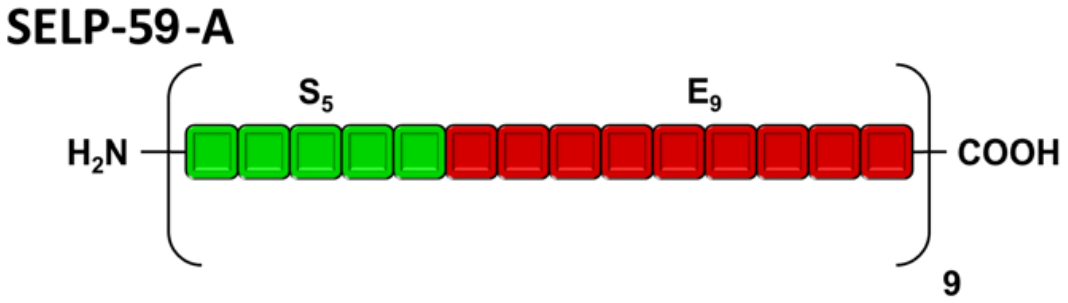



Figure 4. Schematic representation of silk-elastin-like protein (SELP-59-A) used in this study[24].

In the present study we will investigate one of these novel SELPs, namely SELP-59-A which contains five blocks of the silk motif, GAGAGS and nine blocks of the elastin-like sequence, VPAVG. The aim of the project will be to optimise the production levels of this novel SELP.

*SELP-59-A* has already been subcloned into pCM13 (a modified pET-25b(+) expression vector) and produced in *E. coli* BL21 (DE3)[24]. This is an IPTG (β-D-1-thiogalactopyranoside)/lactose inducible system in which a highly active T7 RNA polymerase is used and in which the target gene expression is under the control of a T7 RNA polymerase promoter (Figure 5). The T7 promoter is only

recognised by T7 RNA polymerase which is able to transcribe genes five times faster than the *E. coli* RNA polymerase[30.31]. The gene encoding the T7 RNA polymerase, under the control of a *lacUV5* promoter-operator, has been inserted in the chromosome of the host bacterium *E. coli* BL21(DE3). Furthermore, a *lacI* gene encoding a lac repressor has been inserted in both the bacterial chromosome and the plasmid, thereby ensuring high amounts of repressor molecules. Therefore, under non-inducing conditions, expression of the T7 RNA polymerase and hence also the target protein is repressed by the binding of the lac repressor to the *lacO* operator sequences upstream of both the T7 RNA polymerase and target genes. On the other hand, adding isopropyl-*β*-*D*-*thiogalactoside* (IPTG) or lactose to the medium will allow for the transcription of the T7 RNA polymerase and target protein by inhibiting lac repressor binding to the *lacO* sequences (Figure 5) and thereby lead to high level production of the protein of interest[32.33].

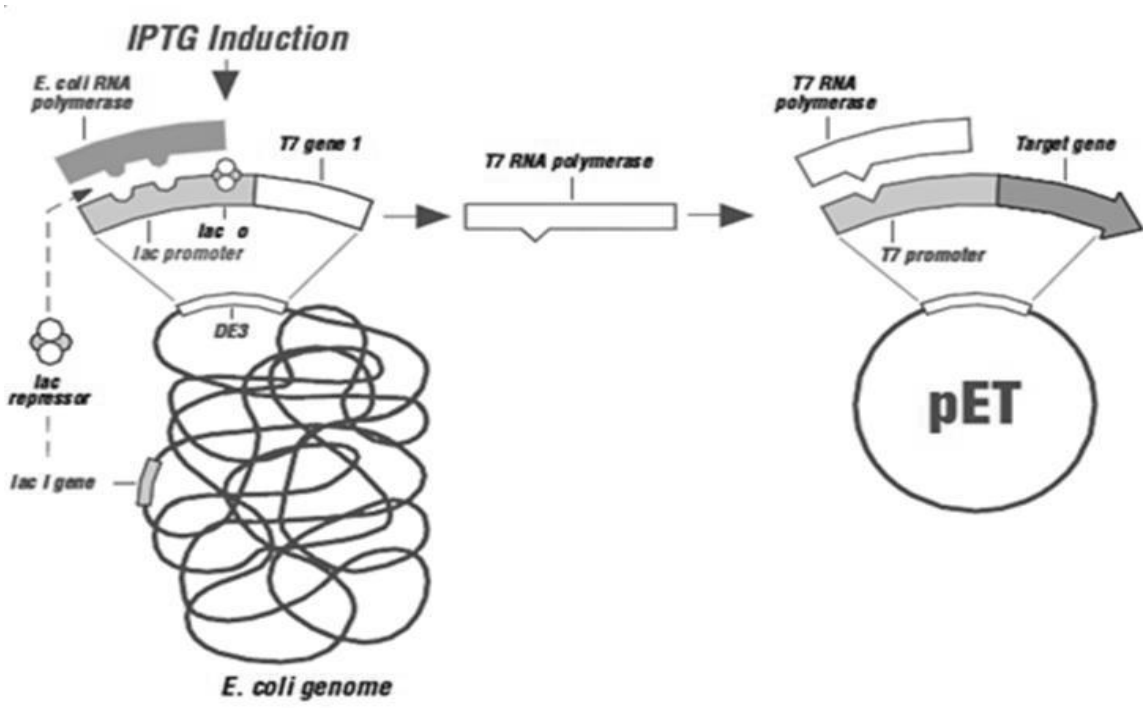



Figure 5. Representation of *E. coli* BL21(DE3)/pET expression system[33].

Previous studies with SELPs reported volumetric productivities of up to only 20 mg/L when using rich media for batch production in *E. coli* with the *lacI* regulated T7 system[33]. More recently, a systematic empirical optimisation of all

process parameters for the batch production in shake flasks of *SELP-59-A* with the pCM13-*E. coli* BL21(DE3) system allowed for the production of approximately 500 mg/L after purification[1]. Here, rich, buffered Terrific Broth (TB) at 37 ºC and pH 6.8, and with high agitation and oxygenation as well as induction early in the stationary phase allowed for maximum production[1]. However, analysis of the optimised process indicated that the selection agent (ampicillin) was rapidly degraded during the first hour of cultivation at a dry cell weight (DCW) of about 0.1 g/L and that consequently plasmid stability decreased dramatically on induction, thereby reducing polymer production levels. Furthermore, acetate was found to accumulate during the bioprocess to levels which were shown to be inhibitory to the host cells and hence reduced the maximum attainable cell density and hence also further reducing polymer levels.

Acetate accumulation during *E. coli* growth occurs when the carbon flux into the cells exceeds the capacity of the central metabolic pathway and in aerobic fermentations is typically a result of an uncontrolled cell growth in the presence of excess carbon source. Use of the fed-batch approach for the controlled growth of *E. coli* allows for the control and reduction of acetate accumulation and initial studies have already been carried out to investigate this approach for the production of SELP-59-A. With the fed-batch approach, controlled addition of a limiting essential ingredient (e.g. the carbon source, glucose) enables control of the growth rate and consequently also the rate of acetic acid production. The fed-batch production approach is typically composed of three phases: a batch phase, a pre-induction fed-batch phase and a post-induction fed-batch phase. In the batch phase, *E. coli* is allowed to grow in the presence of excess nutrients until all carbon sources (i.e. glucose and all organic acids produced during the process) are exhausted and allows for the attainment of an stable initial culture density before the fed-batch phase. During the fed-batch phase the limiting carbon source is fed at a desired rate so as to maintain cell growth at a desired rate as defined by the mass balance equation for cell growth[34.35]. Finally, when the cells have reached a desired cell density, protein production is induced by addition of IPTG or lactose and feeding is continued post induction until the fermentation is stopped. Initial investigations of the fed-batch approach with the *SELP-59-A*/pCM13/*E. coli* BL21(DE3) system using growth rates of 0.2, 0.4 or

0.6 $h^{-1}$ both pre and post-induction allowed for the successful reduction of acetate accumulation levels and hence enabled high cell density fermentations (approx. 50 g/L DCW, approximately 10 times higher than with the batch production approach) and an accompanying improved polymer production. Induction at a DCW of 50 g/L allowed for the obtention of approximately 3 g/L of SELP-59-A after purification. However, this optimisation is incomplete. No attempts were made to optimise or identify the optimal pre and post induction growths rates which would allow for improved production levels or reduced process times, furthermore, the maximum cell density and optimal inducer concentration for maximising production were not investigated. Therefore, in the present study, we attempted to address this.

As already mentioned, a second problem noted for the batch production process was plasmid instability on induction, with total loss of plasmid being observed 2 hours after induction. The initial studies with the fed-batch approach also highlighted this problem with again rapid plasmid loss being noted on induction. Plasmid containing host cell resistance to ampicillin is achieved by the production of plasmid encoded β-lactamase which degrades the antibiotic and hence allows for the host cell survival. However, during cell growth the extracellular levels of this can accumulate to levels which can result in total degradation of all ampicillin present, thereby eliminating the selection pressure and thereby also facilitating the rapid plasmid loss observed as a result of the increased metabolic burden placed on the cells during induction. In the present study we will attempt to overcome the limitation of plasmid instability by investigating the use of kanamycin as the selection agent and by investigating the use of a toxin/antitoxin post segregational suicide system.

Kanamycin acts by interacting with the 30S subunit of prokaryotic ribosomes and inhibit translation. A commonly used system for resistance to this antibiotic is by the intracellular production of a phosphotransferase (Neomycin phosphotransferase II) which inactivates the antibiotic by converting them to the inactive 3'-phosphate form. Here it is hypothesised that the intracellular location and action of this enzyme may lead to an improved persistence of kanamycin in the medium. Furthermore, in the vector used in the present study, the kanamycin resistance gene is in the opposite orientation to the T7 promoter, so induction of

the T7 promoter should not result in an increase in kanamycin gene product. This property can reduce the metabolic burden and may lead to a positive effect on protein production[33].

Toxin/antitoxin post segregational suicide systems are composed of two or more genes encoding a toxin and another gene encoding a respective antidote. Cells producing the toxin cannot survive unless the antitoxin is introduced and therefore plasmids containing the gene for the antitoxin will be preferentially maintained. Only those cells with a plasmid allowing for antitoxin production and their daughter cells that inherit the plasmid will survive i.e. the plasmid confers an evolutionary advantage to cells. Toxin/antitoxin post segregational suicide systems are divided into three Types. Type I is based on antitoxin inactivation by the base pairing of an RNA antitoxin with toxin mRNA. With Type II, used in this study, an antitoxin protein strongly binds a toxin protein, thereby inhibiting it. Type III has not been well studied but is believed to inactivate the toxin protein directly with antitoxin RNA. Here we will look at a Type II toxin/antitoxin system, namely the ccdB/ccdA poison/antidote system, for improved plasmid stability during *SELP-59-A* production. The StabyExpress 1.2 kit will be used, here a toxin gene (*ccdB*) has been introduced in the chromosome of an expression host (*E. coli* SE1) and the antidote gene (*ccdA*) under the control of a constitutive promoter has been introduced in the expression vector (pStaby 1.2) (Figure 6)[36].

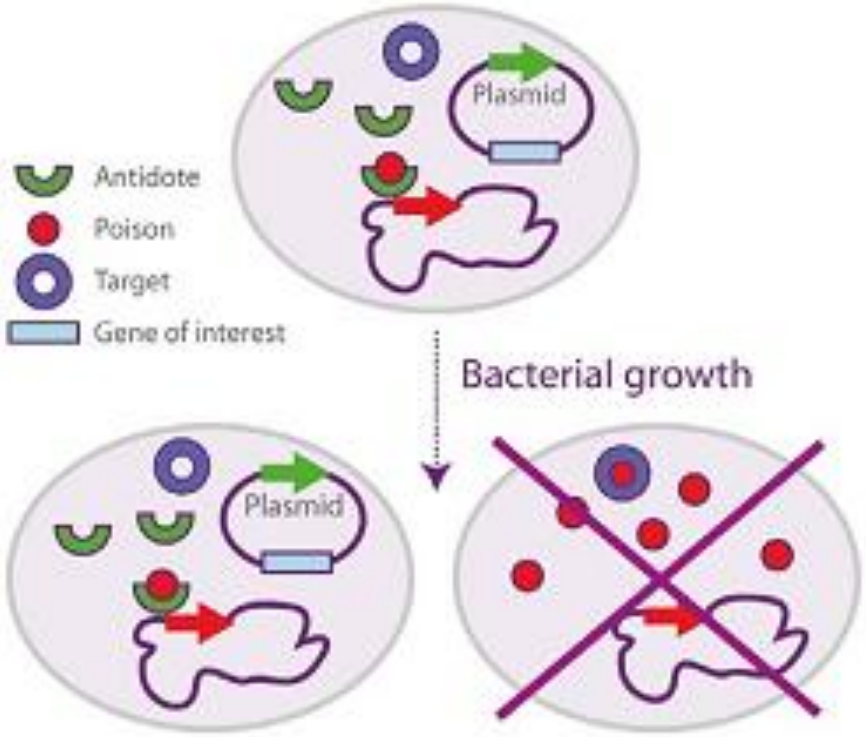



Figure 6. Representation of pStaby Express toxin/antitoxin system[36].

The pStaby 1.2 expression vector used in this study for the investigation of toxin/antitoxin post segregational suicide systems for plasmid stabilisation is also characterised by the constitutive production of a β-lactamase. While this allows for a simplified selection and screening of cells containing this vector, it also places an increased metabolic burden on the cells. Here we will examine the effect of removing the promoter for this gene, thereby preventing β-lactamase production, as well as the effect of removing both the promoter and gene, thereby preventing β-lactamase production and also reducing vector size[37].

## 1.4. Objectives

The principal objective of this study is to maximise production levels and reduce the production process time for *SELP-59-A*. To achieve this, the work was divided into two main sections:

- Optimisation of the fed-batch process
    - Optimise pre-induction growth rate
    - Optimise post-induction growth rate
    - Determine the maximum achievable cell density with the fermentation system used
    - Optimise IPTG concentration for maximum SELP-59-A production

- Overcoming plasmid instability
    - Examine and compare the use of various plasmid stabilisation systems: ampicillin resistance ($amp^R$), kanamycin resistance ($kan^R$) and the ccdB/ccdA toxin/antitoxin post segregational suicide system
    - Reduce the metabolic burden on the cell by elimination of ampicillin production from the production plasmid
        - Examine the effect of removal of the $Amp^R$ promoter (*SELP-59-A*/pStaby Δ*blap*)
        - Examine the effect of removal of the whole $Amp^R$ gene and promoter (*SELP-59-A*/pStaby Δ*bla*)

# 2. Materials and Methods

The present project is divided into two main sections:

2.1) Optimisation of the fed-batch process

2.2) Overcoming plasmid instability

# 2.1. Optimisation of the fed-batch process

For optimisation of the high cell density fed-batch production of *SELP-59-A* the *SELP-59-A*/pCM13/*E. coli* BL21(DE3) expression host as provided for the study was used. The variables investigated were: pre-induction growth rate (0.2 – 0.8 $h^{-1}$), post-induction growth rate (0.1 – 0.6 $h^{-1}$), dissolved oxygen ($dO_2$) concentration (15 – 55 %), cell density (50 – 100 g/L DCW) and inducer concentration (1 – 6 mM IPTG). Comparative fed-batch studies, with a complete on-line (pH, temperature, $dO_2$, agitation), off-line (*SELP-59-A* production levels, glucose, organic acids, phosphate and nitrogen levels, total number of viable cells, plasmid stability) and at-line (biomass concentration) analysis of all process parameters, were carried out and compared for all variables investigated.

## 2.1.1. Fed-Batch Production of SELP-59-A

Fermentation was carried out in a BioFlo 110 Modular Benchtop 3 litre Fermentor (New Brunswick Scientific-NBS) with centralised monitoring, control and data collection via a BioCommand General Purpose Multi-Loop controller (NBS) and the BioCommand Plus 3.3 software (NBS).

Two precultures were used in the preparation of the inoculum for the fermentation culture; a first preculture in 50 mL lysogeny broth with ampicillin (LB) and a second preculture in 100 mL of a modified M9 minimal medium with ampicillin (MMLBM). This latter medium, MMLBM, was based on the minimal media described in[38.39]. A mid log phase pre-culture was used to inoculate 900 mL of the MMLBM fermentation culture to an initial dry cell weight (DCW) of 0.05 g $L^{-1}$.

Temperature was maintained at 37 ºC, pH at 6.8 with 25 % $NH_4OH$ and 3 M $H_3PO_4$, and dissolved oxygen ($dO_2$) at 25 % for the batch phase and 15 - 55 % for the fed-batch phase by cascading to agitation (200 - 1100 rpm) and $O_2$ (only at fed-batch phase). The air flow rate was maintained at 5 L $min.^{-1}$ and foaming was controlled with antifoam Y-30 emulsion (Sigma).

The fed-batch phase was started when all the carbon sources present had been exhausted after an overnight batch phase and as indicated by a sustained increase in the $dO_2$ levels above 50 %. During the fed-batch phase, the growth rate was controlled at a desired rate by the glucose feeding rate as determined by the mass balance equation[39]:

$$\mathrm{M}_S(\mathrm{t}) = \mathrm{F}_F(\mathrm{t})\mathrm{S}_\mathrm{F}(\mathrm{t}) = \left(\frac{\mu(t)}{Yx/s} + m\right) V(t)X(t) \qquad (1)$$

where $M_S(t)$ is the mass feeding rate of the glucose substrate (g $h^{-1}$) at time t, $F_F$ is the volumetric feeding rate of the feed solution (L $h^{-1}$), $S_F$ the glucose concentration of the feed solution (g $h^{-1}$), μ the desired specific growth rate ($h^{-1}$), $Y_{X/S}$ the yield coefficient (0.4 g $g^{-1}$), m the specific maintenance coefficient (0.025 g $g^{-1}$ $h^{-1}$), V the cultivation volume (L) and X the biomass concentration (g $L^{-1}$).

The actual biomass concentration was measured every hour and the culture volume calculated from a determination of the volumes sampled, the acid or base added and the feeding rate, and used to calculate the required rate of glucose feeding using equation (1). Furthermore, every 15 minutes, the feeding rate was calculated from equation (1) and from estimates of the biomass concentration and culture volume as follows

$$X(t)V(t) = Xt_F Vt_F e^{\mu(t-t_F)} \qquad (2)$$

where $t_F$ is the time when the biomass concentration and culture volume were last measured, X(t) is the biomass concentration at time t, V(t) the culture volume at time t, $Xt_F$ is the biomass concentration at $t_F$ and $Vt_F$ is the culture volume at time $t_F$.

In addition to a glucose feed, a second feeding solution of 600 mM phosphate and 1200 mM nitrogen was also applied so as to maintain phosphate at

concentrations between 50 and 300 mM and nitrogen between 50 and 100 mM throughout the fermentation.

Induction of SELP-59-A production was carried out with filter sterilised isopropyl β-D-1-thiogalactopyranoside (IPTG) with an induction period of 4 hours.

Fed-batch Protocol: See Annexe I

#### 2.1.1.1. Analytical Methods for Monitoring of Fed-batch process

At-line (biomass concentration) and off-line (SELP-59-A production levels, glucose, organic acids, phosphate and nitrogen levels, total number of viable cells, plasmid stability) analyses of process parameters were carried out and compared for each fed-batch production. Cultures were aseptically sampled every hour and used directly for analysis of dry cell weight (DCW), optical density measurements ($OD_{600nm}$), enumeration of total viable cell numbers and plasmid stability. Extracellular samples were taken every hour and maintained at -20 ºC until analysis of sugar and organic acid content (by HPLC) as well as phosphate concentration and nitrogen concentration. Cell pellets from 0.5 mL culture samples were collected before induction and every hour thereafter and stored at -20 ºC until SDS-PAGE analysis.

*The optical density* at 600nm was measured on a Genesys 20 spectrophotometer (ThermoSpectronics).

*Biomass dry cell weights* (DCW, g/L) were determined from the weights of washed pellets of 2 mL culture samples dried overnight at 70 °C.

*Organic acids (citric, tartaric, malic, succinic, lactic, acetic and formic), carbohydrate (glucose) and ethanol levels* in culture supernatants collected every hour were monitored using a Rezex™ 8 µm ROA-organic acid H+(8%) high performance liquid chromatography column (Phenomenex). 2.5 mM $H_2SO_4$ was used for the mobile phase, the column was maintained at 60°C and detection was

by refractive index measurement with an Elite LaChrom L-2490 RI detector (VWR Hitachi) at 40°C. An Elite LaChrom (VWR Hitachi) chromatography system was used with the EZChrom Elite 3.3.2 SP2 software for data collection and analysis. Culture supernatant samples were treated with 10 % TCA to remove protein contaminants, centrifuged for 5 minutes at maximum speed and the supernatant filtered through a 0.22 µm filter before HPLC analysis. The HPLC protocol used is detailed in Annexe II.

To ensure that adequate phosphate was present throughout the fermentations, the *phosphate concentration* in the culture supernatants of samples taken every hour was monitored. The phosphate assay described in[40] was used, see Annexe III for details.

*Ammonia-nitrogen concentration* in the culture supernatants was determined by the Berthelot colour reaction as previously described[41], see Annexe IV for details.

*Viable cell numbers* were determined by serially diluting culture samples in sterile phosphate buffered saline (8 g/L NaCl, 0.2 g/L KCl, 1.44 g/L $Na_2HPO_4$, 0.24 g/L $KH_2PO_4$, pH 7.4) and spread plating 25 µL of the $10^{-5}$ – $10^{-7}$ dilutions in duplicate on LB agar plates (10 g/L bacto tryptone; 5 g/L yeast extract, 5 g/L NaCl, 20 g/L agar). After incubation overnight at 37 ºC, cell numbers were enumerated.

*Plasmid stability in viable cells* was monitored by repicking 100 isolated colonies from the LB plates to LB + ampicillin plates (10 g/L bacto tryptone; 5 g/L yeast extract, 5 g/L NaCl, 20 g/L agar, 100 µg/mL ampicillin) and incubating overnight at 37 ºC. Plasmid stability is expressed as the percentage of the cells growing on LB which are able to grow on the antibiotic containing media.

Intracellular (from cell pellets of 0.5 mL cultures) and extracellular *SELP-59-A production levels were monitored* by means of SDS-PAGE with a 10 % SDS-PAGE gel and staining with 3M $CuCl_2$. See Annexe V for details of the protocol used.

*Purified SELP-59-A production levels* for the various fed-batch productions was determined from the weight of purified polymer. The purification protocol[24] is described in detail in Annexe VI. Briefly, the contaminating *E. coli* proteins were

removed by acid pH treatment (pH 3.5) and *SELP-59-A* then precipitated from the supernatant by treatment with 22 % ammonium sulphate. Following extensive dialysis in water the purified *SELP-59-A* was then lyophilised and weighed.

### 2.1.2. Effect of growth rate on fed-batch process

Initial attempts in developing the fed-batch process involved the investigation of constant pre- and post-induction growth rates (μ) of 0.2, 0.4 and 0.6 $h^{-1}$. The process was carried out as described in 2.1.1 above and analysed as described in 2.1.1.1. The dissolved oxygen concentration was maintained at 35 % during the fed-batch phase and induction was with 3 mM IPTG at a DCW of 50 g/L.

Table 3. Conditions used for optimisation of the growth rate.

| **Pre-Induction Growth Rate** | **Post-Induction Growth Rate** | **$dO_2$ (%)** | **Induction DCW** | **[IPTG] (mM)** |
|---|---|---|---|---|
| **0.2** | **0.2** | 35 | 50 | 3 |
| **0.4** | **0.4** | 35 | 50 | 3 |
| **0.6** | **0.6** | 35 | 50 | 3 |

### 2.1.3. Optimisation of post-induction growth rate

Fed-batch productions were carried out as described in 2.1.1 above with pre-induction growth rates of 0.2 and 0.6 $h^{-1}$ and post-induction growth rates (μ) of 0.1, 0.2, 0.4 or 0.6 $h^{-1}$. All process parameters were monitored as described in 2.1.1.1 above and each fermentation was repeated at least two times. The dissolved oxygen concentration was maintained at 35 % during the fed-batch

phase and induction was carried out at a dry cell weigh of 50 g/L with 3 mM IPTG in all cases.

Table 4. Conditions used for optimisation of the post-induction growth rate

| **Pre-Induction Growth Rate ($h^{-1}$)** | **Post-Induction Growth Rate ($h^{-1}$)** | **$dO_2$ (%)** | **Induction Dry Cell Weight (g/L)** | **[IPTG] (mM)** |
|---|---|---|---|---|
| 0.2/0.6 | **0.1** | 35 | 50 | 3 |
| 0.2/0.6 | **0.2** | 35 | 50 | 3 |
| 0.2/0.6 | **0.4** | 35 | 50 | 3 |
| 0.2/0.6 | **0.6** | 35 | 50 | 3 |

### 2.1.4. Optimisation of pre-induction growth rate

The effect of the pre-induction growth rate on the fed-batch process was investigated by carrying out the fed-batch process as described in 2.1.1 above with the desired specific growth rate (μ) set at 0.2, 0.4, 0.6 or 0.8 $h^{-1}$ during the fed-batch phase prior to induction. Following induction at a DCW of 50 g/L with 3 mM IPTG the desired specific growth rate was reduced to 0.1 $h^{-1}$. The dissolved oxygen concentration was maintained at 35 % during the fed-batch phase, all process parameters were monitored as described in 2.1.1.1 above and each fermentation was repeated at least two times.

Table 5. Conditions used for optimisation of the pre-induction growth rate

| **Pre-Induction Growth Rate** | **Post-Induction Growth Rate** | **$dO_2$ (%)** | **Induction DCW** | **[IPTG] (mM)** |
|---|---|---|---|---|
| **0.2** | 0.1 | 35 | 50 | 3 |
| **0.4** | 0.1 | 35 | 50 | 3 |
| **0.6** | 0.1 | 35 | 50 | 3 |
| **0.8** | 0.1 | 35 | 50 | 3 |

### 2.1.5. Optimisation of dissolved oxygen concentration

The effect of the dissolved oxygen concentration on the fed-batch process was investigated by carrying out the fed-batch process as indicated in 2.1.1 with dissolved oxygen concentration during the fed-batch phase being maintained at 15, 35 or 55 % and monitoring and comparing all process variables as indicated in 2.1.1.1 above. A pre-induction growth rate of 0.4 $h^{-1}$ and post-induction growth rate of 0.1 $h^{-1}$ was used and induction was carried out at a dry cell weigh of 50 g/L with 3 mM IPTG.

Table 6. Conditions used for optimisation of the dissolved oxygen concentration

| **Pre-Induction Growth Rate ($h^{-1}$)** | **Post-Induction Growth Rate ($h^{-1}$)** | **$dO_2$ (%)** | **Induction Dry Cell Weight (g/L)** | **[IPTG] (mM)** |
|---|---|---|---|---|
| 0.4 | 0.1 | **15** | 50 | 3 |
| 0.4 | 0.1 | **35** | 50 | 3 |
| 0.4 | 0.1 | **55** | 50 | 3 |

## 2.1.6. Optimisation of the dry cell weight at induction

The fed-batch process was carried out as described in 2.1.1 with a pre-induction growth rate of 0.4 $h^{-1}$ - 0.1 $h^{-1}$, a post-induction growth rate of 0.1 $h^{-1}$, $dO_2$ set at 35 % during the fed-batch phase and with induction with 3 mM IPTG at dry cell weighs of 50, 75 and 100 g/L. The lower pre-induction growth rates were used at the higher dry cell weights in an attempt to maintain $dO_2$ levels above 0 %. All process variables were monitored as indicated in 2.1.1.1 above.

Table 7. Conditions used for optimisation of the dry cell weight at induction

| **Pre-Induction Growth Rate ($h^{-1}$)** | **Post-Induction Growth Rate ($h^{-1}$)** | **$dO_2$ (%)** | **Induction Dry Cell Weight (g/L)** | **[IPTG] (mM)** |
|---|---|---|---|---|
| 0.4 | 0.1 | 35 | **50** | 3 |
| 0.4 | 0.1 | 35 | **75** | 3 |
| 0.4 | 0.1 | 35 | **100** | 3 |

## 2.1.7. Optimisation of IPTG concentration

Induction at a DCW of 75 g/L with 1, 3 or 6 mM IPTG was investigated with the fed-batch process described in 2.1.1 above, with a pre-induction growth rate of 0.4 $h^{-1}$ and post-induction growth rate of 0.1 $h^{-1}$. The oxygen concentration was maintained at 35 % during the fed-batch phase and all process variables were monitored as indicated in 2.1.1.1.

Table 8. Conditions used for optimisation of the IPTG concentration for induction

| **Pre-Induction Growth Rate ($h^{-1}$)** | **Post-Induction Growth Rate ($h^{-1}$)** | **$dO_2$ (%)** | **Induction Dry Cell Weight (g/L)** | **[IPTG] (mM)** |
|---|---|---|---|---|
| 0.4 | 0.1 | 35 | 75 | **1** |
| 0.4 | 0.1 | 35 | 75 | **3** |
| 0.4 | 0.1 | 35 | 75 | **6** |

## 2.2. Overcoming plasmid instability

For overcoming problems with plasmid instability observed during part 2.1 of this study we investigated various means of plasmid stabilisation and their effect in the fed-batch process for SELP-59-A production. Various vector constructs were prepared by molecular biology approaches (restriction digestion, fragment purification and ligation), transformed to their appropriate expression hosts and their use in the fed-batch production of SELP-59-A investigated and compared using the standard fed-batch approach described in 2.1.1 above. A complete on-line (pH, temperature, $dO_2$, agitation), off-line (SELP-59-A production levels, sugar, organic acid, phosphate and nitrogen levels, total number of viable cells, plasmid stability) and at-line (cell density) analysis of all process variables as described in 2.1.1.1 was carried out and compared in all cases.

Four different means of plasmid stabilisation were investigated:

- ampicillin resistance (control, pCM13/*E. coli* BL21(DE3) as used in 2.1 above),
- kanamycin resistance (pET29a(+)/*E. coli* BL21(DE3))
- *ccdB/ccdA* toxin/antitoxin post segregational suicide system (pStaby/*E. coli* SE1)
- reduction of the metabolic burden on the cell by elimination of ampicillin production from the production plasmid (pStaby Δ*blap*/*E. coli* DE1 and pStaby Δ*bla*/*E. coli* SE1).

See Tables 9 and 10 for details of plasmids and hosts used.

Table 9. Details of plasmids used in this study

| **Plasmid** | **Selection Marker** | |
|---|---|---|
| **pCM13/*SELP-59-A*** | Ampicillin | Modified pET-25b(+) (Novagen) vector containg the gene for SELP-59-A. Provided for this study[24]. |
| **pET29a(+)** | Kanamycin | pET Expression System (Novagen)) |
| **pStaby 1.2** | Ampicillin and *ccdB/ccdA* | Modified pET-21b(+) vector containing gene for antitoxin *ccdA* (Delphi Genetics). |
| **pStaby 1.2 Δ*blap*** | *ccdB/ccdA* | pStaby 1.2 with promoter for ampicillin resistance gene removed. Prepared in this study |
| **pStaby 1.2 Δ*bla*** | *ccdB/ccdA* | pStaby 1.2 with ampicillin resistance gene and promoter removed. Prepared in this study |

Table 10. Details of *E. coli* strains used in this study

| Strains | Information | |
|---|---|---|
| **XL1Blue** | *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F' proAB$^+$ lac L$^q$ZΔM15 Tn 10(Tet$^r$)]* | Used as cloning host for pET vector constructs |
| **BL21(DE3)** | *F- ompT gal dcm lon hsdSb($r_b$-$m_b$-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 Sam7 nin5])* | Used as expression host for pET vector constructs |
| **BL21(DE3)/pCM13 *lSELP-59-A*** | BL21(DE3) transformed with expression vector pCM13 with *SELP-59-A* gene inserted[24] | Provided for this study[24] |
| **CYS21** | *F-, Cm$^R$, mcrA, endA1, Δ(mrr-hsdRMS-mcrBC) (*restriction-, modification-), *Φ80lacZΔM15, ΔlacX74, recA1, Δ(ara, leu)7697, araD139, galU, galK, nupG, rpsl, ccdB$^+$* | Cloning host for pStaby vector constructs. Contains chromosomal copy of gene for *ccdB* toxin. |
| **SE1** | *F-, Cm$^R$, ompT, lon, hsdS$_B$*(restriction-,modification-), *gal, dcm,DE3 (lacI*, T7 polymerase under the control of the *PlacUV5* promoter), *ccdB$^+$* | Variant of *E. coli* BL21(DE3) containing chromosomal copy of gene for *ccdB* toxin. Expression host for pStaby vector constructs. |

### 2.2.1. Plasmid Isolation

*SELP-59-A*/pCM13 plasmid was isolated and purified from *SELP-59-A*/pCM13/*E. coli* BL21(DE3) by use of the GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich) as recommended by the manufacturer (see Annexe VII). Purified plasmids pET29a(+) (Novagen) and pStaby 1.2 (Delphi Genetics) were obtained from the manufacturers.

### 2.2.2. Plasmid Quantification

Plasmid concentration was determined by absorbance measurements at 260 nm using a NanoDrop ND 1000 spectrophotometer (Thermo Scientific) as recommended by the manufacturer (see Annexe VIII for protocol details). The nucleic acid quantification is made using the Beer-Lambert equation, modified to use an extinction coefficient with units of ng-cm/µL (50 ng-cm/µL for double-stranded DNA).

The 260 nm/280 nm and 260 nm/230 nm absorbance values ratios were used to determine sample purity and the absence of protein or other organic compound contamination, respectively.

### 2.2.3. Restriction Digestion

Restriction digestion was used in the construction of *SELP-59-A*/pET29a(+), and *SELP-59-A*/pStaby 1.2 (see Figure 7 for an illustration of the approach used) and in the removal of the $Amp^R$ promoter as well as the whole $Amp^R$ gene and promoter from *SELP-59-A*/pStaby 1.2 (see Figure 8 for an illustration of the approaches used).

Restriction digestion with *Nde*I (Thermo Scientific) and *Blp*I (Thermo Scientific) was used to remove the *SELP-59-A* gene from *SELP-59-A*/pCM13 before

insertion into similarly digested pET29a(+) and pStaby 1.2. The double digestion was carried out at 37 ºC in Tango buffer via a two step protocol as described in Annexe IX.

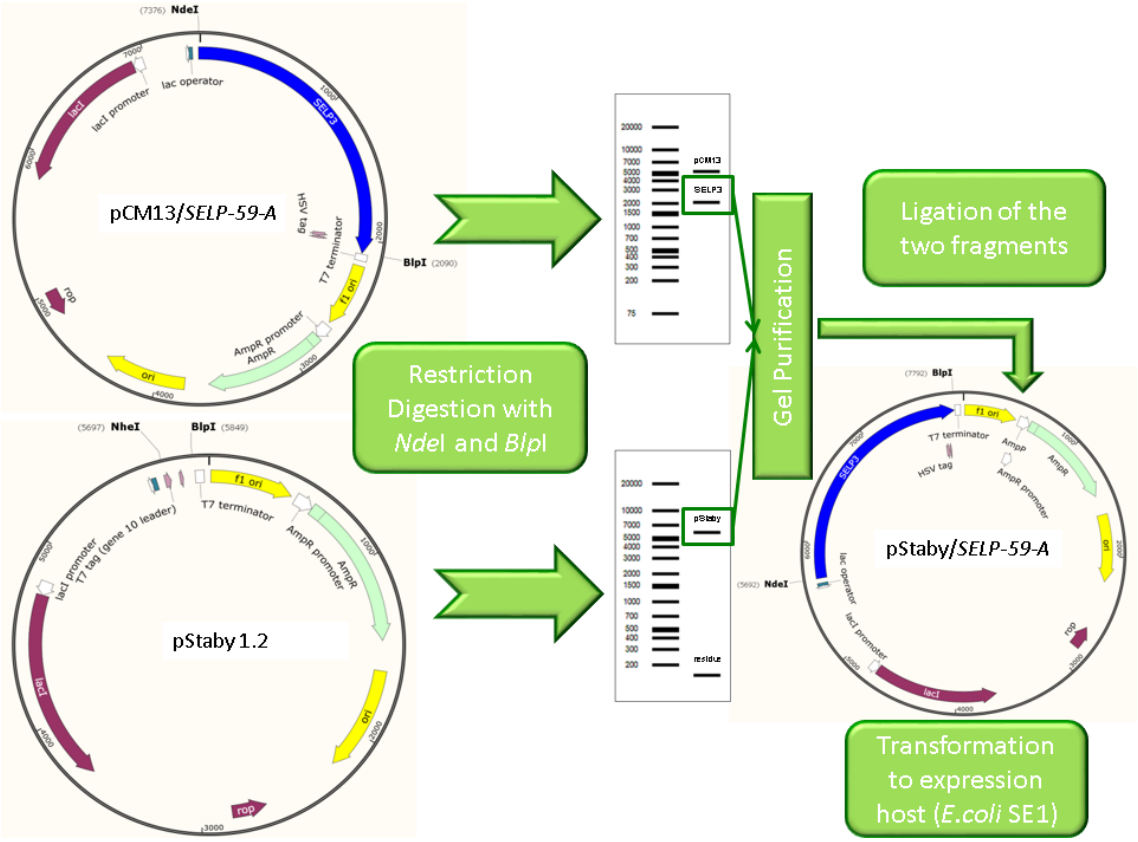



Figure 7. Schematic illustration of the approach used for preparation of *SELP 59 A*/pStaby.

For removal of the Amp$^R$ promoter from *SELP-59-A*/pStaby 1.2 (i.e. preparation of *SELP-59-A*/pStaby Δ*blap*) double digestion with *Ssp*I (Thermo Scientific) was carried out as recommended by the manufacturer. See Annexe IX for details of the protocol used.

For removal of both the Amp$^R$ gene and promoter from *SELP-59-A*/pStaby 1.2 (i.e. preparation of *SELP-59-A*/pStaby Δ*bla*) double digestion with *Ssp*I (Thermo Scientific) and *Eam*1105I (Thermo Scientific) was used. See Annexe IX for details of the protocol used.

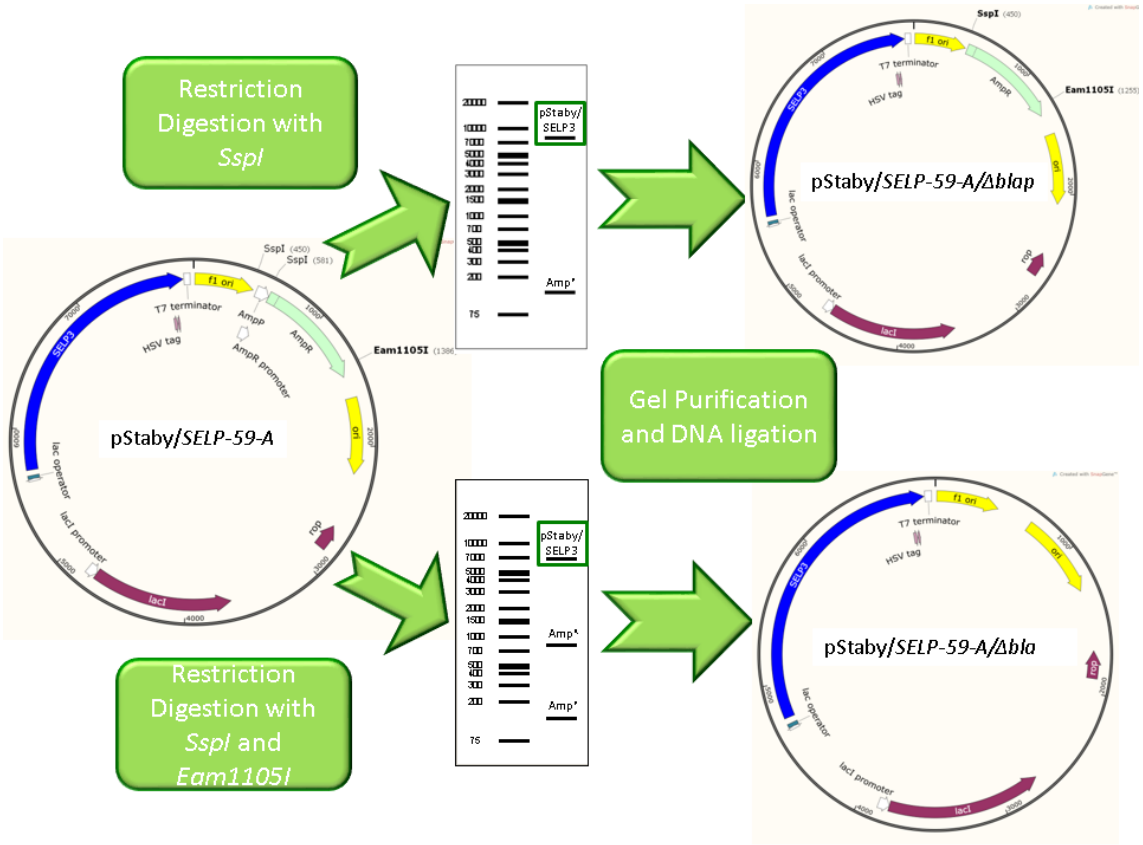



Figure 8. Schematic illustration of the approach used for preparation of *SELP 59 A*/pStaby Δ*blap* and *SELP-59-A*/pStaby Δ*bla*.

## 2.2.4. Fragment Purification

Following restriction digestion of the *SELP-59-A*/pCM13, pET29a(+), pStaby 1.2 and *SELP-59-A*/pStaby 1.2 vectors, the samples were then run on a 1.5 % agarose gel so as to allow for separation and purification of the digest fragments and further analysis (see Annexe X). The gel fragments containing the DNA bands of interest were removed and purified by use of the Ezway™ Gel Extraction Kit (Komabiotech) as recommended by the manufacturer (Annexe XI).

### 2.2.5. Preparing blunt-ended fragments

During preparation of *SELP-59-A*/pStaby Δ*bla*, filling-in of the *Ssp*I - *Eam*1105I digest fragment was required so as to ensure blunt ends and allow for recircularisation. Here the accuzyme DNA polymerase mix (Bioline Ltd) was used as described by the manufacturers (see Annexe XII).

### 2.2.6. Fragment Ligation

The purified *SELP-59-A* fragment was inserted into pET29a(+) and pStaby 1.2, and the *SELP-59-A*/pStaby Δ*blap* and *SELP-59-A*/pStaby Δ*bla* fragments were re-circularised by ligation with T4 DNA ligase (Thermo Scientific) as recommended by the manufacturer (Annexe XIII).

### 2.2.7. Transformation

*SELP-59-A*/pET29a(+) was transformed to *E. coli* strains XL1Blue (cloning strain) and BL21(DE3) (expression strain) using standard protocols (Annexe XIV). Vectors containing the *ccdA* antidote gene (i.e. *SELP-59-A*/pStaby, *SELP-59-A*/pStaby Δ*bla*, *SELP-59-A*/pStaby Δ*blap*) were transformed to *E. coli* strains CYS21 (cloning strain) and SE1 (expression strain) using the protocols provided with the Staby Expression Kit (Annexe XIV).

### 2.2.8. Confirmation of Constructs

Confirmation of correct preparation of constructs was performed by plasmid isolation (2.2.1 above) followed by restriction digestion (2.2.3) and agarose gel analysis (2.2.4) of fragments.

### 2.2.9. Fed-Batch Production

Comparative fed-batch productions with complete process analyses were carried out for all five constructs using the protocols described in 2.1.1 and 2.1.1.1 above. A pre-induction growth rate of 0.5 $h^{-1}$ and post-induction growth rate of 0.1 $h^{-1}$ was used, $dO_2$ was maintained at 35 % during the fed-batch phase and induction with 3 mM IPTG was carried out at a DCW of 25 g/L. The post induction period was extended to 8 hours.

# 3. Results and Discussion

# 3.1. Optimisation of Fed-batch Process

The objective of this first section of the project was to develop and optimise a fed-batch process for the high cell density production of SELP-59-A in the BioFlo 110 benchtop 3L fermentor with the *SELP-59-A*/pCM13/*E. coli* BL21(DE3) expression system. To achieve this objective we established a fed-batch process and optimised process variables such as pre- and post-induction growth rates, dissolved oxygen ($dO_2$) concentration, cell density and inducer concentration. Each of these parameters was investigated and optimised in the fed-batch process with, in all cases, a full analysis of process parameters.

## 3.1.1. Effect of growth rate on fed-batch process

A fed-batch process was developed for the high cell density production of SELP-59-A and applied and monitored as described in the Materials and Methods section. Initially the effect of pre and post induction growth rates (μ) of 0.2, 0.4, 0.6 and 0.8 $h^{-1}$ with $dO_2$ at 35 % were investigated for the attainment of cell densities of 50 g/L before induction with IPTG. With growth rates up to 0.6 $h^{-1}$ a DCW of 50 g/L was successfully achieved whereas with 0.8 $h^{-1}$ the maximum cell density achievable was only 40 - 45 g/L. Indeed, with this latter growth rate, continued feeding leads to a decrease in DCW, indicative of cell disruption. A DCW of 50 g/L is already 10 times higher than that achieved with the optimised batch production approach previously reported with the same expression system[1] and points to the fed-batch process developed here being suitable for overcoming some of the limitations to cell growth inherent to batch production approaches.

Figure 9 shows the results for the monitoring of the control parameters for the fed-batch process with pre- and post-induction growth rates of 0.2 $h^{-1}$ and it can be seen that these were successfully maintained at the desired levels throughout the process i.e. 37 ± 1 ºC, pH 6.8 ± 0.1, $dO_2$ = 25 ± 10% during the batch phase and ~35 ± 10% during the fed-batch phase (with ‘spikes’ at the time points of

changes in feeding rate), phosphate concentration between 50 and 300 mM and nitrogen concentration between 50 and 100 mM. This points to appropriate feeding and control regimes having been developed and successfully applied here and indeed similar profiles were observed for all studies reported in this thesis, hence only the results for the 0.2 – 0.2 $h^{-1}$ study will be discussed here.

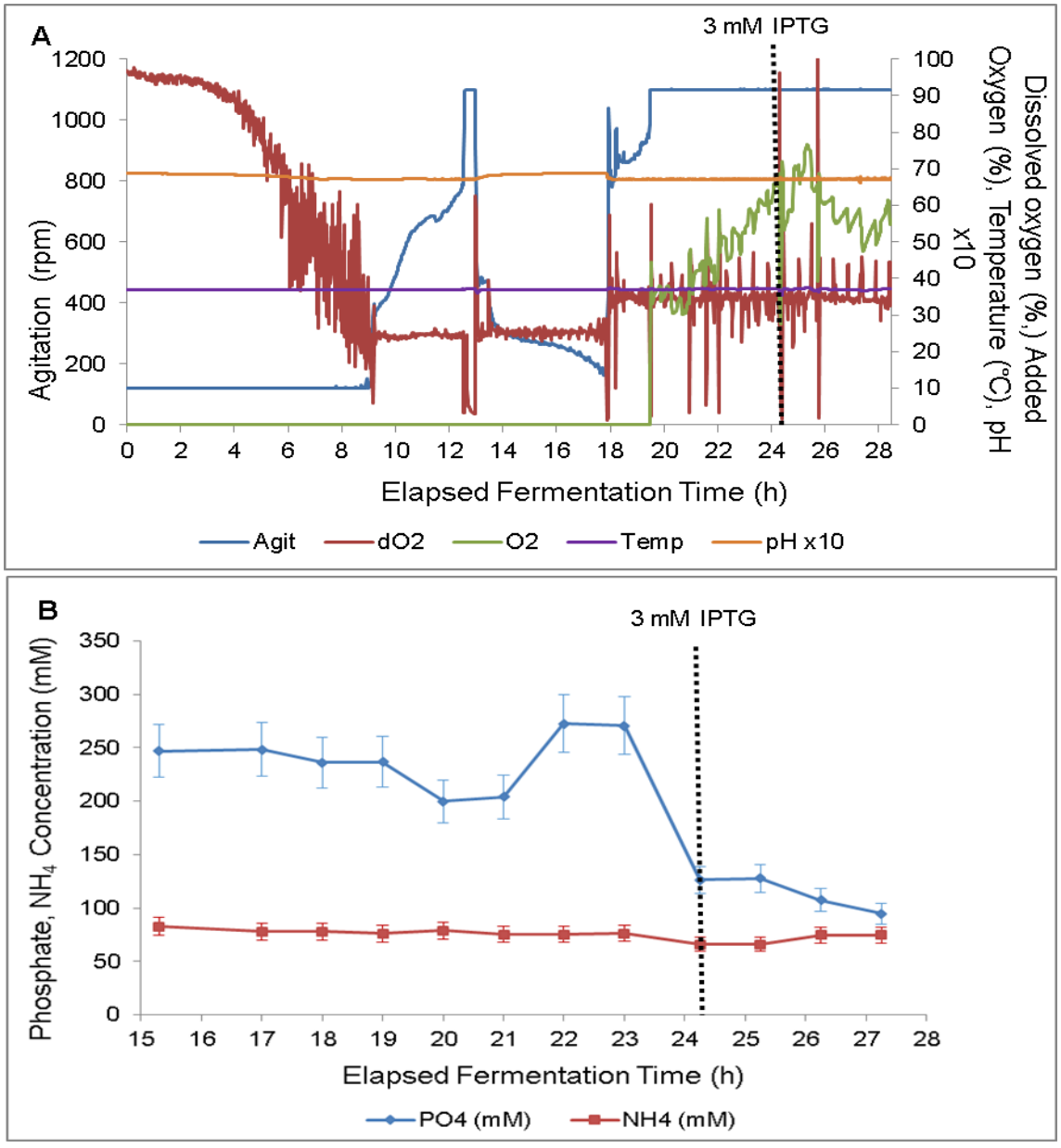



Figure 9. Variations in control parameters during the fed-batch process with pre- and post-induction growth rates of 0.2 $h^{-1}$, 35 % $dO_2$ and induction with 3 mM IPTG at 50 g/L dry cell weight. A) shows the results for the on-line monitoring throughout both the batch (0 - 18 h) and fed-batch (18 - 29 h) phases and B) the variations in the phosphate ($PO_4$) and nitrogen (N) concentrations during the fed-batch phase only. The vertical dotted line indicates the time point of IPTG induction.

From Figure 9A it can be seen that the $dO_2$ level varies throughout the fermentation as a result of cell requirements and of the response of the control system (i.e. response of agitation rate and $O_2$ addition) to $dO_2$ levels below the permitted values (i.e. ≥ 25 – 35 % $dO_2$). Initially the $dO_2$ level decreases as the cells grow until the lower set $dO_2$ limit of 25 % is reached (batch-phase) whereupon the agitation rate increases to the maximum allowed so as to ensure maintenance of a sufficient level of $dO_2$. Following further incubation, an increase in $dO_2$ concentration leads to a rapid decrease in the agitation rate and is believed to be due to reduced cell growth as a result of an exhaustion of all the glucose present. The cells then grow on organic acids produced during the batch phase, the complete utilisation of which is indicated by a further rise in $dO_2$ levels. Hereupon, the fed-batch phase of controlled glucose feeding and cell growth is initiated with the expected response of a decrease in $dO_2$ levels and a rapid compensatory increase in agitation rate and $O_2$ addition being noted. Finally, on induction with IPTG at a DCW of 50 g/L the $O_2$ consumption increases slightly before slowly decreasing until the fermentation is stopped four hours after induction.

From Figure 9B it can be seen that phosphate and nitrogen levels were maintained throughout the process at levels known to be non-limiting and non-toxic to the cells[7] and hence indicating that the feeding protocol developed for our second feed solution (phosphate/nitrogen feed) was successful during the process. Interestingly, throughout this project, while it was found that phosphate consumption was maintained throughout the fed-batch process, nitrogen consumption decreased following induction. Indeed this has been previously reported for *E. coli* cells following induction and is suggested to be indicative of an impaired cell metabolism following induction[7].

Figure 10 shows the results for the analysis of the process parameters for all three growth rates investigated. It can be clearly seen that 50 g/L DCW was successfully achieved before induction, upon which an arrestation of cell growth is observed. In fact, at the higher growth rates investigated a decrease in the DCW is observed following induction. Indeed this has been previously reported for recombinant protein production in *E. coli* and is believed to be due to an

increased metabolic stress on the cells following induction and a diversion of cell resources to protein expression[1,42].

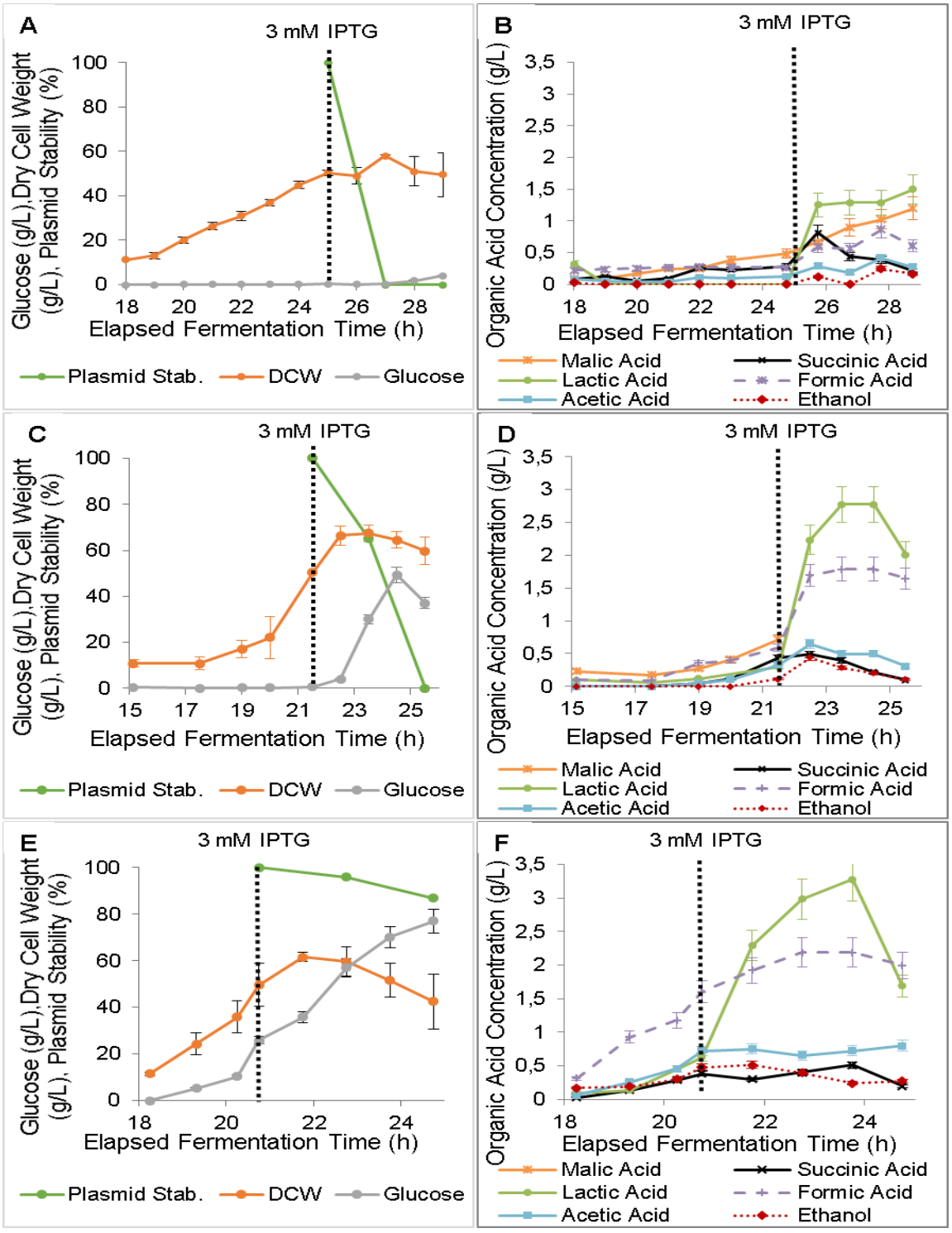



Figure 10. Monitoring of process parameters during the fed-batch phase of production. Results for pre- and post-induction growth rates of 0.2 $h^{-1}$ (A, B), 0.4 $h^{-1}$ (C, D) and 0.6 $h^{-1}$ (E, F) with 35 % $dO_2$ and induction at 50 g/L dry cell weight with 3 mM IPTG are shown. A, C and E show the results for the variation of the plasmid stability, dry cell weight

(DCW) and glucose concentration with time. B, D and F show the variation in organic acid levels as monitored by HPLC. The vertical dotted line indicates the time point of IPTG induction.

From Figure 10 it can also be seen that the rates of extracellular glucose and organic acid accumulation vary with the growth rate used. Glucose accumulates to higher levels in the extracellular environment at higher growth rates and this accumulation increases further following induction. Glucose levels at induction were approximately 0.3, 0.6 and 20 g/L for growth rates of 0.2, 0.4 and 0.6 $h^{-1}$, respectively whereas following induction, respective maximum values of approximately 4, 50 and 80 g/L were observed. Similar trends are observed for formic acid where up to approximately 2.5 g/L was detected at the highest growth rate examined. Interestingly lactic acid is found to be maintained at low level before induction but concentrations drastically increase following induction, with up to 3 g/L being detected. Malic acid was found to accumulate up to 1.2 g/L over the course of the process at the lowest growth rate investigated but the high levels of glucose accumulation at the higher growth rates investigated impeded analysis by HPLC (peak overlap) for these growth rates. Finally, acetic acid, succinic acid and ethanol accumulation remain at low levels throughout the processes and indeed the latter two compounds appear to level off or even decrease in concentration following induction. The acetic acid concentration reached is below levels shown to have a strong negative effect on cell growth[1] and is also lower than the level previously reported with the same expression system for batch production in shake flasks, where up to 5 g/L was accumulated. This highlights the benefits of using the fed-batch approach for the controlled growth of cells with a reduced acetic acid accumulation and hence allowing for the attainment of high cell densities as observed here. Finally, with a growth rate of 0.8 $h^{-1}$ glucose had accumulated to 50 g/L when cell growth was arrested (45 g/L DCW) whereas the concentrations of extracellular organic acids were similar to those observed at 0.6 $h^{-1}$. Further studies are required to better understand the reasons for the reduced DCW attainable at this growth rate.

Figure 11 shows the results for the SDS-PAGE analysis of the intracellular protein production when a growth rate of 0.2 $h^{-1}$ was used. This result is representative of all conditions investigated here where production of SELP-59-A only becomes

visible 1 hour after induction and continues to increase in concentration up to 4 hours after induction.

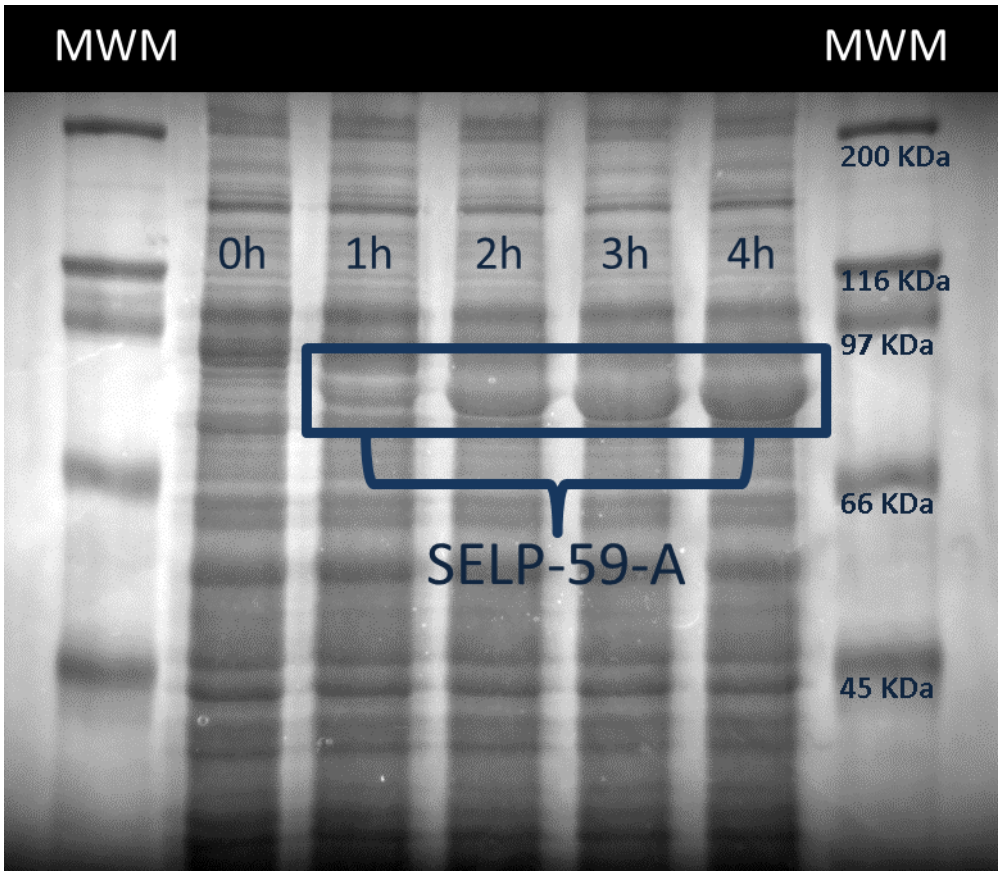



Figure 11. SDS-Page analysis of intracellular production at 0, 1, 2, 3 and 4 hours of IPTG induction of fed-batch production with pre- and post-growth rates of 0.2 $h^{-1}$, 35 % $dO_2$ and with 3 mM IPTG induction at a dry cell weight of 50 g/L. A Broad Range Molecular Weight Standards (Bio-Rad) can be seen in first and last lanes.

SELP-59-A productions as high as approximately 3 g/L following purification were observed with growth rates of 0.2 and 0.4 $h^{-1}$ (Figure 12). This is approximately 6-fold greater than the highest production level reported for SELPs to date[1] and points to the appropriateness of the approach used for the high level production of this polymer. The production level was reduced to approximately 2 g/L when a growth rate of 0.6 $h^{-1}$ was used with, interestingly, a concomitant increase in plasmid stability being also observed. Indeed, while it is possible that the accumulated organic acids, and in particular the 1.5 g/L of formic acid observed before induction, may have a negative effect on SELP-59-A production. It is also possible that the high glucose concentrations observed also negatively affect SELP-59-A production. Glucose concentrations above 10 g/L with low cell density

cultures are known to have an inhibitory effect on production with the *lacUV5* promoter-operator based system used in this study[43]. Thus the high glucose concentrations observed could reduce production from this promoter, hence leading to the observed reduced production levels but also as a consequence resulting in a reduced metabolic stress on the cells and hence the reduced loss of plasmid observed.

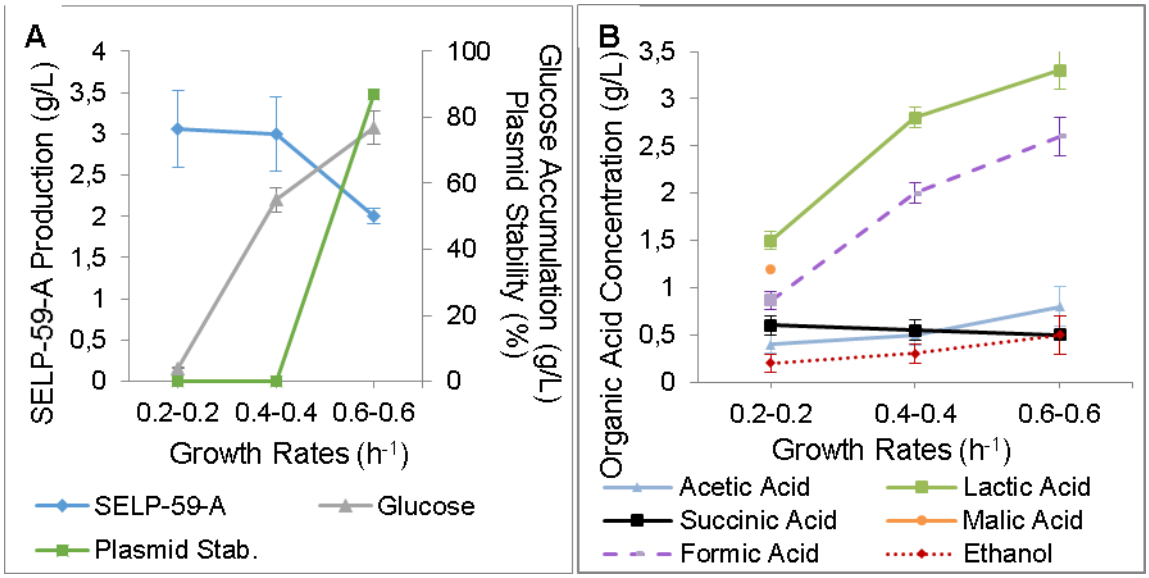



Figure 12. Principal results obtained for fed-batch productions with pre- and post-induction growth rates of 0.2, 0.4 and 0.6 $h^{-1}$. Induction was carried out at a dry cell weight of 50 g/L with 3 mM IPTG for 4 hours and the dissolved oxygen concentration was maintained at 35 % throughout the fed-batch phase. The SELP-59-A and plasmid stability values shown are those measured at the end of the fermentation while the glucose and acetic acid concentrations are the maximum values recorded.

### 3.1.2. Optimisation of post-induction growth rate

In an attempt to further improve production levels and to better understand the fed-batch process we investigated the effect of the post induction growth rate. Various post-induction growth rates were investigated with pre-induction growth rates of 0.2 and 0.6 $h^{-1}$ and, with the exception of the 0.6 – 0.1 $h^{-1}$ study, the trends in glucose and organic acid accumulation were found to be similar to those described for the constant growth rate study in Figure 10 above. Glucose and

formic acid again accumulated throughout the process, with an increased accumulation following induction and at higher feeding rates. In contrast, both were found to be reduced following induction during the 0.6 – 0.1 $h^{-1}$ study. Lactic acid again accumulated at a high rate following induction whereas succinic acid and ethanol remained at low levels with a slight decrease being observed after induction. Finally, acetic acid accumulation was found to be accelerated following induction in all cases, except for the 0.6 – 0.1 $h^{-1}$ study, and this increase accelerated at the higher pre-induction rates studied, probably as a result of the higher glucose concentrations observed. In contrast, during the 0.6 – 0.1 $h^{-1}$ study a rapid decrease in acetic acid concentrations was observed following induction, possibly as a result of the reduced glucose feeding rate used here.

Figure 13 shows the principal results for the comparison of post-induction growth rates at the two pre-induction rates investigated. It can be seen that a maximum of 3 g/L of purified SELP-59-A was obtained and that this was negatively affected by increasing growth rates. Furthermore, as would be expected, no production was observed when glucose feeding was stopped following induction, probably as a result of a lack of glucose and hence plasmid stability was high in this case. Glucose accumulated more at the higher growth rates investigated and, as discussed above, may have had a negative effect on SELP production. Interestingly, plasmid stability of cultivable cells remained low (0 % after 4 hours) for the study with a pre-induction growth rate of 0.2 $h^{-1}$, even at a post-induction growth rate of 0.6 $h^{-1}$ where a maximum glucose concentration of approximately 70 g/L was measured. This value is similar to the maximum glucose concentrations observed with the 0.6 – 0.4 $h^{-1}$ (60 g/L) and 0.6 – 0.6 $h^{-1}$ (80 g/L) studies where high plasmid stability was observed. These differences may be due to the fact that the data presented is for the maximum glucose concentration whereas at induction the concentration of this was much lower with a pre-induction growth rate of 0.2 $h^{-1}$ (0.3 g/L) than with 0.6 $h^{-1}$ (20 g/L) and hence could lead to the difference in the observed plasmid stability of cultivable cells. Finally, it can be seen that organic acid accumulation remains low, with only formic and lactic acid accumulating to relatively high levels and may also negatively influence SELP-59-A production levels[44]. Hence, it can be concluded that a post-induction

growth rate of 0.1 $h^{-1}$ should be optimal for the high cell density production of SELP-59-A and this was used in all further studies.

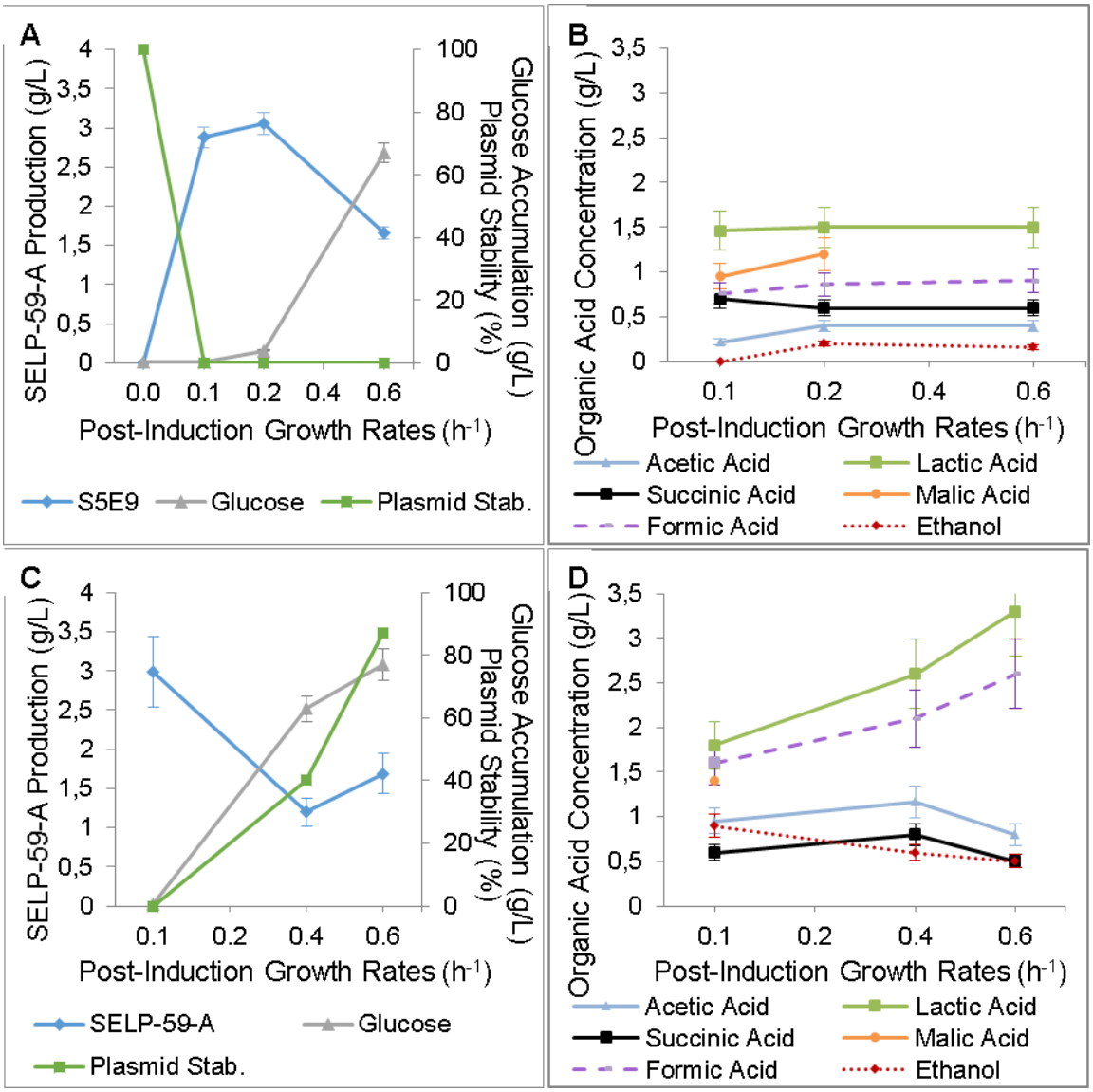



Figure 13. Principal results obtained for fed-batch productions with a pre-induction growth rate of 0.2 $h^{-1}$ (A and B) and 0.6 $h^{-1}$ (C and D) and post-induction growth rates 0.0 to 0.6 $h^{-1}$. Induction was carried out at a dry cell weight of 50 g/L with 3 mM IPTG for 4 hours and the dissolved oxygen concentration was maintained at 35 % throughout the fed-batch phase. The SELP-59-A and plasmid stability values shown are those measured at the end of the fermentation while the glucose and acetic acid concentrations are the maximum values recorded.

### 3.1.3. Optimisation of pre-induction growth rate

Here pre-induction growth rates of between 0.2 and 0.8 $h^{-1}$ were investigated with a post-induction feeding rate equivalent to 0.1 $h^{-1}$. From Figure 14 it can be seen that SELP-59-A production remained at the maximum level when pre-induction growth rates between 0.2 and 0.6 $h^{-1}$ were used. As previously mentioned, a maximum DCW of only 45 g/L was attained with a growth rate of 0.8 $h^{-1}$ and induction at this DCW resulted in a greater than 4-fold reduction in SELP-59-A production. Glucose concentrations as high as 50 g/L were observed before induction and while this did not alter following induction it may be responsible for the high plasmid stability observed. An acetic concentration of 0.9 g/L was observed at induction when the 0.8 $h^{-1}$ growth rate was used and this increased rapidly following induction to levels previously shown to have a bacteriostatic on the expression host used, hence potentially further reducing production levels. In conclusion, a pre-induction growth rate of 0.6 $h^{-1}$ can be applied for SELP-59-A production (with a post-induction feeding rate equivalent to 0.1 $h^{-1}$) and thereby should allow for a reduced process time as compared to lower pre-induction growth rates. Indeed the process time can be reduced almost 6 hours as compared to when using a pre-induction growth rate of 0.2 $h^{-1}$.

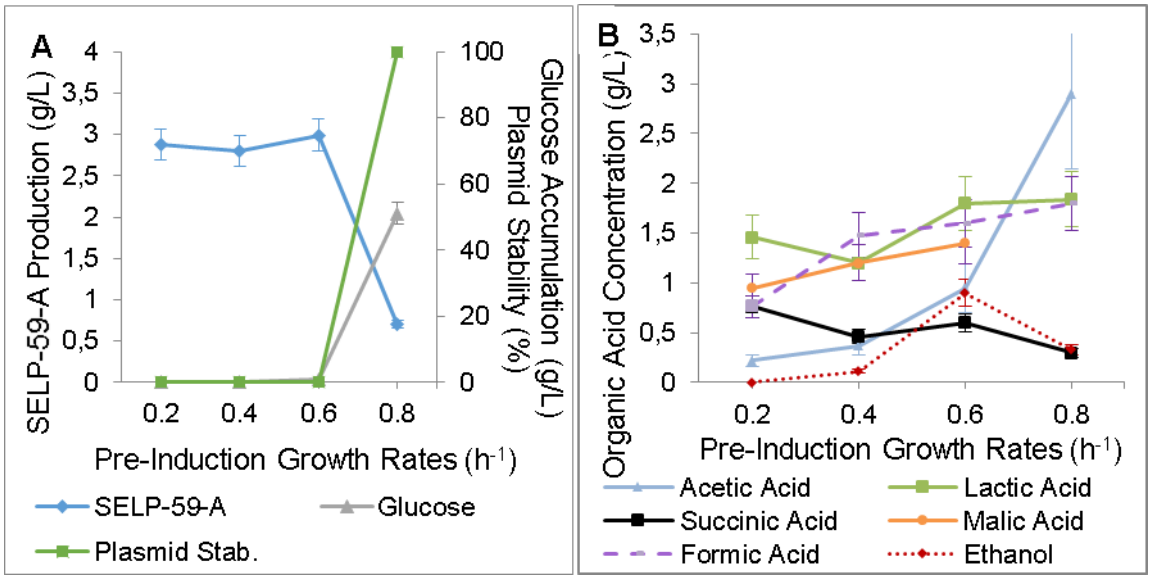



Figure 14. Principal results obtained for fed-batch productions with pre-induction growth rates of 0.2, 0.4, 0.6 and 0.8 $h^{-1}$ and a post-induction growth rate of 0.1 $h^{-1}$. Induction was carried out at a dry cell weight of 50 g/L with 3 mM IPTG for 4 hours and the dissolved oxygen concentration was maintained at 35 % throughout the fed-batch phase. The SELP-59-A and plasmid stability values shown are those measured at the end of the fermentation while the glucose and organic acids concentrations are the maximum values recorded.

### 3.1.4. Optimisation of dissolved oxygen concentration

Dissolved oxygen ($dO_2$) concentrations between 15 and 55 % during the fed-batch phase were investigated and showed that process parameters and SELP-59-A production varied little under these conditions. Hence maintenance of $dO_2$ at approximately 15 % should be applied so as to allow for reduced process costs.

### 3.1.5. Optimisation of dry cell weight at induction

An increase in DCW before induction would be expected to allow for higher SELP-59-A production and indeed this was observed here where increasing this

from 50 to 75 g/L resulted in a corresponding increase in SELP-59-A production from 3 to approximately 4 g/L (Figure 15). When attempting to increase the DCW to 100 g/L complications were encountered in maintaining aerobic conditions. Even with addition of 100 % $O_2$ the $dO_2$ level decreased to near 0 % and a corresponding increase in extracellular organic acids and ethanol to high levels. It is believed that these high levels have a negative effect on the host cell leading to the reduced SELP-59-A production and increased plasmid stability. Hence it can be seen that SELP-59-A production is limited here by the oxygen transfer rate of the fermenter used. In fact, dry cell weights up to 200 g/L have been previously reported with use of specialised fermenters with high oxygen transfer rates and hence could theoretically allow for an almost 3-fold further increase in SELP-59-A production to approximately 12 g/L[45].

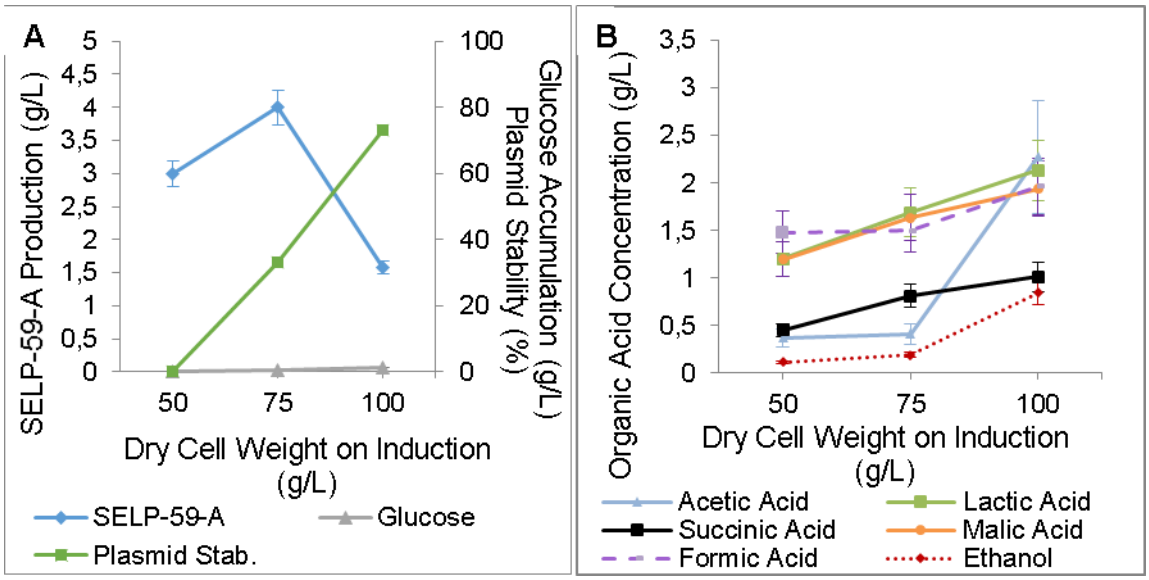



Figure 15. Principal results obtained for fed-batch productions with induction at dry cell weights of 50, 100 and 200 g/L. A pre-induction growth rate of approximately 0.4 $h^{-1}$ and post-induction growth rate of 0.1 $h^{-1}$ were used. Induction was carried out with 3 mM IPTG for 4 hours and the dissolved oxygen concentration was set at 35 % during the fed-batch phase. The SELP-59-A and plasmid stability values shown are those measured at the end of the fermentation while the glucose and organic acids concentrations are the maximum values recorded.

### 3.1.6. Optimisation of IPTG concentration

As a final step in our optimisation process we investigated the optimal IPTG concentration for induction at a DCW of 75 g/L. As can be seen from Figure 16, no significant differences in SELP-59-A production levels were observed with IPTG concentrations between and 1 and 6 mM. In contrast, a decrease in plasmid stability and increase in lactic acid concentrations is observed. Here it is believed that the higher IPTG concentrations investigated lead to a higher induction of production and hence also an increased metabolic burden with a consequent higher plasmid loss and possibly also leading to the observed increased lactic acid production.

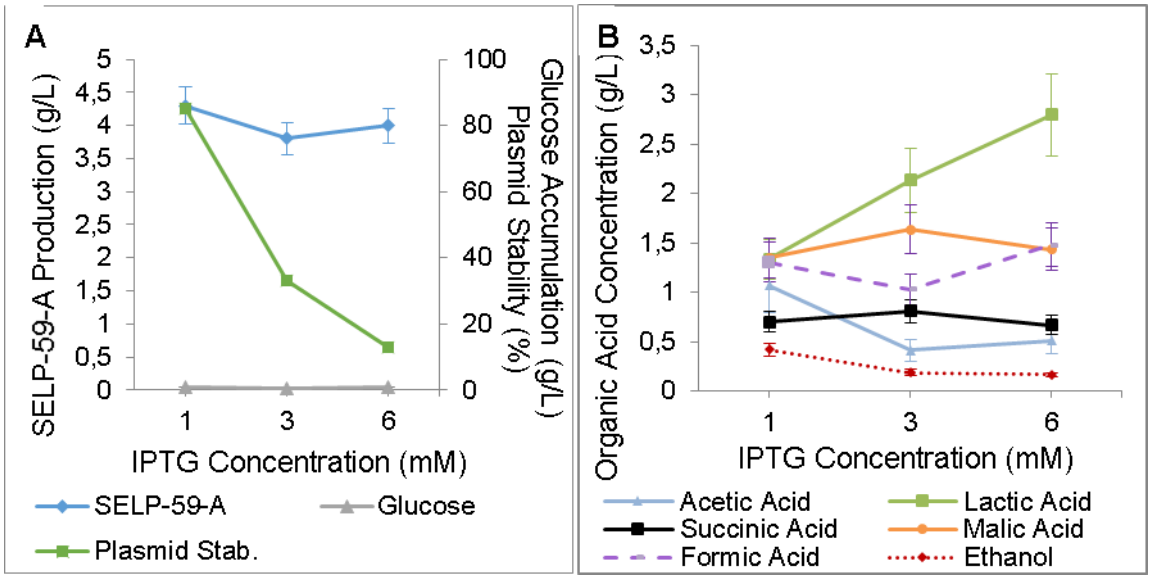



Figure 16. Principal results obtained for fed-batch productions with 1, 3 and 6 mM IPTG for induction. A pre-induction growth rate of 0.4 $h^{-1}$ and post induction growth rate of 0.1 $h^{-1}$ were used. Induction was carried out at a dry cell weight of 75 g/L for 4 hours and the dissolved oxygen concentration was set at 35 % during the fed-batch phase. The SELP-59-A and plasmid stability values shown are those measured at the end of the fermentation while the glucose and organic acids concentrations are the maximum values recorded.

Hence, to conclude this first section of our study, it can be seen that we have successfully developed and optimised a fed-batch process allowing for an almost 9-fold increase in SELP-59-A production as compared to a previously optimised batch production approach[1]. The optimised conditions are: pre-induction growth

rate = 0.6 $h^{-1}$, post-induction growth rate = 0.1 $h^{-1}$, induction at a DCW of 75 g/L with 1 mM IPTG with $dO_2$ maintained at 15 % during the fed-batch phase. This process allows for a reduced cost process with a fed-batch phase of approximately 7 hours. Furthermore, it is believed that the factors limiting further increased production levels include: plasmid loss on induction, the reduced oxygen transfer rate of the fermentation system used and an increased stress on cells following induction and leading to an increased metabolic burden.

## 3.2. Overcoming Plasmid Instability

Previous studies with a batch production approach indicated plasmid instability on induction as a major limiting factor in SELP-59-A production[1] and this was again observed in our fed-batch optimisation study described above. In the second section of our study we attempted to address this problem by investigating various plasmid stabilisation systems as well as by attempting to reduce the metabolic burden placed on the cells by the expression vectors used.

Five different plasmid constructs were successfully constructed and investigated for SELP-59-A production with the previously optimised fed-batch approach developed earlier. Here induction was carried at a lower DCW for practicality reasons and the post-induction phase was extended to 8 hours so as to obtain a better measure of plasmid stability over time.

Our initial SDS-PAGE studies indicated that SELP-59-A production appeared to initiate earlier with the *SELP-59-A*/pStaby/*E. coli* SE1 system (and its variants) as a SELP band was already visible at 30 minutes. In contrast, the SELP-59-A SDS-PAGE band only became visible at 2 hours with the *SELP-59-A*/pCM13/*E.coli* BL21(DE3) control under the conditions used.

An overview of the results for the analysis of process parameters is shown in Figure 17. From 17A and 17B it can be seen that use of the ccdB/ccdA plasmid stabilisation system (pStaby vector) allows for maximum plasmid stability in cultivable cells after 8 hours induction as measured by the approach described in

the materials and methods. Indeed this increased stability was accompanied by a 50 % increase in SELP-59-A production as well as an increase in lactic acid production. Use of kanamycin resistance allowed for a slight increase in plasmid stability as measured by the approach used but was not translated into a significant increase in SELP production or to differences in any of the other variables measured. Furthermore, removal of the *bla* gene or prevention of its transcription from the pStaby vector did not allow for increases in SELP production but were characterised by reduced lactic acid accumulation (Figures 17 C and 17D). Indeed, as previously discussed, lactic acid accumulates rapidly following induction and may be a response to the metabolic stress placed on the cells during recombinant protein production. Hence, the increased levels observed with the pStaby vector (ccdB/ccdA) may be resultant of the increased SELP production as a result of improved plasmid stability while the reduced levels following removal of the *bla* gene or its promoter may indicate a reduced metabolic stress with these constructs. This would also indicate that other factors, probably as a result of the metabolic stress placed on the cells, are limiting further improved production levels. Nevertheless, use of the ccdB/ccdA plasmid stabilisation system allows for approximately 5 g/L of SELP-59-A to be obtained after induction at a DCW of 25 g/L for 8 hours. As previously discussed, a DCW before induction of 75 g/L can be successfully reached with the fermentation equipment used in this study and thereby suggests a potential SELP-59-A production of up to 15 g/L. Furthermore, the absence of antibiotics during production would have obvious advantages for use of the SELP in biomedical applications where antibiotics should be avoided.

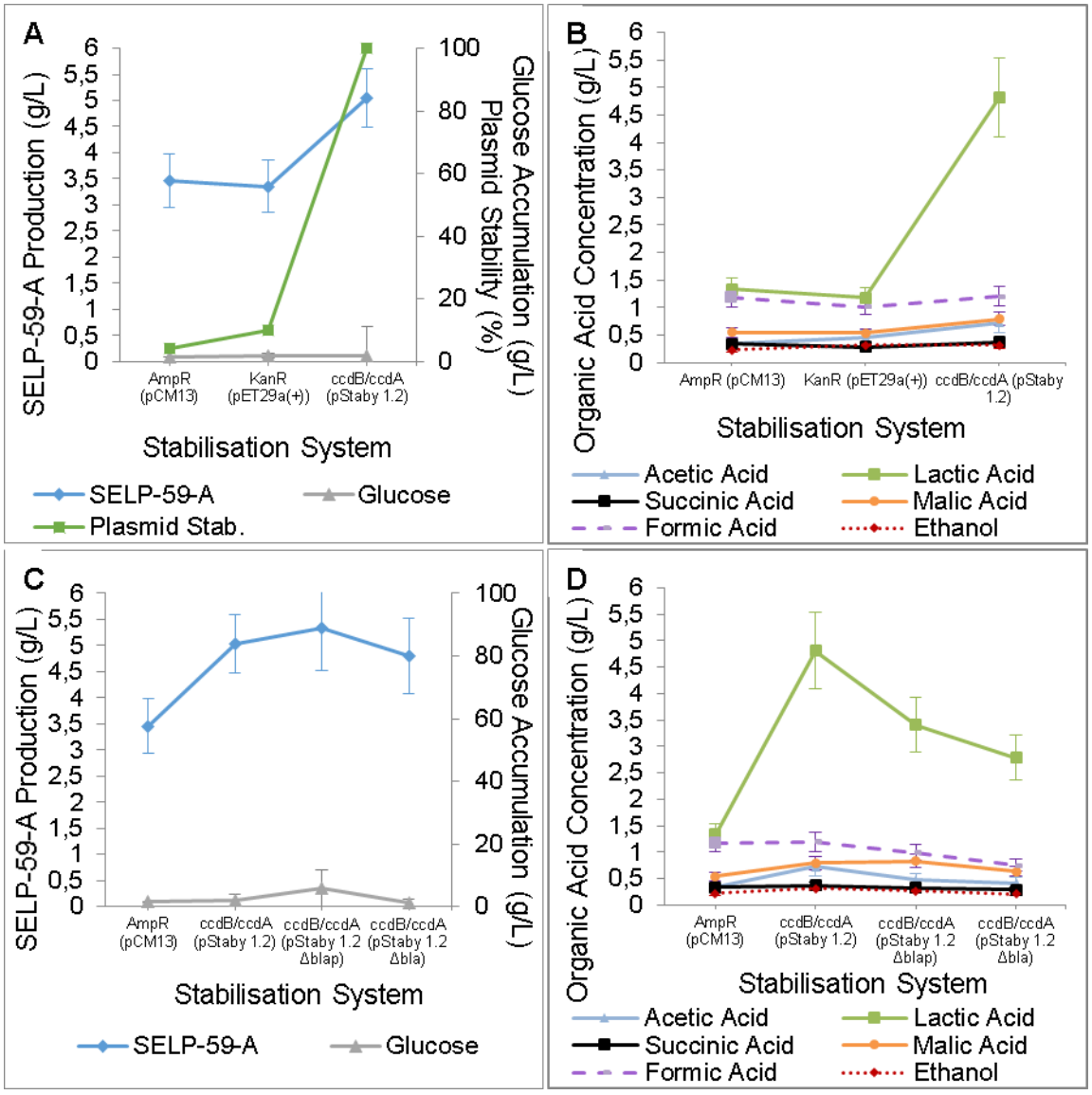



Figure 17. Principal results obtained for fed-batch productions with the various vector constructs investigated. A) and B) compare different plasmid stabilisation systems while C) and D) investigate the removal of the *bla* gene and/or *bla* promoter from pStaby. Due to the removal of ampicillin resistance in these constructs plasmid stability could not be measured by the approach used in this study. A pre-induction growth rate of 0.5 $h^{-1}$ and post induction growth rate of 0.1 $h^{-1}$ were used. Induction was carried out at a dry cell weight of 25 g/L for 8 hours and the dissolved oxygen concentration was set at 35 % during the fed-batch phase. The SELP-59-A and plasmid stability values shown are those measured at the end of the fermentation while the glucose and organic acids concentrations are the maximum values recorded.

# 4. Final Remarks and Future Perspectives

The main long term objective of the project to which the present work forms part is to maximise SELP-59-A production and reduce production costs. Prior to commencement of the project the highest reported SELP production levels were in the order of 20 – 50 mg/L[28.29.46]. Initial studies by the group where the present work was carried out involved the optimisation of a batch production in shake flask approach and allowed for a 10 fold increase in SELP production to approximately 500 mg/L[1]. During the study both accumulation of acetic acid to toxic levels and rapid plasmid loss on induction were identified as factors preventing the attainment of a higher cell density and higher SELP production. In the present study we investigated the use of the fed-batch approach and of alternative plasmid stabilisation systems in an attempt to overcome the limiting factors to SELP production previously identified.

From the first section of our study it can be seen that we have successfully developed and optimised a fed-batch process allowing for an almost 9-fold increase in SELP-59-A production as compared to the previously optimised batch production approach[1]. The optimised conditions are: pre-induction growth rate = 0.4 $h^{-1}$, post-induction growth rate = 0.1 $h^{-1}$, induction at a DCW of 75 g/L with 1 mM IPTG and with $dO_2$ maintained at 15 % during the fed-batch phase. This process allows for a reduced cost process with a fed-batch phase of approximately 8 hours. A pre-induction growth rate of 0.6 $h^{-1}$ was shown to allow for the successful high level production of SELP-59-A following induction at a DCW of 50 g/L. Unfortunately, due to time limitations, this growth rate was not investigated for induction at 75 g/L, the successful application of which would further reduce the process time and hence also production costs. Furthermore, a previous study[7] indicated the advantages of using a variable pre-induction growth rate for the rapid high cell density production of recombinant proteins in *E. coli*. Here, the pre-induction growth rate is maintained at the maximum possible and slowly reduced over time so as to maintain organic acid concentrations at low levels. The use of this approach was shown to not only reduce the process time but also allowed for improved recombinant protein production[7]. It is suggested that such a variable pre-induction growth rate approach be investigated in future studies.

A further interesting future study in relation to the first section of the present project is to better understand the arrestation of growth at 45 g/L DCW when using a growth rate of 0.8 g/L. Extracellular organic acid concentrations were not found to be much higher than those observed with a growth rate of 0.6 $h^{-1}$ whereas glucose was found to accumulate to approximately 50 g/L as compared to 20 g/L with a growth rate of 0.6 $h^{-1}$. Perhaps intracellular organic acid concentrations were already high when cell arrestation was observed and it is therefore suggested that a study with a pre- and post- induction growth rate of 0.8 $h^{-1}$ with induction at 45 g/L be carried out to investigate this.

Following the studies described in the first section of this work it was concluded that the factors limiting further increased production levels include: plasmid loss on induction, limitations in the oxygen transfer rate of the fermentation system used and an increased stress on cells following induction. We decided to address the problem of plasmid instability on induction by investigating and comparing three plasmid stabilisation systems, namely: ampicillin resistance, kanamycin resistance and a Type II toxin/antitoxin post segregational suicide system (ccdB/ccdA). Furthermore, we investigated the effect of inhibition of β-lactamase production from the ccdB/ccdA stabilisation system expression vector. Here it was hypothesised that this would reduce the stress on the host cells, giving rise to 'fitter' cells and potentially increasing SELP production. Our study indicated a 50 % increase in SELP production with the ccdB/ccdA system but with no further increase following removal of the ampicillin resistance marker from the expression vector. Use of the ccdB/ccdA plasmid stabilisation system was shown, with the protocol used, to allow for all cultivable cells to retain the plasmid up to 8 hours following induction. Nevertheless, it must be noted that the protocol used here for measuring plasmid stability has many limitations as it only measures the ability of cells to grow in the presence and absence of the antibiotic selection marker. Obviously, with the ccdB/ccdA system only those cells with plasmid will be able to grow and hence should always indicate 100 % plasmid stability with the approach used. Furthermore, it is possible that cells producing SELP are unable to grow under the conditions used and hence would not be detected with the approach used. Therefore, to get a truer measure of plasmid stability or, more specifically, of the actual concentration of plasmid it is suggested that a

quantitative PCR (qPCR) approach be developed in the future. We originally attempted to quantify plasmid by means of plasmid isolation from culture samples but variations in yields prevented the use of this approach, hence the need for the development of a qPCR approach.

An interesting observation from this study was the high production of lactic acid directly following induction and the variation in the extracellular concentration of this. It was found to increase at higher growth rates, higher IPTG concentrations and interestingly also it was found to increase with the ccdA/ccdB stabilisation system where higher SELP production was obtained. This production of lactic acid is probably a response to the stress placed on the host cells during SELP production and further investigation of this is warranted to better understand those conditions that lead to such an accumulation and how this can be overcome. Furthermore, it was found that this accumulation of lactic acid decreased when β-lactamase production from the expression vector was removed. This suggests a reduced induced stress in these cells but nonetheless was found to be insufficient for further improved SELP production.

In conclusion, it can be seen that our study has allowed for an increase in SELP-59-A production to approximately 5 g/L which is 10-fold higher than the highest production previously reported[1]. This high production was achieved with induction at 25 g/L and as shown in this study a DCW before induction of 75 g/L is attainable with the system used in the study. This suggests a further 3-fold increase in SELP-59-A production with this system. Future studies should be focused on investigating this.

# 5. References

# References


1. Collins, T., Azevedo-Silva, J., da Costa, A., Branca, F., Machado, R. & Casal, M. (2013). Batch production of a silk-elastin-like protein in E. coli BL21(DE3): Key parameters for optimisation. *Microb. Cell Fact.* Feb 27; **12:21**. doi: 10.1186/1475-2859-12-21.
2. Mohanty, A. K., Misra, M. & Hinrichsen, G. (2000). Biofibres, biodegradable polymers and biocomposites: An overview. *Macromolecular Materials and Engineering* **276-277**, 1-24.
3. Rodríguez-Cabello, J. C., Martín, L., Alonso, M., Arias, F. J. & Testera, A. M. (2009). "Recombinamers" as advanced materials for the post-oil age. *Polymer* **50**, 5159-5169.
4. Kiick, K. L. (2007). Biosynthetic methods for the production of advanced protein-based materials. *Polymer Reviews* **47**, 1-7.
5. Sakiyama-Elbert, S. E. & Hubbell, J. A. (2001). Functional biomaterials: Design of novel biomaterials. *Annual Review of Materials Science* **31**, 183-201.
6. Sørensen, H. P. & Mortensen, K. K. (2005). Advanced genetic strategies for recombinant protein expression in Escherichia coli. *Journal of Biotechnology* **115**, 113-128.
7. Babaeipour, V., Shojaosadati, S. A., Robatjazi, S. M., Khalilzadeh, R. & Maghsoudi, N. (2007). Over-production of human interferon-γ by HCDC of recombinant Escherichia coli. *Process Biochemistry* **42**, 112-117.
8. Van Hest, J. C. M. & Tirrell, D. A. (2001). Protein-based materials, toward a new level of structural control. *Chemical Communications*, 1897-1904.
9. Tatham, A. S. & Shewry, P. R. (2000). Elastomeric proteins: Biological roles, structures and mechanisms. *Trends in Biochemical Sciences* **25**, 567-571.
10. Alper, J. (2002). Stretching the Limits. *Science* **297**, 329-331.
11. Song, H. & Parkinson, J. (2012). Modelling the self-assembly of elastomeric proteins provides insights into the evolution of their domain architectures. *PLoS Computational Biology* **8**.
12. Gosline, J., Lillie, M., Carrington, E., Guerette, P., Ortlepp, C. & Savage, K. (2002). Elastic proteins: Biological roles and mechanical properties. *Philosophical Transactions of the Royal Society B: Biological Sciences* **357**, 121-132.
13. Heim, M., Keerl, D. & Scheibel, T. (2009). Spider silk: From soluble protein to extraordinary fiber. *Angewandte Chemie - International Edition* **48**, 3584-3596.
14. Vendrely, C. & Scheibel, T. (2007). Biotechnological production of spider-silk proteins enables new applications. *Macromolecular Bioscience* **7**, 401-409.
15. Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts & Walter, P. (2002). Proteins. In *Molecular Biology of the Cell* 5 edit., pp. 125-193. Garland Science, Taylor & Francis Group, New York.
16. Pechar, M., Brus, J., Kostka, L., Koňák, Č., Urbanová, M. & Šlouf, M. (2007). Thermoresponsive self-assembly of short elastin-like polypentapeptides and their poly(ethylene glycol) derivatives. *Macromolecular Bioscience* **7**, 56-69.
17. Floss, D. M., Schallau, K., Rose-John, S., Conrad, U. & Scheller, J. (2010). Elastin-like polypeptides revolutionize recombinant protein expression and their biomedical application. *Trends in Biotechnology* **28**, 37-45.
18. MacEwan, S. R. & Chilkoti, A. (2010). Elastin-like polypeptides: biomedical applications of tunable biopolymers. *Biopolymers* **94**, 60-77.
19. Rodríguez-Cabello, J. C., Alonso, M., Caballero, M. I. & Herguedas, M. M. (2000). Experimental study of the assignment of non-amide Raman bands of the polypentapeptide of elastin, poly(VPGVG). *Journal of Raman Spectroscopy* **31**, 523-527.

20. Araújo, R., Silva, C., Machado, R., Casal, M., Cunha, A. M., Rodriguez-Cabello, J. C. & Cavaco-Paulo, A. (2009). Proteolytic enzyme engineering: A tool for wool. *Biomacromolecules* **10**, 1655-1661.
21. Herrero-Vanrell, R., Rincón, A. C., Alonso, M., Reboto, V., Molina-Martinez, I. T. & Rodríguez-Cabello, J. C. (2005). Self-assembled particles of an elastin-like polymer as vehicles for controlled drug release. *Journal of Controlled Release* **102**, 113-122.
22. Demura, M. & Asakura, T. (1991). Porous membrane of Bombyx mori silk fibroin: Structure characterization, physical properties and application to glucose oxidase immobilization. *Journal of Membrane Science* **59**, 39-52.
23. Cappello, J., Crissman, J., Dorman, M., Mikolajczak, M., Textor, G., Marquet, M. & Ferrari, F. (1990). Genetic engineering of structural protein polymers. *Biotechnology Progress* **6**, 198-202.
24. Machado, R., Azevedo-Silva, J., Correia, C., Collins, T., Arias, F. J., Rodríguez-Cabello, J. C. & Casal, M. (2013). High level expression and facile purification of recombinant silk-elastin-like polymers in auto induction shake flask cultures. *AMB Express* **3**, 1-15.
25. Qiu, W., Teng, W., Cappello, J. & Wu, X. (2009). Wet-spinning of recombinant silk-elastin-like protein polymer fibers with high tensile strength and high deformability. *Biomacromolecules* **10**, 602-608.
26. Qiu, W., Huang, Y., Teng, W., Cohn, C. M., Cappello, J. & Wu, X. (2010). Complete recombinant silk-elastinlike protein-based tissue scaffold. *Biomacromolecules* **11**, 3219-3227.
27. Teng, W., Cappello, J. & Wu, X. (2009). Recombinant silk-elastinlike protein polymer displays elasticity comparable to elastin. *Biomacromolecules* **10**, 3028-3036.
28. Nagarsekar, A., Crissman, J., Crissman, M., Ferrari, F., Cappello, J. & Ghandehari, H. (2002). Genetic synthesis and characterization of pH- and temperature-sensitive silk-elastinlike protein block copolymers. *Journal of Biomedical Materials Research* **62**, 195-203.
29. Haider, M., Leung, V., Ferrari, F., Crissman, J., Powell, J., Cappello, J. & Ghandehari, H. (2005). Molecular engineering of silk-elastinlike polymers for matrix-mediated gene delivery: Biosynthesis and characterization. *Molecular Pharmaceutics* **2**, 139-150.
30. Chamberlin, M. & Ring, J. (1973). Characterization of T7-specific ribonucleic acid polymerase. 1. General properties of the enzymatic reaction and the template specificity of the enzyme. *Journal of Biological Chemistry* **248**, 2235-2244.
31. Golomb, M. & Chamberlin, M. (1974). A preliminary map of the major transcription units read by T7 RNA polymerase on the T7 and T3 bacteriophage chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* **71**, 760-764.
32. Tabor, S. & Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 1074-1078.
33. (2002-2003). pET System Manual 11th edit. (Novagen, ed.). EMD Biosciences, Inc, Germany.
34. Doran, P. M. (1995). *Bioprocess Engineering Principles*, Elsevier Ltd.
35. Van't Riet, K. (1979). Review of measuring methods and results in nonviscous gas-liquid mass transfer in stirred vessels. *Industrial and Engineering Chemistry Process Design and Development* **18**, 357-364.
36. (2006). Staby Express™ T7 Kit Manual v1.7 edit. Delphi Genetics SA, Belgium.
37. Heyland, J., Blank, L. M. & Schmid, A. (2011). Quantification of metabolic limitations during recombinant protein production in Escherichia coli. *Journal of Biotechnology* **155**, 178-184.
38. Horn, U., Strittmatter, W., Krebber, A., Knupfer, U., Kujau, M., Wenderoth, R., Muller, K., Matzku, S., Pluckthun, A. & Riesenberg, D. (1996). High volumetric yields of functional dimeric miniantibodies in Escherichia coli, using an optimized expression vector and

high-cell-density fermentation under non-limited growth conditions. *Appl Microbiol Biotechnol* **46**, 524-32.

39. Korz, D. J., Rinas, U., Hellmuth, K., Sanders, E. A. & Deckwer, W. D. (1995). Simple fed-batch technique for high cell density cultivation of Escherichia coli. *Journal of Biotechnology* **39**, 59-65.
40. Chen, P. S., Toribara, T. Y. & Warner, H. (1956). Microdetermination of Phosphorus. *Analytical Chemistry* **28**, 1756-1758.
41. Weatherburn, M. W. (1967). Phenol-hypochlorite reaction for determination of ammonia. *Analytical Chemistry* **39**, 971-974.
42. Flickinger, M. C. & Rouse, M. P. (1993). Sustaining Protein Synthesis in the Absence of Rapid Cell Division: An Investigation of Plasmid-Encoded Protein Expression in Escherichia coli during Very Slow Growth. *Biotechnology Progress* **9**, 555-572.
43. Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-34.
44. Kirkpatrick, C., Maurer, L. M., Oyelakin, N. E., Yoncheva, Y. N., Maurer, R. & Slonczewski, J. L. (2001). Acetate and formate stress: opposite responses in the proteome of Escherichia coli. *J Bacteriol* **183**, 6466-77.
45. Lee, S. Y. (1996). High cell-density culture of Escherichia coli. *Trends in Biotechnology* **14**, 98-105.
46. Xia, X.-X., Xu, Q., Hu, X., Qin, G. & Kaplan, D. L. (2011). Tunable Self-Assembly of Genetically Engineered Silk–Elastin-like Protein Polymers. *Biomacromolecules* **12**, 3844-3850.
47. Ngo, T. T., Phan, A. P. H., Yam, C. F. & Lenhoff, H. M. (1982). Interference in determination of ammonia with the hypochlorite-alkaline phenol method of Berthelot. *Analytical Chemistry* **54**, 46-49.
48. Inoue, H., Nojima, H. & Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. *Gene* **96**, 23-28.

# 6. Annexes

# Annexe I

## Fed-Batch Production

Fed-batch production is carried out with a BioFlo 110 3 litre fermentor (NBS) controlled via a Primary Control Unit (NBS) and with a BioCommand General Purpose Multi-Loop controller (NBS) and the BioCommand Plus 3.3 software (NBS) for automated data collection and supervision.

Protocol:

- Inoculate 50 mL of LB (+ antibiotic if required) in a 500 mL Erlenmeyer (first preculture) with a cfu of the production host and incubate at 37 °C/200 rpm until a dry cell weight (DCW) of 0.5 g/L is reached (approx. 4 hours).
- Inoculate 100 mL of MMLBM (+ antibiotic if required) in a 500 mL Erlenmeyer (second preculture) with the first preculture to an initial dry cell weight of 0.05 g/L and incubate at 37 °C/200 rpm until a dry cell weight (DCW) of 0.5 g/L is reached (approx. 3 hours).
- Inoculate 960 mL of MMLBM (+ antibiotic if required) in the fermenter (production culture) with the second preculture to an initial dry cell weight of 0.005 g/L.
- Set fermenter batch phase parameters: temperature at 37 °C, pH 6.8 with 3 M $H_3PO_4$ and 25 % $NH_4OH$, air flow rate at 5 L/min., $dO_2$ controlled at 25 % with cascade to agitation (200-1100 rpm), $dO_2$ PID 0.05-0.25, foaming controlled with a conduction probe (NBS) and antifoam Y-30 emulsion (Sigma).
- Determine initial acid, base and antifoam weights.
- Leave batch phase to run overnight (approx 15 hours) until all carbon has been completely exhausted and as indicated by a sustained increase in the $dO_2$ level above 50 %.
- Start fed-batch phase: start glucose feed feeding rate as deifned by the desired growth rate (μ) and the mass balance equation (equation 1), start phosphate/nitrogen feeding, change $dO_2$ setting to desired value (i.e. 15 – 55 %) and cascade to both agitation (200-1100 rpm) and $O_2$, connect $O_2$ at 1.0 bar, change $dO_2$ PID to 0.1-0.25.

$$\mathrm{M_S(t)} = \mathrm{F_F(t)S_F(t)} = \left(\frac{\mu(t)}{Yx/s} + m\right)V(t)X(t) \qquad (1)$$

- Measure weights of acid, base and antifoam and calculate volumes added and hence also total culture volume.

- Calculate the cell density from equation 2

$$X(t)V(t) = Xt_F Vt_F e^{\mu(t-t_F)} \quad (2)$$

- Every 15 minutes, use calculated cell density (equation 2) and calculated total culture volume to determine the required feeding rate according to equation 1.
- Every hour measure the actual $OD_{600nm}$ and calculate DCW ($OD_{600nm}$ x 0.51), calculate culture volume and use to determine the required feeding rate according to equation 1.
- Every hour aseptically sample culture (note volume sampled) for $OD_{600nm}$ readings, centrifuge 4 mL for 5 minutes at maximum speed, retain supernatant at – 20 ºC for later analysis (HPLC, phosphate, nitrogen etc.).
- When desired DCW is reached induce with IPTG (typically 15 mL of concentrated solution), may need to remove culture volume before induction, note volume removed. Alter desired growth rate (μ) as required.
- At 0 hours induction and at every hour thereafter asceptically remove approx. 8 mL samples: use directly for $OD_{600nm}$ readings and for viable cell counts; for SDS-PAGE analysis centrifuge 0.5 mL samples (x 4) for 5 minutes at maximum speed and retain pellets at -20 ºC until analysis; for HPLC, phosphate and nitrogen analyses retain 4 mL supernatant samples at –20 ºC until analysis.
- Stop fermentation after desired induction period, measure the final volume and centrifuge at 9000 rpm for 30 minutes. Retain cell pellets at -20 °C and supernatantat at 4 °C until later SELP purification and analysis.

# Annexe II

## HPLC analysis

HPLC was carried out with a Rezex™ 8 µm ROA-organic acid H+(8%) high performance liquid chromatography column (Phenomenex) on an Elite LaChrom (VWR Hitachi) chromatography system. This separates sugars and organic acids on the basis of a differential partitioning between the stationary phase (H+) and the mobile phase ($H_2SO_4$). Here it was used to detect the presence of and to quantify carbohydrates (glucose, lactose, fructose, glycerol), organic acids (citric, tartaric, malic, succinic, lactic, acetic and formic) and ethanol. Detection was by refractive index measurement with an Elite LaChrom L-2490 RI detector (VWR Hitachi) at 40°C. An Elite LaChrom (VWR Hitachi) chromatography system was used with the EZChrom Elite 3.3.2 SP2 software for data collection and analysis.

Protocol:

- Prepare 50 mL 100 % TCA.
- Let culture supernatant samples (~1.5 mL) reach room temperature.
- Add TCA to final concentration of 10 % (~160 µL).
- Leave at room temperature overnight.
- Centrifuge at max speed 15 min.
- Discard pellet and filter supernatant through 0.22 µm filter to HPLC vials.
- Preheat HPLC column to 60 ºC.
- Load 20 µL to the preequilibrated (in 2.5 mM $H_2SO_4$) column.
- Use 2.5mM $H_2SO_4$ as the mobile phase with isocratic elution as follows:

    0 - 14 mins.: 0.15 mL/min

    14 - 35 mins.: 0.15 - 0.2 mL/min

    35 - 90 mins.: 0.2 mL/min

Retention times and peak areas are compared to standard curves (peak area versus concentration) for each of the various carbohydrates and acids analysed for calculation of supernatant composition and component concentration.

# Annexe III

## Determination of Phosphate Concentration

The phosphate concentration of culture supernatants was determined by a method first described in[40]. Here molybdenum complexes with any phosphate present to form a phospho-molybdate which is then reduced by ascorbic acid to give a blue colour.

Protocol:

Reagents Stocks (store for one month):

- 10 % ascorbic acid: 10 g ascorbic acid in 100mL $H_2O$, store at 4 °C
- 2.5 % ammonium molybdate: 2.5 g ammonium molybdate in 100mL $H_2O$, store at room temperature
- 6 N sulphuric acid: 18 mL concentrated acid in 108 mL $H_2O$ and store at room temperature

Reagent C, prepare fresh each day, keep at 4°C until use:

- 2 volumes $H_2O$
- 1 volume 6 N sulphuric acid
- 1 volume 2.5 % ammonium molybdate
- 1 volume 10 % ascorbic acid

Assay:

- Dilute culture supernatant samples 1 in 2000 in $H_2O$.
- Add 500 µL reagent C and 500 µL diluted samples to a cuvette. Cover with parafilm and invert several times to mix.
- For blank, mix 500 µL reagent C with 500 µL $H_2O$.
- Also include control samples with known molarities of phosphate ranging from 1 µM to 200 µM. These standards will be used to create a standard curve that plots absorbance vs. molarity for calculation of the phosphate concentrations of samples.
- Incubate samples for 2 hours at room temperature in the dark.
- Measure absorbance at 820 nm.
- Determine the phosphate concentration from the standard curve.

# Annexe IV

## Ammonia-nitrogen Concentration Determination

Ammonia-nitrogen concentration of culture supernatant was determined by the Berthelot colour reaction[47]. This method is based on the reaction of $NH_3$ in alkaline solution with phenate to produce a blue colour (indole blue) in the presence of a strong oxidizing agent such as hypochlorite.

Protocol:

All steps should be executed in the extraction hood, including the absorbance reading if possible. Wear gloves and dispose of all samples and cuvettes appropriately.

| Solution A: | *Reagent* | *Final Concentration* |
|---|---|---|
| | Phenol | 10 mg/mL |
| | Sodium Nitroprusside | 50 µL/mL |
| Solution B: | *Reagent* | *Final Concentration* |
| | NaOH | 10 mg/mL |
| | Sodium Hypochlorite | 0.84 % (v/v) |

- Prepare $NH_4$ standards of 1 mM, 2.5 mM, 5 mM and 10 mM with $(NH_4)_2SO_4$.
- Dilute samples 10 fold.
- Add 600 µL of Solution A to 30 µL of diluted sample (or standard) in disposable cuvettes. Use 30 µL of deionised water for blank.
- Add 600 µL of Solution B.
- Incubate for 2 hours at room temperature.
- Measure absorbance at 625 nm, starting with the blank, standards and then the samples.
- Determine the Ammonia-nitrogen concentration from the standard curve.

**Phenol is highly toxic. Carry out all work in the extraction hood and immediately have residuals removed from the laboratory (arrange with appropriate person prior to commencing work). All residues and material in contact with phenol should be placed in a sealed container placed inside 1 to 2 bags. Gloves should be worn at all times and immediately removed, discarded (inside 2 sealed bags) and changed on contact with liquid. Leave extraction hood and laboratory fans on following work.**

# Annexe V

## SDS-PAGE

10 % SDS-PAGE was used for monitoring of SELP-59-A production. This technique separates proteins based on their size as they migrate across an acrylamide/bis-acrylamide matrix gel. Proteins are denatured and linearised by SDS and β-mercaptoethanol treatment and imparted with a net negative charge by the SDS before being separated on the gel by use of an applied electric field. Separated protein bands are then visualised by negative staining with copper chloride.

Protocol:

Gel preparation (2 gels):

| **10% Page** | **Running** | **Stacking** |
|---|---|---|
| **Acrilamide 40%** | 2.7 mL | 432 µL |
| **Bis-acrilamide 2%** | 1.5 mL | 234 µL |
| **0,25 M Tris-HCl, pH 6,8 + SDS 0,2%** | 0 | 2.2 mL |
| **0,75 M Tris-HCl, pH 8,8 + SDS 0,2%** | 5.6 mL | 0 |
| **$H_20$** | 1.3 mL | 1.5 mL |
| **APS 10%** | 60 µL | 25 µL |
| **TEMED** | 10 µL | 8 µL |

Sample Preparation:

*Intracellular proteins, cell pellet from 500µL culture sample:*

- Add 50 mM Tris + 1 mM EDTA, pH 8.0 to a final volume of 1 ml and resuspend.
- Add 25 µL SDS-PAGE loading solution (10 % SDS, 10 mM β-mercaptoethanol, 20 % glycerol, 0.2 M Tris at pH 6.8 and 0.05% bromophenol blue) and vortex well.
- Centrifuge at max speed for 25 min.
- Run 10 µL of supernatant on a 10 % SDS-PAGE gel at a constant current flow of 25 amps per gel.

*Extracellular Proteins, Culture Supernatants*:

- Add 5 µL of SDS-PAGE loading solution (10 % SDS, 10 mM β-mercaptoethanol, 20 % glycerol, 0.2 M Tris at pH 6.8 and 0.05% bromophenol blue) to 20 µL of supernatant
- Run 5 µL of supernatant on a 10 % SDS-PAGE gel at a constant current flow of 25 amps per gel.

Gel Staining

- Stain gels with 3 M solution of freshly prepared $CuCl_2$, 50 mL per gel.
- Immediately mix at room temperature for approx. 5 - 10 minutes.
- Capture image with ChemiDoc$^R$ XRS+ system (BioRad).
- Analyse band intensities with ImageJ$^R$.

# Annexe VI

## SELP-59-A Purification

The protocol for SELP-59-A purification was described by[24] and is based on the high stability of SELP-59-A at acidic pHs as compared to *E. coli* proteins as well as its precipitation at relatively low ammonium sulphate concentrations.

Protocol:

- Add 2 volumes of 50 mM Tris + 1 mM EDTA, pH 8.0 to 1 volume (frozen wet weight) of frozen cell pellets. Resuspend by agitation overnight.
- Reduce suspension pH to 3.5 with 37 % HCL.
- Incubate at 4 ºC overnight with constant mixing.
- Centrifuge at 9000 rpm/30 min./4 ºC
- Measure supernatant volume, retain at 4 ºC.
- Add ammonium sulphate slowly to a final concentration of 22 % saturation (calculate using: http://www.encorbio.com/protocols/AM-SO4.htm) with constant mixing at 4 ºC. Leave overnight at 4 ºC with constant mixing.
- Remove large insoluble protein particles from solution.
- Centrifuge solution at 9000 rpm/30 min./4ºC.
- Resuspend pellets and insoluble particles in $H_2O$ to ~ 5% saturation at 4 ºC
- Incubate overnight at 4 ºC with agitation
- Dialyse suspension in water with 12000-14000 kDa MWCO dialysis tubing for 5 days at 4 ºC with constant mixing and with daily changes of water.
- Centrifuge dialysed suspension to separate soluble and insoluble fractions.
- Lyophilise soluble and insoluble fractions in a Christ Alpha 2-4 LD Plus (Bioblock Scientific) lyophiliser for 5 days at 0.012 mbar pressure and -85 ºC.

# Annexe VII

## Plasmid Isolation

Plasmid isolation from *Escherichia coli* was carried out with the GenElute™ Plasmid miniprep kit (Sigma). This kit provides a simple method based on DNA affinity to silica and the convenience of a spin column format for isolating plasmid DNA. Following cell lysis by alkaline-SDS treatment, DNA is selectively bound to the silica column in the presence of high salt concentrations and then, following column washing, the purified DNA is eluted with a solution of low salt concentration.

Protocol:

### Experienced User Protocol

*All spins at ≥12,000 × g, except as noted.*

**1 Harvest & lyse bacteria**

- ❒ Pellet cells from 1–5 ml overnight culture *1 minute* (1 ml from TB or 2xYT; 1–5 ml from LB medium). Discard supernatant.
- ❒ Resuspend cells in 200 µl Resuspension Solution. Pipette up and down or vortex.
- ❒ Add 200 µl of Lysis Solution. Invert gently to mix. Do not vortex. Allow to clear for ≤5 minutes
- * *Prior to first time use, be sure to add the RNase A to the Resuspension Solution.*

**2 Prepare cleared lysate**

- ❒ Add 350 µl of Neutralization Solution (S3). Invert 4–6 times to mix.
- ❒ *Pellet debris 10 minutes at max speed.*

**3 Prepare binding column**

- ❒ Add 500 ml Column Preparation Solution to binding column in a collection tube.
- ❒ *Spin at ≥12,000 × g, 1 minute.* Discard flow-through.

**4 Bind plasmid DNA to column**

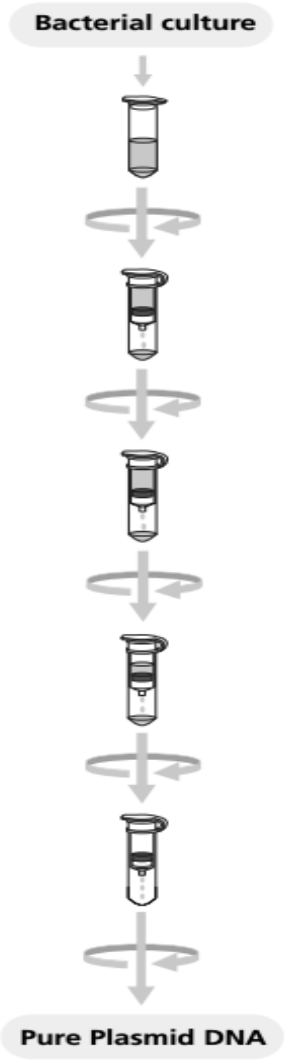
- ❒ Transfer cleared lysate into binding column.
- ❒ *Spin 30˝,1 minute.* Discard flow-through.

**5 Wash to remove contaminants**

- ❒ *Optional (EndA⁺ strains only):* Add 500 µl Optional Wash Solution to column. *Spin 30˝, 1 minute.* Discard flow-through.
- ❒ Add 750 µl Wash Solution to column. *Spin 30˝, 1 minute.* Discard flow-through.
- ❒ *Spin 1 minute* to dry column.
- * *Prior to first time use, be sure to add ethanol to the concentrated Wash Solution.*

**6 Elute purified plasmid DNA**

- ❒ Transfer column to new collection tube.
- ❒ Add 100 µl Elution Solution. *Spin 1 minute.*
- * *If a more concentrated plasmid DNA prep is required, reduce the elution volume to a minimum of 50 µl.*



# Annexe VIII

## DNA Quantification

The aromatic heterocyclic ring structures of DNA nucleobases absorb with a maximum near 260 nm (extinction coefficient used for double-stranded DNA = 50 ng-cm/µL), hence allowing for their quantification. Furthermore, proteins, phenol and other contaminants absorb strongly at 280 nm while phenol and carbohydrates absorb at 230 nm, hence measurements of the 260 nm/280 nm and 260 nm/230 nm absorbance ratios allows for analyses of sample purity. A 260 nm/280 nm ratio of about 1.8 is generally accepted as pure DNA. If the ratio is less than 1.8 it is indicative of the presence of contaminating proteins or other compounds that absorb near 280 nm. 260 nm/230 nm absorbance ratio values between 2.0 - 2.2 can be considered as pure. Ratios considerably below these values indicate the presence of contaminants which absorb near 230 nm.

Here DNA was quantified by absorbance measurements with a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific).

Protocol:

- Connect the NanoDrop and select the option ‘nucleic acids’.
- Clean the sample loader with 2 µL of ultrapure water.
- Set blank with 1 µL of ultrapure water.
- Measure absorbances of 1 µL sample.
- Analyse absorbance ratios to determine presence of contaminants.

# Annexe IX

## Restriction Digestion

Restriction digestion is based on the use of specific enzymes which recognise and cut at specific DNA sequences thereby allowing for the controlled digestion of a DNA sequence. All enzymes and buffers were from Thermo Scientific.

Protocol:

*SELP-59-A*/pCM13, pET29a(+) and pStaby 1.2 digestion

- 1 – 1.5 µg of template in minimum volume possible (15 – 50 µL).
- 1.5 µL of *Blp*I (1 unit/µl).
- Tango Buffer at final concentration of 1x.
- Incubate at 37 ºC for 2 hours.
- 1.5 µL of *Nde*I (1 unit/µl).
- Tango Buffer at final concentration of 2x.
- Incubate at 37 ºC for 3 - 4 hours

*SELP-59-A*/pStaby digestion for *SELP-59-A*/pStaby Δ*blap* preparation

- 20 µL of *SELP-59-A*/pStaby (56 ng/µL)
- 1 µL of *Ssp*I (1 unit/µL)
- 2 µL of 10x Fast Digest Buffer
- Incubate at 37 ºC for 2 hours

*SELP-59-A*/pStaby digestion for *SELP-59-A*/pStaby Δ*bla* preparation

- 20 µL of pStaby/SELP3 (56 ng/µL)
- 1 µL of *Ssp*I (1 unit/µL)
- 1 µL of *Eam*1105I (1 unit/µL)
- 2.5 µL of 10x Fast Digest Buffer
- Incubate at 37 ºC for 2 hours

# Annexe X

## DNA Gel Electrophoresis

DNA gel electrophoresis allows for the separation of DNA fragments according to their size as they migrate accross an agarose matrix towards the positive anode of an applied electric field. This method is used to separate, visualise and determine the size of various DNA fragments. Following electrophoresis, the gel is coloured with Midori Green which interacts with the DNA present and allows for their visualisation with UV light.

Protocol:

- Prepare 1.5 % agarose gel (1.5 g of agarose in 100 mL of 1x TEA Buffer (4.84 g/L Tris Base, 1.142 mL Glacial Acetic Acid and 0.372 g/L EDTA)), heat to dissolve, pour into electorophoresis cassette and allow to polymerise.
- Load 10 - 20 µL of sample with 6x loading buffer (50 % glycerol, 0.2 M EDTA at pH 8.3, 0.05 % (w/v) bromophenol blue).
- Load 3 – 4 µL of 1 kb Plus DNA ladder (Thermo Scientific)
- Run the gel at 75 V for 45 min.
- Stain the gel with Midori Green (1 µL/10 mL).
- View DNA bands under a UV illuminator.

# Annexe XI

## DNA purification by Gel Extraction

Once the appropriate DNA fragments have been separated and identified on the agarose gel they can then be purified by use of the *Ezway™ Gel Extraction Kit*. This kit is based on DNA affinity to silica and the convenience of a spin column format for purifying the DNA fragments from excised agarose gel bands.

Protocol:

**7. Principle**

After addition of Extraction Buffer to agarose gel slice sample and heating at 50-55°C to melt gel, DNA is bound to a silica gel membrane in the presence of high concentration of chaotropic salt. The purified DNA product on the membrane is eluted in elution buffer.


**8. Protocol**

1. Excise the agarose gel containing relevant DNA fragments with a clean scalpel. Remove the extra agarose gel to minimize the size of the gel slice.
2. Transfer up to 200 mg of the gel slice into a 1.5 ml microcentrifuge tube (not provided). The maximum volume of the gel slice is 200 mg.
3. Add 3 volume of Extraction Buffer to the sample and mix by vortexing.
   For example, add 300 ul of Extraction Buffer to 100mg of gel.
   For >2% agarose gels, add 6 volume of Extraction Buffer.
4. Incubate at 55°C for 10-15 minutes and vortex the tube every 3 min until the gel slice dissolved completely.

   **Note:** During incubation, interval vortex can accelerate the gel dissolved. Make sure that the gel slice has been dissolved completely before proceed the next step.
5. Cool down the sample mixture and place a DNA-prep column into a Collection Tube. Transfer to DNA-prep Column. Centrifuge for 1 min then discard the flow-through. If the sample mixture is more than 850 ul, repeat the step 5 for the rest sample mixture.
6. Add 750 ul of Wash Buffer (ethanol added) to the DNA-prep Column. Centrifuge for 1 min, then discard the flow-through.

   **Note:** Make sure that ethanol (96-100%) has been added into Wash Buffer when first open.
7. Centrifuge for an additional 3 minutes to dry the column.

   **Note:** Important step. This step will avoid the residual liquid to inhibit subsequent enzymatic reactions.
8. Place the DNA-prep Column into Elution Tube (provided).
9. Add 40 ul of Elution Buffer or $ddH_2O$ (pH7.0 - 8.5) to the membrane center of the DNA-prep Column. Stand the DNA-prep Column for 2 min.

   **Note:** Important step. For effective elution, make sure that the elution solution is dispensed onto the membrane center and is absorbed completely.
10. Centrifuge for 1 min to elute the DNA.

# Annexe XII

## Preparing of blunt ended fragments

Sticky ended DNA fragments produced by two different restriction digestion enzymes cannot be ligated together and the conversion of these to blunt ends is required. Here the 5' to 3' polymerase activity and 3´to 5´exonuclease activity in addition to an absence of a 5' to 3' exonuclease activity of a thermostable DNA polymerase is used for removal of 3' overhangs and filling in of 5' overhangs and thereby leading to a blunt ended DNA fragment.

The commercial mix *ACCUZYME*™ Mix (Bioline) was used here, this is a 2x reaction mix designed for high fidelity and maximum experiment reproducibility. ACCUZYME Mix contains Accuzyme DNA Polymerase, $MgCl_2$ and ultra-pure dNTPs manufactured.

Protocol:

- 30 µL of purified *SELP-59-A*/pStaby *Δbla* digest fragment (19.2 ng/mL).
- 30 µL of 2X concentrated $Mg^{2+}$ (4 mM) Accuzyme DNA polymerase mix (Bioline Ltd.)
- Heat to 72 ºC for 30 min with a *Thermal Cycler*$^{(R)}$ (Biometra Uno II)

# Annexe XIII

## Ligation

Ligation of overlapping sticky and blunt ends was carried out with T4 DNA ligase. This catalyses the ligation of 5' phosphate ends with 3' hydroxyl group ends.

Protocol:

Insertion of *SELP-59-A* in pET29a(+)

- 12 µL of purified insert (*SELP-59-A*) (11 ng/µL).
- 10 µL of purified pET29a(+) (11.1 ng/µL).
- 2.5 µL of 10x T4 DNA Buffer (Thermo Scientific).
- 1 µL of T4 DNA ligase (Thermo Scientific).
- Incubate at room temperature for 4 hours.

Insertion of *SELP-59-A* in pStaby 1.2

- 12 µL of purified insert (*SELP-59-A*) (11 ng/µL).
- 6 µL of purified pStaby 1.2 (21.3 ng/µL).
- 2.5 µL of 10x T4 Buffer (Thermo Scientific).
- 1 µL of T4 DNA ligase (Thermo Scientific).
- 3.5 µL of UP H2O.
- Incubate at room temperature for 4 hours.

Recircularisation of *SELP-59-A*/pStaby Δ*blap*

- 12.5 µL of purified *SELP-59-A*/pStaby Δ*blap* fragment.
- 1 µL of T4 DNA ligase (Thermo Scientific).
- 1 µL of T4 Buffer (Thermo Scientific).
- Incubate at room temperature for 4 hours.

Recircularisation of *SELP-59-A*/pStaby Δ*bla*

- 8 µL of *SELP-59-A*/pStaby Δ*bla* fragment.
- 1 µL of T4 DNA ligase (Thermo Scientific).
- 1 µL of T4 Buffer (Thermo Scientific).
- Incubate at room temperature for 4 hours.

# Annexe XIV

## Transformation

Transformation of *E.coli* strains XL1Blue and BL21(DE3):

Transformation was made with an adaptation of a commonly utilized protocol from Inoue and co-workers[48] Here the competent cells are made susceptible to uptake of DNA by a thermal shock treatment[48].

Protocol:

- Defreeze 200 µL of competent cells on ice.
- Add approximately 100 ng of circular DNA to each tube.
- Leave tubes on ice for 30 minutes.
- Heat shock for 45 seconds at 42 ºC with gentle agitation of tubes.
- Leave tubes on ice for 10 minutes and add 800 µL of preheated LB.
- Incubate for 1 hour at 37 ºC with 200 rpm agitation.
- Centrifuge the mixture at 14500 rpm for 1 minute.
- Reject 800 µL of the supernatant and ressuspend the remaining 200 µL of culture.
- Plate cells on solid LB medium supplemented with ampicillin (100 µg/mL) and incubate over night at 37 ºC.

Transformation of *E.coli* strains CYS21 (cloning strain) and SE1 (expression strain)

Here the protocol suggested in the *StabyExpress™ T7 kit Manual (v1.7)* was used. This protocol is also based on a heat shock treatment of competent cells for plasmid uptake.

Protocol: Transformation using chemically competent cells:

- Prepare LB plates containing 100 µg/mL Ampicillin. Let the plates dry and then warm them up at 37 ºC.
- Set a water bath or a heating-bloc to 42 ºC
- Thaw (bring to room temperature) one vial of regeneration medium (provided by manufacturer) per cloning reaction.

- For each cloning reaction, place one vial of the CYS21/SE1 chemically-competent cells (self-standing tube with pink cap) on ice. Allow the cells to thaw on ice for 5-10 minutes.
- Add 5 µL of the ligation product to one vial of the CYS21/SE1 chemically competent cells (self-standing tube with pink cap). Stir gently to mix. Do not mix by pipetting up and down.
- Incubate on ice for 30 minutes.
- Heat-shock the bacteria by placing the vial at 42 ºC for 30 seconds without shaking.
- Immediately transfer the tubes to ice.
- Add 250µl of room-temperature regeneration medium (provided by manufacturer) and mix well.
- Spread immediately 10, 20 and 100 µL of the product (from previous step) on different pre-warmed plates. If you wish to have more clones, incubate the product (from previous step) at 37 ºC for one hour for regeneration of the bacteria before spreading 10, 20 and 100 µL on different pre-warmed plates.
- Incubate the plates overnight at 37 ºC.
- Pick about 10 colonies and culture them overnight in 10 mL of LB medium with or without ampicillin (100 µg/mL).

**Note:** The stabilization is now effective; the *ccdB* gene is activated. Consequently, the plasmid is stabilized in the CYS21/SE1 strain and no antibiotic is needed to select bacteria containing the plasmid. However, the ampicillin resistance is still available. The stabilization system will insure high yield of plasmid DNA.